Title
Essays On Risk, Uncertainty, and Information

Author
Lane, Gregory

Publication Date
2019

Peer reviewed|Thesis/dissertation

Essays On Risk, Uncertainty, and Information

By

Gregory V Lane

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Agricultural and Resource Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Elisabeth Sadoulet, Chair
Professor Jeremy Magruder
Professor Paul Gertler

Spring 2019

Abstract

Essays On Risk, Uncertainty, and Information

by

Gregory V Lane

Doctor of Philosophy in Agricultural and Resource Economics

University of California, Berkeley

Professor Elisabeth Sadoulet, Chair

This dissertation presents a three-part study in modern empirical development economics. In each part, I tackle an important economics issue that affects the developing world—agricultural risk, principle-agent problems in firms, and poor public service delivery. I use randomized control trails (RCT) to uncover the causal impact of interventions designed to relieve constraints and improve outcomes for stakeholders.

The first chapter evaluates a novel microfinance product that is designed to help small-holder farmers manage uninsured agricultural risk. I run a large-scale RCT involving 300,000 microfinance clients in Bangladesh with one of the country's largest microcredit institutions. Microfinance clients were randomly given pre-approval to take a loan if they experienced flooding in their local area. I show that this unique type of microcredit improves household welfare through two channels: an ex-ante insurance effect, where households increase investment in risky production, and an ex-post effect, where households are better able to maintain consumption and asset levels after a shock. I also document that households value this product, taking costly action to preserve their guaranteed access. Importantly, the extension of this additional credit improves loan repayment rates and MFI profitability, suggesting that this product can be sustainably extended to households already connected to microcredit networks.

The second chapter examines the problem of moral hazard in employer–employee contracting and how this may be an important barrier to firm efficiency and growth in the developing world. To do so, we run an RCT with a fleet of 255 minibuses (matatus) in Nairobi, Kenya, where we introduce monitoring devices that track real-time vehicle location, daily productivity, and safety statistics. We randomize whether minibus owners have access to these monitoring data using a novel mobile app. This information allows owners in the treatment group to observe a more precise signal of driver effort, the amount of revenue drivers collected in fares, and the extent to which the driver engages in reckless driving. We find that treated vehicle owners modify the terms of the contract by decreasing the rental price they demand. Drivers respond by working more hours, decreasing behaviors that damage the vehicle, and under-reporting revenue by less. These changes improve

firm profits and reduce management costs, thereby helping treated firms grow. The device also improves owners' trust in their drivers, which drivers say makes their job easier. Finally, we investigate whether these gains to the company come at the expense of passenger safety, in an environment where accidents are common. While we do not find any evidence that conditions deteriorate, offering detailed information on driving behavior also does not *improve* safety. Only by incentivizing drivers through an additional cash treatment do we detect safety improvements.

Finally, the third chapter investigates the problem of poor road safety that plagues the developing world. Road traffic accidents are rapidly increasing, claiming more than 1.35 million lives annually and causing up to 50 million injuries. However, little is known about effective ways to improve road traffic safety in contexts where policing and government capacity are limited. To address this problem, we use an RCT to examine whether dissemination of safety information to passengers of public transit vehicles can reduce unsafe driving. We find that passengers are not more likely to choose a safe bus after receiving safety information, but are willing to switch their bus choice in response to monetary incentives. However, there is some evidence that passengers who care about safety are more likely to choose a bus they *already* believed was safe ex-ante, suggesting that the failure to respond to safety information may be due to skepticism about the quality of safety information provided. This emphasizes the importance of disseminating information through channels that every-day passengers are familiar with and trust.

Jointly, the three parts of this dissertation aim to apply careful research strategies, in conjunction with economic theory, to provide policy-relevant evidence towards improving outcomes in the developing world.

Contents

F | Figures

T | Tables

Acknowledgements

Coauthor, data, and funding Chapter 2 and Chapter 3 of this dissertation are coauthored with Erin Kelley and David Schönholzer. I thank BRAC Bangladesh, Echo Mobile, and the amazing teams at each that made this research possible. I also thank and acknowledge PEDL, the Weiss Foundation, and BASIS for funding these research projects. All errors are my own.

Personal thanks I am immensely grateful to Elisabeth Sadoulet and Jeremy Magruder for their guidance and kindness throughout my entire time at UC Berkeley. I thank Alain de Janvry, Paul Gertler, Aprajit Mahajan, Edward Miguel, and Fred Finan for their time, insights, and encouragement. Thank you all.

1 | Credit Lines as Insurance: Evidence from Bangladesh

Chapter abstract: In the absence of insurance markets, theory suggests that households can use credit to protect themselves against adverse income shocks. However, in many developing countries, access to credit in the aftermath of shocks is scarce as negatively affected households are frequently denied loans. In this paper, I test whether a new financial product that offers guaranteed credit access after a shock allows households to insure themselves against risk. To this end, I run a large-scale RCT involving 300,000 subjects in Bangladesh with one of the country's largest microcredit institutions. Microfinance clients were randomly pre-approved for loans that are made available in the event of local flooding. I show that this unique type of microcredit improves household welfare through two channels: an ex-ante insurance effect, where households increase investment in risky production, and an ex-post effect, where households are better able to maintain consumption and asset levels after a shock. I also document that households value this product, taking costly action to preserve their guaranteed access. Importantly, the extension of this additional credit improves loan repayment rates and MFI profitability, suggesting that this product can be sustainably extended to households already connected to microcredit networks.

1.1 Introduction

Poor households throughout the developing world struggle with income risk. Such households primarily rely on agriculture and small business profits, which are vulnerable to shocks such as harvest failure after extreme weather events, price volatility, and a sudden death in the family (Dercon, 2002). Research has demonstrated that these high levels of income variability prevent households from accumulating wealth and exiting poverty. Moreover, the set of risk coping and mitigation strategies that are available to households can often leave them worse off in the long run. Many turn to low-risk production technologies and under-invest in inputs, which negatively affects their future returns (Donovan, 2016). Others are often forced to implement damaging coping strategies such as lowering food consumption, selling productive assets, and reducing health and educational investments (Hoddinott, 2006; Janzen and Carter, 2018). Traditional insurance markets are designed to help households cope with these risks, but they are often absent or incomplete in developing countries because of moral hazard and adverse selection, while alternative tools such as index insurance have been hampered by low demand (Jensen and Barrett, 2017; Cole and Xiong, 2017).

Theory suggests that a realistic alternative to these tools is to provide households with a credit line so they can self-insure. Credit and savings models have long highlighted the precautionary value of credit access, which can serve to insure households against income fluctuations (Deaton, 1991, 1992). However, there is little empirical evidence demonstrating that households can in fact use credit access in this way. What evidence does exist comes from developed countries (Gross and Souleles, 2002), even though the benefits of credit access will likely be larger in developing countries where insurance markets are lacking. In developing countries, the largest providers of formal credit to the poor are microfinance institutions (MFI), which severely curtail credit access in the aftermath of large aggregate income shocks (Demont, 2014). MFI's requirements that households be financially evaluated at the time of loan disbursal, and that households cannot borrow if they have any outstanding debt, severely restricts households' use of credit access as a buffer against risk in developing countries.[1]

This paper provides the first empirical evidence that guaranteed access to credit after negative shocks increases productive investment, improves households' welfare, and ultimately, is profitable for the MFI. To this end, I partner with BRAC, a large MFI in Bangladesh, and extend a *guaranteed* credit line to poor, rural households. The product, marketed as an Emergency Loan, is a pre-approved loan that is made available to clients when an aggregate local shock (in this case a flood) occurs. I randomize the availability of the Emergency Loan across 200 rural BRAC microfinance branches serving over 300,000 clients with over one million loans during the study period. Clients in the 100 treatment branches were informed, before the beginning of the planting season, that they were pre-approved to take the new loan product should a flood shock occur in their area.[2] Control branches continued their normal microfinance operations. Loans were then extended upon request by eligible households after a flood had occurred (and been externally validated).

The experiment documents four primary results: First, I find that households value access to guaranteed credit and respond as theory would predict – indeed, some households are willing to forgo credit in the pre-period in order to preserve access to the state-contingent Emergency Loan, suggesting that at least a subset of clients value the precautionary benefits of credit access. Rough estimates suggest that these households value credit access after a shock at least 1.8 times more than credit access in the pre-period.

[1]Many MFIs have a dual mission of profit making and increasing social welfare, which theoretically makes them more likely to extend credit to households after an income shock. However, the field staff responsible for approving individual loans are almost always evaluated primarily on repayment metrics, making them likely to avoid lending to risky households.

[2]The Emergency Loan was offered to approximately 40% of clients within each treatment brach based on an individual credit score; see section 1.2 for more details.

Second, I find that informing households that they are pre-approved for credit in the event of a flood is associated with a significant rise in risky investments. Treated households increase the amount of land dedicated to agricultural cultivation by 15% and increase non-agriculture business investments. Both of these effects are concentrated among the most risk-averse households. These findings suggest that households view guaranteed liquidity access as reducing their exposure to flood risk, and respond by increasing their investment in riskier, potentially more profitable investments.

Third, I document that emergency credit, unlike many other microcredit products, improves household welfare outcomes. When there is no flood, the larger ex-ante investments translate into higher revenues. When flooding does occur, households are better able to maintain consumption and asset levels. Furthermore, we find that the most severely affected households were the most likely to use this additional liquidity. This finding means that the largest gains associated with guaranteed credit could be concentrated among those who need it the most.

Finally, I find that extending guaranteed credit to clients in the aftermath of shocks does not harm (and marginally improves) overall MFI performance. Borrowers with access to the Emergency Loan improve their overall repayment rate, driven by improvements in repayment rates after a flood shock. Overall, evidence suggests that branch profits increase, with the largest increases in profits coming from “marginal" clients. This result is encouraging for MFIs that have traditionally withheld credit in the aftermath of aggregate shocks. Nevertheless, it is worth highlighting that these results may not generalize to contexts where repayments rates are already low.

The provision of guaranteed credit lines combines aspects of traditional microcredit and insurance products, both of which have been extensively studied in developing countries. The provision of traditional (loss-indemnity) insurance is almost completely absent among low-income households due to high administrative costs, adverse selection, and moral hazard (Jensen and Barrett, 2017). In recent years, index insurance has been promoted as a viable alternative. By linking payouts to easily measurable and exogenous indices, such as rainfall, index insurance removes moral hazard concerns and reduces the need to collect additional data on household-specific losses. Index insurance has been found to generate positive results by inducing more investment in agricultural production and reducing the sale of assets after shocks (Karlan et al., 2014; Janzen and Carter, 2018). Despite these benefits, demand for index insurance remains very low across many developing countries when offered without heavy subsidies (Cole and Xiong, 2017). Low demand appears to be linked to the requirement that insurance payments be collected ex-ante, which can be difficult for households that are credit constrained, are present-biased, face basis risk that the index will not correspond to their own personal shock, and lack trust in their insurers’ ability to pay out when the time comes (Cole et al., 2013; Clarke, 2016). In recent work, Serfilippi, Carter, and Guirkinger (2018) show that preferences for certainty drive down demand for insurance contracts where premiums are always paid but payouts are uncertain. In some contexts, low demand can be overcome by allowing the upfront insurance premium to be paid after harvest. However, this solution is only feasible when there is the possibility of an interlinked transaction; specifically, this can take the form of a monopsony buyer that can credibly (and cheaply) collect payment from farmers after the fact, as in Casaburi and Willis (2018), or by tying insurance payments to credit contracts, as in McIntosh, Sarris, and Papadopoulos (2013).

This research demonstrates that emergency credit can function as a viable alternative to insurance products while offering several key advantages. Specifically, the Emergency Loan overcomes the challenges associated with the timing of insurance payments while maintaining many of the positive features associated with index insurance. Like index insurance, the availability of the additional credit is contingent on an exogenous indicator (floodwater height) to avoid high administrative costs and moral hazard. However, unlike index insurance, no purchase (or any binding decision) is required by the household during the planting season (which is similar to the innovation explored in Casaburi and Willis (2018)). Providing coverage under a guaranteed credit scheme simply requires notifying a household that they are eligible for the product. As long as a household understands the offer and trust that it will be executed if needed, the household is “treated." This feature ensures that credit-constrained or present-biased households that stand to benefit from the product will not be deterred from adopting it. As a result, guaranteed credit lines have the ability to provide coverage to a large number of households that might not otherwise choose to purchase insurance. Critically, households can benefit from the security of the credit line even if they choose *not* to take a loan after a shock. This arises because the decision to take credit is postponed until after uncertainty has resolved, which means households can opt in or out depending on realized damages from the shock and any alternatives that may be available. Indeed, I see in my experiment that many households increase ex-ante investment, suggesting a reduction in perceived risk, even though ex-post take-up of the Emergency Loan is low.

There are, however, several limitations associated with using guaranteed credit as a risk management tool. As with insurance, households may be reluctant to rely on the product in times of need if they are concerned about default by the provider (a fact that is mitigated in this context by working with BRAC Bangladesh, a well-established and trusted MFI in the region). Unlike insurance, the sequence of shocks can have an impact on the usefulness of credit for income smoothing. If a household experiences multiple successive shocks under a guaranteed credit scheme, they may accumulate excess debt or exhaust their available credit line.[3] Finally, extending credit to households after a shock is inherently risky for MFIs. While I find good repayment rates in this setting, if repayment rates are lower elsewhere, providing guaranteed credit may not be sustainable from the lender’s perspective. It follows that, while guaranteed credit provides clear advantages to some households, it may not be a panacea.

This research also contributes to the large literature on microcredit. Developed

[3]With insurance, a household that experiences several shocks in a row will simply receive the fixed insurance payout each period (provided they purchased the product).

in Bangladesh in the 1970s, microcredit institutions have since rapidly expanded, reaching over 137 million households worldwide (Maes and Reed, 2011). Unfortunately, despite this extensive growth and early enthusiasm for microcredit,[4] the majority of research shows only modest impacts on households' well-being (Karlan and Zinman, 2011; Angelucci, Karlan, and Zinman, 2015; Banerjee et al., 2015; Banerjee, Karlan, and Zinman, 2015). This observation may be partly attributable to the fact that microcredit only solves the problem of credit access, without remedying the underlying risks that prevent households from optimally investing (Karlan et al., 2014). Indeed, early microcredit products typically featured group lending with joint liability and frequent, rigid repayment schedules designed to overcome high transaction costs and asymmetric information; however, such characteristics come at the cost of making repayment difficult for those with uncertain income (Karlan, 2014). In response to these results, a line of research has focused on easing these constraints by matching repayment schedules to borrowers' cash flows. Field and Pande (2010) and Field et al. (2013) show that reducing payment frequency and delaying the start of installment payments reduce borrower transaction costs and encourage greater investments and profits. Similarly, Beaman et al. (2014) study agricultural loans that allow repayments to come in a lump sum after harvest and find higher investments in the planting season, and Barboni (2017) shows that more productive borrowers opt into flexible repayment contracts even when they are more expensive. This paper builds on these results by showing that credit products that increase the flexibility of households' access to credit (not just repayment flexibility) can lead to important improvements in outcomes.

Lastly, additional research has focused on understanding how new credit products affect MFI profits. Field et al. (2013) develop a structural model to show that longer grace periods are not sustainable for MFIs due to adverse selection and moral hazard concerns. In contrast, Barboni (2017) uses theory and lab-in-the-field experiments to show that offering flexible repayment schedules could increase profits for lenders. An advantage of our relatively large experiment is that it allows for an *empirical* examination of the effects of this new product on overall MFI profitability, which is difficult in settings where risk-averse MFIs are hesitant to experiment. In this setting, I find significant positive effects on MFI profits, with significant heterogeneity among borrowers.

The rest of the paper is organized as follows: Section 1.2 describes the context of the experiment and describes the new credit product in detail. Section 1.3 lays out a theoretical framework which provides predictions. Section 1.4 describes the main research design and execution of the experiment and section 1.5 describes the data used in the analysis. Finally, section 1.6 presents the results of the experiment and section 1.7 concludes.

[4]In 2006, Mohammad Yunus and the Grameen Bank, which he founded, were awarded the Nobel Prize for Peace.

1.2 Context and Product Description

Bangladesh and Income risk

This project takes place in Bangladesh, a country with over 165 million people that is covered by the Bengal delta (a confluence of the Ganges, the Brahmaputra and the Megna rivers). Approximately 70 percent of Bangladesh's population lives in rural areas and more than 80 percent of rural households rely on agriculture for some part of their income (World Bank, 2016). While the country's economy has grown rapidly in recent years, GDP per capita still stands at $2,363 and approximately 43% of the population earns less than $1.25 per day (UNDP, 2015).

Many types of extreme weather events are frequent, and are projected to worsen with the advent of climate change. Approximately 80% of the country is located on floodplains, and floods occur yearly with varying degrees of severity (Brammer, 1990). Moreover, recent projections estimate that flood areas could increase by as much as 29% in Bangladesh (World Bank, 2016). Therefore, the experiment focuses on flood risk over other shocks because of the high frequency of flooding across the entire country. As such, the randomized control trial was conducted in areas located close to the major rivers where frequent flooding occurs. In these areas, most productive investments are exposed to flooding risk. Due to the risk of crop failure, investments made by the household in agriculture - such as renting land for cultivation, using synthetic fertilizer, or purchasing improved seeds - offer greater upside potential but also increase losses in the event of a flood. Furthermore, even investments made in non-agriculture businesses are exposed to flooding risk. This is due to the fact that physical businesses assets may be lost or damaged after flooding and that demand may fall after a flood shock because of the local economy's overall dependence on agriculture.

I work with a subset of households that are active microfinance borrowers. These households are primarily engaged in agriculture: 50% grow their own crops and 22% work as day laborers. This group is also active in starting their own businesses (27% reported owning a small shop). Education is low in these areas, and approximately two-thirds of the sample have less than a primary school education.

BRAC Microcredit

BRAC was founded in 1972 in Bangladesh and is currently one of the largest NGOs in the world. Their microfinance operations began in 1974 and have expanded to serve the entire country. They operate over 2000 branches where each branch serves anywhere from 20 to 60 village organizations (VO's). These organizations are designed to facilitate coordinated activities between borrowers at the village level, including the distribution of information about new micro-finance products and creating a convenient space for BRAC loan officers to collect loan payments and instill some social pressure on borrowers to make their loan payments. VO meetings occur either weekly or monthly depending on branch policy. At each meeting the loan officer collects the scheduled loan repayments from each active borrower and answers enquiries about desired new loans from members without existing debt.

The most common loan provided by BRAC is called the *Dabi* loan, which is only given to women and is targeted at poor households[5]. Dabi loans are typically small in value (approximately 15,000 taka ($187)), and are required to be repaid within a year. Microfinance interest rates are regulated in Bangladesh and BRAC charges 25% interest on the Dabi loan, which is near the legal maximum and similar to other MFI's. During the repayment period, borrowers are not allowed to apply for any other BRAC loans, and are discouraged from taking any additional loans from other microfinance institutions or local money lenders. There is, however, one exception. Clients who make every loan payment on-time for the first six months of their loan cycle are eligible to take a top-up loan called the "Good Loan". The Good Loan is capped at 50% of the principal amount of the currently held loan. The offer is only available for two months after they become eligible at the 6 month mark of their current loan cycle.[6] In every other respect, Good Loans are identical to normal Dabi loans, with the same 25% interest rate and one year repayment timeline. Taking the Good Loan does not delay the normal loan cycle, and the client can take another normal Dabi loan as soon as she has repaid her old one.

Product Description

The Emergency Loan was designed together with BRAC to improve its utility for borrowers exposed to flood risk while also limiting BRAC's exposure to risky loans. Clients were eligible to access the Emergency Loan provided they had a credit score above a fixed threshold. The credit score was created specifically for this product, and was calculated from each borrower's past repayment behavior on previously held BRAC loans. Specifically, the score was based on four metrics: past percentage of missed payments, average percent behind on loan payments, maximum percent behind on any loan, and the number of months as an active BRAC microfinance member. Each variable received a linear weight determined by a regression of these variables on a binary indicator for loan default. This weighted sum was then normalized to a 0-100 scale. The variables themselves were chosen based on several criteria, including a) ease of calculation due to record keeping and computation limitations, b) relevance for predicting future default, and c) ease of explanation for transparency.[7] The threshold was set so that approximately 40% of borrowers were eligible at any given branch (77 out of a maximum score of 100). It is worth highlighting that targeting based on credit score does not select richer households

[5]While the Dabi loans are given only to women, it is common that these loans are used for broader household investments such as agriculture or a business that is run by the husband of the official borrower.

[6]Good Loans are also subject to the Loan Officer and Branch Manager approval (i.e. they can be denied even if the borrower is technically eligible)

[7]To determine relevance for predicting default, the complete set of possible variables was assessed in two historical training samples and then confirmed using more recent data. Only variables that were consistently predictive were kept in the final credit score. Additionally, linear regression was used rather than more complex techniques such as machine learning due to the desire to make the credit score transparent, and easily adjustable in the future.

over poorer ones. Table 1.1 examines differences in observables between the eligible and ineligible borrowers. The two groups look fairly similar, but differ along a few dimensions. Eligible borrowers have slightly less annual income, they are a few years older, have fewer years of education, and own more livestock and savings.

Client eligibility was assessed for every borrower in April, just before planting of the Aman season and several months before the flooding season. Borrowers could retain access to their Emergency Loan eligibility for the duration of the Aman cropping season regardless of their repayment behavior in the interim. Eligible clients were guaranteed to be able to borrow up to 50% of the total principal amount of their last regularly approved loan. For example, a borrower who took a 10,000 taka loan ($125) in May from BRAC was guaranteed to borrow up to 5,000 taka ($63) should a flood occur regardless of her existing loan balance. No further evaluations of the client's ability to repay, or any other checks, were conducted before disbursing the Emergency Loan. Emergency Loans were then made available to eligible clients if flooding occurred. This was validated in two ways. First, the river gauge associated with the branch area had to be reporting water level above the pre-determined danger level for at least one day.[8] Second, a non-microfinance BRAC employee had to confirm that at least 20% of the branch service area had experienced flooding.[9] On a case by case basis, loans were also made available if the local Branch Manager notified BRAC headquarters of local floods and this report was confirmed by the BRAC employee (even if the matched river water gauge had not passed the official danger level).

Two additional features of the Emergency Loan are important to review. First, the eligibility list created by the credit score was provided directly to branch managers, who could veto up to 10% of the names on the list based on their private knowledge of a borrower's credit worthiness.[10] The final list was then shared with BRAC headquarters. These steps were put in place to minimize the risk that BRAC would lend to borrowers that would fail to repay the loan.[11] For the purposes of the experimental results, I do not consider Branch Manager vetos and include all clients who were determined to be eligible based on the credit score alone.

Second, it is important to note how the Emergency Loan interacts with the existing Good Loan product. The Good Loan product differs from the Emergency Loan in the timing that is made available to clients (6-8 months into their normal Dabi rather than post-flood), and in how it is disbursed (by asking branch managers who can deny the request, instead of pre-approval based on credit scores). Looking through historical data, this means that the Good Loans were much less likely to be disbursed in the aftermath of aggregate income shocks because most borrowers were either not in their 6-8 month timeframe or branch managers did not want to

[8]The danger level is not the water height at which the river overflows it banks, but the height at which there is estimated to be a high probability of significant property damage in the area.

[9]This second check was deemed necessary after piloting showed that for some branch service areas, a higher river water level was necessary to cause any risk of flood damage.

[10]Branch managers in the control group performed this same veto process for a future product rather than the emergency loan itself.

[11]This is also a standard practice for every other loan.

approve additional top-up loans.

Clients in the sample could be *eligible* for the Good Loan or the Emergency Loan, both, or neither. However, borrowers were informed that they could not have both a Good Loan *and* an Emergency Loan together — if they take a Good Loan they lose future eligibility for the Emergency Loan should a flood occur.[12] This limitation was introduced because of BRAC's concerns that borrowers would carry too much debt. Therefore, clients who were *eligible* for the Emergency Loan and the Good Loan then faced a tradeoff: they could take the Good Loan before the flood season occurred and forgo the option of accessing additional liquidity in the event of a flood, or they could preserve their credit access as a buffer against future flood risk. Clients who had access to the Good Loan but not the Emergency Loan did not face this tradeoff. This creates an additional feature of the experiment that I can exploit. Namely, I can compare these two groups to determine whether households choose to preserve credit access as a buffer stock against risk. Figure 1.1 summarizes borrower choices related to the Good Loan and Emergency Loan.

1.3 Theory

Framework For Effect of Guaranteed Credit

Clients are informed about their eligibility for the Emergency Loan in April, before decisions need to be made on inputs for the coming Aman season (e.g. land to cultivate, inputs to use, business investments), and how much to borrow to finance these choices. After making these decisions, the cropping season commences and flooding either does or does not occur. If flooding does occur, each eligible borrower is informed that the Emergency Loan is available for them to access. During this period, borrowers make decisions on whether or not to take the Emergency Loan (if it is available), and whether or not to repay existing loans. Later, borrowers move into the dry season and choose to repay any loans taken after flooding.

Therefore, it is useful to categorize client decisions into three periods: choices made after being informed about their Emergency Loan eligibility but before the realization of any flooding (first period decisions), choices made after any flooding has occurred (second period decisions), and choices made in the dry season (third period decisions).

First Period Decisions

1. *Productive Investments:* Households decide how much to invest in production, whether in agricultural land and inputs, or in other business investment.
2. *Dabi Loan Uptake:* Each member will decide whether and how much they wish to borrow before the start of the Aman season.

[12] Of the 350,000 individuals in the data, approximately 165,000 (47%) were eligible for a Good Loan at some point during the experiment. Of these, 66,000 (40%) were also eligible for the Emergency Loan.

3. *Good Loan Uptake:* For members who are eligible to take a Good Loan, they will decide whether or not to take this additional credit to invest for the Aman season.

Second Period Decisions

1. *Emergency Loan Uptake:* In the event of a flood, borrowers will make the decision about whether to take an Emergency Loan.
2. *First Period Loan Repayment:* Once borrowers choose whether or not to take the Emergency Loan, they will need to decide how (or whether) to repay the loans they have.

Third Period Decisions

1. *Second Period Repayment:* Borrowers choose whether to repay the Emergency Loan if they took one in the second period. This decision will also depend on whether or not the household defaulted in the second period.

Below, I present a simple model that seeks to provide a framework for understanding how the extension of guaranteed credit could impact each of these decisions in turn.

Baseline Model

The model[13] has three periods that correspond to planting, harvest, and post-harvest periods, incorporates risky production and a credit market with constraints, and assumes that no insurance is available. For ease, I limit the harvest realization to two possible states, $s \in \{G, B\}$ that are realized in time period two and occur with probability π_s (later defined as $\pi_B = q$ and $\pi_G = (1-q)$). Further, I assume that the only source of credit available to a household comes from the MFI. Preferences are over consumption (c) with discount factor β:

$$u(c^1) + \beta \sum_{s \in G,B} \pi_s u(c_s^2) + \beta^2 \sum_{s \in G,B} \pi_s u(c_s^3)$$

A household starts with exogenous cash on hand Y and also has access to a risk free asset b^1 which it can buy (up to a limit) or sell on the market at interest rate R (therefore positive values of b represent net borrowing while negative values represent net saving). The household also has access to a concave production function $m_s f(x)$, which takes input x and provides output in the second period. The production function has a state dependent marginal product m_s which changes with the realized state s. In period two, the state of the world is resolved and the household decides whether to repay its initial loan with interest (Rb^1) or default by paying zero. I also allow for borrowing in the bad state of the world b_B^2, which is made available with the introduction of the Emergency Loan (to simplify the

[13] Based on a model from Karlan and Udry (2015)

problem, I do not allow savings from period two to three, but this assumption does not change the core results). In period three, the household pays (or receives) return R on any period two loans, provided they have not already defaulted, and also receive exogenous risk free income (I). Finally, households that default are penalized K, which is the household-specific loss in utility from losing access to future dealings with the MFI. The basic household problem can then be stated as:

$$\max_{x,b^1,b^2_B,D,ND}\{u(c^1)+\sum_{s\in G,B}\max\{\beta\pi_s u(c^2_s|ND)+\beta^2\pi_s u(c^3_s|ND),$$
$$\beta\pi_s u(c^2_s|D)+\beta^2\pi_s u(c^3_s|D)-K\}\} \quad s.t.$$

$$\begin{aligned}
c^1 &= Y - x + b^1\\
c^2_G &= \mathbb{1}\,[ND]\left[m_G f(x) - Rb^1\right] + \mathbb{1}\,[D]\,[m_G f(x)]\\
c^2_B &= \mathbb{1}\,[ND]\left[m_B f(x) - Rb^1 + b^2_B\right] + \mathbb{1}\,[D]\left[m_B f(x) + b^2_B\right]\\
c^3_G &= I\\
c^3_B &= \mathbb{1}\,[ND]\left[-Rb^2_B + I\right] + \mathbb{1}\,[D]\,[\,I\,]\\
x &\geq 0\\
b^1 &\leq \bar{B}_1\ ,\ (\lambda_1)\\
b^2_B &\leq \bar{B}_2\ ,\ (\lambda_2)
\end{aligned}$$

where D and ND stand for default and no default respectively, c^t_s and b^t_s are consumption and borrowing choice in the corresponding time period and state, x is inputs, Y is exogenous first period wealth, and I is exogenous third period income.

A household can borrow up to $\bar{B}_j$ in each period where borrowing is possible (where if $\bar{B}$ is equal to zero there would be no access to credit). To begin, I will assume $\bar{B}_2 = 0$, meaning there is no credit available in the bad state. I also make a few additional simplifying assumptions. First, I assume that it is never optimal for a household to default on its loan when the good state is realized ($s = G$). This assumption rules out households that always default and therefore took first period loans in bad faith. Second, I normalize the marginal product of x as zero in the bad state, i.e. $m_B = 0$[14].

The rest of this section is organized as follows. First, I start by separately describing the optimal borrowing and input choices assuming households do not default and then again assuming households will default in the event of a shock. Second, I compare these two scenarios and find the condition that will lead the

[14]Note that this normalization also implies a shift in the utility function such that the utility of a negative value does not imply zero or negative utility.

household to choose to repay or to default. Third, I allow for bad state borrowing and observe how the relaxation of this constraint changes household choices of inputs, borrowing, and the choice to default. Finally, given the expected effect on households' decisions, I examine the implications of extending bad state borrowing for the performance of the lending MFI.

No Default

In this section, I derive the optimal choice of first period input use and borrowing assuming that the borrower will not default in the event of a shock. The household's problem is:

$$\begin{aligned}\max_{x,b^1} \; & u(Y - x + b^1) + q\beta u\left(-Rb^1\right) + (1-q)\beta u\left(m_G f(x) - Rb^1\right) + \\ & q\beta^2 u(I) + (1-q)\beta^2 u(I) + \lambda_1[\bar{B}_1 - b^1]\end{aligned} \tag{1.1}$$

where λ_1 is the Lagrange multiplier associated with the first period borrowing constraint. Optimizing equation 1 implies that the input x is purchased until the following condition is satisfied:

$$m_G \frac{\partial f}{\partial x} = R\left[\frac{q}{1-q}\frac{u'(c_B^2)}{u'(c_G^2)} + 1\right] + \frac{\lambda_1}{\beta(1-q)u'(c_G^2)} \tag{1.2}$$

Under a scenario without risky production or credit constraints, the agent would invest in x until the marginal product equaled the return on the risk-free asset R. The equation above shows us that there are two potential sources of distortion away from that standard result. The first term in brackets above will be greater than one and reflects the presence of the risky production technology that has a zero marginal product in the event of the bad outcome. Second, the first period credit constraint could bind, in which case $\lambda_1 > 0$, which will also drive a wedge between the marginal product of the input and R. Therefore, both potential distortions will lower the choice of x relative to the unconstrained optimum.

Now, I move to examine the borrowing choice in period one. The first order condition implies that the first period borrowing is chosen such that:

$$u'(c^1) = \beta R\left[qu'(c_B^2) + (1-q)u'(c_G^2)\right] + \lambda_1 \tag{1.3}$$

Again, there are two potential distortions away from the optimum without risky production or credit constraints. First, the gap between second period consumption in the bad and good state ($qu(c_B^2)$ and $(1-q)u(c_G^2)$) will increase the size of the second term (due to concavity), and imply reduced consumption in period one relative to a choice without risky production, which, combined with the reduction in inputs purchased, implies an overall reduction in borrowing as well. Second, as before, if the first period borrowing constraint binds, λ_1 will be positive and will also imply a reduction in borrowing relative to the unconstrained optimum.

Default

In this section, I assume that the household will choose not to repay their period one loans if the bad state occurs in the second period. Under this assumption, the household problem changes to:

$$\max_{x,b^1} \; u(Y - x + b^1) + q\beta u(0) + (1-q)\beta u\left(m_G f(x) - Rb^1\right) + \\ q\beta^2 u(I) + (1-q)\beta^2 u(I) + \lambda_1[\bar{B}_1 - b^1] + \lambda_1[0 - b_B^2] \tag{1.4}$$

The fact that the household knows they will not repay their loans in the event of a shock changes the optimal use of inputs and borrowing in the first period. First, I can see that the optimal choice of inputs is defined by

$$m_G \frac{\partial f_G}{\partial x} = R + \frac{\lambda_1}{\beta(1-q)u'(c_G^2)} \tag{1.5}$$

This condition implies that households that know they will default in the bad state of the world will use inputs until the marginal return is equal to the interest rate R. The only distortion comes from the borrowing constraint in period one (λ_1). Similarly, borrowing will be chosen such that the only consideration is equalizing marginal utility in period one with discounted marginal utility in period two:

$$u'(c_1) = (1-q)\beta R u'(c_2^G) + \lambda_1 \tag{1.6}$$

Repayment Decision

To examine the borrower's repayment decision, I compare the utility for the household when they choose to repay to the utility they receive under default. If a household chooses to repay, their utility under repayment must be higher than their utility under default:

$$U^{repay} \geq U^{default}$$

which is given by:

$$\begin{aligned} & u(c_r^1) + q\beta u\left(-Rb_r^1\right) + (1-q)\beta u\left(m_G f(x_r) - Rb_r^1\right) + \\ & \quad q\beta^2 u(I) + (1-q)\beta^2 u(I) \\ & \geq \\ & u(c_d^1) + q\beta u(0) + (1-q)\beta u\left(m_G f(x_d) - Rb_d^1\right) + \\ & \quad q\beta^2 u(I) + (1-q)\beta^2 u(I) - qK \end{aligned} \tag{1.7}$$

where an index of d or r signifies the optimal value of the variable given repayment or default respectively. To understand this decision, I examine the switch point where a household is indifferent between repayment and default by setting these two expressions equal to each other. In order to declutter the expression, it is useful to define new terms. First I define M as the difference in utility between default and repayment in the first period and in the second period under the good state, where $M > 0$.[15]

Using this simplification and rearranging the initial condition, I can define K^*:

$$K^* = \frac{M}{q} + \beta\left[u(0) - u(-Rb_r^1)\right] \tag{1.8}$$

where K^* is the cost of lost access to microfinance that would make a household indifferent between repayment and default.[16] If a household's actual K is larger than K^*, they will repay; if it is lower, they will default. Therefore, if I assume that K is a random variable defined by the CDF F_K, the proportion of households that will default after a shock is given by $F_K(K^*)$.

Adding Liquidity in the Bad State

I now explore how the optimal choices of x and b^1 change when the option to borrow in the bad state in period two is added. Starting with the no-default case, the household's problem is now expanded to include the choice b_B^2:

$$\begin{aligned} \max_{x,b^1,b_B^2} \; & u(Y - x + b^1) + q\beta u\left(-Rb^1 + b_B^2\right) + (1-q)\beta u\left(m_G f(x) - Rb^1\right) + \\ & q\beta^2 u(I - Rb_B^2) + (1-q)\beta^2 u(I) + \lambda_1[\bar{B}_1 - b^1] + \lambda_2[\bar{B}_2 - b_B^2] \end{aligned} \tag{1.9}$$

In order to understand how the introduction of borrowing after a shock influences decisions, I assume that the first period borrowing constraint does not bind (i.e. $\lambda_1 = 0$), which allows first period choices of x and b^1 to adjust rather than being fixed at the constraint. Under this assumption, the optimal choice of x is determined by:

---

[15]

$$M = \underbrace{\left[u(c_d^1) - u(c_r^1)\right]}_{First\ Period} + \underbrace{\left[(1-q)\beta u\left(m_G f(x_d) - Rb_d^1\right) - (1-q)\beta u\left(m_G f(x_r) - Rb_r^1\right)\right]}_{Second\ Period\ Good\ State}$$

The difference in these terms is *only* due to the different optimal choices of x and b^1 in the first period, rather than the repayment (or non-repayment) of loans. Therefore, because I know that $x_d > x_r$ and $b_d^1 > b_r^1$, the utility received when a client defaults is higher than the repayment utility. Therefore $M > 0$.

[16] Note that K^* is monotonically increasing in b^1, implying the more indebted a household, the higher value of K necessary to ensure repayment.

$$m_G \frac{\partial f_G}{\partial x} = R\left[\frac{q}{1-q}\frac{u'(c_B^2)}{u'(c_G^2)} + 1\right] \tag{1.10}$$

Allowing borrowing after a shock in the second period will increase consumption in this state (c_B^2) relative to the constrained case. Thus, $u'(c_B^2)$ decreases as does the ratio $\frac{u'(c_B^2)}{u'(c_G^2)}$ in equation 10. This implies the entire RHS of the equation falls and that therefore that optimal first period input use will rise.[17]

I use a similar argument for first period borrowing, where the gap between $u'(c_B^2)$ and $u'(c_G^2)$ is reduced in equation 11 below, which causes the entire RHS of the equation to fall. This in turn implies an increase in period one consumption (and therefore an increase in borrowing).

$$u'(c^1) = \beta R\left[qu'(c_B^2) + (1-q)u'(c_G^2)\right] \tag{1.11}$$

Last, I examine what factors determine the choice of b_B^2. Because there is no uncertainty moving into the third period, the optimal choice of bad state borrowing is defined by the standard condition:

$$u'(c_B^2) = \beta R u'(c_B^3) + \lambda_2$$

Households will be more likely to borrow in the bad state if they have a particularly low value of c_B^2 or have a high value of c_B^3. Therefore, I would expect more demand for the Emergency Loan from households that are hit hardest by a flood shock and those that have high expected future income I.

Therefore, the model gives four predictions that result from extending a credit line in the bad state of the world:

1. Consumption increases after a shock
2. First period investment increases
3. First period borrowing increases
4. Probability of taking the Emergency Loan will be higher among households that experience heavy damage from flooding or those with good post-Aman income opportunities

If I consider the case of households that default after a shock, it is easy to see that only prediction 1 will carry through. These households will indeed choose to borrow in the bad state and therefore increase their consumption as they do not plan to repay the loan. However, because they already planned to default if a shock occurred, neither ex-ante input choice or first period borrowing will be impacted by changes in the level of c_B^2. Further, I can see that the optimal bad state borrowing

[17] Appendix A shows a more formal derivation of the comparative statics of x and b^1 with respect to b_B^2.

amount will always be to take the maximum allowed, $b_B^2 = \bar{B}_2$, as there is no cost of repayment when already under default.

Interaction with the Good Loan

I now consider the situation faced by clients who also have access to the Good Borrower loan. Without the Emergency Loan, these households solve the same baseline model as above, with the only difference being that their first period borrowing constraint is $1.5\bar{B}_1$.[18] However, with the introduction of the pre-approved Emergency Loan, which is mutually exclusive with the Good Loan, the problem facing these households changes. The borrowing constraints facing a household in this situation are:

$$\begin{aligned} b^1 &\leq 1.5\bar{B}\ , \\ b_B^2 &\leq 0.5\bar{B} \\ b^1 + b_B^2 &\leq 1.5\bar{B}\ , \end{aligned}$$

Now, any borrowing above $\bar{B}$ in the first period (i.e. using the Good Loan) comes at the expense of available liquidity after a shock. It is for borrowers in this position that the problem of credit line preservation becomes salient - households must now consider the value of preserving their credit line for a time of need, and whether or not this is worth forgoing current period investment. The constrained maximization problem changes to:

$$\begin{aligned} \max_{x, b^1, b_B^2} \ & u(Y - x + b^1) + q\beta u(-Rb^1 + b_B^2) + (1-q)\beta u(m_G f(x) - Rb^1) + \\ & q\beta^2 u(I - Rb_B^2) + (1-q)\beta^2 u(I) + \lambda_1[1.5\bar{B} - b^1] + \\ & \lambda_2[0.5\bar{B} - b_b^2] + \lambda_3[1.5\bar{B} - b^1 - b_B^2] \end{aligned}$$

To simplify expressions, I assume that the Emergency Loan credit availability ($0.5\bar{B}$) would be enough so that the borrower will not be credit constrained in the bad state of the world if they maintain their full credit line (i.e. $\lambda_2 = 0$). Under this assumption, the ex-ante input choice optimality is now determined by:

$$\frac{\partial f_G}{\partial x} = R\left[\frac{q}{1-q}\frac{u'(c_B^2)}{u'(c_G^2)} + 1\right] + \frac{\lambda_1}{\beta(1-q)u'(c_G^2)} + \frac{q}{1-q}\left[\frac{u'(c_B^2) - \beta u'(c_B^3)}{u'(c_G^2)}\right] \tag{1.12}$$

The first two terms of the equation are the same as we have seen previously

[18] Note that this result relies on the implicit model restriction that households cannot both borrow and save in the same period.

in equation 2. However, the last term is new and reflects the fact that with the combined constraint, any borrowing in period one now limits the ability to smooth consumption in the future bad state by using some of the household's credit line. If this cross-period constraint binds ($\lambda_3 > 0$), then $u'(c_B^2)$ and $\beta u'(c_B^3)$ will not be equalized and the numerator in the last term will be positive. This has the effect of increasing the value of the right hand side of the equation, implying that the increase in ex-ante inputs will be lower than for a Good Loan eligible client who did not have access to the Emergency Loan.

Turning to the first period borrowing choice, the condition is now:[19]

$$u'(c^1) = \beta R\left[qu'(c_B^2) + (1-q)u'(c_G^2)\right] + \lambda_1 + q\beta\left[u'(c_B^2) - \beta u'(c_B^3)\right] \tag{1.13}$$

Here, again, there is an additional term reflecting the potential gap between period two and three consumption in the bad state. As before, if the combined borrowing constraint binds, then the third term will be positive and this will imply that the increasing first period borrowing will be lower relative to a Good Loan eligible client who does not have access to the Emergency Loan.

These results imply that when we consider the impact of the new product, we should see on average a *stronger* effect of the Emergency Loan product on ex-ante input use and borrowing among clients who are *not* eligible for the good borrower loan than among those who are. Additionally, the emergency loan product will *reduce* the probability that eligible clients actually take the Good Loan if they expect to be credit constrained in the bad state of the world.[20] Therefore, we get two further predictions:

5. The treatment effect on first period investment will be lower among Good Loan eligible clients.
6. The offer of the Emergency Loan will reduce the probability that eligible clients take the Good Loan.

Repayment Decision with Guaranteed Credit

The goal here is to understand how allowing second period borrowing in the bad state changes borrowers' loan repayment decisions. Recall equation 8 that defined the value K^*, which is the benefit of future access to microfinance that would make a household indifferent between repayment and default. With the introduction of the Emergency Loan, this expression expands to include the option to borrow in the second period bad state and therefore also repay, or not, in the third period:

[19] Again, assuming $\lambda_2 = 0$

[20] Because in reality the uptake of the Good Loan is a binary choice, the effect of the Emergency Loan on Good Loan uptake will be weakly negative.

$$K^* = \frac{M}{q} + \beta \left[u(b_B^2) - u(-Rb_r^1 + b_B^2) \right] + \beta^2 \left[u(I) - u(I - Rb_b^2) \right] \tag{1.14}$$

To see how the repayment rates change with the introduction of the Emergency Loan, we need to sign $\frac{\partial K^*}{\partial b_B^2}$ when evaluated at $b_B^2 = 0$.

$$\frac{\partial K^*}{\partial b_B^2} = \underbrace{\frac{1}{q}\frac{\partial M}{\partial b_B^2}}_{-} + \beta \underbrace{\left[u'(0) - u'(-Rb_r^1)\left(1 - R\frac{\partial b_r^1}{\partial b_B^2}\right)\right]}_{-} + \underbrace{\beta^2 R u'(I)}_{+} \tag{1.15}$$

The first and second term above are negative; they capture improved good state outcomes and the reduced cost of repayment respectively when the Emergency Loan is available. However, the last term is positive and captures the added benefit of default when given more credit. Therefore, the overall effect on repayment is ambiguous.

MFI Problem

I now move beyond the household and consider the implications of offering guaranteed credit after a shock from the MFI's perspective. I assume that the lender is maximizing interest revenue minus the cost of defaults. For simplicity, I ignore the cost of capital and assume loans are either repaid in full (earning the MFI $b(R-1)$ in interest), or lost completely, costing the branch the full loan amount b. When a shock occurs, $F(K^*)$ gives the proportion of borrowers who will default on their loan. As before, I assume that there is no default under the good state. The MFI's expected profit from lending to a particular household (defined by parameters Y and I) is therefore given by:

$$\Pi = q\left[(1 - F(K^*))(R-1)b - F(K^*)b\right] + (1-q)(R-1)b \tag{1.16}$$

We are interested in whether it is profitable for the MFI to extend additional, guaranteed liquidity to borrowers after a shock has occurred. To explore what happens to expected profits with this policy change, we can simply explore how equation 16 changes when the amount borrowed (b) is allowed to move from b^1 to $(b^1 + b_B^2)$.[21] The MFI will want to offer the Emergency Loan if $\Pi_E \geq \Pi_{NE}$, where E and NE stand for Emergency Loan and No Emergency Loan respectively. This is given by

$$\begin{aligned} &q\left[(1 - F(K_E^*))(R-1)(b_E^1 + b_B^2) - F(K_E^*)(b_E^1 + b_B^2)\right] + (1-q)(R-1)b_E^1 \\ \geq &q\left[(1 - F(K_{NE}^*))(R-1)(b_{NE}^1) - F(K_E^*)(b_E^1)\right] + (1-q)(R-1)b_{NE}^1 \end{aligned} \tag{1.17}$$

[21] I assume households will take the Emergency Loan when offered, as otherwise the expected profits do not change

Where K_E^*, K_{NE}^* and b_E^1, b_{NE}^1 represent the indifference points for repayment and optimal first period borrowing choice with and without the Emergency Loan respectively. Rearranging equation 17, we can write that profits will increase if

$$
\begin{aligned}
&\underbrace{q(R-1)\left[(1-F(K_E^*)(b_E^1+b_B^2)-(1-F(K_{NE}^*)(b_{NE}^1)\right]}_{A}+\\
&\underbrace{q\left[F(K_{NE}^*)b_{NE}^1-F(K_E^*)(b_E^1+b_B^2)\right]}_{B}+\\
&\underbrace{(1-q)(R-1)(b_E^1-b_{NE}^1)}_{C}\geq 0
\end{aligned}
\tag{1.18}
$$

In equation 18, term A captures the change in profits from repayments. Because we know that b_E^1 is at least as large as b_{NE}^1, then $b_E^1+b_B^2 \geq b_{NE}^1$ unambiguously.[22] However, as we saw in equation 15, the effect of the Emergency Loan on K^* is ambiguous, therefore it is unclear whether $(1-F(K_E^*))$ is greater or less than $(1-F(K_{NE}^*))$. If the offer of the Emergency Loan improves repayment rates ($\frac{\partial K^*}{\partial b_B^2}<0$) then A is clearly positive. However, if the offer worsens repayment rate, then the sign of A is ambiguous.

Similarly, term B captures the lost capital from defaults. We know that $b_E^1+b_B^2 \geq b_{NE}^1$, but it is unclear whether $F(K_{NE}^*)$ is greater or less than $F(K_E^*)$. Therefore, as before, the sign of term B depends on what the effect of the Emergency Loan is on repayment rate (i.e. the sign and magnitude of $\frac{\partial K^*}{\partial b_B^2}$). If $\frac{\partial K^*}{\partial b_B^2}$ is positive, then this term is clearly negative and there will be larger losses from default. However, if $\frac{\partial K^*}{\partial b_B^2}$ is negative, then the overall sign of B is ambiguous.

Finally, term C captures profits when there is no shock. Again, this term is ambiguous. For households without access to the Good Loan in the pre-period, $b_E^1 \geq b_{NE}^1$. However, for households *with* access to the Good Loan, then b_E^1 could be less then b_{NE}^1 for clients who choose to preserve their access to the Emergency Loan. The size of these effects and the number of households that are in each situation will determine the overall sign of term C. Taking all three terms together, the overall effect on MFI profits from offering the Emergency Loan is ambiguous, and will be determined by i) the extent that the Emergency Loan changes households' repayment rates positively and ii) how the number of loans the MFI extends (including Dabi, Good, and Emergency Loans) changes as a result.

[22] This is clear for households without access to the Good Loan; however for households *with* access to the Good Loan, the situation is less clear. Because the Good Loan and Emergency Loan are the same size by design, households with a preexisting Dabi loan will either be able to take a Good Loan or the Emergency Loan, leading to the same total borrowed amount. However, treated households may optimally increase their Dabi loan size (this is unlikely in the first year of the program due to the timing of the pre-approval notification), in which case the borrowing amount will again be larger.

1.4 Research Design

The impact of the Emergency Loan was tested using a randomized control trial with a sample of 200 BRAC branches. These 200 branches were randomly selected from a group of branches that satisfied several criteria. First, I only included branches located in flood-prone areas. Second, I limited the sample to branches that were located within 15 kilometers of a river gauge run by the government's Flood Forecasting and Warning Center (FFWC) so that flooding could be monitored remotely. Next, I analyzed 15 years of historical data from the FFWC river gauges and selected areas of the country where flooding had exceeded the danger height levels at least twice. Last, I consulted the Bangladeshi branch of the International Rice Research Institute (IRRI) and the BRAC branches themselves to confirm that each branch's service area had experienced flood damage in the past six years. Figure 1.2 shows a map of the selected branches, their treatment status, and the matched water level gauges. The selected branches are concentrated in four main regions, including the Jamuna (Brahmaputra) basin, the Atrai river and Padma (Ganges) river basin, the Meghna river basin, and the Feni river basin. A total of 100 branches were assigned to the treatment group, and the remaining 100 branches were placed in the control group stratified by district. Table 1.3 provides descriptive statistics from households sampled from the treatment and control branches and p-values for the differences between these groups. The table shows that the randomized branches are largely balanced on baseline observables.

The experiment began in April 2016 when the Emergency Loan eligibility lists were created in each of the 200 experimental branches. Each branch manager could then review the lists and remove up to 10% of the eligible borrowers based on their knowledge of borrowers' behavior. The final eligibility lists were then sent to BRAC headquarters for data keeping and to verify that no more than 10% of borrowers had been removed from the original lists. Once finalized, referral slips (see Figure 1.3) were created for each eligible borrower in the branch. Each slip contained the borrower's name, BRAC identification numbers, and details on the Emergency Loan including the amount they had been pre-approved to borrow, the conditions when the loan would be made available, and the fact that they would lose their eligibility status if they took a Good Loan. The top half of the slip was kept by the borrowers to serve as "proof" of their eligibility status and to serve as a reference about the details of the loan. The bottom half of the slip was filled out with the borrower's information and phone number to help the branch management contact eligible borrowers after a flood.

The referral slips were distributed throughout the month of April during the normal VO meetings for each branch. At the end of each meeting, the loan officer distributed the referral slips to each eligible borrower and read a script that explained the purpose and the key features of the product. The concept of pre-approval was emphasized repeatedly because the idea was new within BRAC microfinance operations. Borrowers were asked questions about the Emergency Loan to confirm their understanding and time was given to answer any questions that eligible clients had about the product. Random branch visits during June of 2016 confirmed relatively good execution of pre-approval notification. Almost all borrowers had received the referral slips and understood that the Emergency Loan was available in case of flooding. There was some heterogeneity in borrowers' understanding of the more nuanced details of the loan, including pre-approval and conflict with the Good Loan. This was largely driven by differences in the quality of branch management.

During the Aman season, the FFWC flood water gauges were monitored every day for high water levels. When a gauge showed water levels crossing the danger level, a BRAC research employee (designated a "sector specialist") was asked visit the branches matched with the gauge. They mapped the area within each branch that had been affected by flooding based on conversations with local officials. If the reported amount exceeded 20%, the branch was activated (importantly, the sector specialists did not know about the 20% threshold needed to activate each branch). The branch manager was instructed from headquarters to notify all eligible borrowers that Emergency Loans were available. Borrowers were notified through their normally scheduled VO meetings or, in cases where VO meetings were suspended because of flooding, by calling clients directly and passing information through BRAC's social network. Additionally, eligible clients were reminded about the Emergency Loan's availability at every subsequent VO meeting until the expiration of the offer in November.

Over the course of the 2016 Aman season, 92 branches were activated: 40 control and 51 treatment.[23] However, 2016 was not a major flooding year and the water levels in the majority of activated branches did not cause widespread damage. As a result, BRAC decided to continue piloting the Emergency Loan for a second year in 2017. From 2016 to 2017, the experimental protocol remained the same. Only small improvements were made to the loan officer's description of the product. However, 14 branches (7 treatment, 7 control) were removed from the experiment from 2016 to 2017 due to changes in the local topography (new dams and roads) that reduced the probability of local flooding in these regions to almost zero. These 14 branches were replaced with back-up branches that had been pre-selected in the initial selection process described above. The new branches were randomized into treatment and control in February 2017. In 2017, 136 branches were activated, 73 control and 63 treatment. Flooding in 2017 was in general more severe than the previous year, and several locations suffered significant damage to crop land and physical structures.

1.5 Data

The data used in this analysis comes from two primary sources. First, I use BRAC's administrative loans and savings records for all clients in the experimental branches.

[23]The discrepancy in activation rates is due to random chance (the difference in means is not statistically significant). Much effort was put in to ensure that flood activation procedures were followed in the same way in control and treatment branches, and this policy way reinforced when the difference in activation rates emerged early in 2016.

This dataset reveals every client's borrowing behavior, including decisions to take loans, loan repayments and savings activities. Detailed repayment and savings data are available from April 2016 until January 2018, while loan disbursals data extends back for 1-7 years depending on the branch.[24] Within the loans data set, we observe approximately 300,000 unique individuals and 1.3 million unique loans. Eligibility for the Good Loan, which was not included in this data set, was compiled separately by BRAC for the purposes of this research.[25] Figure 1.4 shows a timeline of uptake of the three loan types studied in this research (Dabi Loan, Good Loan, and Emergency Loan) over the periods for which data is available, with the Aman growing season in 2016 and 2017 shaded in gray.

Second, I use survey data collected from 4,000 BRAC clients and 800 BRAC staff drawn from the 200 experimental branches. For the borrower survey, three village organizations (VOs) were randomly selected from each branch. Fifteen eligible borrowers and five ineligible borrowers were randomly selected within each VO[26]. Three rounds of data collection took place: a baseline survey was conducted in April 2016 before borrowers in treatment branches were informed about their eligibility status; a follow-up survey was implemented in December 2016 after the first flooding season; and a second follow-up took place in December 2017 after the second flooding season. Survey rates were very good, 99% at the first follow-up and 98.9% at the second follow-up.[27] The household surveys focused on both agricultural and non-agriculture business investments and outputs, consumption, asset holdings, and household perceptions of and response to any flooding that occurred in the area. The surveys of BRAC's administrative staff included four branch-level managers (both in and outside of microfinance operations) and asked about their perceptions of flood risk, the most important local income generating activities, and their perceptions of overall local flood damage in the branch service area.

Households in the sample on average have approximately five members, own small plots of land (0.44 acres), and earn an annual income of about \$1,600 (\$320 per capita). Education levels are low, with the head of household only having two and a half years of formal schooling on average. Electrification is relatively high with approximately 70% of the sample reporting electricity access. Approximately fifty percent of the sample reported growing crops in the previous aman growing season, with the average land dedicated to of about 0.4 acres (including rented and sharecropped land). In the past five years, 55% of the sample reported experiencing a flood shock that damaged their crops or assets.

[24]Certain BRAC branches began digitizing data earlier than others and some branches in the experiment were founded relatively recently.

[25]Due to uncertainty about whether the project would continue for a second year, this data is missing for five months between November 2016 and March 2017 while the decision to extend the experiment was being made.

[26]Appendix C reports on spillovers to ineligible borrowers. In general, I find no evidence of spillovers; therefore the main analysis discussed in this paper focuses only on eligible BRAC members.

[27]Survey rates were helped tremendously by BRAC's network, which enabled easy tracking of households that relocated within and between communities.

1.6 Results

Emergency Loan Take Up

I first examine the decision to take the Emergency Loan after a flood shock. In both years, uptake of the Emergency Loan among eligible households was quite low. In 2016, only 2.9% of households chose to take the loan, likely because the floods that year were not particularly severe in most locations. In 2017, floods were much more damaging and uptake of the Emergency Loan increased to 5.4%. It is important to note that these low take-up rates do not necessarily imply that households did not value or benefit from the Emergency loan's availability. While I address this point in more detail below, households can respond to the offer of a loan before flooding has even occurred. Indeed, the Emergency Loan stimulates higher investments and greater output, suggesting it offers important protection in the pre-period against low probability shocks. Furthermore, low ex-post uptake of this product is not entirely unexpected because flood damage is highly idiosyncratic within a branch service area, such that certain villages may be dramatically affected while other villages within the same branch will not be hit at all.

Table 1.5 reports which household characteristics correlate with take-up among the set of households that were offered the Emergency Loan (i.e. those that were in a treatment branch after a flood). Considering first baseline characteristics, column 1 shows that households that took the Emergency Loan are observably quite similar on most dimensions to households that did not (risk aversion, time preferences, flooding history and income). Column 2 explores correlations between uptake and households status after a flood. I see higher take-up among households that were less well prepared for a flood and among those that experienced higher levels of distress. Furthermore, figure 1.5 highlights lower yields among households that took the Emergency Loan. Finally, figure 1.6 shows that there is no significant difference in the probability of Emergency Loan uptake by borrower credit score. Overall, the results suggest that the more vulnerable and worst affected households are more likely to take advantage of the guaranteed credit offer, results that are largely consistent with the theory.

Estimation Strategy

To estimate the effects of guaranteed credit lines on household level outcomes, I will compare *eligible* BRAC microfinance members across treatment and control branches. Eligible clients in control branches are those who had credit scores high enough to qualify for the Emergency Loan had they been in a treatment branch. The baseline specification for household outcomes is therefore:

$$Y_{ibdt} = treatment_{ibd}\beta + \alpha_d + \phi_t + \mathbb{X}_{ibd}\gamma + \varepsilon_{ibdt}$$

Where Y_{ibdt} is an observed outcome for an eligible household i in branch b and district d during year t. I regress each outcome on an indicator for treatment, a district fixed effect (the stratification variable), a year fixed effect, and a vector of baseline controls to increase precision.[28] Data from both years of the experiment are pooled together (unless noted otherwise) and standard errors are always clustered at the branch level.[29] For "ex-post" outcomes that occur after the flood season, I run the same regression with an additional indicator for flooding during the growing season.

The same basic procedure is largely followed for MFI level outcomes (e.g. loan uptake decisions, repayments) but with a few modifications. Because I examine observations at the branch-month level, I add month m fixed effects in addition to year and district fixed effects to the estimating equation.[30]

$$Y_{bdmt} = treatment_{ibd}\beta + \alpha_d + \phi_t + \rho_m + \varepsilon_{bdmt}$$

Credit Line Preservation

As mentioned above, low take-up does not necessarily reflect the value that households attribute to the Emergency Loan. Gains from the loan can be reaped even if households decide not to take the loan after the uncertainty about the flood is resolved. Access to the loan improves welfare by reducing households' exposure to the downside risks associated with severe flooding. To test whether households recognize this crucial feature of the Emergency loan, I investigate two phenomena. First, I document whether households choose to preserve their credit access to insure themselves against bad times. Next, I investigate whether households invest more in the pre-period because they know they have access to an additional loan in the event of a flood.

To investigate credit preserving behavior, I take advantage of the tension between the Emergency Loan and the Good Borrower Loan: households that take the Good Loan in the pre-period lose access to the Emergency Loan. Eligible households have to choose whether to take a Good Loan and forgo the Emergency Loan should a flood occur, or decline the Good Loan in order to preserve the option to take the Emergency Loan after a shock. According to the theoretical model, forward looking households will want to preserve credit access as a buffer against this risk. I test this prediction by comparing the probability of taking a Good Loan among Good Loan eligible clients in control branches compared to treatment branches in the pre-flood period. For this analysis I use BRAC's administrative data that captures loan disbursals and repayments.

Table 1.6 shows the results from the cross-branch comparison of Good Loan eligible borrowers (the regressions are run at the branch MFI level). Column 1 shows that the availability of the Emergency Loan reduces the probability of taking

[28]Controls include land owned by the household, household size, and head of household age and education unless specified otherwise

[29]Appendix B accounts for possible differential selection into eligibility in 2017. Results are stable when excluding 2017 data or when instrumenting for eligibility using branch treatment status.

[30]Some regressions have only a single observation per year, in which case month fixed effects are dropped. Note that this dataset does not contain baseline controls and hence they are not included in the regression

a Good Loan by two percentage points, or 15% in treatment branches. Column 2 and 3 examine the extent to which this effect changes based on a measure of the need for liquidity, and by the perceived risk of local flooding. I proxy the need for liquidity by areas that report farming as the primary occupation, where significant investments are needed in the pre-period to prepare the seedbeds for cultivation. I do not see any significant change in the treatment effect in areas where the primary occupation is farming. However, areas that have a higher perceived flood risk are even less likely to take the good loan. Together these results show that, among households that would have exhausted their available credit absent the Emergency Loan, a significant number choose instead to preserve it. Furthermore, areas that face higher risk are more likely to preserve their credit access, confirming that at least some households view guaranteed credit access as offering effective insurance against shocks.

The fact that households are willing to give up investment in the pre-period suggests that the value of preserving the guarantee is substantial for at least this subset of households. Those that forgo the Good Loan in order to preserve access to the Emergency Loan are giving up certain credit today in order maintain credit in the future that will only be made available with some probability. I calculate what this implies about the value these households assign to the Emergency Loan first under conservative assumptions and then under more realistic assumptions. If I assume conservatively that households were able to correctly predict the probability that a loan would be offered (54% over the two years of the study), planned to take the loan whenever it was made available, and that they do not discount the future, this implies that the marginal utility to households of access to credit after a flood is at a minimum of 1.85 times more than the marginal utility of certain credit in the pre-period. If I instead use a more realistic assumption that households expected to use the Emergency Loan at the rates actually observed in the experiment (5%) and that they have an annual discount rate of 6%, then the marginal utility of a loan after a flood is 20.5 times greater than in the pre-period[31].

In order to understand which borrowers are more likely to actively preserve their credit access, I estimate a local average treatment effect across bins of the Emergency Loan credit score. Figure 1.7 plots the treatment effect on Good Loan uptake by credit score bin for eligible clients. There appears to be some evidence of heterogeneous treatment effects: the reduction in the probability of taking a good loan among the eligible population is highest among those with especially high credit scores. Column 1 of Table 1.18 fits a linear trend to this relationship and shows that this effect is (marginally) statistically significant. This suggests that clients with the best repayment histories are more likely to preserve credit access to hedge against future shocks. We might expect this result if clients with higher credit scores have lower discount rates or if they are less present biased. Such households would likely make more timely payments (hence the higher credit scores) and be willing to preserve credit access.

[31] This assumes a waiting time of five months between the decision to forgo the Good Loan and the decision to take the Emergency Loan.

Ex-Ante Household Investment

Recall from theory that the extension of a guaranteed credit line is designed to mitigate the adverse consequences of a shock, thereby encouraging households to invest more in the pre-period. I focus on changes to agricultural investments because it is the most important income generating activity for the majority of rural households in Bangladesh. Moreover, these investments are more likely to be exposed to flood shocks and sensitive to interventions that reduce household flood risk. I also investigate the impacts on non-agricultural business investments because the sample is drawn from microfinance clients that are more likely to be business owners and less likely to own land than the general rural population (48% of surveyed households planted crops during the 2015 Aman season).

I begin with Table 1.7, which shows the amount of land devoted to agriculture during the rainy season. The first three columns separately identify the impact for three different types of land tenure, namely owned, rented, and sharecropped land, while column 4 aggregates these three measures. The last column is a binary indicator for planting any crops during the season. Households that knew they were eligible for the loan increased the amount of land they rented by 30%, and overall cultivation by 15%. Neither owned nor sharecropped land showed any significant change. This result is not altogether surprising because finding additional land to rent is relatively straightforward. Conversely, expanding the cultivation of owned land would require farming previously fallow land or purchasing additional crop land, which is more costly and requires more planning. Similarly, sharecropping contracts become relatively less attractive because farmers' ability to reduce their exposure to risk can now be fulfilled by the Emergency loan. Finally, along the extensive margin, the number of eligible households planting crops increased by approximately 4 percentage points. This represents a 10% increase in the probability that a household cultivates crops during the Aman season.

With an expansion in cultivated land, total input use is likely to increase mechanically. However, households might also increase the intensity of input usage in response to reduced exposure to risk. The first four columns of Table 1.8 present the effects of the intervention on inputs applied to cultivated farm land. Columns 1 and 2 show the amount of fertilizer and pesticides applied per acre of land. While both variables have positive point estimates, neither are statistically significant. Similarly, columns 3 and 4 show that the amount of money spent on seeds and all other inputs per acre increases but remains insignificant. At a minimum, these results indicate that treatment households are maintaining normal levels of input usage per acre despite the overall expansion of cultivated land. Finally, column 5 of Table 1.8 examines the amount of investments in non-agriculture business. We see a marginally significant increase of 29% ($11 USD) over the control group.[32] However, this last result is sensitive to the regression specification and only becomes significant in the second year of the experiment.

These initial results are consistent with the theory that guaranteed credit lines

[32] Business investment was measured by the total value of newly purchased (or repaired) business assets.

can increase investments by providing effective insurance against floods. However, to confirm that the product is operating on farmers' perceptions of risk, I investigate whether the treatment effects are higher among the most risk averse households (as measured at baseline).[33] These households represent a meaningful share of the sample (27% of households exhibit the highest level of risk aversion), and generally invest less at baseline. Tables 1.9 and 1.10 report these results, where the measure for risk aversion is normalized to a 0-1 scale (one representing the most risk averse households and zero the most risk loving). From Table 1.9 we see that all the point estimates on the interaction terms between risk aversion and treatment are positive. However, they are only significant for rented and total land cultivated (which is where I documented the strongest impacts previously). I also investigate the impact for risk averse households specifically, by running a linear combination test on the sum of the treatment and interaction terms. Here I find significance for the treatment effect on rented land in addition to total land. Similarly, in Table 1.10 the interaction term is positive for fertilizer, pesticide, and non-agricultural investment. The tests indicate that pesticide application and non-agricultural investment have increased significantly for most risk averse households. Overall these results suggest that guaranteed credit lines are encouraging investments by reducing households' exposure to risk. The fact that risk averse households tend to underinvest in general suggests this product is particularly valuable at correcting a negative distortion for this subgroup.

The results on increased investment and credit preservation suggest that households perceive the Emergency Loan as reducing their exposure to risk. However, it is possible that households choose not to take the Emergency Loan because they learn ex-post that it would be unlikely to be useful to them (in this season or in any future one). This might be driving the low take-up results we saw above. If households do learn this, we would to expect see their 2017 Aman season investments decrease to pre-treatment levels because they now do not perceive any risk reduction benefit from access to guaranteed credit. To test this theory, I examine how investment decisions change in the second year of the experiment based on whether households experienced a flood shock in the first season. If flood-afflicted households learn that the Emergency Loan does not insure them against negative outcomes, these households should have a smaller treatment effect on investment relative to treatment households that did not experience a flood shock. However, if households still perceive the offer of guaranteed credit as reducing the downside risk of flooding, then investment should be the same (or larger) when compared to households that were not flooded in the first year.[34]

[33] Risk aversion was measured by asking borrowers a series of choices between a certain payout and a larger but uncertain payout. Each successive choice increased the probability that the uncertain payout would be realized (see Sprenger 2015 for more details). The resulting risk aversion spread was normalized to a zero to one scale so that the most risk averse households have a value of one and the most risk loving a value of zero.

[34] A possible confounding factor is that the extra credit afforded by access to the Emergency Loan in the first year could itself impact investment decisions in year two. However, this effect will have a minimal role due to the fact that only 2.9% of eligible households took the Emergency Loan when

Table 1.11 demonstrates how flooding in the first year affects different investment categories. First, we can see that being flooded in the previous year does seem to have negative consequences for control households' investments in the current year. In particular, control households are ten percentage points less likely to cultivate crops in areas that were flooded in 2016. However, the treatment effect on investments does not appear to be different for treatment households that were flooded in the first year relative to treatment households that were not. The interaction term is generally small in magnitude and not statistically significant for any outcome. Overall, this suggests that households that experienced flooding in 2016 still perceive the Emergency Loan as offering viable protection against some flood risk.

Ex-Post Household Outcomes

Next, to examine how the Emergency Loan affects households after the Aman season, both in areas that experience flooding and those that do not. Recall from the model that offering the Emergency Loan will affect households differently depending on the state of the world. In the event of a flood, the emergency loan becomes available and treatment households will have access to more liquidity than will control households. If a flood does not a occur, increases in investment before the Aman season will translate into improved output.

I examine the effect of treatment on four household outcomes: log weekly consumption per capita, log income during the previous month, crop production from the Aman season, and the number of livestock animals owned by the household.[35] Table 1.12 shows the results of regressing these outcomes on an indicator for treatment, an indicator for experiencing a flood shock during the growing season, and an interaction between the two.[36] The coefficient on treated captures the effects on household outcomes from increases in ex-ante investment only. Absent a flood, the only difference in outcomes between treatment and control households will stem from changes in investment in the pre-period. In contrast, the interaction term will capture the effect of *both* channels on household outcomes. After a flood, treatment households will have access to any output the flood did not destroy, and to the Emergency Loan should they choose to use it for recovery.

In branches that did *not* experience flooding, treated households display the same levels of consumption, income, and livestock ownership as control households (Table 1.12). However, there is a significant 28% increase in crop production which aligns with the pre-period results documenting additional crops being sown by treatment households. This suggests that households reap the benefits of greater investments absent a flood even though this does not translate into higher levels of measured consumption or asset holdings. In branches that *did* experience a flood, treated households experience a rather large 10% increase in consumption com-

offered in the first year.

[35]The estimation for log consumption adds week interviewed fixed effects because of holidays that occurred over the survey period which changed consumption patterns for some households.

[36]See Appendix B for ex-post results pooling both flooded and non-flooded branches.

pared to control households.[37] However, their production is affected by the flood, losing almost 80% of the gains they reap when a flood does not occur (Column 3). These losses are proportionally much larger than those observed in the control group, suggesting that treatment households expand cultivation on land that is particularly susceptible to floods. Finally, we see a large increase in the number of livestock among treatment households (Column 4). This suggests that the availability of the Emergency Loan allows households to maintain their asset levels after an income shock[38].

There is a concern that multiple shocks may reduce the usefulness of credit as a risk mitigation tool if households use their entire available credit line, thereby eliminating the products consumption smoothing benefits. Table 1.13 examines this hypothesis. To do so, I expand the regression specification from Table 1.12 to include an indicator for whether the household experienced flooding in both years, and an interaction of this indicator with treatment. The experiment was only conducted over two years, so multiple shocks can only be picked up for households that experience flooding in both 2016 and 2017. The results confirm that experiencing successive shocks reduces food consumption by 19%, but has little observable impact on the other three outcomes. This suggests that multiple shocks are indeed harmful to households' well being, even if the channels through which this occurs are unclear. Next, to determine whether the usefulness of guaranteed credit is reduced after successive shocks, I examine the interaction of the double flood indicator and the treatment indicator. These coefficients are all statistically insignificant, but a joint test of all the treatment coefficients shows that treatment households are still better off after a double shock. Overall, this suggests that the gains in consumption and asset preservation due to treatment are not completely eliminated by successive shocks. However, it is worth interpreting these results with some caution because the 2016 shock was not particularly damaging, and may not reflect responses to larger shocks.

Finally, Table 1.14 explores how the availability of the Emergency Loan changes outcomes for households that earn income through the labor market rather than (or in addition to) agriculture or a small business. The first column shows that laborers' daily wage does not change in flooded or non-flooded areas. In non-flooded areas, however, the number of days employed elsewhere as a day laborer decreases by 10%, likely a result of more households spending time on their own fields. In flooded areas, we see a 30% reduction in day labor, which is substantially reversed in treatment households (by approximately 20%). This result could be consistent with the fact that households often use the Emergency Loan to re-plant the fields destroyed by floods, increasing the number of opportunities for other laborers to find work. However, given the low rates of Emergency Loan take-up, this channel is unlikely to explain the entirety of this effect.

[37] The p-value for the joint test is found in the bottom row of table 1.12

[38] In this context, livestock are a common form of household savings and are often sold when the household has a need for liquidity.

Impact on MFI Operations

I conclude the analysis by investigating how BRAC branches perform when the Emergency Loan is made available. As discussed in the theory section, it is unclear a-priori whether extending guaranteed credit after a shock will help or harm overall branch performance. There are two key outcomes that determine branch profitability: the number of loans disbursed and the repayment rates of those loans. Therefore, to understand the effect of the Emergency Loan product, I will examine each of these outcomes in turn (recall we have already seen that the Emergency Loan reduces the number of Good Loans disbursed).

I begin by examining whether offering the Emergency Loan increases the likelihood that borrowers take an initial loan from BRAC. To do so, I examine the probability that a normal dabi loan is taken in the pre-flood period among all members of the branch.[39] The results in Table 1.15 show that treatment causes the probability of taking a dabi loan to increase by 11% (0.7 percentage points) in the pre-period. However, it is possible that the increase in loan disbursement during the pre-period came at the expense of future loans (for example, if households simply move up their previously planned investment timeline). Figure 1.8 examines whether this is occurring by plotting the monthly probability of dabi loan up-take by treatment status from 2015 until the end of the study period. We can see that the probability of taking a new dabi loan is higher in the treatment branches during the pre-period, but is otherwise fairly similar. This suggests that the extra dabi loans disbursed in the pre-period represent additional loans that would not otherwise have been disbursed. Finally, as with the Good Loan analysis, I examine whether the increase in Dabi Loan uptake differs across credit scores. Figure 1.9 and column 2 in Table 1.18 shows that the increase in Dabi Loans (unlike the reduction in Good Loan uptake) does not differ by credit score.

In addition to loan disbursals, impacts on repayment rates are critical to establish the sustainability of the Emergency Loan. Table 1.16 shows how the probability of a missed payment differs between treatment and control branches in the pre-period and after a flood. The coefficient on treatment shows that access to the Emergency Loan has no effect on repayment rates in the absence of a shock. Looking at the coefficient on flooding, we see that flooding increases the number of missed payments by approximately 3.9 percentage points (40% percent) in control branches. However, in treatment branches this effect is overcome by a reduction in missed payments of 4 percentage points, thereby returning repayment rates to approximately normal rates. Furthermore, the repayment rate of the Emergency Loan itself is almost identical to other loans during the same period (10% missed payments for the Emergency Loan as compared with 9.6% on all loans). This result is even more meaningful when we remember that households that took the Emergency Loan experienced greater damages from the flood. Overall, these results demonstrate that the availability of the Emergency Loan improved repayment for

[39] All members were included in the analysis so that the denominator of eligible borrowers remained constant throughout the study time period and did not change in response to endogenous loan take-up decision.

the MFI in the aftermath of the flood (on a branch wide basis).

Next, I look for heterogeneity in repayments rates by borrower credit score. Figure 1.10 and 1.11 illustrate how repayment rates differ by credit score and by treatment status. First, Figure 1.10 shows that the treatment effect on repayment rates[40] is largest among clients with scores that are close to the eligibility threshold of 77. The effect falls quickly at higher credit scores (column 3 of Table 1.18 shows that this heterogeneity is statistically significant). This decrease is likely explained by the fact that borrowers with high credit scores already repay at such high rates that further improvements are difficult. In Figure 1.11, we see that approximately 6% of payments are missed among those with high credit scores, which is low enough that it may be difficult to improve repayment rates significantly.

Overall branch profitability is derived from the number of loans disbursed and the repayment rates on those loans. So far, we have seen that the effect on total loans disbursed is ambiguous – a decrease in the number of Good Loans taken, but an increase in the number of regular dabi Loans and new Emergency Loans – while the effect on repayment rates appears to be positive. To capture the overall effect on the branch, we can directly compare the profitability of branches that offered the Emergency Loan to those that did not. Table 1.17 shows the estimated effects of treatment on three measures of MFI profitability: the net present value of each loan disbursed, the monthly profitability of the branch in aggregate, and the per-member monthly profitability of each branch.[41] The first two results show positive point estimates, but neither is statistically significant. However, column 3 shows a 4% increase in the per-person profits in treatment branches. In sum, these results suggest a modest increase in branch profitability and allow us to say that branch profitability was likely not harmed.

We can examine the extent to which the effects on profitability vary by borrower credit score. Figure 1.12 plots the treatment effect on per-person profitability by credit score decile. We see that the treatment effect is highest for clients with credit scores closer to the eligibility cutoff and decreases steadily until it is negative for those with higher credit scores (column 4 of Table 1.18 show that this heterogeneity is statistically significant). These results are consistent with the effects I have shown previously. Clients with scores near the cutoff both have the highest improvements in repayment rates and the lowest reductions in the probability of taking a Good Loan. In contrast, high credit score clients make only modest improvements to their repayment rates while experiencing the largest reductions in the probability of taking a Good Loan.

These results have interesting implications for the targeting of the Emergency Loan. The Emergency Loan was targeted at the top 40% of borrowers based on a credit score reflecting their past loan behavior. This system was designed to reduce the downside risk for the MFI in case repayment rates from the Emergency Loan

[40] The estimated treatment effect is from regressions pooling both flooded and non-flooded branches.

[41] To calculate net present value for each loan, I assume an annual cost of capital of 6%. Branch profit is calculated as the sum of discounted repayments minus the cost of new disbursements, while per-member profitability takes this measure and divides it by the number of branch members.

were low. However, the results suggest that BRAC could do even better by lowering the eligibility threshold. Assuming the measured treatment effects are continuous across the threshold, this would extend access to clients who are most likely to improve MFI profitability. In contrast, restricting access to the Emergency Loan to clients with the highest credit scores could lead to an overall reduction in branch profitability because they are less likely to take the Good Loan, and their repayment rates do not have room to improve.

As a final check on MFI performance, we can look at saving rates. BRAC benefits directly from the amount of savings stored by clients at the branch. Table 1.19 shows how the savings rates differ between treatment and control branches and their differential response to flooding. Column 1 shows that in the pre-period, where we might have expected a draw down in liquid assets, savings rates do not differ between the two branches. However, column 2 shows that in the aftermath of a flood, eligible households are able to maintain higher savings rates by 45 taka on average (which represents a 62% increase on the average transaction amount, but less than a 1% increase on *total* savings). Column 3 shows that this effect does not vary by the level of localized damage inflicted by the flood[42].

1.7 Conclusion

Millions of households across the world are exposed to severe income risk. In many cases, these households live in areas where insurance markets are non-existent and they have to resort to costly coping mechanisms in order to survive. Under these circumstances, it becomes important to develop tools that can decrease households' exposure to risk and help them self-insure. I build on recent literature, which suggests that uninsured risk is an important constraint and that existing insurance products are inadequate, by offering a guaranteed credit line as a potential solution. I run a large scale RCT to investigate whether guaranteed credit can successfully function as insurance in rural regions of Bangladesh where annual flood risk is high. First, I show that households value this product: when given the choice, many households choose to preserve their access to guaranteed credit at the expense of additional liquidity in the pre-period. This behavior is consistent with a model where households utilize their credit access as a buffer against the risk of future shocks. Households that were informed about their guaranteed credit access also increased their investments in productive (but risky) activities in the pre-period. These effects were concentrated among more risk-averse households. This increase in investments translated into more production absent a flood, and higher consumption and asset levels when a shock did occur.

I also show that the extension of a guaranteed credit line after a shock has modest but largely positive effects for MFIs. More members take loans in the pre-period in response to the added security, repayment rates after a shock are improved, and savings rates increase. Therefore, at least in this context, a product like the

[42]Flood damage at the branch level was only collected in 2017; therefore column 3 only uses data from this year.

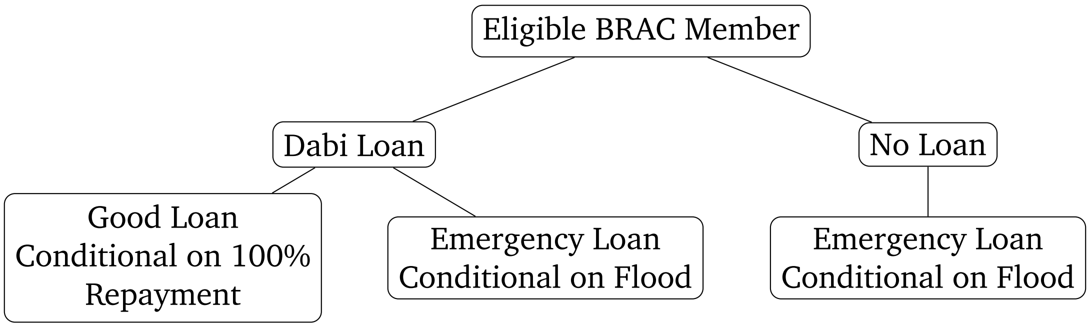
Emergency Loan slightly improves branch performance in addition to benefiting clients. Provided that loan repayment rates remain similar in other settings, this suggests that guaranteed credit can be offered by MFIs and does not require any third party subsidies. This is appealing because MFIs are ubiquitous in low income countries and can easily offer the product to many households using their existing infrastructure.

One question raised by the results is why a product like the Emergency Loan that seems to benefit both households and the MFI has not already been adopted by the microfinance industry. I suggest two obstacles that may prevent adoption. First, many MFIs do not keep adequate records and lack the lending history necessary to create a credit score to target responsible borrowers. The results are unlikely to generalize to lower performing clients and it is important to be able to identify who these households are. Second, a guaranteed credit product does not necessary align with the incentives facing branch level officials. Branch managers are commonly incentivized to disburse a certain number of loans and to maximize repayment rates. However, in the aftermath of an aggregate shock, a branch manager may be concerned that households are going to miss payments on their existing loans, and a product like the Emergency Loan will compound these losses. This would increase the downside risk facing the branch, and could potentially jeopardize the manager's job. Therefore, there is likely be resistance from branch level staff to adopt similar guaranteed products.

From a policy perspective, this research suggests that credit can be a useful tool to address uninsured risk in places where traditional insurance markets have failed. With the growing frequency and severity of weather shocks due to climate change, adding an easily accessible tool that helps households reduce exposure to risk is important. The tool I explore here (guaranteed credit) is appealing because it is already understood in these environments and is widely used worldwide. While the household impacts are similar to those documented in the index-insurance literature, pre-approved credit has the advantage that it can be extended without requiring any commitment from the beneficiary, bypassing many of the drivers of low demand for insurance. Moreover, because the decision to utilize the additional credit is made after any shock damages have been realized, households can optimally opt-in based on the ex-post costs and benefits. Therefore, guaranteed credit can crowd-in ex-ante investment even if households choose not to use the product in the aftermath of a flood.

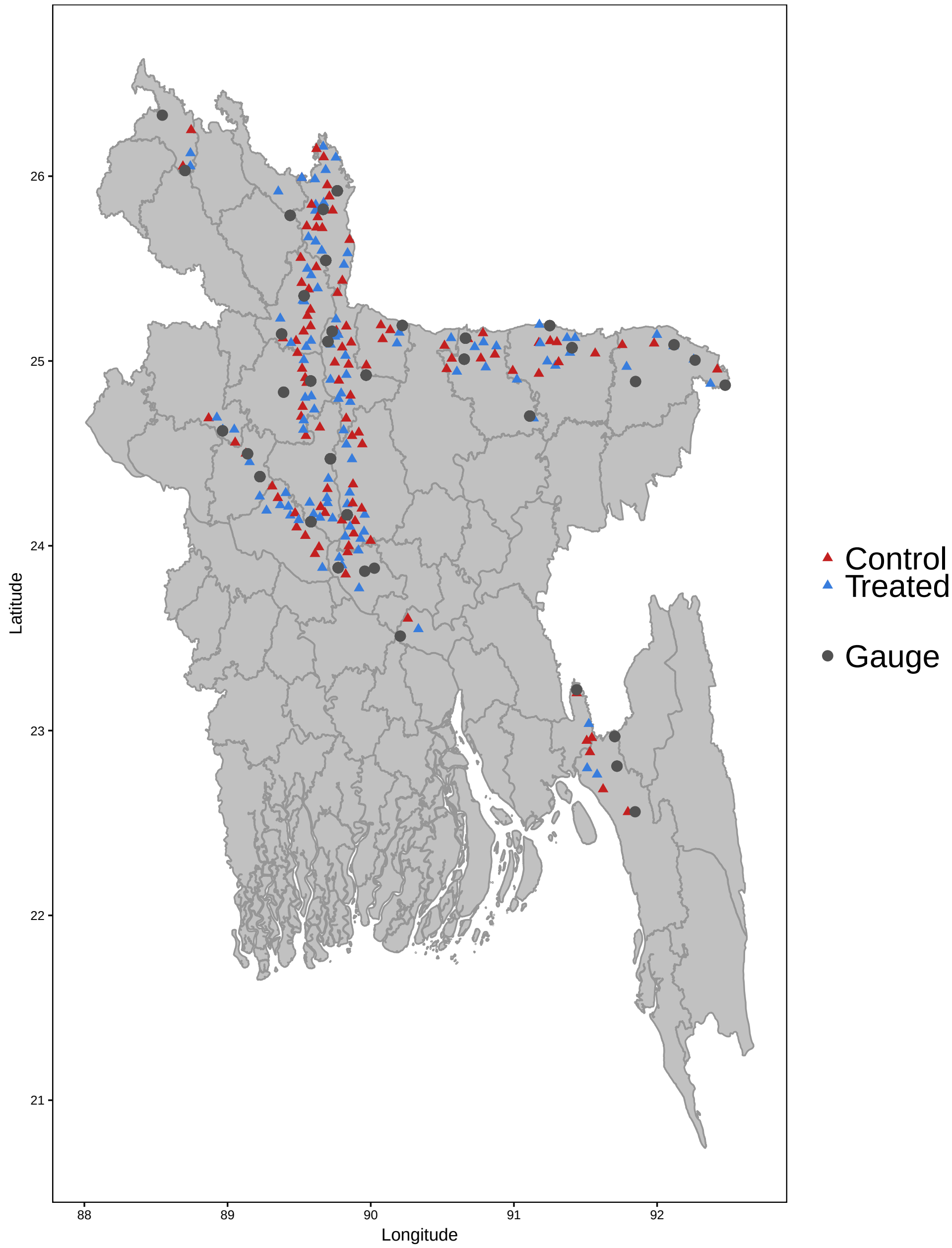
Figures

Figure 1.1: Loan Choices for Eligible Members



Notes: The Figure above shows a schematic representation of the loan choices facing a BRAC microfinance member. There are three types of loans: the normal Dabi loan, the Good Loan, and the Emergency Loan. The Good Loan is only available to borrowers who have taken a Dabi Loan and have made all on-time payments through the first six months of the original loan. The offer of a Good Loan expires after two months. The Emergency Loan is only available after a flood has occurred, but it is offered whether or not the member currently has an active Dabi Loan. Members who take a Good Loan cannot also take an Emergency Loan when a flood occurs.

Figure 1.2: Map of Sample Branches



Notes: Map shows the locations of BRAC branches that participated in the experiment (triangles) as well as the water level gauges used to monitor flood water levels (circles). Branches were selected based on their history of flooding and proximity to a water level gauge maintained by the Bangladeshi government.

Figure 1.3: Referral Slip

brac

Referral Slip – Emergency Loan

Member Copy: Please keep

Branch Name:... Code: ☐☐☐☐ Branch contact #: ☐☐☐☐☐☐☐☐
Member Name:... Member No: ☐☐☐☐ VO Code: ☐☐☐☐
PO Name: Sign: Branch Manager Sign:

If you have a completed form with a signature then you are guarenteed eligiblity for Emergency Loan

Loan Conditions:

- River overflow and local area flooding confirmed by BRAC

Loan Amount

- Can take up to 50% of current or last loan
- Maximum of 50,000 taka

Things to bring when getting Emergency Loan

- Referral slip
- Identification card

Ineligibility condition

- If you take a Good Loan
- Your branch area is not affected by flooding

- - - - - - - - - - - - - - - - - Tear here - - - - - - - - - - - - - - - - -

brac

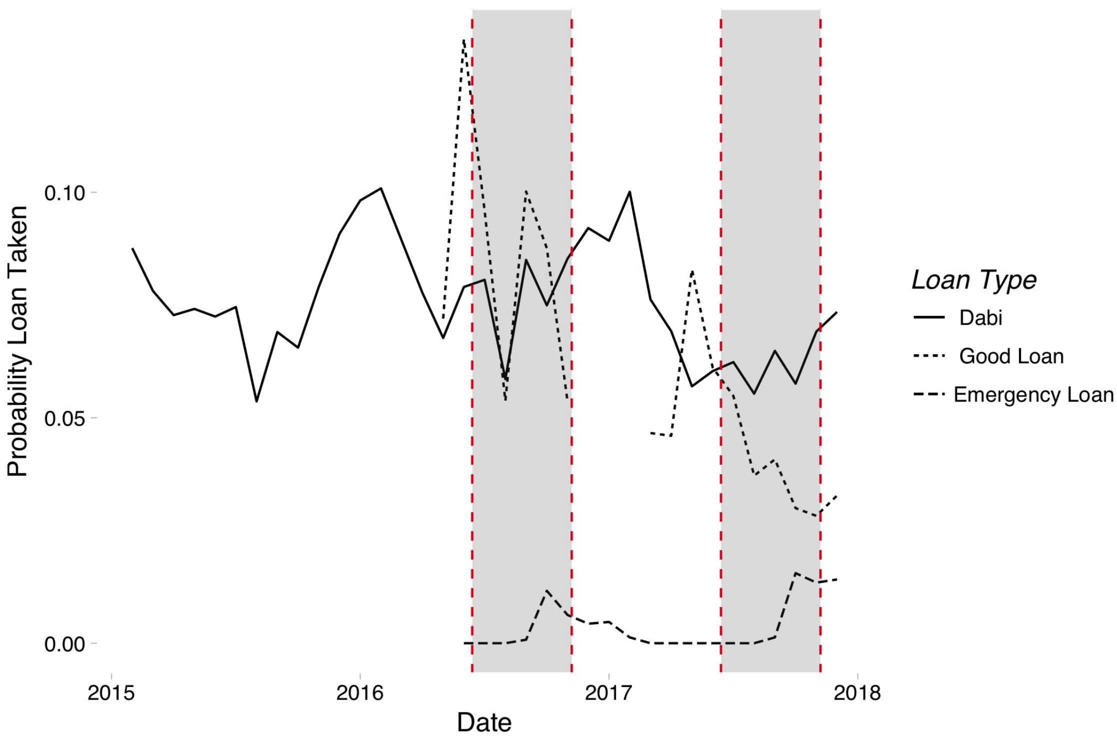
Referral Slip – Emergency Loan

Office Copy: Please keep

Branch Name:... Code: ☐☐☐☐ Member contact #: ☐☐☐☐☐☐☐☐
Member Name:... Member No: ☐☐☐☐ VO Code: ☐☐☐☐
PO Sign: Branch Manager Sign: Accountant Sign:

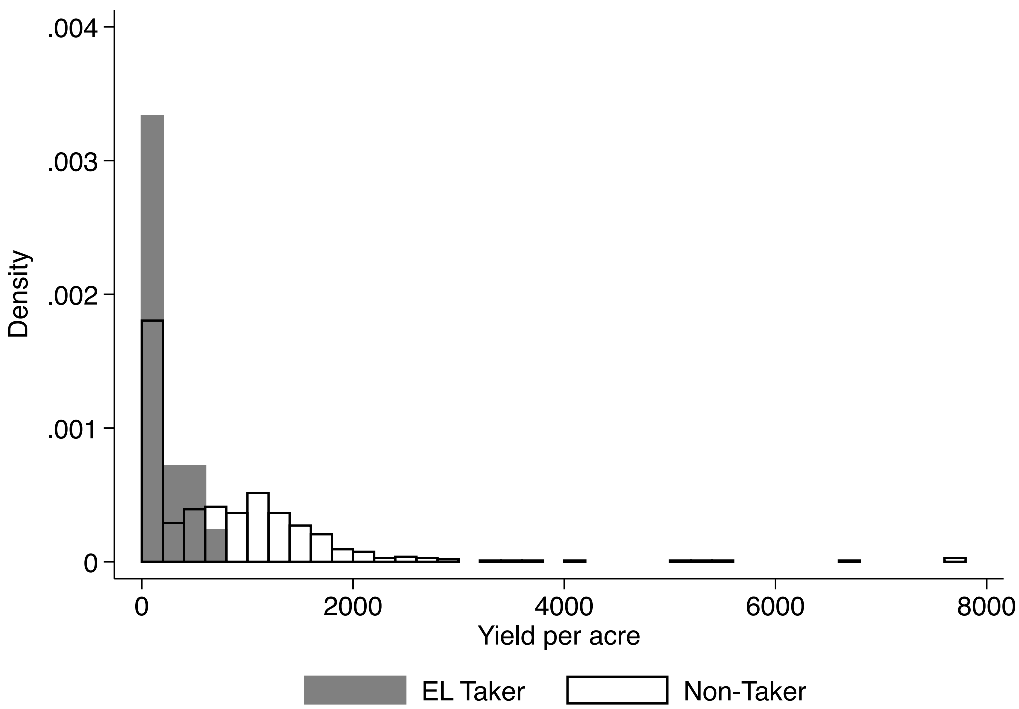
Notes: The Figure shows the referral slip (translated from Bangla) given to BRAC microfinance members eligible for the Emergency Loan. The slip records a client's name and BRAC identifiers, the maximum pre-approved loan size, as well as a brief description of the loan product. The bottom of the slip also contained the borrower's information and was kept by the branch manager to facilitate easy follow-up should a flood occur in the area.

Figure 1.4: BRAC Loans



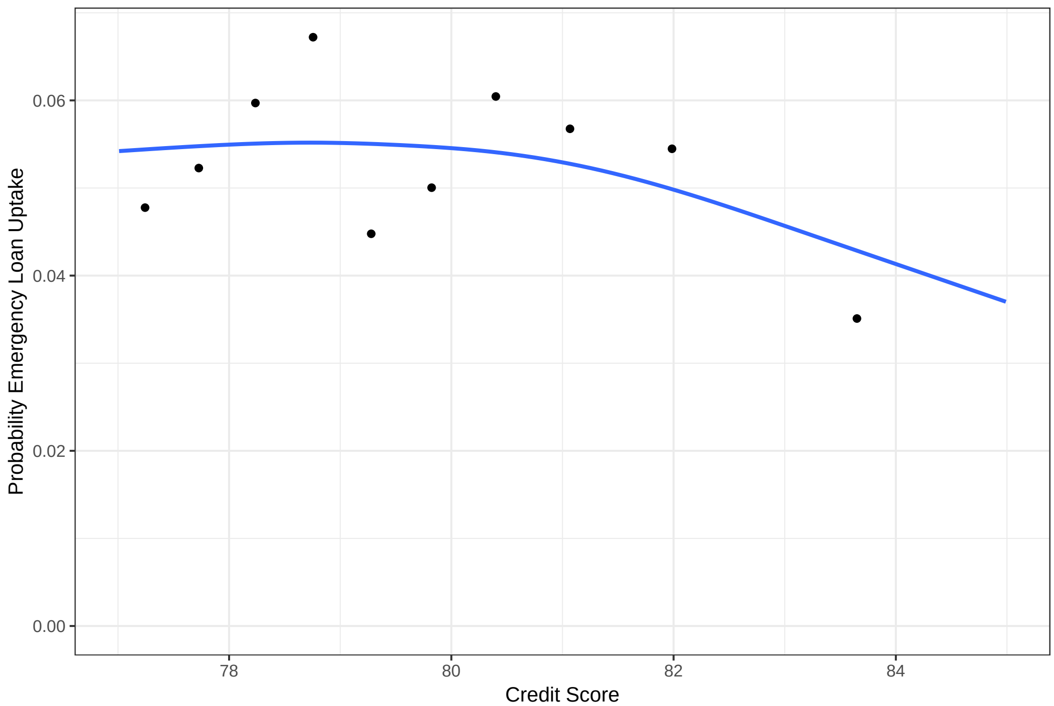
Notes: Figure shows the uptake of the three different BRAC loan products examined in the experiment. The solid line shows Dabi loan uptake as a proportion of overall branch membership. The Short-dashed line shows Good Loan uptake as a proportion of Good Loan eligible clients. The long-dashed line shows Emergency Loan uptake as a proportion of eligible clients. The shaded regions show the Aman cropping season. The Good Loan eligibility data set is not usually recorded by BRAC, therefore there is a gap in this data between the 2016 and 2017 Aman seasons when this data was not recorded because of uncertainty about the continuation of the experiment.

Figure 1.5: Yield Per Acre by Emergency Loan Uptake



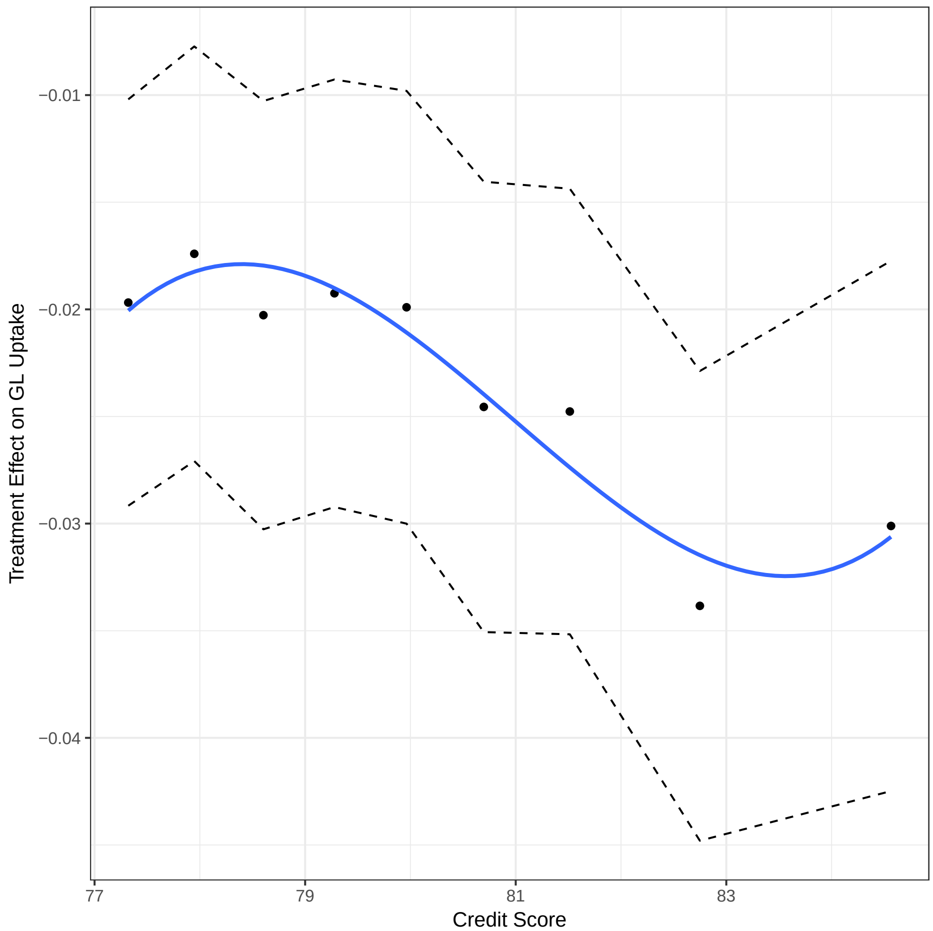
Notes: Histogram of the yield per acre for Emergency Loan takers and non-takers separately. Sample pools data from both 2016 and 2017 and is limited to respondents who were Emergency Loan eligible and located in flooded branches.

Figure 1.6: Emergency Loan Uptake by Credit Score



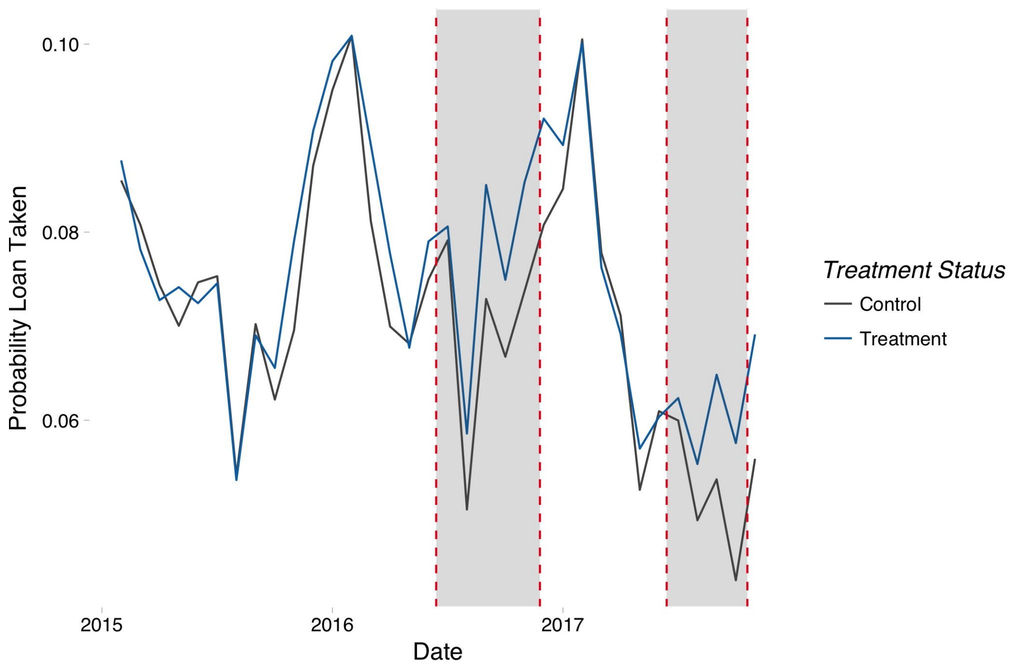
Notes: Plots the probability of Emergency Loan uptake by borrower credit score deciles. The cutoff for Emergency Loan eligibility is a score of 77. Sample pools data from both 2016 and 2017 and is limited to respondents who were Emergency Loan eligible and located in flooded branches.

Figure 1.7: Good Loan Uptake Heterogeneity



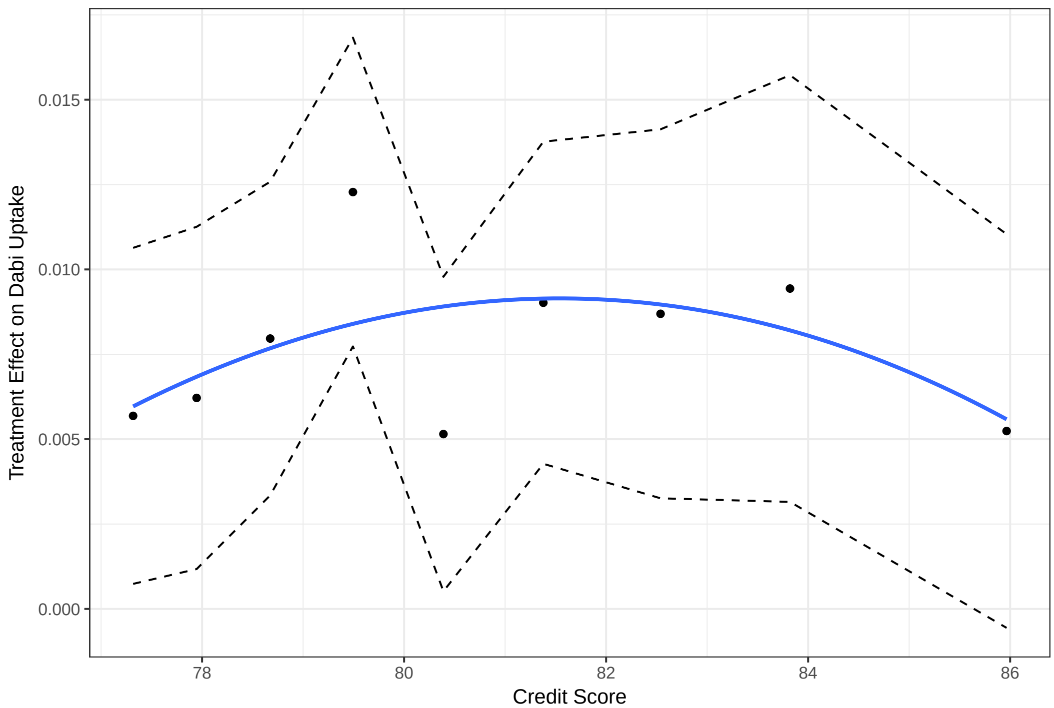
Notes: Plots the treatment effect on the uptake of the Good Loan in treatment branches by decile of borrower credit score. The regression run on each decile includes year and district fixed effects. Sample is comprised of Emergency Loan eligible borrowers who were also eligible for a Good Loan in the pre-flood period. Standard errors are clustered at the branch level. Table 1.18 tests whether the treatment effect heterogeneity is significant.

Figure 1.8: Dabi Loan Uptake Over Time



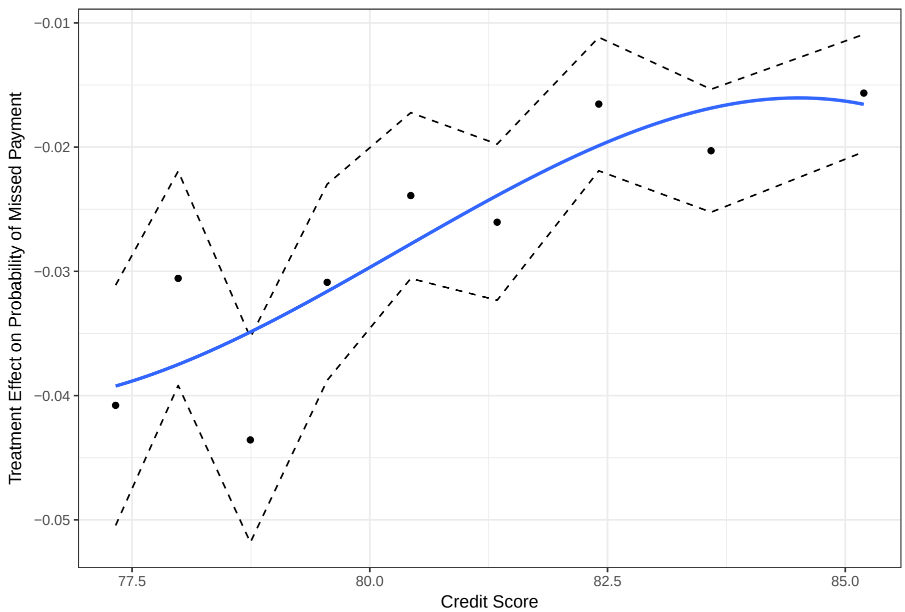
Notes: Plots the probability that an Emergency Loan eligible BRAC member takes a dabi loan in a given month in treatment and control branches separately. Probability of loan uptake is calculated using the complete number of BRAC members in each branch, regardless of whether or not they have a current dabi loan. This is to ensure that the denominator does not endogenously change based on pervious loan uptake decisions. The shaded regions are the “pre-period" before the beginning of the flood season in 2016 and 2017.

Figure 1.9: Dabi Loan Uptake Heterogeneity



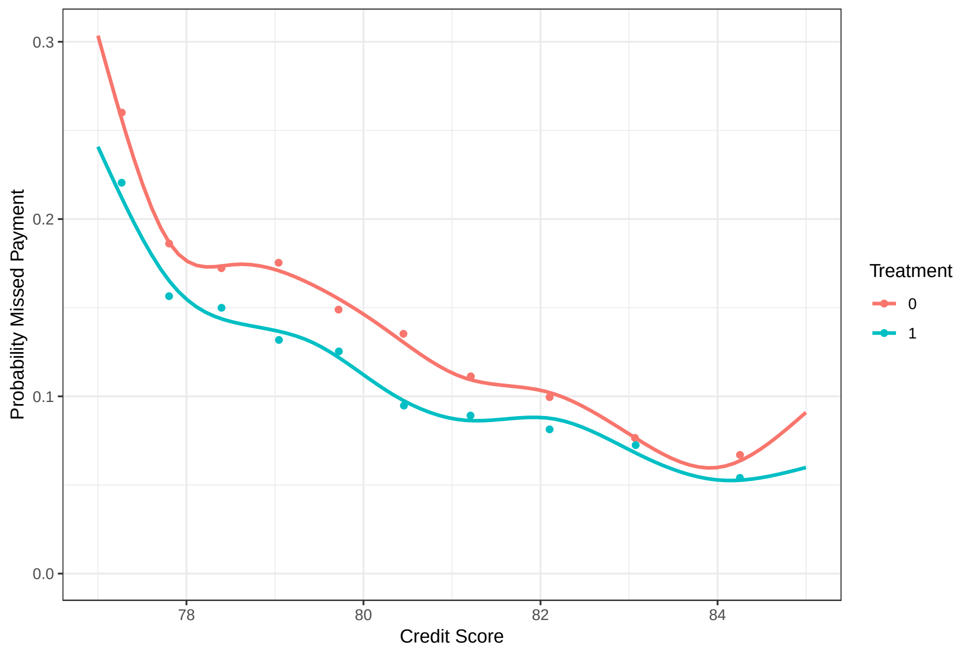
Notes: Plots the treatment effect on the uptake of the Dabi Loan by decile of borrower credit score. The regression run on each decile includes year, month, and district fixed effects. The sample includes only Emergency Loan eligible borrowers. Standard errors are clustered at the branch level. Table 1.18 tests whether the treatment effect heterogeneity is significant.

Figure 1.10: Missed Payment Treatment Effect Heterogeneity



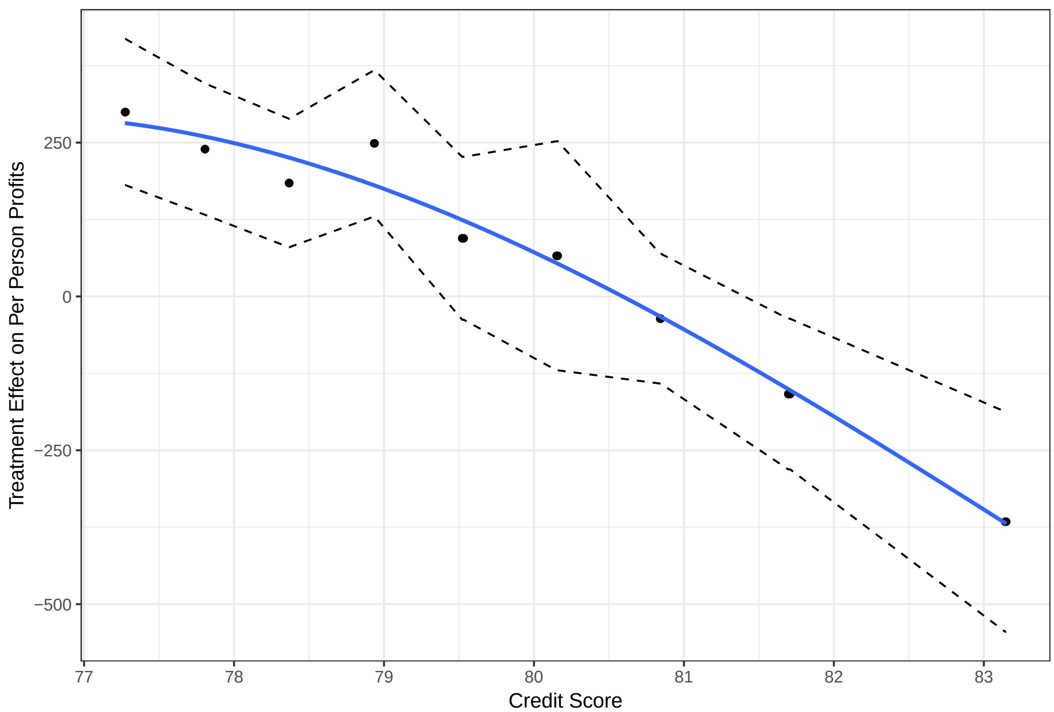
Notes: Plots the treatment effect on the probability of a missed payment by decile of borrower credit score. The estimated treatment effect is the average change in repayment rate across both flooded and non-flooded branches. The regression run on each decile includes year, month, and district fixed effects. The sample includes only Emergency Loan eligible borrowers. Standard errors are clustered at the branch level. Table 1.18 tests whether the treatment effect heterogeneity is significant.

Figure 1.11: Missed Payment Heterogeneity



Notes: Plots the probability of a missed payment by decile of borrower credit score separately for treatment and control branches. The sample is comprised of only Emergency Loan eligible borrowers.

Figure 1.12: Per-Person Profits Heterogeneity



Notes: Plots the treatment effect on per-person MFI branch profits by decile of borrower credit score. Profits are measured in Bangladeshi taka ($1 = 84tk). The regression run on each decile includes year, month, and district fixed effects. The sample includes only Emergency Loan eligible borrowers. Standard errors are clustered at the branch level. Table 1.18 tests whether the treatment effect heterogeneity is significant.

Table 1.1: Eligible Compared to Ineligible

	(1) Ineligible	(2) Eligible	(3) p-value of equality
Household Size	4.788	4.893	0.010
	(0.030)	(0.027)	
Age Head of Household	39.831	40.763	0.004
	(0.246)	(0.208)	
Educ. Head of Household	2.772	2.497	0.001
	(0.069)	(0.053)	
Acres of Land Owned	0.461	0.454	0.868
	(0.021)	(0.032)	
Household Income	1627.133	1560.817	0.042
	(26.429)	(20.100)	
Weekly Expenditure	22.256	22.330	0.873
	(0.344)	(0.305)	
Flooded in Past	0.537	0.543	0.598
	(0.009)	(0.007)	
Electricity Access	0.706	0.717	0.265
	(0.008)	(0.007)	
Asset Count	1.659	1.678	0.418
	(0.018)	(0.015)	
Cows Owned	0.741	0.916	0.000
	(0.023)	(0.021)	
Risk Aversion	0.499	0.513	0.147
	(0.007)	(0.006)	

Notes: Table compares households that were eligible for the Emergency Loan to those who were ineligible in both treatment and control branches at baseline in April 2016. Asset count is the number of items a household reported owning of a gas or electric stove, radio, television, refrigerator, bicycle, and motorcycle. Risk aversion was measured by asking households to choose between a certain payoff and a lottery with increasing odds. The measure ranges from zero to one, where 0=most risk loving and 1=most risk averse.

Table 1.2: Research Timeline

Oct 2015 - Jan 2016	Development of product.
Feb 2016	200 experimental branches selected.
Apr 2016	Baseline survey of 4,000 households; Year one credit scores created; Clients informed about eligibility.
Jun - Oct 2016	Flood monitoring and Emergency Loans made available as necessary.
Dec 2016	Follow-up survey of 4,000 households.
Apr 2017	Year two credit scores created; Clients informed about eligibility.
Jun - Oct 2017	Flood monitoring and Emergency Loans made available as necessary.
Dec 2017	Endline survey of 4,000 households.

Table 1.3: Balance Table

	(1) Control	(2) Treatment	(3) p-value of equality test
Household Size	4.867	4.874	0.910
	(0.047)	(0.046)	
Age Head of Household	40.883	40.374	0.339
	(0.371)	(0.381)	
Educ. Head of Household	2.542	2.464	0.564
	(0.095)	(0.095)	
Acres of Land Owned	0.394	0.436	0.202
	(0.021)	(0.025)	
Household Income	1594.585	1537.005	0.244
	(34.486)	(35.453)	
Weekly Expenditure	21.989	22.191	0.779
	(0.485)	(0.531)	
Flooded in Past Five Years	0.527	0.548	0.250
	(0.013)	(0.013)	
Electricity Access	0.707	0.724	0.326
	(0.012)	(0.012)	
Asset Count	1.724	1.658	0.076
	(0.026)	(0.027)	
Cows Owned	0.887	0.922	0.497
	(0.035)	(0.039)	
Risk Aversion	0.509	0.511	0.905
	(0.010)	(0.010)	

Notes: Table compares households in treatment and control branches at baseline conducted in April 2016 before treatment status was revealed. Asset count is the number of items a household reported owning of a gas or electric stove, radio, television, refrigerator, bicycle, and motorcycle. Risk aversion was measured by asking households to choose between a certain payoff and a lottery with increasing odds. The measure ranges from zero to one, where 0=most risk loving and 1=most risk averse.

Table 1.4: Flood Summary

	Flooded 2016	
Treatment	No	Yes
No	60	40
Yes	49	51

	Flooded 2017	
Treatment	No	Yes
No	27	73
Yes	37	63

Table 1.5: Emergency Loan Uptake

	(1) Took Emergency Loan	(2) Took Emergency Loan
Baseline HH Income	-0.005	
	(0.003)	
Risk Aversion	0.007	
	(0.013)	
Baseline Time Preference	-0.003	
	(0.002)	
Number of Past Floods	-0.008	
	(0.005)	
Ex-post Investment Opportunity		0.021
		(0.016)
Preparation for flood (1=low, 5=high)		-0.026*
		(0.014)
Distress from flood (1=low, 5=high)		0.054***
		(0.014)
Controls	Yes	Yes
District FE	Yes	Yes
Mean Dep. Var	0.03	0.05
Observations	1193	525

Notes: Sample includes only treatment BRAC members who were eligible to take an Emergency Loan in an activated branch. The outcome variable is an indicator for the borrower taking the offered Emergency Loan. Standard errors clustered at branch level. Column 1 shows results predicting Emergency Loan take-up using data collected at baseline. Yearly household income is measured in thousands of dollars. Risk aversion ranges 0 to 1, where 0=most risk loving and 1=most risk averse. Time preference ranges from 1 to 9, where 1 = most impatient and 9 = most patient. Number of past floods is the number of flood shocks experienced by the household over the previous five years (2011-2016). Column 2 predicts Emergency Loan take-up using data gathered at endline and only has observations from 2017. Flood preparation was measured at baseline. Ex-post investment opportunity is an indicator for whether the household reported having a good investment opportunity after the flood. Preparation for flood and distress from flood were self-reported by households.

Table 1.6: Uptake of Good Loan by Emergency Loan Availability

	Took Good Loan		
Treatment	–0.020**	–0.022**	–0.020**
	(0.008)	(0.009)	(0.008)
Farming x Treatment		0.006	
		(0.016)	
Farming Main Activity		–0.007	
		(0.010)	
Flood Risk x Treatment			–0.015***
			(0.006)
Flood Risk			0.011***
			(0.004)
Year F.E.	Yes	Yes	Yes
District F.E.	Yes	Yes	Yes
Mean of Dependent Var	0.130	0.130	0.129
Unique Borrowers	66,232	66,232	63,744
Observations	75,818	75,818	73,282

Notes: Sample is comprised of Emergeny Loan eligible clients who were offered a Good Loan in the pre-flood period. Observations at the month-person level. Data is pooled from both 2016 and 2017. Standard errors clustered at branch level. The outcome variable is an indicator for whether or not the borrower took the offered Good Loan. Farming is a branch level indicator for farming being the major source of income for BRAC members in that branch. Flood risk is measured at the branch level on 1-5 scale where 1 = least risk and 5 = high risk.

Table 1.7: Land Farmed

	(1) Own land	(2) Rented land	(3) Sharecrop land	(4) Total land	(5) Any Cult.
Treatment	0.000	0.063^{***}	-0.004	0.058^{**}	0.044^{*}
	(0.013)	(0.016)	(0.004)	(0.026)	(0.024)
Controls	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var	0.13	0.20	0.02	0.35	0.46
Observations	4744	4740	4743	4739	4745

Notes: Sample includes only eligible BRAC members from both treatment and control groups. Data is pooled from both the 2016 and 2017 Aman season. Controls are included for precision, and are comprised of baseline measures of total land owned, household size, and the age and education of the head of household. Standard errors clustered at the branch level. Land measured in acres. Total land is the sum of own land, rented land, and sharecropped land. Any Cult. is an indicator for whether or not a household planted any crops during the season.

Table 1.8: Ex-Ante Investments

	(1) Fert. Applied	(2) Pest. Applied	(3) Cost Seeds per acre	(4) Input Cost per Acre	(5) Non-Ag Invest
Treatment	6.51	0.26	0.32	2.06	12.13*
	(5.30)	(0.17)	(0.76)	(2.17)	(6.64)
Controls	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var	140.47	1.58	16.18	65.85	38.69
Observations	2183	2140	2058	2017	4745

Notes: Sample includes only eligible BRAC members from both treatment and control groups. Data is pooled from both the 2016 and 2017 Aman season. Controls are included for precision, and are comprised of baseline measures of total land owned, household size, and the age and education of the head of household. Standard errors clustered at the branch level. Fertilizer and pesticide measured in kg/L per acre. Input cost per acre is the sum of the cost of fertilizer, pesticide, and seeds. Cost and investment are measured in dollars.

Table 1.9: Ex-Ante Land by Risk Aversion

	(1) Own land	(2) Rented land	(3) Sharecrop land	(4) Total land	(5) Any Cult.
Treatment	-0.014	0.035	-0.007	0.007	0.037
	(0.021)	(0.025)	(0.006)	(0.036)	(0.031)
Risk Aversion X Treatment	0.020	0.061*	0.006	0.097**	0.013
	(0.031)	(0.036)	(0.009)	(0.049)	(0.041)
Risk Aversion	0.182**	-0.003	-0.008	0.163*	0.075
	(0.071)	(0.053)	(0.011)	(0.089)	(0.078)
Controls	Yes	Yes	Yes	Yes	Yes
Controls X Risk Aversion	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var	0.13	0.20	0.02	0.36	0.47
Observations	4479	4475	4478	4474	4480
p-value Treat + Risk X Treat	0.756	0.000	0.830	0.004	0.131

Notes: Sample includes only eligible BRAC members from both treatment and control groups. Data is pooled from both the 2016 and 2017 Aman season. Controls are included for precision, and are comprised of baseline measures of total land owned, household size, and the age and education of the head of household. Standard errors clustered at branch level. Land is measured in acres. Total land is the sum of own land, rented land, and sharecropped land. Any Cult. is an indicator for whether or not a household planted any crops during the season. Risk aversion was measured at baseline and ranges 0 to 1, where 0=most risk loving and 1=most risk averse.

Table 1.10: Ex-Ante Inputs by Risk Aversion

	(1) Fert. Applied	(2) Pest. Applied	(3) Cost Seeds per acre	(4) Input Cost per Acre	(5) Non-Ag Invest
Treatment	6.44	0.05	1.12	1.68	3.44
	(7.80)	(0.30)	(1.24)	(3.77)	(11.77)
Risk Aversion X	1.64	0.41	-1.34	0.65	16.06
Treatment	(13.18)	(0.43)	(1.78)	(5.41)	(16.62)
Risk Aversion	2.31	-0.96	-4.95	-17.61*	17.31
	(23.93)	(0.79)	(3.65)	(10.18)	(32.25)
Controls	Yes	Yes	Yes	Yes	Yes
Controls X Risk Aversion	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var	138.71	1.53	16.08	65.50	33.08
Observations	2089	2048	1971	1932	4480
p-value Treat + Risk X Treat	0.358	0.060	0.833	0.463	0.028

Notes: Sample includes only eligible BRAC members from both treatment and control groups. Data is pooled from both the 2016 and 2017 Aman season. Controls are included for precision, and are comprised of baseline measures of total land owned, household size, and the age and education of the head of household. Standard errors clustered at branch level. Fertilizer and pesticide are measured in kg / L per acre. Input cost per acre is the sum of the cost of fertilizer, pesticide, and seeds. Cost and investment measured in dollars. Risk aversion ranges 0 to 1, where 0=most risk loving and 1=most risk averse.

Table 1.11: Investment After Shock

	(1) Fert. Applied	(2) Pest. Applied	(3) Total land	(4) Any Cult.	(5) Non-Ag Invest
Treatment	6.689	0.323*	0.055**	0.035	12.559*
	(5.795)	(0.192)	(0.028)	(0.025)	(6.397)
Flood Last Year X Treat	0.053	-0.339	0.021	0.063	0.358
	(23.333)	(0.556)	(0.044)	(0.046)	(24.457)
Flood Last Year	-4.615	-0.383	-0.033	-0.099**	-21.348
	(20.213)	(0.488)	(0.042)	(0.045)	(23.778)
Controls	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var	140.47	1.58	0.35	0.46	38.69
Observations	2183	2140	4739	4745	4745
p-value Treat + Interaction	0.757	0.974	0.069	0.029	0.591

Notes: Sample includes only eligible BRAC members from both treatment and control groups. Data is pooled from both the 2016 and 2017 Aman season. Controls are included for precision, and are comprised of baseline measures of total land owned, household size, and the age and education of the head of household. Standard errors clustered at the branch level. Fertilizer and pesticide measured in kg/L per acre. Total land is the sum of own land, rented land, and sharecropped land. Any Cult. is an indicator for whether or not a household planted any crops during the season. Investment is measured in dollars.

Table 1.12: Ex-Post Outcomes

	(1) Log Cons PerCap	(2) Log Income	(3) Crop Prod. (Kg)	(4) Livestock
Treatment	0.050	-0.024	92.104**	-0.075
	(0.046)	(0.044)	(41.259)	(0.106)
Flood X Treatment	0.058	0.002	-83.157	0.353**
	(0.062)	(0.063)	(51.968)	(0.144)
Flood	-0.046	0.030	-0.831	0.058
	(0.059)	(0.057)	(38.074)	(0.109)
Controls	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Week Interviewed FE	Yes	No	No	No
Mean Dep. Var	5.92	10.77	277.07	1.51
Observations	4699	4489	4701	4701
p-value Treat + Flood X Treat	0.011	0.609	0.800	0.007

Notes: Sample includes only eligible BRAC members from both treatment and control groups. Data is pooled from both the 2016 and 2017 Aman season. Controls are included for precision, and are comprised of baseline measures of total land owned, household size, and the age and education of the head of household. Week interviewed fixed effects are included for the log consumption regression due to the presence of holidays over the course of the survey period that changed standard consumption patterns. Standard errors clustered at branch level. Income is measured in dollars. Flood is an indicator that equals one if flooding occurred and the Emergency Loan was activated.

Table 1.13: Ex-post After Successive Shocks

	(1) Log Cons PerCap	(2) Log Income	(3) Crop Prod. (Kg)	(4) Livestock
Treatment	0.036	-0.023	93.639**	-0.083
	(0.046)	(0.044)	(41.287)	(0.107)
Flood X Treatment	0.107	-0.003	-99.495*	0.379**
	(0.067)	(0.072)	(54.868)	(0.146)
Flood Current Year	-0.051	0.032	5.382	0.056
	(0.059)	(0.060)	(38.331)	(0.108)
Flood Both X Treat	-0.100	0.017	54.321	-0.055
	(0.095)	(0.096)	(44.995)	(0.171)
Flood Both Years	-0.199***	-0.000	-0.260	-0.100
	(0.069)	(0.072)	(41.944)	(0.131)
Controls	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Week Interviewed FE	Yes	No	No	No
Mean Dep. Var	5.92	10.77	277.07	1.51
Observations	4699	4489	4701	4701
p-value Sum Treatment Coef.	0.004	0.904	0.229	0.161

Notes: Sample includes only eligible BRAC members from both treatment and control groups. Data is pooled from both the 2016 and 2017 Aman season. Controls are included for precision, and are comprised of baseline measures of total land owned, household size, and the age and education of the head of household. Week interviewed fixed effects are included for the log consumption regression due to the presence of holidays over the course of the survey period that changed standard consumption patterns. Standard errors clustered at branch level. Income is measured in dollars. Flood Current Year is an indicator that equals one if flooding occurred in the current year. Flood Both Years is an indicator that captures the additional effect of successive shocks for branches that experienced flooding in 2017 and that also experienced flooding in 2016.

Table 1.14: Day Labor

	(1) Daily Wage	(2) Days Worked
Treatment	0.45	-1.57**
	(9.77)	(0.78)
Flood X Treatment	-0.90	2.02*
	(11.76)	(1.17)
Flood	0.91	-4.50***
	(8.40)	(0.94)
Controls	Yes	Yes
District FE	Yes	Yes
Mean Dep. Var	306.75	14.79
Observations	928	2776
p-value Treat + Flood X Treat	0.941	0.573

Notes: Sample includes only eligible BRAC members from both treatment and control groups. Data is pooled from both the 2016 and 2017 Aman season. Controls are included for precision, and are comprised of baseline measures of total land owned, household size, and the age and education of the head of household. Standard errors clustered at branch level. Daily wage is in taka per day. Days worked is the number of days working for others providing farm labor or local construction.

Table 1.15: Dabi Loan Uptake by Emergency Loan Availability

	Loan Uptake
Treatment	0.007***
	(0.002)
Year & Month F.E.	Yes
District F.E.	Yes
Mean of Dep. Var.	0.062
Unique Borrowers	108,446
Observations	462,172

Notes: Sample is comprised of all Emergency Loan eligible clients in the pre-flood period. Observations at the month-person level. Data is pooled from both the 2016 and 2017. Standard errors clustered at branch level. The outcome variable is an indicator for whether or not the client took a new dabi loan in the period before the flood season.

Table 1.16: Repayment by Emergency Loan Availability

	Missed Payment
Treatment	0.011
	(0.024)
Treat x Flood	–0.040*
	(0.020)
Flood	0.039*
	(0.023)
Year & Month F.E.	Yes
District F.E.	Yes
Mean of Dep. Var.	0.096
Unique Borrowers	109,647
Observations	378,216

Notes: Sample includes only Emergency Loan eligible clients. Standard errors clustered at branch level. Observations at the loan-month level. The outcome variable is an indicator for whether or not the client missed a loan payment in a given month. The variable flood is an indicator for anytime after a flood until the following March.

Table 1.17: Branch Profit by Emergency Loan Availability

	Profit (Taka)		
	Per Loan	Monthly Branch	Monthly Per Person
	(1)	(2)	(3)
Treatment	161	76,312	96**
	(233)	(95,405)	(46)
District F.E.	Yes	Yes	Yes
Month F.E.	No	Yes	Yes
Mean of Dep. Var.	2,823	1,745,794	2202
Observations	106,695	3,706	3,706

Notes: Sample for column 1 includes loans made only to Emergency Loan eligible clients. Standard errors clustered at branch level. The outcome for column 1 is the measured profit in Bangladeshi taka ($1 = 84 taka) for a given loan assuming an annual cost of capital of 6% for the MFI. The outcome for column 2 is overall branch profitability. The outcome in column 3 is overall branch profitability divided by the number of branch members. Observations in column 1 are at the loan level and for column 2 and 3 are at the branch-month level.

Table 1.18: MFI Outcomes by Credit Score

	Good Loan Uptake	Dabi Uptake	Missed Payment	Per Person Profit
	(1)	(2)	(3)	(4)
Treatment	-0.020^{*}	0.008^{***}	-0.027^{**}	169^{**}
	(0.011)	(0.002)	(0.013)	(78.034)
Credit Score x Treatment	-0.003^{*}	0.000	0.004^{*}	-25^{*}
	(0.002)	(0.0002)	(0.002)	(14.7)
Credit Score	0.004^{***}	−0.0001	-0.010^{***}	13^{**}
	(0.001)	(0.0002)	(0.002)	(5.740)
District F.E.	Yes	Yes	Yes	Yes
Month F.E.	No	Yes	Yes	No
Year F.E.	Yes	Yes	Yes	No
Mean of Dep. Var.	0.13	0.062	0.096	2202
Observations	37,392	396,228	910,862	40,514

Notes: Sample includes only Emergency Loan eligible clients. Standard errors clustered at branch level. The outcome in column 1 is the probability of taking an offered Good Loan among Good Loan eligilbe clients in the pre-flood period. The outcome in column 2 is the probability of taking a Dabi Loan in the pre-flood period. The outcome in column 3 is the probability of missing a loan payment in a given month. The outcome in column 4 is the measured profit in Bangladeshi taka per branch member assuming an annual cost of capital of 6% for the MFI.

Table 1.19: Savings Transactions by Emergency Loan Availability

	Savings Transctions		
	Pre-Period	All	All (2017)
	(1)	(2)	(3)
Treatment	8.85	−14.58	−55.73
	(9.34)	(18.57)	(43.11)
Treat x Flood		45.37**	34.75*
		(20.67)	(20.75)
Flood		−53.75**	−50.19**
		(24.60)	(22.19)
Flood Damage x Treatment			11.58
			(10.05)
Flood Damage			−17.15***
			(6.42)
Year & Month F.E.	Yes	Yes	Yes
District F.E.	Yes	Yes	Yes
Mean of Dep. Var.	82.6	71.8	64.5
Unique Accounts	108,446	109,647	75,477
Observations	622,551	1,150,895	711,184

Notes: Sample includes only Emergency Loan eligible clients. Standard errors clustered at branch level. Observations at the person-month level. The variable flood is an indicator for anytime after a flood until the following March. Column 1 uses observations only from the pre-flood period in both 2016 and 2017. Column 2 uses all observations. Flood damage data at the branch level is only available for 2017, therefore column 3 shows results only for this year. Flood damage is measured at the branch level and ranges from [1-5] with 1=least damage and 5=most damage.

2 | The Impact of Monitoring Technologies on Contracts and Employee Behavior: Experimental Evidence from Kenya's Transit Industry

Chapter abstract: Agency theory suggests that moral hazard in employer–employee contracting constrains firm profits. We use a randomized controlled trial to empirically evaluate how information and communication technologies (ICT) can mitigate moral hazard and enable firms to design more efficient contracts which increase profits and engender business growth. Specifically, we study a fleet of 255 minibuses (matatus) in Nairobi, Kenya, where we introduce monitoring devices that track real-time vehicle location, daily productivity, and safety statistics. We randomize whether minibus owners have access to these monitoring data using a novel mobile app. This information allows owners in the treatment group to observe a more precise signal of driver effort, the amount of revenue drivers collected in fares, and the extent to which the driver engages in reckless driving. We find that treated vehicle owners modify the terms of the contract by decreasing the rental price they demand. Drivers respond by working more hours, decreasing behavior that damages the vehicle, and under-reporting revenue by less. These changes improve firm profits and reduce management costs, thereby helping treated firms grow. The device also improves owners' trust in their drivers, which drivers say makes their job easier. Finally, we investigate whether these gains to the company come at the expense of passenger safety, in an environment where accidents are common. While we do not find any evidence that conditions deteriorate, offering detailed information on driving behavior also does not *improve* safety. Only by incentivizing drivers through an additional cash treatment do we detect safety improvements.

2.1 Introduction

Firms design contracts to ensure their employees exert the profit-maximizing level of effort. In the presence of moral hazard, however, firms cannot condition the terms of the contract on important dimensions of employee behavior, including effort and output. Firms respond by relying on “second-best" self-enforcing incentive contracts or coercive measures to align agents’ interests with their own (Hölmstrom, 1979; Grossman and Hart, 1983; Hart and Holmstrom, 1987; Shapiro and Stiglitz, 1984). In theory, firms can overcome these frictions by investing in monitoring technologies that reduce information asymmetries and reveal the performance of their workers more accurately (Harris and Raviv, 1979; Hölmstrom, 1979; Hubbard, 2003). In practice, however, the impact of these technologies on contracts is unclear: the presence of institutional and managerial frictions may limit employers’ ability to leverage the additional information needed to change the contract and employees’ behavior.

This paper studies the impact of moral hazard on labor contracting, and productivity, and the extent to which improved monitoring eases these frictions. We also investigate whether monitoring technologies subsequently improve firm profits, and worker well-being. We implement a randomized control trial where we introduce a novel monitoring device to a subset of firms operating in Kenya’s transit industry. The industry is dominated by thousands of small-scale entrepreneurs who own a few minibuses (“matatus") that run on designated routes. These matatus are the only reliable form of transportation and serve 70% of Nairobi’s four million commuters daily.[1] We recruited 255 owners operating along 9 major commuter routes to participate in the study, and we randomly selected 125 to be part of the treatment group. The monitoring device was fitted to all the matatus in our sample, but only transmitted data to minibus owners in the treatment group.[2] We developed our own device because available alternatives on the market were either too costly, or not sophisticated enough. Our device records and transmits via a mobile app the location of the vehicle, the number of kilometers driven, and the number of hours the ignition was on. While the owner does not know the number of passengers that boarded the vehicle, they can use this information to monitor drivers’ operations throughout the day and to gain a more precise estimate of total daily revenue.

The contracting environment we study here is not unique to Kenya or transportation, as the dynamics that characterize this space are prevalent in many other settings, including agriculture and the service industry. First, employers (owners) cannot observe the amount of revenue their employees (drivers) collect, nor the amount of effort drivers invests. Second, drivers in this setting are from relatively poor households, and they cannot afford to walk away without pay on days when total revenue is low nor can they pay for repairs when the vehicle is damaged,

[1] Similar transit systems are present in Mexico (peseros), the Philippines (jeepney’s), Indonesia (tuk-tuks), India (rickshaws), and Tanzania (dala-dala’s), among others.

[2] The drivers were present during the installations of the devices, but they were not informed about whether the owner was in the treatment group or not.

meaning that in practice they have limited or no liability to the owners. Drivers are known to run away from accidents so they can avoid being held accountable by owners or the police. In light of these constraints, firms have overwhelmingly opted for fixed-rent contracts (locally referred to as a “target" contract) with limited liability. The owner specifies an amount of revenue that the driver must deliver by the end of the day, net of fuel expenses. According to the contract, the driver should deliver the fixed rent (“target") amount if the revenue they collect exceeds the target (keeping any revenue they earn above it). If the earnings are below the target, the contract stipulates that the driver must hand over all of the revenue they earned.

In order to understand the impact of the device, we adapt a standard principal-agent framework to reflect the actual employer-employee relationships within this network. In the absence of a monitoring technology, the model predicts that drivers will engage in a number of behaviors that are sub-optimal from the firm's perspective. First, drivers under-report revenue so they can be sure to walk away with slightly more income than they would otherwise. Note: rampant cheating is kept in check by owners who threaten to punish and ultimately fire drivers who are caught under-reporting. Second, drivers under-supply effort on days when they are unlikely to make the target price. On these particular days, drivers know they will not be the residual claimant and reap the benefits of higher effort. Third, drivers engage in more damaging driving than what owners would optimally choose. Damaging driving refers to the maneuvers drivers make that may damage the vehicle. These actions include driving on the shoulder of the road, or veering off the designated route onto roads that are more bumpy and damaging to the vehicle. The driver engages in such damaging behavior because he reaps the benefits in terms of higher revenue, without bearing any of the downside risk (the limited liability constraint binds).

We model the introduction of the new technology as increasing: 1) the precision of the owner's signal about total revenue and 2) the probability the owner detects damaging driving. This has implications for the owners' choice of contract and drivers' behavior, which we test in our data. According to the model, the monitoring technology reduces drivers' information rents and lowers their utility. Owners recognize this outcome and compensate by reducing the target. Empirically, we have some suggestive evidence that owners steadily reduce the target throughout the study period. By the last month of the study, the target set by treatment owners is approximately 4.1% lower than the target set by control owners.

The model then predicts that drivers' behavior is affected along three key dimensions. First, the drivers' incentive to lie about the total amount of revenue they collected is reduced, which means that we should see lower under-reporting. We confirm this prediction in our data: under-reported revenue falls by approximately 100 shillings (1 USD) per day (a 16% decrease). Second, the model predicts that drivers will increase their effort to compensate for the income they lose from lower under-reporting. In parallel, as the target falls, drivers also have an incentive to increase their effort because they can become the residual claimant more easily. We capture a precise measure of effort through the tracking device, which powers on and off with the matatu, and consequently find that the number of hours the vehicle is on the road increases by 1.4 hours per day (a 9.9% increase). Finally, we expect the driver to reduce instances of damaging driving because they are more likely to be caught by the owner. We proxy damaging driving by the amount of repair costs the owner incurs, and we find these decrease by 200 shillings (2 USD) per day (46%) by the last month of the study. We have evidence to suggest this outcome comes from fewer instances of driving on alternate routes that are bumpy.

Next, we investigate the effect of these changes on firm profitability. We find that profits increase by 13%, which is primarily driven by lower repair costs. 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. Owners also report that monitoring their drivers has become significantly easier, and they trust their drivers more, consistent with a reduction in management costs. Subsequently, we ask whether this improved profitability and better management fueled business growth. We find that treatment owners have 0.145 more vehicles (11% increase) on average than control owners by the end of the study. This suggests that inadequate monitoring may represent an important barrier to firm growth in low-income countries.

It is also important to investigate whether these benefits to the firm come at the expense of their workers. The introduction of ICT has generated some debate in low-income countries. On the one hand, there is a concern that these technologies concentrate all of the bargaining power in the hands of the employer. We do see some evidence of this as the amount of revenue drivers can under-report falls, and drivers work more (although their salary per hour stays the same). However, proponents of these new technologies suggest that they increase employers' trust in their employees, which makes for better managers (Pierce, Snow, and McAfee, 2015). We have some suggestive evidence of this in our data. In a qualitative survey we conducted six months after the experiment concluded, we find that 65% of drivers said the tracking device made their job easier (26% said nothing changed). This suggests that the effects of new technologies on worker well-being are nuanced, despite the net benefit they represent for firms.

These first results demonstrate how alleviating moral hazard affects operations *within* the firm. However, the presence of monitoring devices can also have effects *outside* of the firm, as profit-maximizing behavior by firms and their employees may impose negative externalities. In public transportation systems, monitoring technologies are often used to check and limit instances of unsafe driving. Kenya's matatu sector is notorious for its poor safety standards: drivers often over-accelerate, speed, stop suddenly, and turn sharply in order to collect more passengers.[3] Our monitoring technology records these four instances, in addition to maximum and average speeds, and conveys them to owners through a separate tab in the mobile app. It is important to note that these traditional measures of

[3]Matatus account for 11% of registered vehicles but 70.2% of passenger casualties (Macharia et al., 2009). Buses in the US account for 1% of registered vehicles and 0.4% of casualties (BTS, 2016).

unsafe driving need not be perfectly correlated with the damaging driving behavior outlined above. For example, drivers often choose to bypass slow-moving traffic by taking rough, unpaved roads that harm the vehicle but pose no safety danger. Alternatively, driving quickly through crowded pedestrian areas may be unsafe, but it is unlikely to cause significant damage to the vehicle.

A priori, it was not clear how the owners would use the safety information we provided. On the one hand, owners may internalize the dangers associated with unsafe driving because, unlike their drivers, they bear the cost if their vehicle gets into an accident or receives a fine.[4] In that case, they might use the information to induce drivers to limit behaviors that result in these accidents and fines. On the other hand, owners may perceive the risk of accidents to be low and disregard this information. Alternatively, while owners care about damages to the vehicle, these may only be weakly correlated with unsafe driving (as detailed above). If they use the device to incentivize effort only, the number of safety violations could worsen over the study period, producing a negative externality for Nairobi's commuters.

Despite all the safety information we provided, the frequency of unsafe driving events flagged by the device does not change significantly, and instances of speeding remain the same. It follows that the gains to firms do not come at the expense of commuters. However, these results also suggest that external intervention may be necessary to improve safety. We tested the efficacy of one such intervention by providing small cash incentives to drivers conditional on safe driving. This treatment was designed to mirror the actions that a regulatory body could potentially take in this setting.[5] Our objective was to determine the effectiveness of an intervention that encouraged the employees (drivers) rather than the employers (owners) to internalize the negative externalities produced by the business. We find that the cash incentives meaningfully reduce safety violations committed by drivers, confirming that third-party intervention can successfully address these firm externalities. However, these effects do not persist after the removal of the cash incentives, suggesting that further action or permanent regulation is needed to induce long-lasting change.

This paper contributes to three different literatures. First, the paper speaks to the vast theoretical work on principal-agent relationships and contract formation, which predominantly focuses on deriving the optimal contract subject to various constraints (Hölmstrom, 1979; Grossman and Hart, 1983; Hart and Holmstrom, 1987; Shapiro and Stiglitz, 1984). Our paper, by contrast, empirically demonstrates how contracts change when these constraints are alleviated via monitoring technologies. Measuring the impact of monitoring is challenging because shirking behavior is hard to detect by design, a firm's decision to monitor is often not random, and data on firm operations are difficult to obtain. There are only two other papers to our knowledge that overcome these limitations. Baker and Hubbard

[4]The owner may also have limited liability as they are protected by insurance. However insurance does not always cover full repairs, and premiums will rise if the owner continues filing insurance claims. The owner therefore always absorbs much more liability in the event of an accident, which means that owners should prefer less accidents and fines on average than the driver.

[5]South Africa's Ethekwini municipality is testing one such intervention in the coming months (Payet, 2018).

(2003, 2004) investigate how the introduction of onboard diagnostic computers (OBCs) change ownership patterns in the U.S trucking industry. Baker and Hubbard (2004) demonstrate that shipping companies respond to the introduction of OBCs by hiring drivers to operate their vehicles (rather than working with drivers who already own their own trucks). Our paper differs from this existing work in a number of ways. We generate exogenous variation in the usage of monitoring technologies by randomizing which companies receive data from a tracking device. We also capture high frequency data on contracts and worker behavior. This allows us to monitor how different dimensions of the contract, and worker performance, change over time. We can then document the impact on firm profits and worker well-being (salary per hour, hours worked, sense of trust).

These papers are also concentrated in developed countries, and we have reason to believe that the impacts of monitoring could be different in low-income countries. Management quality is different, and employers may not use the information effectively (Bloom et al., 2013; Bloom, Sadun, and Van Reenen, 2017). Employers also face additional frictions that may limit their ability to use the information: law enforcement is weak and limited liability constraints bind. Contrary to Baker and Hubbard (2004), we do not find that ownership patterns change as a result of monitoring, precisely *because* of existing constraints (drivers cannot afford their own vehicles). Nevertheless, we do find that firms use the technology to change the terms of the contract, and to induce their employees to behave in a way that aligns with the firms' best interest. There is only one other paper to our knowledge that investigates the impact of monitoring technologies within firms in a developing country: de Rochambeau (2018) studies the use of GPS devices by managers in Liberia's long-range trucking industry. She finds that monitoring technologies crowd out high-performing workers' intrinsic motivation.[6] Our analysis builds on this work by investigating how monitoring technologies alleviate other key dimensions of moral hazard (including damaging driving, and lying about total revenue). We also focus on the impact of these devices on contracts.

Second, our findings add to the literature investigating the barriers to firm growth in low-income countries. Identifying the constraints to firm growth is a question of great policy relevance given the large contribution these firms make to emerging economies. Empirical research on small firm growth has identified three key challenges facing firms: credit constraints, labor-market frictions, and managerial deficits (Bloom et al., 2014). Our paper most closely resembles the work on managerial deficits, which refers to the difficulties firms face managing the day-to-day operations (including financial accounts and inventories), and incentivizing and monitoring workers. Most of the work in this field studies the impact of business training programs (Bloom et al., 2013; Bloom, Sadun, and Van Reenen,

[6]Note there are additional studies that document the impacts of monitoring in low-income countries - but they do not focus on the employer-employee relationship within the firm. Duflo, Hanna, and Ryan (2012) find that teacher absenteeism in India decreases when their attendance is monitored; Björkman and Svensson (2009) demonstrate that community health workers exert more effort when their performance is scrutinized by the community; and Duflo et al. (2013) find that incentives for third-party auditors can improve their reporting.

2017; McKenzie and Woodruff, 2016; Berge, Bjorvatn, and Tungodden, 2014; de Mel, McKenzie, and Woodruff, 2014; Valvidia, 2012). These interventions provide information about how to manage aspects of the business that do not involve employees (maintaining business records, separating finances, inventory, controlling for quality, marketing). In contrast, our paper focuses on the role of moral hazard, and how providing information specifically about *employees' behavior* can change firm operations. We find monitoring technologies improve firm profits and reduce management costs, which helps the treated firms grow. As prices fall, these technologies are becoming increasingly prevalent, making their impacts important to understand.

Finally, our results on damaging driving and traditional safety metrics contribute to a growing empirical literature on policies that promote compliance with government regulation - in this case with safety regulation. In recent years, international institutions have provided funding, knowledge and technical assistance to build systems aimed at reducing the number of traffic injuries and deaths worldwide (World Bank, 2014).[7] These efforts are difficult to evaluate because the investments are multi-faceted and typically rolled out across an entire city. One exception is a program that was launched in Kenya, which placed stickers inside Nairobi's matatus to encourage passengers to complain to their drivers about unsafe driving (Habyarimana and Jack, 2015). They find that the intervention reduced accidents by 25-30%. Our intervention complements their approach by asking whether drivers, in addition to passengers, can be incentivized to improve safe driving. We find that owners with access to the monitoring device do not internalize the negative externalities produced by their drivers. We only document improvements in safety when we directly incentivize drivers. This result suggests that investments in technologies that monitor unsafe driving may be more effective when combined with external incentives.

The remainder of this paper is organized as follows: Section 2.2 discusses Kenya's transportation system, the prevalence of moral hazard, and the scope for monitoring. Section 2.3 details the field experiment, and Section 2.4 reviews the data. We present a simple theoretical framework in Section 2.5. Section 2.6 discusses each of our results. We then discuss the implications of the findings and conclude in the final section.

2.2 Context

2.2.1 Nairobi's Matatus

Nairobi's transportation system was developed after Kenya's independence in 1963 (Mutongi, 2017). Small-scale entrepreneurs responded to the growing demand for mobility by retrofitting old vehicles and transporting passengers from the suburbs

[7]According to the Global Status Report on Road Safety, 1.24 million people are killed in traffic accidents each year and 90 percent of these deaths occur in low- and middle-income countries (LMICs)

to the urban center. The buses were labelled “matatus", meaning three in Kikuyu, in reference to the early ticket price in Kenyan Shillings (KES) of a matatu ride (where 100 KES = 1 USD). These private businesses were legalized in 1973, but remained largely unregulated until 2003 when the government passed the Michuki rules, requiring that buses install speed limiters, safety belts, and ensure that all drivers exhibit valid licenses (Michuki, 2003). To date, these regulations are rarely enforced. In 2010, the Ministry of Transport issued a new directive to further formalize the industry and eliminate the presence of gangs that were becoming increasingly active in the sector. This required that all minibus owners form or join transport Savings and Credit Cooperatives (SACCOs) or transport companies licensed to a particular route (McCormick et al., 2013). At present, industry newcomers must first register with a SACCO or transport company before they can put their vehicle on the road. Transport companies are rare in Nairobi and manage buses on behalf of individual investors. SACCOs on the other hand leave the daily management of the vehicle to the owner, but facilitate centralized organizational activities including scheduling, resolving internal disputes between owners, ensuring compliance with the National Transport and Safety Authority (NTSA) regulations, and providing financial services to owners and drivers.

This informal network of buses constitutes the only dependable transit system in Nairobi, and the city comes to a near standstill on days when drivers strike. Rough estimates suggest that 15,000 to 20,000 buses currently circulate throughout the city, swerving on and off the road to collect passengers along their designated route. The industry remains almost entirely locally owned: private entrepreneurs purchase 14 or 33 seat minibuses, and hire a driver to operate the vehicle along their SACCO’s designated route. The presence of severe competition within a route explains the dangerous driving that prevails throughout the industry. According to the World Health Organization’s Global Status Report on Road Safety, approximately 3,000-13,000 people die annually from traffic incidents in Kenya, and at least 30% of cases involve matatus (WHO, 2015). Conditions have not improved measurably in recent years. However, in an effort to combat negative stereotypes, matatu owners are increasingly investing in the comfort of their vehicle, the aesthetic (colorful interior and exterior), the quality of the “experience" (helping passengers on and off the bus), and the perks (TV’s) (Reed, 2018). There are no regulations placed on the aesthetic of the vehicle. Nevertheless, the more attractive and comfortable vehicles can charge up to twice the price of regular ones. Matatu fares vary between 0.5 and 1.5 USD for travel inside the city center, and between 1 and 5 USD for trips to the outskirts.

This transportation industry is appealing to study for a number of reasons. First, it is representative of many other informal transit systems worldwide, including Tanzania’s dala dala’s, Haiti’s tap tap’s and India’s rickshaws, among others. Moreover, the sector is economically meaningful in terms of the number of individuals it employs and the amount of income it generates. In Kenya, estimates suggest that the industry employs over 500,000 people and contributes up to 5% of the country’s GDP (Kenya Roads Board, 2007). Most importantly, this context allows us to overcome major data constraints that have limited previous research in the space. Namely, we collect detailed information on the contract terms set by the employer and the actions of the employee (their choice of effort and lying). We also introduce exogenous variation into the costs of monitoring in order to observe changes to the contract.

2.2.2 Driver and owners

In this study we work exclusively with small firms that meet three basic criteria. First, the owners of these vehicles manage their matatu themselves, as opposed to hiring a third party manager. Second, the owners are not the primary drivers of the vehicle.[8] These two conditions were designed to focus the research on the classical principal-agent relationship. Finally, we only worked with owners that had a single matatu at the time of recruitment. We chose single owner-driver pairs to remove any dynamics that arise from one driver reporting differently from another, sending competing signals for the owner to parse through. According to an exploratory survey we conducted in the pre-pilot phase of the experiment, approximately 25% of matatu owners in the general population meet these three criteria.[9]

In this industry, owners have settled on a fixed rent contract with limited liability that is negotiated daily. Owners rent their vehicles to a driver every morning for an agreed upon “target price" (henceforth referred to as the ‘target’). Unlike the taxi systems in many high-income countries, the driver is expected to deliver this amount at the *end* of the day once all the fares have been collected. This is primarily because drivers have limited capital and cannot afford to pay the amount up front. Drivers are the residual claimants in this contract and keep everything they earn above the target. The owner is not allowed to revise the terms of the contract and claim more revenue if the driver has had a good day. In the event that the driver cannot make the target, they are supposed to provide the total revenue they earned to the owner. In practice, drivers under-report total revenue to make sure they have some income left over. If they fail to make the target too many times, or they are caught under-reporting too frequently, they will be fired. In addition to choosing the amount of revenue they declare (under-reporting), drivers decide the number of hours they work (effort), their driving style (including driving maneuvers that may damage the vehicle). Owners cannot directly observe these three actions by the driver (under-reporting, effort, damaging driving), and must resort to costly monitoring techniques. This includes phone calls, dropping by the terminus of the route and staging someone at various stops to monitor whether the bus drives by.

This negotiation process is repeated daily over the phone (and occasionally in person). Formal documents are not signed because legal recourse is virtually non-existent. Typically owners and drivers have worked with each other for just over two years. The target price is set at approximately 3000 KES (30 USD) every day, and

[8]Owners do not operate the vehicles themselves for two reasons. First, it allows them to pursue side-jobs that are more lucrative than being a driver. Second, driving is a tough job that individuals like to avoid if they have other options.

[9]If we allowed owners to posses two or three matatus, over 50% of matatu owners satisfied these conditions.

drivers make this target 44% of the time. On days when they do not report making above the target, they under-report revenue by approximately 700 KES. Drivers are typically on the road between 12-14 hours per day, and make approximately 10 trips to and from the city center.

This contract structure appears to be one of the only viable alternatives in this industry. A fixed wage payment is unattractive to most owners because drivers face incentives to undersupply effort when they cannot be monitored. The few SACCOs that have adopted this payment scheme have hired full time managers who supervise the drivers closely. Anecdotal evidence suggests that drivers also dislike this remuneration scheme because it eliminates the large windfall they receive on the best days. Next, a fixed rent contract is impossible to enforce because drivers are poor and hence the limited liability constraint always binds. This leaves the traditional sharecropping model or a fixed rent contract with limited liability. A sharecropping contract in this industry would have to take the form of a profit-sharing agreement where owners and drivers are each responsible for their share of the costs. However, mosts of the costs that the vehicle incurs are beyond the means of a matatu driver. A typical service fee is 2000 KES, which the driver simply cannot pay upfront. Similarly, in the extreme case that a matatu gets into an accident, drivers are known for running away from the scene. Moreover, a sharecropping model with unobserved output means that drivers can consistently under-report the amount they collect. The cost of under-reporting is low because drivers can easily hide undisclosed revenue. These limitations may explain why owners find this contract structure is less attractive.

A fixed rent contract with limited liability ensures that drivers face incentives to supply effort when they earn more than the target. It also limits under-reporting to days when the driver does not report making the target (there is no incentive to under-report on good days because they are the residual claimant). Note that due to the limited liability constraint, the supply of effort under this contract scheme will be less than the first-best outcome because the driver does not supply optimal effort on days when they do not expect to reach or exceed the target. This contract structure is prevalent in many informal transportation systems worldwide. It also characterizes relationships in agriculture where absentee landlords cannot supervise their tenants; in the service industry where employers cannot record the number of services provided by their employees; and in businesses where inventory is difficult to monitor.

2.2.3 Device (Hardware and Software)

Monitoring technologies are becoming widely available in many developing countries, including Kenya. The majority of long-range bus companies that travel between the country's main cities are equipped with tracking devices. Moreover, some banks in Nairobi recently announced that they would only issue loans for minibuses whose location could be tracked with a device. Despite their availability, most medium range buses and inner-city public transportation vehicles are not yet using them. When asked why, most vehicle owners cite the high cost of sophisticated tracking systems (approximately 600 dollars for the tracker and additional monthly installments for system access), or the lack of detailed information provided by the cheaper alternatives.

To fill this need, the research team created a new monitoring system for city buses that is considerably cheaper, more flexible and more powerful than traditional tracking devices. The physical tracking units were procured for 125$ from a company in the United States (CalAmp). They feature GPS, internal back-up battery packs, 3-axis accelerometer for motion sense, tilt and impact detection. The device was designed to capture and transmit the information we required, including the 95th percentile and average forward/backward/lateral/vertical acceleration, as well as the 95th percentile and average forward/backward jerk. The device was also calibrated to generate alerts for every instance of vehicle speeding, over-acceleration, sharp braking and sharp turning. These safety alerts were calculated by an internal algorithm built into the CalAmp device with threshold parameters as inputs, using the full sequence of acceleration and speed data to identify unsafe driving actions. Further processing of the CalAmp system data on the server provided additional measures of interest including the total number of kilometers traveled that day, the total time the matatu was running, and a safety index (from aggregating the day's safety alerts). Finally, an API call was generated each time the owner used the app to request data from the server. These calls were recorded in the database and provided a measure of owners' usage of the app. In this way, we could track which types of information the owner found most valuable and how often the owner requested this information.

The data captured by the CalAmp device was transmitted to owners via a mobile application that was specifically designed to present information simply. The app (referred to as "SmartMatatu") provided information in three ways (Figure 2.1). The first tab was a map of Nairobi and presented the real-time location of the vehicle. By entering a specific date and time interval into the phone, the app would display the exact routes traveled by the matatu over this time period. This first tab provided owners with a more accurate measure of driver effort because they could track where the driver was at any point in time. It also conveyed a more accurate measure of damaging driving because they could see if the driver was operating on bumpy routes. The second tab displayed all the safety alerts that were captured by the device. The owner could click on the safety event to find out when and where it had occurred on the map. It is important to remember that these safety alerts are not necessarily correlated with damaging driving behaviors such as off-route driving. The final tab conveyed a summary of the driver's productivity and safety. The productivity section of this page listed the total mileage covered, and the duration the ignition was turned on that day. This could be used by owners to estimate total revenue more precisely. The safety section of this page also provided the owner with an overview of the number of safety violations that occurred that day, as well as the driver's daily safety rating relative to all other drivers on the road that day (where a thumbs up appeared for scores of 60% and above, a sideways thumb for scores between 40% and 60%, and a thumbs down for scores of 40% and below).

2.3 Experimental Design

2.3.1 Sample Recruitment

We conducted an extensive recruitment drive in late 2015 by contacting 61 SACCOs that were operating along various routes across the city. We organized several large meetings with matatu owners in each SACCO, presenting the study's goals and methodology. All owners were informed at the time of recruitment that a monitoring device would be placed in their vehicle free of charge, and they would be required to *provide* daily information about their business operations. We also mentioned that a random subset of owners would be selected to *receive* information from the tracker via a smartphone app for a six month time period, while others would have to wait 6 months before gaining access to the information for a shorter two month period. It took approximately four months to recruit enough participants across 9 major commuter routes (Figure 2.2). Owners who expressed interest in the study during the recruitment drive were contacted again over the phone to confirm their willingness to participate in the experiment, and to check that they met the three study requirements (owners had to own a single matatu, which they rented to a driver, and manage the firm's operations themselves).

2.3.2 Installations

The first installation took place in November 2016, and continued until April 2017. The field team, managed by EchoMobile, was able to fit approximately 15 matatus per week with a device (Figure 2.3). The team scheduled a time to meet each owner individually at a location of their choosing. The owner was compensated for the time their vehicle spent off-road to perform the installation of the device with a one-time payment of 5000 KES (50 USD) and a new Android phone (worth approximately 80 USD) to ensure they could access the SmartMatatu app. The installation process was rather complex, requiring a team of three staff (an enumerator, a field manager, and an engineer). While the mechanic worked on fitting the device in the matatu, the field manager took the owner aside to re-explain the purpose of the research project and the tracking device's functionality. For owners in the treatment group, the field manager conducted an additional training on the app. At the same time, the enumerator administered the baseline survey to the driver in a separate location, outside of the owner's earshot, so that the driver felt comfortable answering the questions honestly. Once the field manager finished the training with the owner, and the enumerator finished administering the survey to the driver, they switched. The field manager then took the driver aside to explain that they would receive a daily SMS to elicit information about the day's operations and to emphasize that all of the data they shared would remain confidential. Meanwhile, the enumerator conducted a 20-minute baseline survey with the owner. This whole installation process took approximately 1 hour to complete. The field manager shared his contact information with the owner and the driver so they could contact him with any further questions they had.

2.3.3 Treatment Assignment

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

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

2.4 Data Collection

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

2.4.1 Non-system application variables

The SmartMatatu app was also designed to collect information from owners. Collecting accurate data can be very challenging in these settings, and this system was created to improve the quality of the data we received. Owners in the study were reminded daily via a notification on their phone to report on that day's business activities through a form located on the app. They were asked to submit data on: the “target" amount assigned to their driver at the beginning of the day; the amount the driver delivered to the owner; any repair costs incurred; an overall satisfaction score for their driver's performance (bad, neutral, good); and whether the driver was fired/quit that day. Once the report was successfully submitted, owners received 40 KES via M-Pesa (a mobile phone-based money transfer service). We collected similar information from drivers through SMS surveys (because the drivers were not provided with smartphones). At the end of every work day around 10pm, drivers would receive a text message asking whether they were ready to respond to the survey. Once they agreed, individual text messages were sent to the driver asking for: the total revenue the matatu collected from fares that day; the amount they spent on fuel; and their “take home salary" (their residual income after expenses and paying the owner). Once the driver responded to all the questions, they were sent 40 KES via M-Pesa to incentivize consistent reporting.[10]

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

[10]Note we do not see different submission rates, or differential reporting, between treatment and control - as detailed in appendix Tables B.1 and B.2).

2.4.2 Monthly Surveys

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

We proceeded with 5 monthly follow-up surveys. The monthly surveys were administered with three purposes in mind. First, they provided an opportunity for enumerators to follow up regularly with matatu owners and drivers and address any questions they might have about the device. Second, they were used to remind both parties to continue submitting the daily reports in the SmartMatatu app. Finally, they were designed to collect some basic data. As owners and drivers reached the 6-month mark, we conducted an endline survey to measure changes in key outcomes, and to assess the impact of the information treatment and the cash incentives.

2.4.3 Tracking data

The CalAmp tracking device transmitted high frequency data on forward/backward/lateral/vertical acceleration, jerk, location and a timestamp. We use the raw measures of acceleration to investigate changes in driver behavior. Specifically, we look at vertical and lateral acceleration to determine whether the driver is operating on bumpier stretches of road. Furthermore, we use the GPS data to calculate how far each vehicle is from the route they are licensed to be on at any point in time. This provides a measure of how far the driver is deviating from the actual route. Figure 2.4 depicts the number of times vehicles licensed to route 126 pass through a particular lon-lan cell. The first panel clearly shows what the route should be, and the second panel overlays the designated route to confirm. The figure illustrates that off-route driving is relatively common practice.

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

2.5 A Principal-Agent model with unobserved output

The purpose of the model is to generate key predictions about how the monitoring device affects the principal-agent relationship. The owner sets a target in KES that the driver must deliver by the end of the day. The driver chooses the amount of effort and damaging driving they engage in. Damaging (i.e risky) driving refers to the set of behaviors that damage the company asset (which are not necessarily correlated with the unsafe driving metrics we will discuss in the results section). The driver also decides when they will under-report revenue, and by how much - because revenue is unobserved to the owner. This results in five key parameters, and we derive comparative statics for each one. For simplicity we assume that both owners and drivers are risk-neutral.

Given that all of the contracts in this transportation sector follow this fixed form, and they do not change over time, we restrict our attention to contracts of this type (detailed in more depth right below). We then derive comparative statics resulting from the introduction of the monitoring technology, which affects the *parameters* of the contract (not the *type* of contract), and driver behavior. There are three additional departures from the classical setting: 1) the driver selects both effort *and* risk levels, 2) production is unobservable to the principle (allowing the driver to misrepresent production), and 3) the principle observes a signal of production.

2.5.1 Status Quo

The model is comprised of four steps that correspond to the owner-driver daily interaction.

1. First, the owner sets the target (T), and asks the driver to deliver it by the end of the day.
2. Second, the driver chooses how much effort (e) and damaging driving (r) to engage in. The days' random events unfold (ε), and total revenue is generated. We define total revenue q as:

   $$q = e + r\varepsilon$$

   where $r \geq 1$ (it is impossible for drivers to avoid risky driving entirely). This production function says that an increase in effort shifts the distribution of output to the right, while an increase in risk reduces the variance of revenue while keeping the mean constant (this is similar to (Ghatak and Pandey, 2000)). We assume drivers choose the amount of effort and risk *before* ε is realized. A driver that chooses to behave in a really risky manner will experience more extreme outcomes depending on the days events. For example, if he chooses to drive off-route a lot (higher risk), he may be caught by the police and impounded (earning very low revenue that day) or he may get away with it and earn high revenue. While $q = q(e, r, \varepsilon)$, I will suppress the arguments for notational convenience throughout the rest of the paper. Note: q is observed to the driver, because they collected passenger fares throughout the day, but it is not unobservable to the owner.

3. Third, the driver chooses how much revenue to report to the owner ($\tilde{q}$).
4. Finally the owner decides whether or not to punish based on 1) whether he detects any damages to the vehicle, and 2) whether he can be sure the driver is under-reporting revenue. Note these punishments are non-monetary (firing, reprimands) because the driver's limited liability constraint binds - they cannot be expected to pay monetary costs.

We solve the model in 4 steps, via backward induction.

Step 1: Owner's punishment

The owner punishes the driver for two types of behaviors. First, they will punish the driver if they detect any damages to the vehicle. The expected punishment is expressed as βr, where β is the probability the owner detects risky driving that damages the vehicle, and r are the damages incurred. Note that without a monitoring technology β is close to zero because the owner has a hard time detecting behavior that damages the vehicle.

Second, the owner will punish based on the driver's reported revenue ($\tilde{q}$). We know the owner (principal) sets a target amount (T) at the beginning of the day, which they want the driver (agent) to deliver by the end of the day. The driver earns revenue (q) from passenger fares, and decides how much to report to the owner ($\tilde{q}$). Owners do not observe actual revenues that are collected by drivers from passenger fares throughout the day (q), but they receive a signal ($\hat{q}$) about what true revenue (q) should be. This signal comes from the information owners gather about driving conditions and driver behavior throughout the day. Absent a monitoring technology, the owners signal will be informed by listening the news, visiting the designated route, or talking to their friends. We assume the owner's signal is noisy but unbiased, so that $\hat{q}$ is defined as follows:

$$\hat{q} = q - \sigma$$
$$\sigma \sim F\left(0, \frac{1}{\alpha}\right)$$

where α is the precision of the owners' signal about true revenue. Any monitoring technology we introduce will provide more information to the owner about driver behavior. This will increase the precision of the owner's signal about what revenue should be, which gives the driver less leeway to significantly under-report revenue on a particular day. In the case of our monitoring technology specifically, the owner can observe the number of kilometers driven, and where the driver was operating at any point in time. Owners can use this information to estimate the number of trips to and from the city center, which provides a more accurate measure of total daily revenue.

In principle the owner's decision to punish the driver could depend on 1) the level of reported revenue ($\tilde{q}$), and 2) the signal of actual revenue ($\hat{q}$). In theory, punishing based on the level of $\tilde{q}$ is rational because owners may be able to infer that their driver invested low effort or took too much risk from a very low $\tilde{q}$. Alternatively, owners could punish based on $(\hat{q} - \tilde{q})$ because this difference reflects how much the owner thinks the driver is lying. In practice, the most effective punishment will depend on owners ability to infer e and r from reported $\tilde{q}$, which will depend on the variance of ε. If the variance of ε is high, then the owner cannot infer e and r from a low $\tilde{q}$. Anecdotally, it appears the variance of epsilon is large, which means the owner will place more weight on $(\hat{q} - \tilde{q})$ (which owners in our sample confirmed). For ease of exposition we assume that the punishment is quadratic in the difference between the owners signal $\hat{q}$ and the driver's offer $\tilde{q}$ - where the distributional assumptions that rationalize this are presented in the Appendix. Note that because $\hat{q}$ is an unbiased signal of q, and σ is independent of $\tilde{q}$, the punishment can also be expressed as an increasing function of the gap between the true revenue (q) and the offer ($\tilde{q}$) (or in this case, a quadratic of the difference). All comparative statics go through as long as this punishment remains an increasing function of the in gap between $q - \tilde{q}$. The punishment is expressed as:

$$E[punishment] = \frac{\alpha}{4}(q - \tilde{q})^2$$

Step 2: Solve the agent's optimal reporting schedule

The driver needs to choose how much revenue to report to the owner ($\tilde{q}$). Broadly, the driver can choose to report truthfully, or to under-report, and there will be some threshold of revenue beyond which they will choose to truthfully report.

Case 1: The driver chooses to reports above the target ($\tilde{q} > T$)
When the agent chooses to report above the target, they do not face any incentive to lie because they keep everything they earn above the target and the owner cannot renegotiate the terms of the contract. Therefore they report truthfully $\tilde{q} = q$ and their utility is

$$U^D = \overbrace{q - T}^{salary} - \overbrace{\beta r}^{damages}$$

Case 2: The driver chooses to report below the target ($\tilde{q} < T$)
When the agent chooses to report below the target, they face an incentive to lie in order to increase their take-home pay. Indeed, on days when the driver truly fails to make the target $q < T$, under-reporting revenue means they get to walk away with some money rather than handing it all over to the owner as stipulated by the

contract. Even when the amount of revenue is slightly above the target $q > T$, the driver faces some incentive to lie in order to walk away with slightly more income than what they would if they reported truthfully and had to hand over the entire target amount. On the days that the driver decides to report below the target, they know they can be punished by the owner. The driver then needs to choose the amount of revenue to report ($\tilde{q}$) to maximize their utility:

$$\max_{\tilde{q}} \qquad U^D = \overbrace{(q - \tilde{q})}^{salary} - \underbrace{\frac{\alpha}{4}(q - \tilde{q})^2}_{punishment} - \overbrace{\beta r}^{damages}$$

Solving for $\tilde{q}$ yields:

$$\tilde{q} = q - \frac{2}{\alpha}$$

where

$$\frac{\partial \tilde{q}}{\partial \alpha} = \frac{2}{\alpha^2} > 0$$

This says that the optimal amount for the driver to under-report is a function of the owner's signal. More specifically, it is optimal for the driver to under-report by a constant amount $\left(\frac{2}{\alpha}\right)$. Figure 2.5 shows that the data confirms this behavior. The graph summarizes values of under-reported revenue at unique/binned values of net revenue above the target.[11] We see that drivers continuously under-report approximately 700 KES (7 USD) until the net revenue they generate exceeds the target by approximately 500-1000 KES. This graph confirms that our model does a good job of predicting the behavior we observe in the data.

Switch point: Next we need to determine the point at which the driver is indifferent between reporting above the target (and telling the truth) and reporting below the target (and under-reporting). When the driver tells the truth i.e $\tilde{q} = q$, they get utility:

$$(q - T - \beta r)$$

When they lie "optimally" i.e $\tilde{q} = q - \frac{2}{\alpha}$, they get utility:

$$(q - \tilde{q}) - \frac{\alpha}{4}(q - \tilde{q})^2 - \beta r = \frac{1}{\alpha} - \beta r$$

[11]On the y-axis we have under-reported revenue, which we compute from driver and owner's daily surveys. When the driver does not make the target (owner's income is less than the target), any salary the driver reports to us is under-reported revenue because they should have handed this over to the owner. When the driver makes the target (the owner's reported income equals the target), there is no under-reporting and this variable is set to zero. On the x-axis we plot net revenue above target, defined as owner income + driver salary – target. Indeed, to get an accurate measure of under-reporting we want to know the share of joint revenue that the driver withholds. In other words, we need to know the income that the owner took home *and* and the salary of the driver.

Setting the two utilities equal and solving:

$$q - T - \beta r = \frac{1}{\alpha} - \beta r$$
$$q^* = T + \frac{1}{\alpha}$$

Where

$$\frac{\partial q^*}{\partial \alpha} = -\frac{1}{\alpha^2} < 0$$
$$\frac{\partial q^*}{\partial T} = 1 > 0$$

This says that the revenue required to truthfully report, (q^*), is a function of the owner's signal and the target. It is perhaps useful to relate this result that $q^* = T + \frac{1}{\alpha}$ to the optimal reporting behavior $\tilde{q}$. In particular, if the driver observes q at the upper bound of $[T, T + \frac{1}{\alpha}]$, he will report $\tilde{q} = q - \frac{2}{\alpha} = T - \frac{1}{\alpha}$, which is indeed less than the target.

Step 4: Driver's optimal choice of effort

The driver chooses two actions, effort (e) and risk (r), which affect the probability distribution of revenue. Following (Ghatak and Pandey, 2000) we define q as:

$$q = e + r \cdot \varepsilon$$

where ε is a random variable that reflects the idiosyncratic events of the day (weather and traffic). The driver chooses effort and risk to maximize his utility

$$\max_{e,r} \quad \underbrace{E\left[(q - T - \beta r) \mid q \geq q^*\right] \cdot Pr(q \geq q^*)}_{Truth} + \underbrace{E\left[(q - \tilde{q}) - \frac{\alpha}{4}(q - \tilde{q})^2 - \beta r \mid q < q^*\right] \cdot Pr(q < q^*)}_{Under-report} - \underbrace{h(e, r)}_{Cost}$$

Where h(e,r) is the driver's private cost for his actions, and we assume that this function is twice continuously differentiable, monotonically increasing and convex in e and r. The optimization problem yields the following F.O.C with respect to e and r, respectively (all the derivations in the paper can be found in the Appendix):

$$1 - F_\varepsilon\left(\frac{q^* - e}{r}\right) - h'_e = 0$$
$$\int_{\frac{q^*-e}{r}}^{\infty} \varepsilon f_\varepsilon(\varepsilon) d\varepsilon - h'_r - 2\beta = 0$$

Where

$$\frac{\partial r}{\partial \alpha} < 0 \ \ \& \ \ \frac{\partial e}{\partial \alpha} > 0$$
$$\frac{\partial r}{\partial T} > 0 \ \ \& \ \ \frac{\partial e}{\partial T} < 0$$
$$\frac{\partial r}{\partial \beta} < 0 \ \ \& \ \ \frac{\partial e}{\partial \beta} > 0$$

This means that the driver's choice of effort and risk is a function of the owner's signal, the target, and the probability the owner detects damaging driving. Intuition for the sign of these partial derivatives will be provided in the next section when we introduce the monitoring technology.

Step 5: Owner's choice of the target

Constrained case

The owner chooses T to maximize his utility:

$$\begin{aligned} &\max_{T} \qquad T \cdot Pr(q \geq q^*) + E[\tilde{q} \mid q < q^*] \cdot Pr(q < q^*) - \gamma(r) \qquad \text{s.t} \\ &E\left[(q - T - \beta r)|q \geq q^*\right] \cdot Pr(q \geq q^*) + E\left[(q - \tilde{q}) - \frac{\alpha}{4}(q - \tilde{q})^2 - \beta r \mid q < q^*\right] \\ &\qquad \cdot Pr(q < q^*) - h(e^*, r^*) > 0 \end{aligned}$$

Where $\gamma(r)$ are the costs owners incur for any damages to the vehicle they detect that need repairs. The constraint is the driver's participation constraint, where we assume for simplicity that they have a reservation wage of zero. This optimization problem yields the following F.O.C with respect to T and λ, respectively:

$$\begin{aligned} \frac{\partial}{\partial T} &= 1 - F_\varepsilon(\cdot) + \frac{\partial e}{\partial T} F_\varepsilon(\cdot) + \frac{\partial r}{\partial T}\int_0^{\frac{q^*-e^*}{r^*}} \varepsilon f_\varepsilon(\varepsilon) d\varepsilon - \frac{\partial}{\partial T}\left(\frac{q^*-e^*}{r^*}\right)\left(\frac{1}{\alpha}\right) f_\varepsilon(\cdot) - \gamma'(r)\frac{\partial r}{\partial T} + \\ &\lambda\left[-(1 - F_\varepsilon(\cdot)) + \frac{\partial e}{\partial T}(1 - F_\varepsilon(\cdot)) + \frac{\partial r}{\partial T}\int_{\frac{q^*-e^*}{r^*}}^{\infty} \varepsilon f_\varepsilon(\varepsilon) d\varepsilon + \frac{\partial r}{\partial T}(-\beta) - h'\left(\frac{\partial e}{\partial T} + \frac{\partial r}{\partial T}\right)\right] = 0 \\ \frac{\partial}{\partial \lambda} &= \int_{\frac{q^*-e^*}{r^*}}^{\infty} (e^* + r^*\varepsilon - T - \beta r^*) f_\varepsilon(\varepsilon) d\varepsilon + \int_0^{\frac{q^*-e^*}{r^*}} \left(\frac{1}{\alpha} - \beta r\right) f_\varepsilon(\varepsilon) d\varepsilon - h(e^*, r^*) = 0 \end{aligned}$$

Unconstrained case

The owner chooses T to maximize his utility:

$$\max_{T} \qquad T \cdot Pr(q \geq q^*) + E[\tilde{q} \mid q < q^*] \cdot Pr(q < q^*) - \gamma(r)$$

Which yields the following F.O.C with respect to T

$$(1 - F_\varepsilon(\cdot)) + \frac{\partial e}{\partial T} F_\varepsilon(\cdot) + \frac{\partial r}{\partial T}\left(\int_0^{\frac{q^*-e^*}{r*}} \varepsilon f_\varepsilon(\varepsilon) d\varepsilon\right) - \frac{1}{\alpha} f_\varepsilon(\cdot) \frac{\partial}{\partial T}\left(\frac{q^* - e}{r}\right) - \gamma'(r)\frac{\partial r}{\partial T}$$

2.5.2 Monitoring Technology

Introducing the monitoring technology increases α and β. In other words, the precision of the owner's signal about output (α), and the probability of detecting damaging driving (β), increase. In what follows we consider how the owner changes the terms of the contract, and how this subsequently affects the driver's choices of effort, risk and under-reporting. We assume throughout that the driver's participation constraint binds.[12] As α and β increase:

Proposition 1. ***(Target)***

$$\frac{\partial T}{\partial \alpha} < 0 \quad \textit{and} \quad \frac{\partial T}{\partial \beta} < 0$$

This says that as the precision of the owner's signal (α) increases, the owner will reduce the target. We know that the driver's utility will fall as α increases. Because the constraint binds, the owner needs to reduce the target to ensure the driver makes their reservation wage. The same reasoning explains why the owner reduces the target as the probability of detecting damaging driving (β) increases.

Proposition 2. ***(Damaging (Risky) Driving)***

$$\frac{dr}{d\alpha} = \overbrace{\frac{\partial r}{\partial \alpha}}^{-} + \overbrace{\frac{\partial r}{\partial T}}^{+}\overbrace{\frac{\partial T}{\partial \alpha}}^{-} < 0$$
$$\frac{dr}{d\beta} = \underbrace{\frac{\partial r}{\partial \beta}}_{-} + \underbrace{\frac{\partial r}{\partial T}}_{+}\underbrace{\frac{\partial T}{\partial \beta}}_{-} < 0$$

This says that risk will unambiguously decrease when α increases. There are two effects at work. As α increases, the driver is going to have to truthfully report more often (q^* falls), which increases the probability they are the residual claimant and must bear the cost of a bad revenue day. Being exposed to greater downside risk makes damaging behavior less attractive, $\frac{\partial r}{\partial \alpha} < 0$. Similarly, as the owner reduces the target in response to a higher α, (q^* falls), and the driver will take on less risk, $\frac{\partial r}{\partial T} > 0$. Next, the driver's immediate response to an increase in probability of detection (β) is to reduce risk, $\frac{\partial r}{\partial \beta} < 0$. Moreover as the owner reduces the target in response to higher β, the driver will further reduce the amount of risk they take.

[12] Anecdotally we know that there are a lot of drivers on the market and we think it reasonable to assume that they would have been bargained down to their constraint. Switching costs are not trivial and the owner would prefer less turnover all else equal.

Proposition 3. ***(Effort)***

$$\frac{de}{d\alpha} = \overbrace{\frac{\partial e}{\partial \alpha}}^{+} + \overbrace{\frac{\partial e}{\partial T}}^{-}\overbrace{\frac{\partial T}{\partial \alpha}}^{-} > 0$$
$$\frac{de}{d\beta} = \underbrace{\frac{\partial e}{\partial \beta}}_{+} + \underbrace{\frac{\partial e}{\partial T}}_{-}\underbrace{\frac{\partial T}{\partial \beta}}_{-} > 0$$

This says that effort will unambiguously increase when α increases. Yet again, there are both direct and indirect effects driving this result. As the precision of the owner's signal, α, increases, the driver will supply more effort, $\frac{\partial e}{\partial \alpha} > 0$. Indeed, for every level of output below q^*, the driver can no longer under-report as much as they used to because the owner has a more precise signal of what output should be. Holding all else constant, their revenue will decrease if they do not respond by increasing effort and generating more revenue. Next, we know that the owner responds to an increase in α by lowering the target. As the target decreases the driver is more likely to meet it, which increases the returns to effort and incentivizes the driver to work more, $\frac{\partial e}{\partial T} > 0$. Turning next to the effects of better detection of risky driving (β), we know from above that the driver will have to reduce the amount of risk they take, which reduces the probability of large windfall days. The driver has to compensate by investing more effort, $\frac{\partial e}{\partial \beta} > 0$.

Proposition 4. ***(Switch Point)***

$$\frac{dq^*}{d\alpha} = \overbrace{\frac{\partial q^*}{\partial \alpha}}^{-} + \overbrace{\frac{\partial q^*}{\partial T}}^{+}\overbrace{\frac{\partial T}{\partial \alpha}}^{-} < 0$$

As the owner's signal becomes more precise, the revenue required to truthfully report and make the target q^* will decrease, $\frac{\partial q^*}{\partial \alpha} < 0$. Similarly, as the target decreases, the revenue required to truthfully report and make the target q^* will decrease, $\frac{\partial q^*}{\partial T} > 0$. Together these two factors drive the reduction in q^* we anticipate with the introduction of the monitoring technology.

Proposition 5. ***(Under-reporting)***

$$\frac{d\tilde{q}}{d\alpha} = \frac{2}{\alpha^2} > 0$$

This says that the optimal amount for the driver to report ($\tilde{q}$) increases with the precision of the owner's signal (α). It follows mechanically that the amount they under-report will fall.

2.6 Results

We test each one of these predictions in our data. We first investigate how the technology affects the target contract. We then examine drivers response along the key dimensions detailed above, including 1) effort, 2) damaging driving, and 3) reporting behavior (probability of making the target, and under-reporting revenue). Before reviewing whether these behaviors change in the way we would expect from the model, we briefly provide some summary statistics and evidence that owners were interacting with the device. Finally, we conclude this section with an overview of the device's impact on safe driving (which is not highly correlated with the our measure of damaging driving).

2.6.1 Baseline Characteristics

We work exclusively with matatu owners with one vehicle, which they do not operate themselves. They are approximately 40 years of age, and have completed 11 years of education. These small-scale entrepreneurs have spent an average of 8 years in the matatu industry, owning a vehicle for the past 4 years. While it is possible to have a salaried job and manage a matatu at the same time, only 20% of our sample juggle these two responsibilities. Typically owners have worked with their current driver for the past 2 years. Drivers have very similar profiles - which we expect - because many owners were previously driving matatus themselves. They are a few years younger (35 years old on average), with slightly lower levels of education (8 years on average). They have worked in the industry for over a decade, driving a vehicle for the past 7-8 years. They have worked with 5 different owners, averaging 1.5 years with each one. Both driver and owner characteristics are balanced across treatment and control groups (Table 2.1 and Table 2.2).

2.6.2 Information Treatment Arm

To study the treatment effect of information on contracts, productivity (which includes effort, damaging driving, under-reported revenue, and profits) and safety, over the 6 month time frame we run the following regression model:

$$y_{ird} = \sum_{m=1}^{6} D_{im}\beta_m + \alpha_d + \tau_r + X_i\gamma + \varepsilon_{ird}$$

where y_{ird} is a daily contract/productivity/safety outcome for owner i on route r, on day since installation d; D_{im} is a treatment indicator equal to 1 if the owner is in the information group in month m (which allows us to examine the treatment effect as it evolves over the six months of the study); α_d is a day fixed effect; τ_r is an assigned route fixed effect; X_i is a set of firm-level baseline specific controls included for precision, and ε_{irmd} is an error term.[13] We cluster our standard errors at the firm level. Note that the study offered the information to control owners in months 7 and 8 (as compensation for participating in the study). As a result all the regressions only include data before month 7.[14]

We also have an endline survey that asked owners about their use of the device, perceptions of drivers' performance, their monitoring strategies, and their firm's size. To study the impact of our device on these outcomes we run the following regression model:

$$y_{ir} = \alpha_d + \tau_r + D_i\beta + X_i\gamma + \varepsilon_{ir}$$

where y_{ir} is an endline outcome for owner i on route r; τ_r is an assigned route fixed effect; D_i is a treatment indicator equal to 1 if the owner is in the information group; X_i is a set of firm-level baseline controls included for precision, and ε_{ir} is an error term.

2.6.3 Usage

First, we monitor owners' usage of the device. We do so by tracking the API calls that are generated every time the owner logs into the app and requests different pieces of information (including historical location, up-to-date summary information, and where the safety violations occurred on the map). Figure 2.6 calculates the share of owners that made at least one API call during the week. We find high rates of take-up. In the early months of the study approximately 80% of owners are checking the app at least once a week. This share decreases but stabilizes at about 70% as the study progresses. A large share of owners are also using the app daily. In the first few months, 60% of owners check the app once a day, and 40% continue their daily usage after 6 months. This suggests that owners are actively engaging with the device throughout the study.

We also check whether owners are internalizing the information we provided. At endline we asked owners to state the revenue earned, the number of kilometers driven, the fuel costs, and the extent of off-road driving on the most recent day their vehicle was active. Owners had the option of answering "don't know". We find that owners in the treatment group are 27 percentage points more likely to know about the number of kilometers driven and 45 percentage points more likely to know about the the instances of off-route driving (Table 2.3 Columns 1 and 2). We do not find any differences between the treatment and control groups regarding knowledge of exact revenue (Column 3). This last result is not altogether surprising. While the device provides a better *estimate* of revenue by revealing the number of kilometers the vehicle has covered, and the location of the vehicle throughout the day, it does not reveal the *exact* revenue because the number of passengers

[13] Controls include the matatu's age and number of features, as well as owner's age, education, gender, tenure in the industry, their raven score and the number of other jobs they have.

[14] All of the results from these regressions are presented through figures because it is easier to see the how owners' and drivers' behavior changes over time. Please see the appendix for the specific point estimates from these regressions, and alternate regression specifications (Appendix Tables B.3 though B.11).

is uncertain. As a final test, we also ask owners to rate how challenging it is to monitor their employees on a scale from 1 (not hard) to 5 (very hard). Having access to the information reduces the reported difficulty level by just over 2 points. In other words, control owners maintain that monitoring is hard while treatment owners reveal that it is easy (Table 2.4).[15] We do not, however, find any significant changes in traditional monitoring behaviors (checking-in with the driver over the phone, at the terminal (locally referred to as “stage"), or through a third party).

Finally, we investigate whether there are any changes in how the owners and drivers interact. We asked drivers to report the frequency with which they were contacted and criticized by the owner that month. Formal reprimands are not frequent but they are used by owners to signal their displeasure with the driver’s behavior. Figure 2.7 suggests that the number of reprimands is marginally higher in the treatment group at the beginning of the study period. This is consistent with the idea that owners use the information to correct behavior early on. The frequency increases by approximately 20-30 % (off of a control mean of 0.5) in months 1-4 before decreasing significantly in month 6. We also investigate whether owners take more extreme actions and fire their drivers more frequently. While the trend in Figure 2.7 suggests that the number of firings increased in the second month of the study and decreased thereafter, this result should be interpreted carefully because there are so few firings in our data (17 in total).

Contracts

We first investigate whether access to the tracking information changes the *terms* of the contract. While the intervention could also have changed the *type* of contract they offered their drivers (fixed wage or sharecropping), extensive interviews with owners suggested this was unlikely to occur. The fixed wage contract is unpopular among owners and drivers, and the sharecropping model is difficult to implement when limited liability constraints bind, and revenue is unobserved/can be easily withheld by the drivers. Moreover, social norms are engrained in this industry, and a change of this magnitude would be unexpected in a 6 month time frame. We further confirmed in our endline survey that every owner maintained the same type of contract.

As detailed in the model, owners can use the information from the device to change the target they set for drivers (Proposition 1). Absent the technology, the target for 14 seater buses is usually set at 3000 KES. Discussions with owners confirm this is an industry standard that only fluctuates with good reason (they know that demand will be high or low that day because the weather or road conditions have changed). Charging much more would alienate drivers, and charging any less would cut into firm revenues. Figure 2.8 depicts the estimated treatment effect on the owners’ daily target across the 6 months of the study. There are no significant changes in the first month, likely because owners were still learning how to use the

[15]Column 2 of Table 2.4 also shows that control owners report no change in the time they spend monitoring during the study period, while 70% of treatment owners reveal they spend less time monitoring.

app and experiment with ways to improve their business operations. In subsequent months, however, we see the target steadily declining. By month 6, the daily target amount is 135 KES (1.35 USD) below the control group, representing a 4.1% decrease (0.2 standard deviation). While the result is not statistically significant (likely because we are underpowered), the downward trend is clear. This steady reduction suggests that the information allows managers to re-optimize the terms of their employees' contracts. Taking this result back to the model, it suggests that the drivers are operating at their participation constraint. When the constraint binds, the owner needs to decrease the target in order to compensate the driver for their lost information rents. Lowering the target also reduces owners' revenue on days where the driver makes the target. As a result it is only profitable for the owner to reduce the target if the owner is compensated in other ways, namely with a higher share of revenue on days when the driver does not make the target, fewer damages to the vehicle, or an increase in the frequency with which drivers make the target. We turn to these results next.

Productivity

We consider three measures of productivity, which correspond to the choices that drivers make throughout the day. This includes how damaging (risky) they will drive, how much effort to supply, and the amount of revenue they disclose to the owner (which is either the target amount, or some amount below).

1. Damaging (risky) driving: We hypothesize that owners prefer less damaging driving than what the drivers would optimally choose. With the technology, owners can observe driving behavior more accurately, and the probability they detect damaging driving increases. This reduces drivers' incentive to drive in ways that damage the vehicle. Similarly, as the precision of the owner's signal about revenue improves, and they decrease the target, the driver will be exposed to greater downside risk on bad revenue days.[16] This should further reduce their incentive to drive in ways that damage the vehicle (Proposition 2). Figure 2.9 confirms this hypothesis in the data. We see damages substantially decrease throughout the entire 6 month period. In month 3, daily repair costs for treatment owners are reduced by 125 KES (1.25 USD), and continue falling until month 6, where they are 226 KES (2.26 USD) less than what control owners incur on average (this represents a 46% decrease in daily repair costs). These repair costs represent a major business expense for owners, which makes the impact of the monitoring technology significant.

We want to confirm that this result stems from less damaging driving behavior. One of the greatest sources of damaging driving is off-route driving. Drivers often take shortcuts on bumpy roads that are notoriously damaging to matatus. These shortcuts are appealing to the driver because they help them travel to the city center

[16]As detailed in the model, the switch point q^* shifts down. This means that there are days when the driver used to under-report, and they now truthfully report. As the residual claimant on these days, they bear the cost of a bad revenue day (which is higher when they make more damaging maneuvers on the road).

more quickly, and avoid traffic jams where they sit idly without picking up any passengers. Typically owners cannot observe off-route driving and drivers cannot be expected to pay for vehicle repairs. This means that damaging driving along these alternate routes is costless to drivers but offers the opportunity for large windfall days. When owners have access to the monitoring technology, however, they can inform drivers about how to take better care of the vehicle, and mandate that they stay on their designated routes. To investigate this hypothesis, we compute the distance between each GPS point recorded by the device, and the vehicle's licensed route. Figure 2.10 demonstrates that treatment drivers are on average 400 meters closer to the designated route than control drivers throughout the study period. To confirm that the reduction in off-route driving is responsible for fewer damages, we investigate whether the distributions of lateral and vertical acceleration differ across treatment and control groups. Taking fewer bumpy roads that jostle the vehicle should be visible in the acceleration data. Lateral acceleration measures tilting from side to side, while vertical acceleration captures movement upwards and downwards. We find suggestive evidence that driving behavior has changed. The distribution of lateral acceleration in the treatment group tightens around 0 (less tilting - Figure 2.11). Similarly, the distribution of vertical acceleration has more mass around gravity (normal driving) for the treatment group. We can reject equality of these distributions across treatment and control by applying a K-S test, which returns a p-value of 0.000 for both measures of acceleration.

It is also important to rule out any alternative explanations for these effects on repair costs. Specifically, it could be the case that drivers tend to inflate repair costs, and the device reduces their incentive to do so because they are more likely to be caught in the lie.[17] This cannot be the case for larger repairs, however, because they are incurred by the owner directly and/or will be validated with the mechanic. We therefore create an indicator for whether the repair costs exceed 1000 KES (80th percentile). The second panel in Figure 2.9 demonstrates that the probability of incurring a large repair cost decreases significantly (7-8 percentage points). This implies that the decrease in the repair costs that we observe cannot be entirely driven by drivers inflating the repair costs. Drivers are also changing *how* they drive as the result of the technology.

2. Effort: Next, we proxy driver effort by the number of hours the tracking device is on (the device powers on and off with the matatu). When the device is installed in the matatu, drivers know that owners have a more precise signal of output, which means they are more likely to get caught if they under-report heavily. This encourages drivers to under-report less, which reduces their take-home pay. As the model demonstrates, this creates an incentive for drivers to invest more effort throughout the day so they can increase total revenue and ensure they maintain similar compensation. In parallel, owners have lowered the target, which means that it is more

[17]Note that drivers pay for some of the smaller repair costs out-of-pocket, if they can be handled quickly by a mechanic during the day. They subsequently report the amount they spent to the owner, which we would then capture in the owner survey.

likely that the driver will become the residual claimant. Finally, the model predicts that drivers will compensate for the reduction in damaging driving by investing more effort. For all of these reasons we expect effort levels to rise (Proposition 3). This prediction is borne out in our data: Figure 2.12 demonstrates the upward trend in effort that we anticipated. The number of hours the tracking device is on increases by 0.98 hours per day on average in month 3 and rises steadily until the end of the study. By month 6, effort levels increase by 1.47 hours per day on average in the treatment group. This represents a 9.9% increase in drivers' labor supply. This is substantial in an environment where drivers are already working 14 hour days. With more hours on the road, we also see the number of kilometers increase by 12 kilometers per day on average, which corresponds to an extra trip to/from the city center (Figure 2.12).

3. Reporting Behavior: Once the driver chooses how much effort and risk to invest, they need to decide whether or not to truthfully report. The model predicts that the optimal switch point for truthfully reporting (q^*) shifts down because 1) the owner's signal of true revenue becomes more precise, and 2) the owner has lowered the target (Proposition 4). Testing this proposition is difficult because we do not have a direct measure of q^*; we only observe whether or not drivers make the target (when owner reported income is equal to the target). However, a lower q^* implies that drivers should make the target more often, primarily on days when revenue is close to q^* to begin with. To investigate this prediction in the data, we first apply our standard regression specification to determine whether we see a significant change in the probability of making the target.[18] Figure 2.13 suggests that from month 3 onwards, the rate at which drivers make the daily target increases by 11 percentage points from a base of 44 percent (significant in month 3 only). It is not altogether surprising that this result is slightly weaker because the analysis considers the full range of revenue rather than focusing on days when drivers are close to q^* to begin with (i.e. close to making the target). This is where the model predicts we should see these effects. To investigate this further we calculate the average revenue above target on a route-month in the control group to get a sense of the usual revenue above target generated for a day.[19] We then compute drivers' daily reported revenue above target and subtract the average expected amount. This is akin to including route fixed effects - because we know that a certain level of revenue above target will be acceptable on certain routes but not on others. We are left with a measure of daily deviation from expected revenue above target, which we plot in the second panel of Figure 2.13. The revenue above target measure has an approximate mean of 4,000 KES. As such, -2000 KES on the graph implies that drivers only have 2,000 KES in revenue to cover their salary and their costs for that day. This results in a take-home pay of 500 to 1000 schillings, which is right where we expect q^*

[18] The driver has "made the target" if the owner's reported income is equal to the target. When the owner's reported income is below the target, the driver has not made the target.

[19] We use gross revenue below average for this outcome instead of net revenue as we did for the shading amount because it only depends on drivers reporting, which means we have more data to work with.

to be (from plotting the amount drivers under-report in Figure 2.5). Figure 2.13 demonstrates that the probability of making the target increases significantly at this point, which is exactly what we would expect. This represents a meaningful increase in “compliance" with the terms of the contract.

Finally, we anticipate that drivers' reporting behavior ($\tilde{q}$) will change as the monitoring devices are introduced. According to the model, we should see drivers under-reporting below some optimum q^*, at which point they will start truthfully reporting and providing the owner with the target amount. Below this optimum, the model predicts that drivers will under-report by a constant amount. This is consistent with the idea that drivers have some reservation wage they do not want to fall below. We predict that the monitoring technology will decrease the amount of under-reporting we see in the data. Owners can use the device to estimate actual revenue more accurately, and they are more likely to detect when the driver is under-reporting. Drivers should respond by lying less everywhere below the threshold. Figure 2.14 depicts under-reporting across treatment and control groups, to which we apply a non-parametric smoothing function. We observe constant under-reporting below some threshold value q^* in both groups (which falls somewhere between 500-1000 KES). Moreover, we observe that the treatment group under-reports less than the control group. To obtain a more precise estimate for the reduction in under-reporting, we regress the amount drivers under-report on treatment status for different possible q^* (between 500-1000). The regression only considers data below q^* because this is where the model predicts shading will occur.[20] The results in Table 2.5 confirm that the amount drivers under-report falls by approximately 70-100 KES ($\approx$ 1 USD) per day depending on the exact location of q^*. As a final check, taking 900 KES as a benchmark value for q^* and keeping all data below this point, Figure 2.15 applies our standard regression specification over study months. We see that the amount drivers under-report falls by approximately 100 KES throughout the study (except for month 1).

Without knowing q^* exactly, our estimate of 70-100 KES technically includes both a reduction in $\tilde{q}$ and a drop in q^*. We want to confirm that both of these behaviors are indeed happening in reality. We do so by imposing a step function in a regression of under-reported amount on treatment. In other words, we allow under-reporting below q^* and impose zero shading above. We run this regression for every reasonable value of q^* for treatment and control groups. We then plot two outcomes in the second panel of Figure 2.14. The dots represent the estimated under-reported amount in the treatment (in red) and control (in black) groups across different choices of q^*. We can see that the treatment groups under-report by approximately 50-70 schillings less than the control group regardless of the q^* we impose on the model. Next, we plot the Mean Squared Error (MSE) of our regressions (dotted lines) to isolate the q^* that minimizes the MSE for the treatment and control groups respectively. The vertical lines represent the optimal q^* using

[20] The regression includes the standard controls and fixed effects. The regression also excludes data from month 1 because we know that owners were unfamiliar with the device in that first month. The magnitude of the results stay the same when we include month 1 but we lose some precision from the noise this month introduces.

this metric. This demonstrates that our best guess of q^* in the treatment group is 150 schillings below our best guess of q^* in the control group. This confirms that both factors explain the overall reduction in under-reporting that we observed in the more flexible regression specification above.

Company Performance and Employee Well-being

We now turn to investigating the impact of the monitoring device on firm performance. Specifically, we are interested in determining whether the information we supplied allows companies to generate higher profits and ultimately expand their operations by adding more vehicles to their fleet. Company profits are measured by subtracting costs (repairs and driver salary) from total revenue. We documented substantial reductions in repair costs and, assuming drivers are at their reservation wage, we expect their salary to stay the same (Figure 2.17 confirms this is true). The model predicts that the impact on revenue, however, is ambiguous. Improved monitoring increases driver effort, and reduces under-reporting. However, it also reduces the amount of damaging driving they engage in (which we confirmed in our data). Depending on which of these effects dominates revenue could increase or decrease. Panel 1 of Figure 2.16 illustrates that revenue does not change substantially throughout the study. Taken together, decreasing costs and stable revenues suggest that firm profits will increase. Panel 2 in Figure 2.16 demonstrates a similar trend to what we've observed to date: profits increase continuously starting in month 3, and peak at month 5. Specifically, treatment owners see their daily profits rise by approximately 12% in month 4 and 5 (440 KES per day = 4.40 USD per day). Taking the average gains over the study period and extrapolating to the full year (assuming matatus operate 25 days a month), we can expect a 120,000 KES (1200 USD) increase in annual firm profits. It is worth mentioning that this profit measure does not capture any additional gains from having to spend less time and effort monitoring the driver. The device cost 125 USD (including shipping to Kenya), which means that it would take less than 3 months for the investment to become cost-effective for the owner. This return on investment (ROI) suggests that these devices are likely to be welfare improving for owners in the short and long run. One of the reasons we do not see more matatu owners adopting them, however, is because they currently do not exist in this form on the market. The options are either much more expensive (approximately 600 USD and monthly installments), or have more limited capacity. Without having tested their efficacy, owners are hesitant to make the investment. It is perhaps worth mentioning that our profit gains are in line with some of the more successful business training programs documented in the literature. The cost of these trainings range from 20 to 740 dollars and last a few weeks at most (Bloom et al., 2013; Bloom, Sadun, and Van Reenen, 2017; McKenzie and Woodruff, 2016; Berge, Bjorvatn, and Tungodden, 2014; de Mel, McKenzie, and Woodruff, 2014; Valvidia, 2012). Our technology has the added benefit of requiring a single up-front payment for continued use. Moreover it requires relatively little coordination and training.

Are treatment firms also 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. A simple regression of this outcome on treatment with the standard controls reveals that treatment owners have 0.145 more vehicles in their fleet on average than control owners (Table 2.7, Column 1). This represents an 11 percent increase in fleet size. While treatment owners were also more likely to make changes to their matatu's interior, this result is not statistically significant (Table 2.7, Column 2). There are a number of reasons why the monitoring device could have encouraged treatment owners to grow their businesses more actively. First, profits have increased and under-reporting has decreased. Second, our results suggest that owners started trusting their drivers more. Table 2.6 presents four different measures of owners' perceptions of their drivers at endline. We see owners sending an additional 30 KES to drivers in the trust game the enumerators administered (Column 1) - a 30% increase. Moreover, treatment owners' assessment of whether their drivers' skills have improved increases by 0.6 points (where they could be assigned a -1 for worse driving, 0 for no change, and 1 for better driving). Finally, treatment owners are more likely to report that their drivers have become more honest (Column 3). We suspect that greater trust in their drivers' abilities/honesty, combined with a reduction in the amount the drivers under-report, makes the process of managing the company easier. Together with higher profits, treatment owners may have seen an opportunity to expand that did not exist before.

Finally, it is important to investigate whether these gains to the company come at the expense of their employees. While it is difficult to measure welfare, we consider three main outcomes that could impact drivers' well-being: the amount of effort they supply, their salary and their relationship with the owner. We know the amount of effort they supply increases (Figure 2.12), and the amount they under-report decreases (Figure 2.14). While their salary per hour remains unchanged (Figure 2.17), they are potentially worse off from working more hours. However, throughout the course of the study we did not receive any complaints from drivers, despite contacting them regularly to conduct our surveys. To investigate this further, we administered a small survey to drivers via SMS 6 months after the original study concluded (at this point we had given control owners 2 months with the information as well, and no distinction can be drawn between treatment and control drivers). Sixty percent of drivers responded (distributed evenly across treatment and control) with very positive experiences about the device: 27% said it improved their relationship with the owner (70% said nothing changed), 65% said it made their job easier (26 % said nothing changed), 96% said they preferred driving with the tracker, and 65% said it changed the way they drove. While we do not want to lean too much on this qualitative evidence, it does suggest that the drivers benefitted from the device as well. Some of the open ended questions reveal that drivers felt a greater sense of security with the device in their car. They also felt that it increased owners' trust in their work, which reduced their stress levels. In an environment where drivers are constantly being second guessed by their employers, this could represent a meaningful improvement in working conditions.

Externalities

The device conveyed information to owners about productivity and safety. A priori, we thought that owners might care more about safe driving than drivers for two reasons. First, owners are the ones to pay for repairs when the vehicle is damaged, and for fines when the car is ticketed or impounded by the police. Second, drivers can increase their take-home pay if they generate higher revenue in fares by breaking safety regulations to pick up more passengers. It follows that the information we provided should have an effect if unsafe driving is correlated with damages, and the owners want to avoid high repair costs. Alternatively, owners should use the information to reform driving behavior if unsafe driving results in more fines than the owner would optimally choose to incur. The Kenyan government also assumed that owners would reform driving behavior when they mandated that long-range buses be equipped with GPS tracking devices in 2016. However, this measure was only marginally successful, and before rolling out the experiment we were conscious that the safety information we provided may not have the intended effects. This may be because owners down-weight the probability of getting into severe accidents, or fined for unsafe driving. Alternatively, the benefits to unsafe driving in terms of increased revenue may be high for the owners as well. If owners care only about profits, and increased effort comes at the expenses of safety, we might expect instances of unsafe driving to increase. Finally, while owners definitely care about damages to their vehicle, these damages may only be weakly correlated with unsafe driving. If this is the case, then the information we provided may not change safe driving practices.

The device collected six pieces of information that correlate with safe driving: maximum speed, average speed, speeding over 80km, over-acceleration, sharp braking and sharp turning. We do not see any meaningful increases in maximum or average speeds as the study progresses. Similarly, instances of over-acceleration and speeding above 80km do not change significantly (Figure 2.19). We see no effect on sharp turns or sharp braking (Figure 2.20). Finally we tracked the number of accidents throughout the project. There are 41 accidents in total throughout the 6 month period, of varying degrees of severity. While the number of accidents trends upwards in months 4 and 5, it is difficult to conclude that accidents increase significantly (Figure 2.21). Overall the evidence points towards safety standards staying the same, despite the emphasis we placed on safety across all tabs in the app. While this highlights that owners can incentivize optimal levels of effort without further compromising passenger safety, we cannot necessarily expect owners to curb unsafe driving with these technologies. This is especially important for local governments in Kenya and South Africa to know, as they continue to take steps to curb unsafe driving by introducing monitoring technologies.

2.6.4 Cash Treatment Arm

Bearing in mind that owners may not act on the safety information we provided, we tested the impact of an intervention that incentivizes drivers to take safety into account. Drivers were offered 600 KES at the beginning of the day, and incurred a penalty for each safety violation they incurred. The experiment was designed to mimic an intervention that a regulatory body could feasibly implement. We find that the cash treatment has no discernible effect on average speed, over-acceleration, and sharp turning. However, we detect large decreases in the instances of speeding and sharp braking. The number of sharp braking alerts deceases by 0.13 events per day, a 17% decrease relative to the control group. Likewise, the number of sharp braking events decreases by 0.24 per day, representing a 35% decrease. These results suggest that drivers can be incentivized to take safety into account. However, the incentives must come from a third party, as owners are unlikely to induce similar changes in driving behavior.

In Table 2.9 we examine driving behavior among the group of drivers whose cash incentives were removed after the first month. The goal of this exercise is to examine whether the behavioral changes induced by the cash treatment persist after the incentives are removed. The variable "One-month post treat" compares drivers who never received cash incentives, to the drivers whose incentives have been removed. We see that the number of speeding events rebounds almost completely to pre-treatment levels, while the number of sharp braking events remains lower but is statistically insignificant. Overall, it appears that the behavioral effects of the cash treatment arm wear off after the removal of the incentives. This suggests that inducing better driving habits for a short time period may not be sufficient to see longer run improvement in safety outcomes.

2.7 Concluding Remarks

In this paper we design a monitoring technology tailored to the minibus industry in Nairobi. The device provides real time information to the owner of the minibus about the productivity and safety of the driver. We find that the monitoring technology eases labor contracting frictions by improving the contract that owners offer their drivers. The drivers respond by supplying more effort, driving in ways that are less damaging to the vehicle, under-reporting revenue by less and meeting the target more often. This results in higher profits for the firm. Treatment owners also report greater trust in their drivers, and find it less difficult to monitor them, which may explain why their businesses grow faster during the study. Despite the breadth of information we supplied on safety, we do not see drivers improving their performance along this margin unless they are explicitly incentivized to do so with small cash grants. While this suggests that gains to the company do not come at the expense of the quality of service they provide, it also highlights that the technology does not remedy the negative safety externalities of the industry.

These results are important for a number of key stakeholders, including small firms operating in the transportation industry, and policy makers working to improve road safety conditions in urban hubs. We know firms struggle to grow in developing countries for a number of reasons, and this paper identifies another important barrier that is relatively understudied empirically: moral hazard in labor contracting. One solution that can potentially ease this friction is improved monitoring. Monitoring is typically difficult in small firms, however, because they cannot hire dedicated staff to oversee employee performance, and it takes time away from regular business operations. In our paper, we demonstrate that introducing cost-effective monitoring technologies can be a worthwhile investment for companies looking to increase their profits and grow their asset base.

We do not find that safety standards improve when information from the device is conveyed to owners. However, when the drivers are incentivized to drive more safely we see instances of speeding and sharp braking fall. This suggests that simply introducing monitoring technologies, without further regulation, might not achieve the desired effects for governments trying to improve road safety. Local transport authorities in Nairobi and South Africa have already started to discuss ways of introducing remote tracking solutions throughout the transportation industry to help monitor and record the behavior of the drivers on the road. Our research suggests that while this will improve firm operations, more targeted interventions requiring regulatory oversight will be necessary if these devices are to induce safer driving.

This analysis highlights the need for further research estimating the longer term impacts, and general equilibrium effects, of these technologies on firm operations, and worker outcomes. Our study included 255 matatus and lasted 6 months, but we hypothesize that we would have seen greater changes in the terms of the contract, and in the *type* of contract being offered had we continued for an additional year and offered more GPS trackers to owners on the various routes. Similarly, the impacts on driver well-being may have changed if all the matatus on the route were fitted with GPS tracking devices. While our results suggest that benefits accrue to both workers and firms in this context, thinking about who gains and who loses as these technologies become more pervasive is an important area for future work.

Tables

Table 2.1: Balance across information treatment (owners)

Variable	Control	Treatment	Difference
Install date (days since July 1, 2016)	211.9	212.9	-1.03
			(5.09)
Owner age	36.3	37.3	-0.99
			(0.99)
Owner gender	0.18	0.18	-0.0056
			(0.048)
Owner highest level of education	2.94	2.97	-0.030
			(0.11)
Owner is employed in salaried job	0.21	0.24	-0.030
			(0.052)
Years the owner is in matatu industry	7.71	7.71	-0.0066
			(0.79)
Years owner has owned matatus	4.65	4.47	0.18
			(0.52)
Number of drivers hired for this matatu	1.26	1.37	-0.12
			(0.13)
Number of other drivers hired in the past	1.77	1.94	-0.17
			(0.22)
Amount given in trust game	117.7	126.2	-8.50
			(12.4)
Owner Raven's score	4.51	4.65	-0.14
			(0.19)
Driver rating: owner's fairness	8.11	8.33	-0.23
			(0.18)

The summary statistics are calculated using the owners baseline data. The first column shows the mean in the control group, while the second column show the mean in the treatment group. The final column shows the difference between treatment and control. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table 2.2: Balance across cash treatment (drivers)

Variable	Control	Treatment	Difference
Driver age	34.4	37	-2.64
			(0.88)***
Driver highest level of education	2.45	2.48	-0.030
			(0.087)
Driver experience	6.95	8.24	-1.29
			(0.72)*
Weeks unemployed before current job	2.96	2.28	0.68
			(0.77)
Number of vehicles driven for before current	6.05	4.97	1.08
			(0.57)*
Number of past accidents	0.90	0.87	0.035
			(0.13)
Number of months the driver has been employed	15.2	14.3	0.91
			(2.49)
Owner rating: driver's honesty	7.78	7.60	0.18
			(0.18)
Owner rating: how hard driver works	8.29	8.07	0.22
			(0.18)
Owner rating: driver's safety	8.32	8.21	0.11
			(0.18)
Owner rating: driver's performance overall	8.09	8	0.092
			(0.17)
Driver days working for owner	411.7	500.8	-89.2
			(61.2)
Driver Raven's score	4.26	4.28	-0.016
			(0.18)
Revenue at baseline	7744.8	7732.3	12.5
			(207.6)
Baseline target	3113.1	3147.6	-34.5
			(56.6)

The summary statistics are calculated using the driver baseline data. The first column shows the mean in the control group, while the second column show the mean in the treatment group. The final column shows the difference between treatment and control. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table 2.3: Knowledge gathered through the device

	(1) Know Km	(2) Know Off-route	(3) Know Revenue
Info Treatment	0.268***	0.451***	0.039
	(0.068)	(0.065)	(0.072)
Control Mean of DV	0.47	0.40	0.61
Controls	X	X	X
Route FE	X	X	X
Matatu N	187	187	187

The data are from the owner endline surveys (where 3% of the sample - 9 owners - were unreachable, balanced between treatment and control). Note that the variables in this table were added to the endline survey after the first wave of endlines had already been completed, which is why we only have 187 observations (balanced across treatment and control). Each of the variables is a binary indicator for whether the owner said he knew the exact number of kilometers, instances of off-route driving and revenue generated by the vehicle on the most recent day the vehicle was on road. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels. Standard errors are clustered at the matatu level.

Table 2.4: Monitoring through the device

	(1) Difficulty Monitor	(2) Monitoring Time	(3) Check (Phone)	(4) Check (Stage)	(5) Check (Third Party)
Info Treatment	-1.845***	-0.721***	0.966	0.184	-0.116
	(0.156)	(0.053)	(0.895)	(0.383)	(0.257)
Control Mean of DV	4.02	-0.01	7.01	1.95	0.95
Controls	X	X	X	X	X
Route FE	X	X	X	X	X
Matatu N	190	190	190	190	190

The data are from the owner endline surveys (where 3% of the sample - 9 owners - were unreachable, balanced between treatment and control). Note that the variables in this table were added to the endline survey after the first wave of endlines had already been completed, which is why we only have 190 observations (balanced across treatment and control). These variables capture monitoring behaviors by the owner. Difficulty monitoring is an indicator from 1 to 5 for the level of difficulty associated with monitoring (5 = very hard). Monitoring time captures whether owners are spending less time monitoring (= −1), more time monitoring (= 1), or have seen no change (= 0) over the last 6 months. The last three columns document the number of times the owner checked up on the driver by phone, at the terminal (stage), and through a third party, respectively. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels. Standard errors are clustered at the matatu level.

Table 2.5: Under-reporting

	(1) 500	(2) 600	(3) 700	(4) 800	(5) 900	(6) 1000	(7) 1100
Treatment	-63.658^{**}	-79.026^{**}	-80.957^{*}	-97.075^{**}	-93.110^{*}	-81.483^{*}	-81.673^{*}
	(32.366)	(36.311)	(44.088)	(46.778)	(47.664)	(45.250)	(46.821)
Control Mean of DV	767.1	751.3	700.8	750.0	620.3	554.4	751.3
Controls	X	X	X	X	X	X	X
Day FE	X	X	X	X	X	X	X
Route FE	X	X	X	X	X	X	X
Observations	3,378	3,822	4,503	5,339	5,866	6,820	7,101

The data are from the daily surveys we collected from owners and drivers. Under-reporting is the amount of revenue the driver withholds from the owner. According to the contract, the driver must deliver the target to the owner by the end of the day. On days where the driver does not make the target (owner's reported income is below the target), the driver should deliver everything they earned in fares to the owner. On these days, any take-home pay the driver reports to us via the SMS survey is the amount they under-report to the owner. On days when the driver makes the target (owner's reported income is equal to the target), under-reporting is set to zero because they made the target. The model predicts that drivers will continue to under-report until revenue exceeds some threshold q^* of revenue above the target. We cannot determine the exact q^* because we don't observe α for each owner. Therefore we run a regression of the under-reported amount on treatment, and a set of controls, for various possible values of q^*, ranging from 500 to 1000 KES above the target. Note, to get an accurate measure of under-reporting we want to know the share of joint revenue that the driver withholds. In other words we need to know the income that the owner took home *and* and the salary of the driver. We we define q^* to be net revenue above target, defined as owner income + driver salary – target. These regressions are not split out by month and we restrict data to include months 2 onwards (in month 1 there is no behavior change as owners are learning how to use the technology. Note the effects stay the same if month 1 is included, but we lose some precision). Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels. Standard errors are clustered at the matatu level.

Table 2.6: Perceptions of trust

	(1) Trust Amount	(2) Better Driving	(3) More Honest	(4) Performance Rating
Info Treatment	33.796**	0.626***	0.708***	0.112
	(15.123)	(0.057)	(0.052)	(0.174)
Control Mean of DV	151.61	0.04	0.04	7.21
Controls	X	X	X	X
Route FE	X	X	X	X
Matatu N	244	190	190	246

The data are from the owner endline surveys (where 3% of the sample - 9 owners - were unreachable, balanced between treatment and control). The first column represents the amount of KES that was transferred from the owner to the driver in a game of trust. The owner was given an envelope with 900 KES and told that anything they placed back in the envelop would be tripled and sent to the driver. The driver would then choose how much to send back to the owner. The following two columns ask owners whether their drivers' driving has improved, and whether they have become more honest (= +1), or less honest (= −1) in the last 6 months. Note that the variables in Column 2 and 3 were added to the endline survey after the first wave of endlines had already been completed, which is why we only have 190 observations (balanced across treatment and control). The final column reflects owners rating of drivers on a scale from 1 to 10 (where 10 is the highest score). Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels. Standard errors are clustered at the matatu level.

Table 2.7: Business decisions

	(1) Number Vehicles	(2) New Interior
Info Treatment	0.145*	0.074
	(0.078)	(0.057)
Control Mean of DV	1.22	0.21
Controls	X	X
Route FE	X	X
Matatu N	246	240

The data are from the owner endline surveys (where 3% of the sample - 9 owners- were unreachable, balanced between treatment and control). The first column represents the number of vehicles the owner possesses, while the second column is an indicator = 1 if the owner refurbished the interior of his vehicle. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels. Standard errors are clustered at the matatu level.

Table 2.8: Effect of cash (immediate)

	(1) Average speed	(2) Maximum speed	(3) Speeding	(4) Sharp braking	(5) Overacceleration	(6) Sharp turning
Cash Treatment	-0.099	-0.214	-0.239**	-0.131*	-0.009	0.041
	(0.247)	(0.874)	(0.108)	(0.074)	(0.015)	(0.035)
Mileage in km	0.007	0.022	0.001	0.001	0.000	0.000
	(0.005)	(0.014)	(0.001)	(0.001)	(0.000)	(0.000)
Control Mean of DV	15.89	52.64	0.69	0.77	0.08	0.40
Controls	X	X	X	X	X	X
Matatu FE	X	X	X	X	X	X
Day FE	X	X	X	X	X	X
Route FE	X	X	X	X	X	X
Matatu-Day N	39,072	39,072	39,072	39,072	39,072	39,072

The data are from the tracking device throughout the study period. Column (1) and (2) capture average and maximum speeds throughout the day, respectively. Column (3)-(6) capture daily alerts for speeding over 80 km/hour, sharp breaking, over-acceleration, and sharp turning, respectively. We control for the number of miles the vehicle was on the road. Cash Treatment = 1 if the driver received cash transfers (whether it be for one month or two). Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels. Standard errors are clustered at the matatu level.

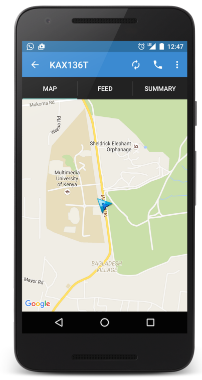
Table 2.9: Effect of no cash (ongoing)

	(1) Average speed	(2) Maximum speed	(3) Speeding	(4) Sharp braking	(5) Overacceleration	(6) Sharp turning
Cash Treatment	-0.120	-0.409	-0.220**	-0.140**	-0.014	-0.000
	(0.220)	(0.756)	(0.096)	(0.058)	(0.013)	(0.031)
One Month Post Treat	-0.039	0.072	-0.058	-0.115	-0.003	-0.017
	(0.260)	(0.971)	(0.135)	(0.089)	(0.012)	(0.031)
Mileage in km	0.008	0.024	0.002	0.001	0.000	0.001
	(0.005)	(0.015)	(0.001)	(0.001)	(0.000)	(0.000)
Control Mean of DV	15.89	52.64	0.69	0.77	0.08	0.40
Controls	X	X	X	X	X	X
Matatu FE	X	X	X	X	X	X
Day FE	X	X	X	X	X	X
Route FE	X	X	X	X	X	X
Matatu-Day N	42,405	42,405	42,405	42,405	42,405	42,405

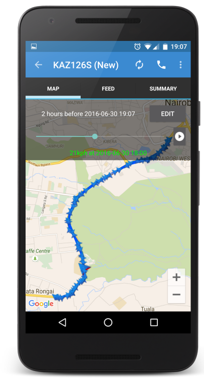
The data are from the tracking device throughout the study period. Column (1) and (2) capture average and maximum speeds throughout the day, respectively. Column (3)-(6) capture daily alerts for speeding over 80 km/hour, sharp breaking, over-acceleration, and sharp turning, respectively. We control for the number of miles the vehicle was on the road. Cash Treatment = 1 if the driver received cash transfers (whether it be one month or two). One Month Post Cash = 1 for drivers in the 1 month cash treatment group, in the month *after* their cash incentives were stopped. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels. Standard errors are clustered at the matatu level.

Figures

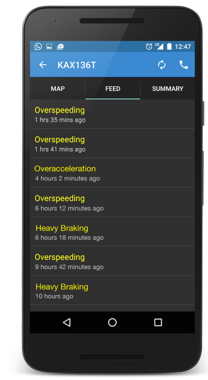
Figure 2.1: Mobile app “SmartMatatu"



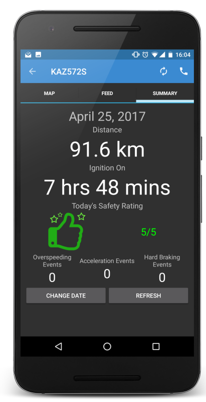
(a) Map Viewer



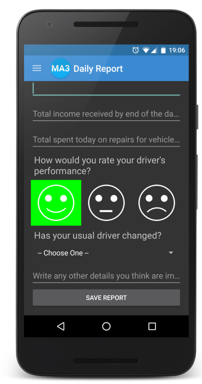
(b) Historical Map Viewer



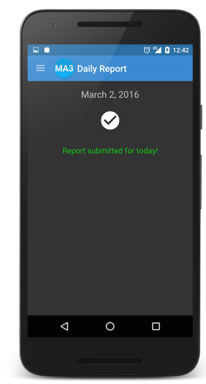
(c) Safety Feed



(d) Productivity Summary



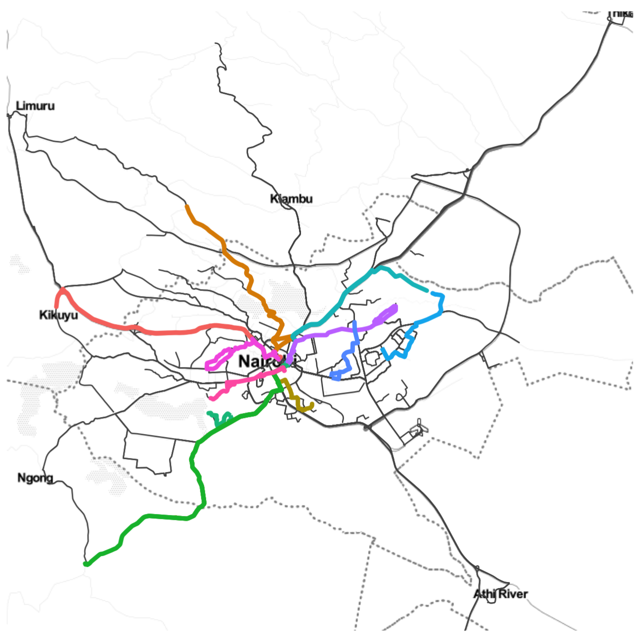
(e) Report Submit



(f) Report Complete

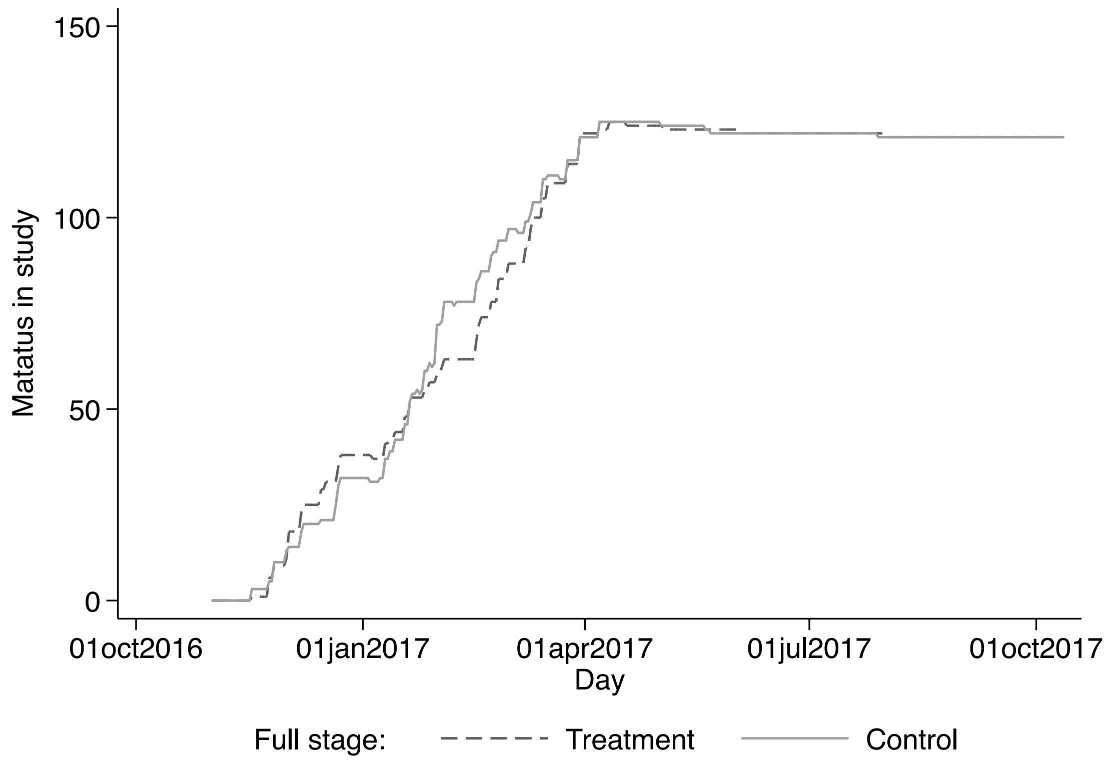
Figure 2.2: Device Location

(a) Designated bus routes in Nairobi (black)



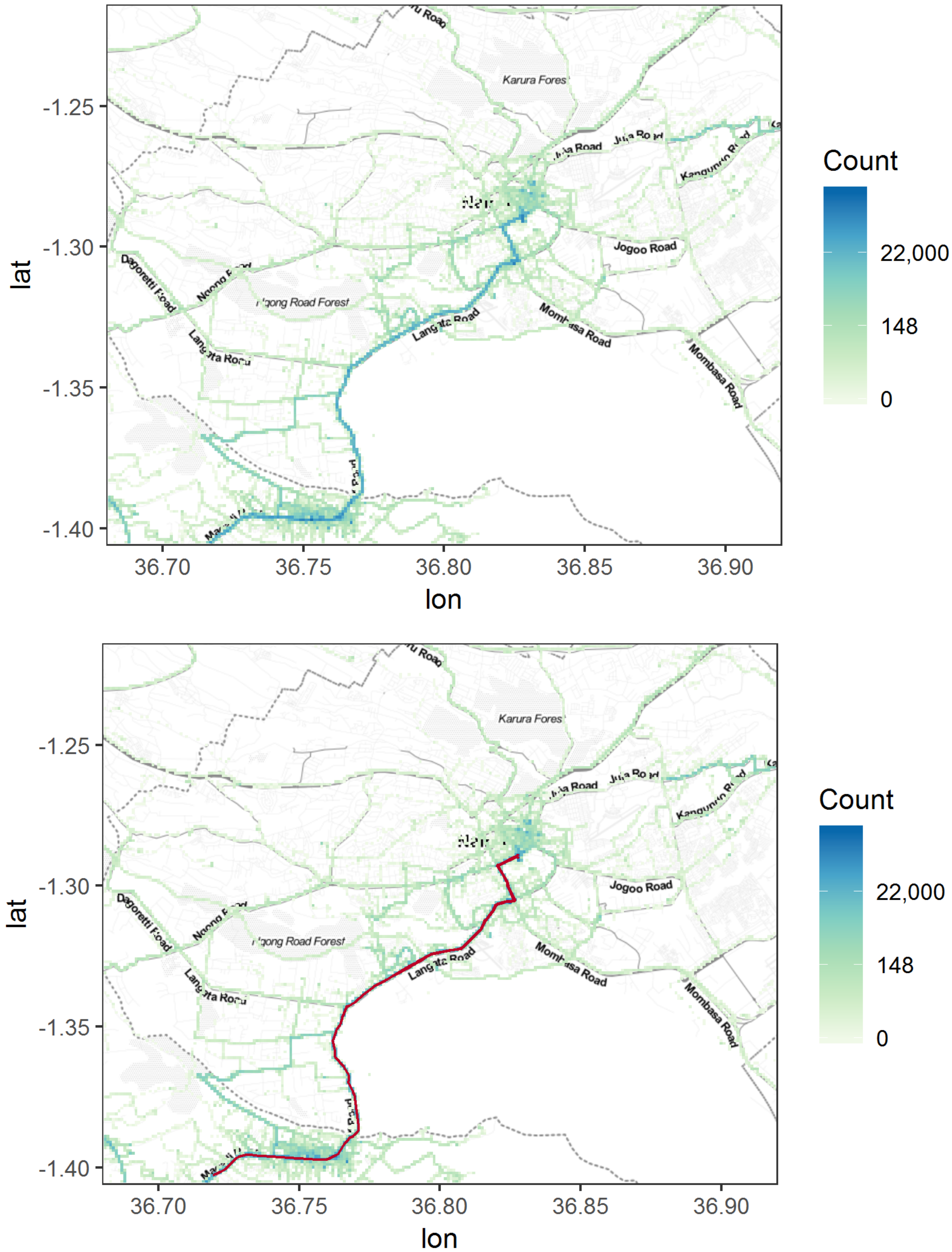
(b) Designated bus routes in Nairobi (black) and routes in our sample (colored)

Figure 2.3: Study Timeline



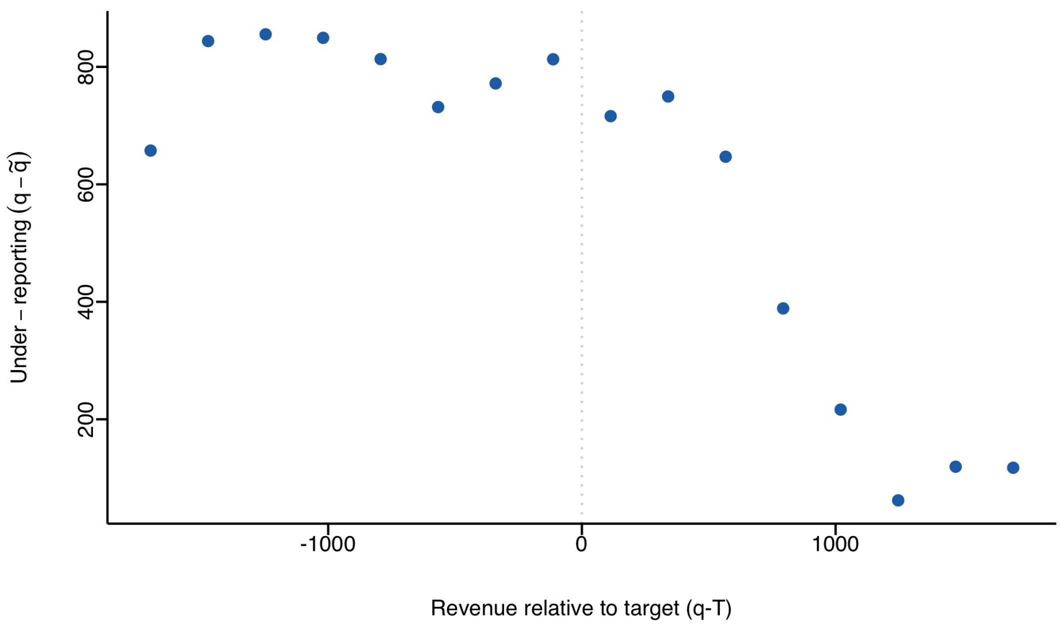
Notes: The figure depicts the number of matatus that were fitted with GPS trackers (and hence were added to the study) per week. The first installation took place in November 2016, and continued until April 2017. On average, the field team was able to fit GPS trackers to 15 matatus per week. As a result it took approximately 5 months to finish installations.

Figure 2.4: Device Location



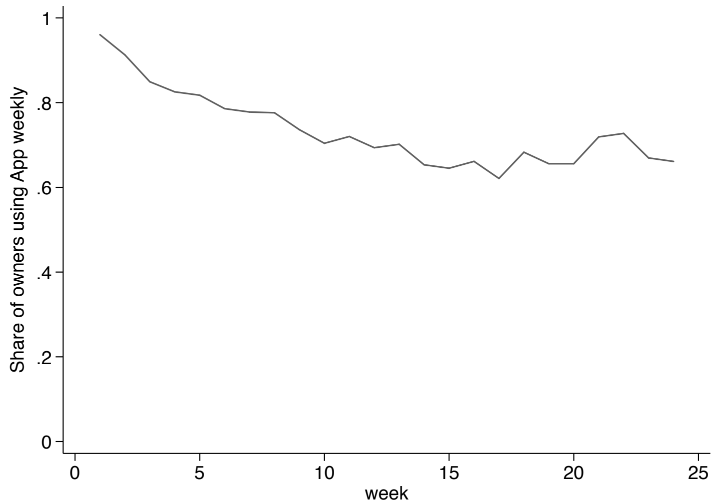
Notes: These maps use data from the trackers that were installed in vehicles licensed to operate on Route 126 (Ongata-Rongai line). Specifically, we count the number of times that vehicles passed through particular longitudinal and latitudinal cells on the map. A deeper shade of blue demonstrates that more vehicles passed through that particular cell. The second panel overlays the designated route that vehicles are supposed to be on (red). Any colored cells outside of the designated route are instances of off-route driving. Some of these are sanctioned by the owner, while others are not.

Figure 2.5: Model Validation (Constant Shading)

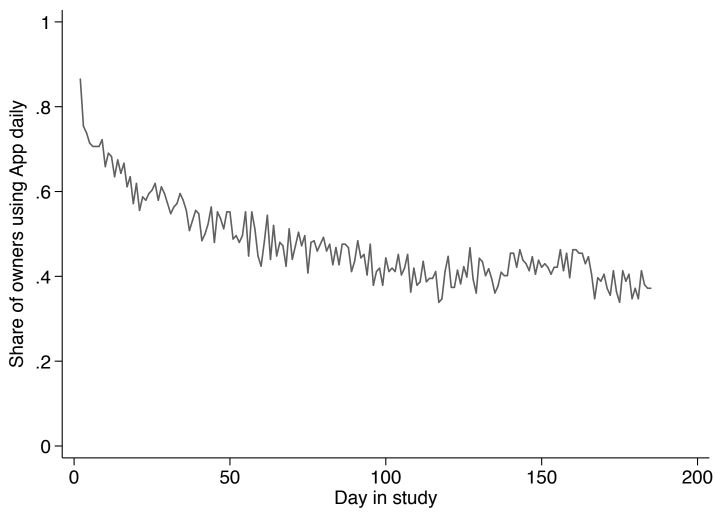


Notes: This figure depicts the amount of under-reporting on the y-axis, and the amount of revenue relative to the target on the x-axis. The data are from the control group. Under-reporting is the amount of revenue the driver withholds from the owner. According to the contract, the driver must deliver the target to the owner by the end of the day. On days where the driver does not make the target (owner's reported income is below the target), they should deliver everything they earned in fares to the owner. On these days, any take-home pay the driver reports to us via the SMS survey, is the amount they under-report to the owner. On days when the driver makes the target (owner's reported income is equal to the target), under-reporting is set to zero because they made the target. Note, to get an accurate measure of under-reporting we want to know the share of joint revenue that the driver withholds. In other words we need to know the income that the owner took home *and* and the salary of the driver. We therefore use net revenue above target on the x-axis, defined as owner income + driver salary − target.

Figure 2.6: Device Usage (API Calls)



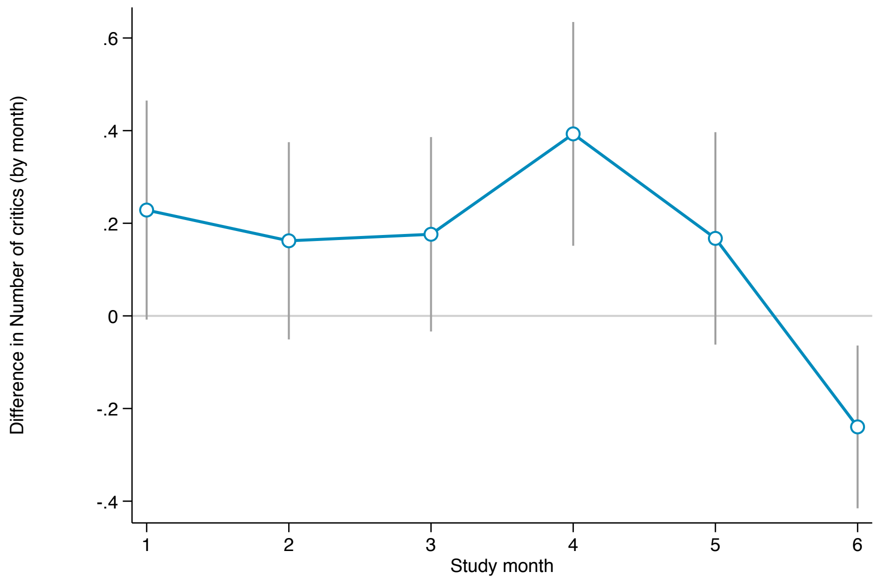
(a) Weekly Usage



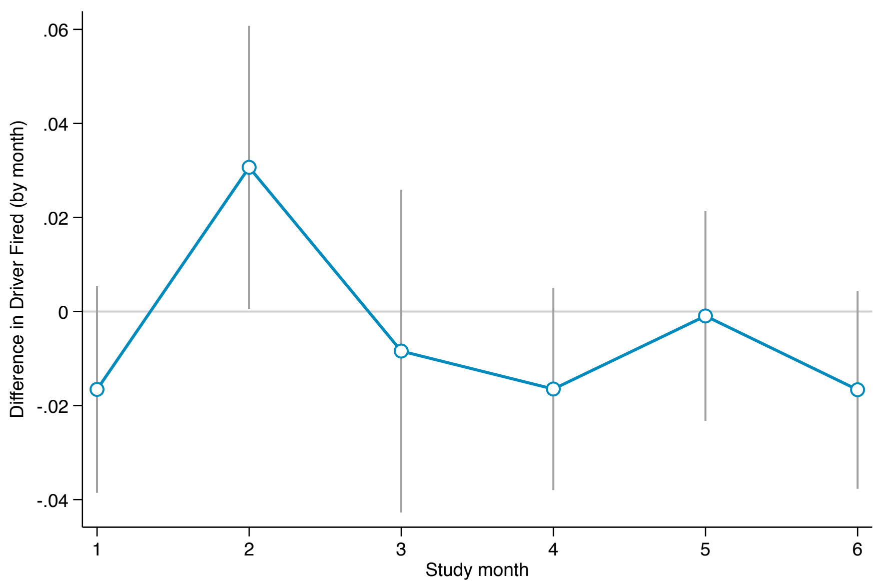
(b) Daily Usage

Notes: To measure device usage, we capture whether any API calls were made in a day. An API call is generated each time the owner logs into the app throughout the study period. The first panel looks at usage by week, whereas the bottom panel looks at usage per day.

Figure 2.7: Reprimands and Firing



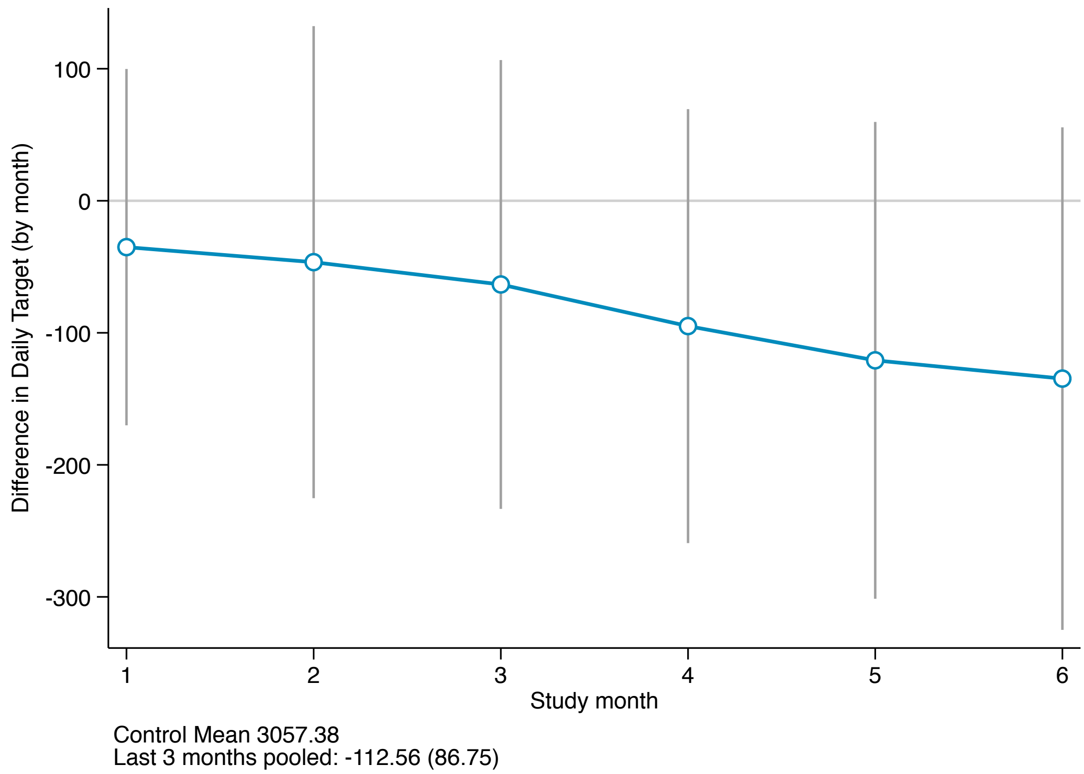
Control Mean: 0.46
Last 3 months pooled: 0.24*** (0.08)



Control Mean: 0.01
Last 3 months pooled: -0.01 (0.01)

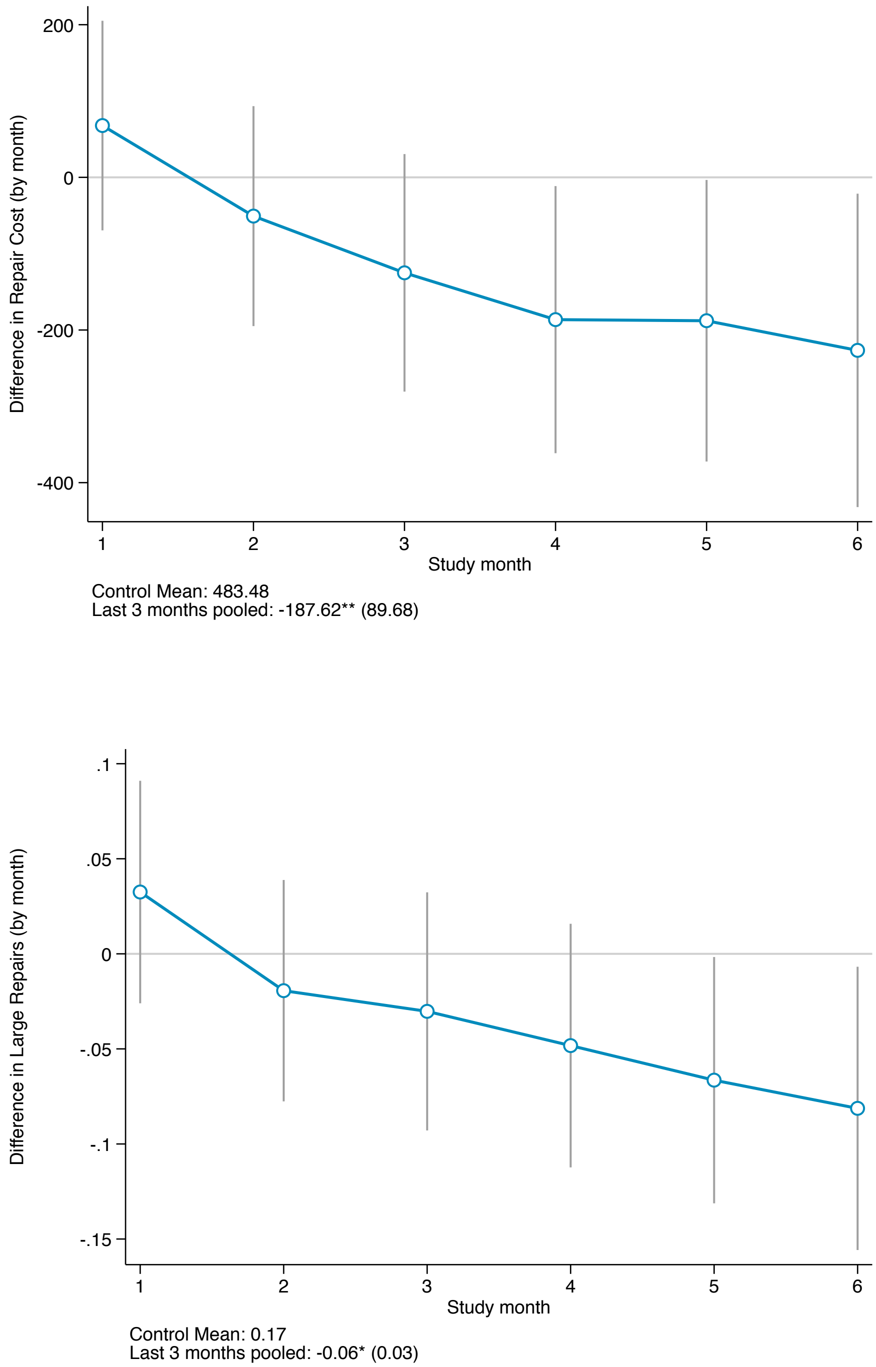
Notes: Each point on the graphs is the average difference between treatment and control in a particular month for a particular outcome. Note that because installations were rolled out over time, Month 1 (through 6) represent the first month (through the sixth month) since *installation*. The outcome in the first panel is number of instances the owner criticized the driver (driver-reported). This data was collected in the monthly driver surveys. The outcome in the second panel is the number of drivers fired. This data was captured by the daily owner/driver surveys, and then validated by an enumerator who called the owner directly to confirm.

Figure 2.8: Prediction 1 → Target



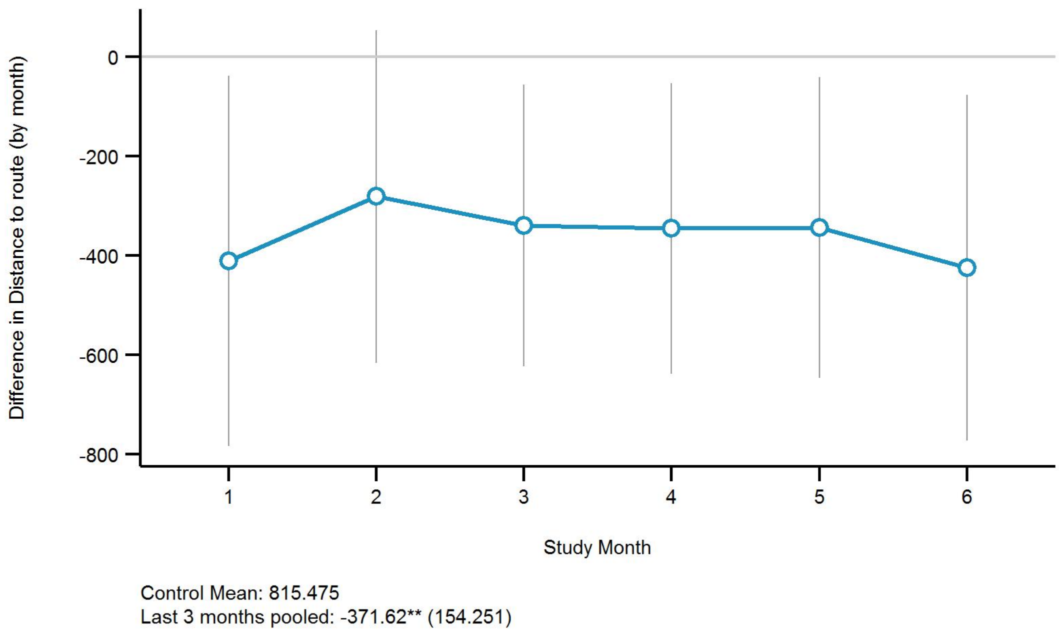
Notes: Each point on the graph is the difference between the average daily target reported by treatment and control owners in a particular month. The data was collected through the owner daily surveys. Note that because installations were rolled out over time, Month 1 (through 6) represent the first month (through the sixth month) since *installation*.

Figure 2.9: Prediction 2 → Damaging (Risky) Driving



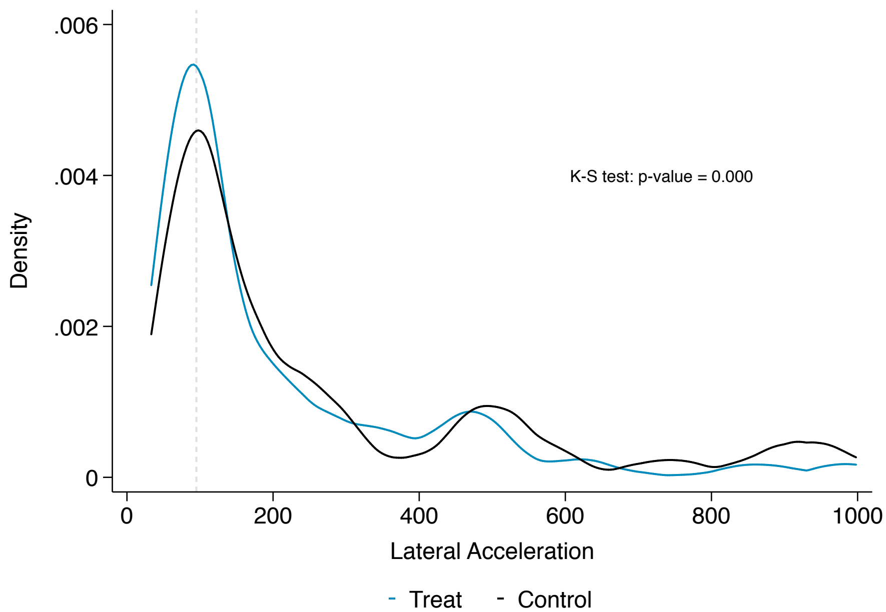
Notes: Each point on the graphs is the average difference between treatment and control in a particular month for a particular outcome. Note that because installations were rolled out over time, Month 1 (through 6) represent the first month (through the sixth month) since *installation*. The outcome in the first panel is the daily repair cost reported by the owner (daily survey). The outcome in the second panel is a binary indicator = 1 if the owner's reported repair was "large" (i.e. in the 80th percentile or above).

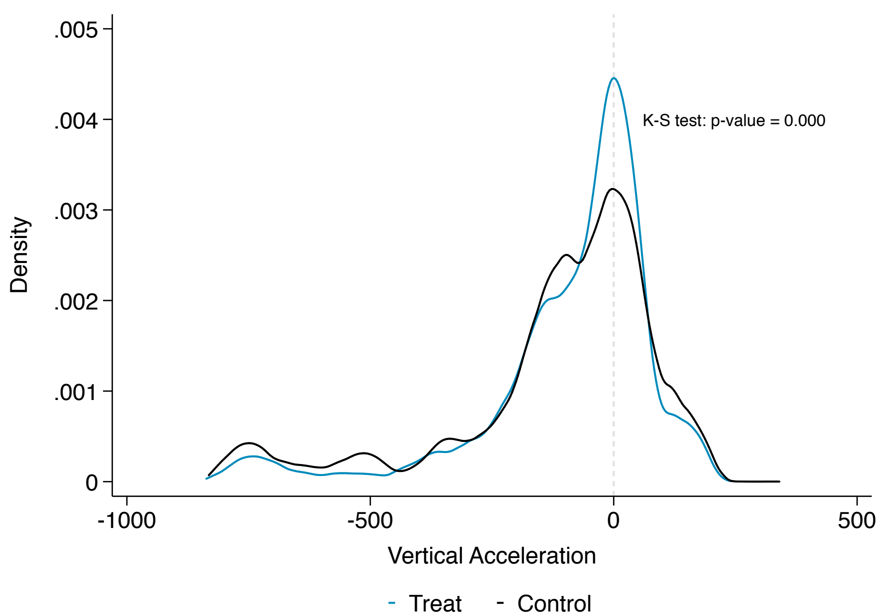
Figure 2.10: Prediction 2 → Damaging (Risky) Driving



Notes: Each point on the graph is the difference between the average distance to designated route calculated for treatment and control owners in a particular month. “Distance to designated route" is the shortest distance between the GPS tracker data point and the line corresponding to the route the vehicle is supposed to be on. A negative coefficient means the vehicles in the treatment group are closer to the designated route than the vehicles in the control group. Note that because installations were rolled out over time, Month 1 (through 6) represent the first month (through the sixth month) since *installation*.

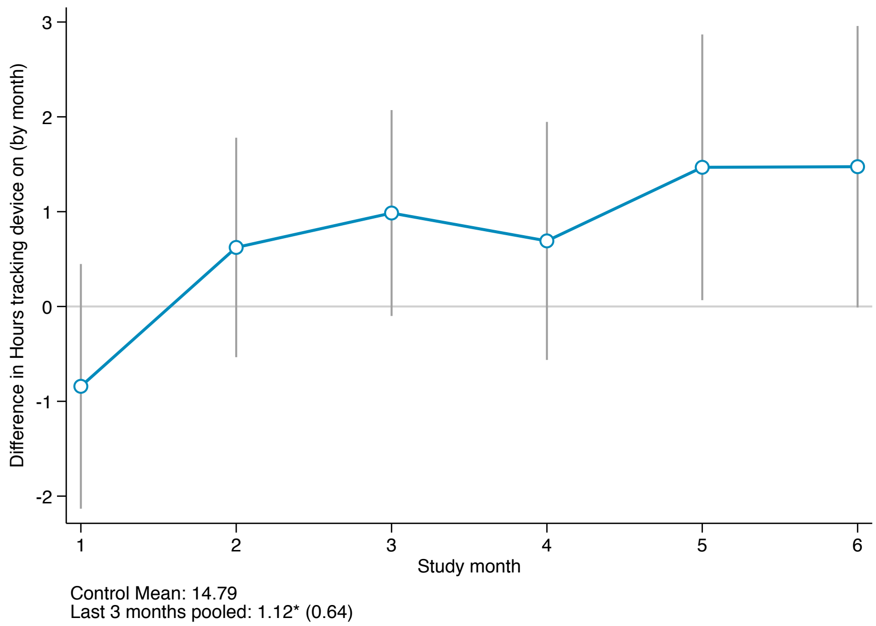
Figure 2.11: Prediction 2 → Damaging (Risky) Driving

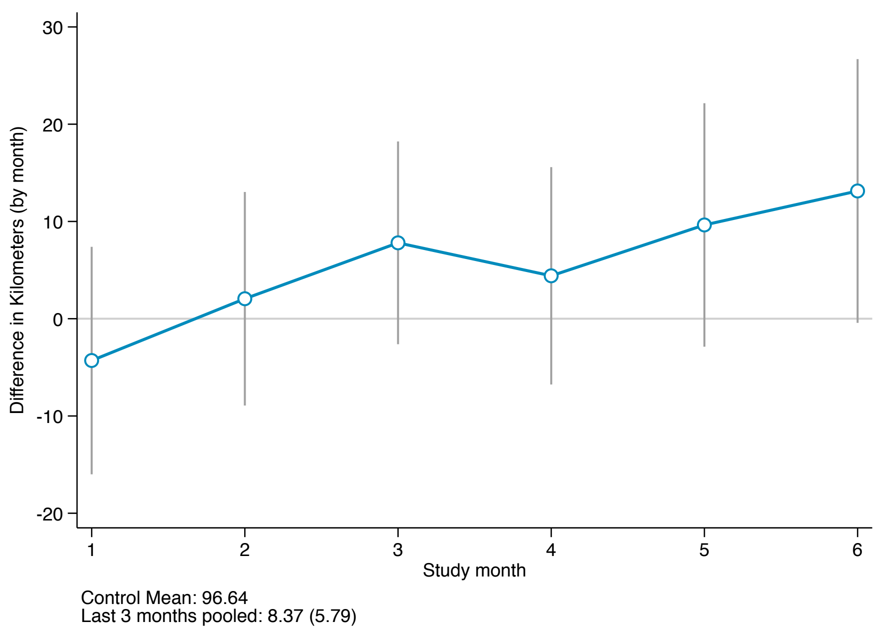




Notes: We plot the distributions of vertical and lateral accelerations for treatment and control vehicles from month 2 onwards (for consistency with the other pooled regressions, but the results are the same if month 1 is included). These acceleration measures are taken from the device directly. The distributions for vertical and lateral acceleration are centered at -200 and 100, respectively, rather than 0 because of some combination of a non exact calibration and the asymmetry of suspension resulting in asymmetrical acceleration.

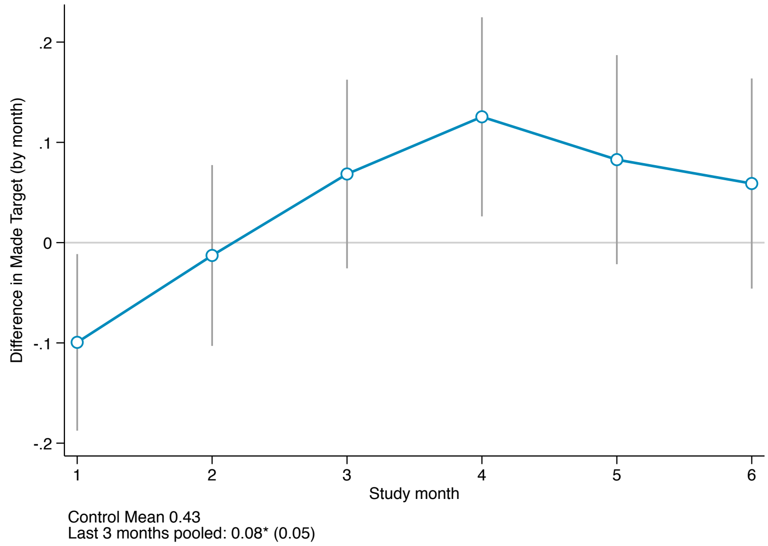
Figure 2.12: Prediction 3 → Effort

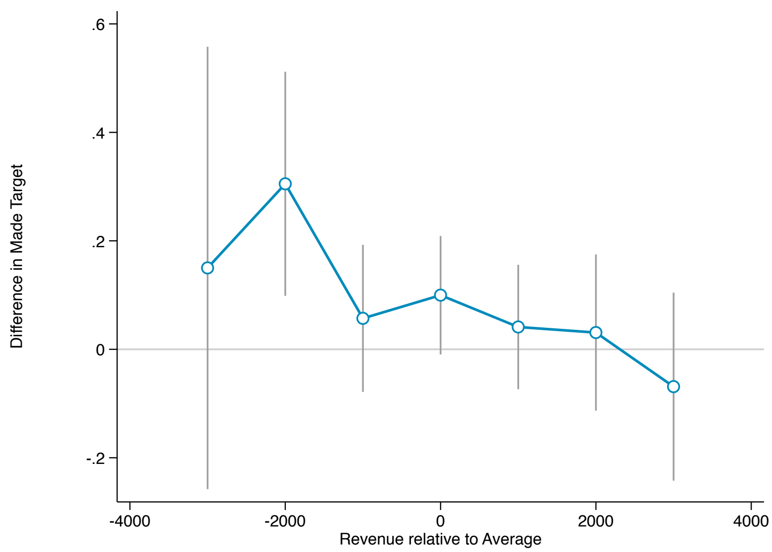




Notes: Each point on the graphs is the average difference between treatment and control in a particular month for a particular outcome. Note that because installations were rolled out over time, Month 1 (through 6) represent the first month (through the sixth month) since *installation*. The outcome in the first panel is number of hours the device was on. The device powers on and off with the vehicle, and thus provides a measure of effort. The outcome in the second panel is the number of kilometers driven. These data points were captured daily by the device.

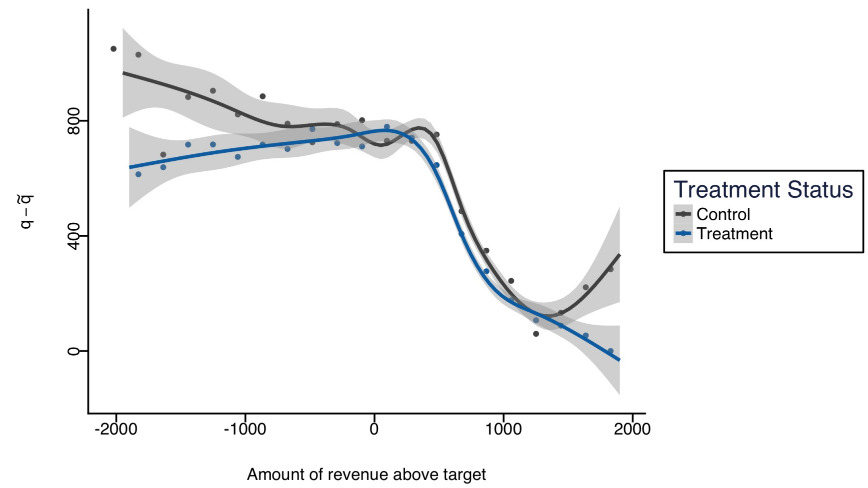
Figure 2.13: Prediction 5 → Achieving Target

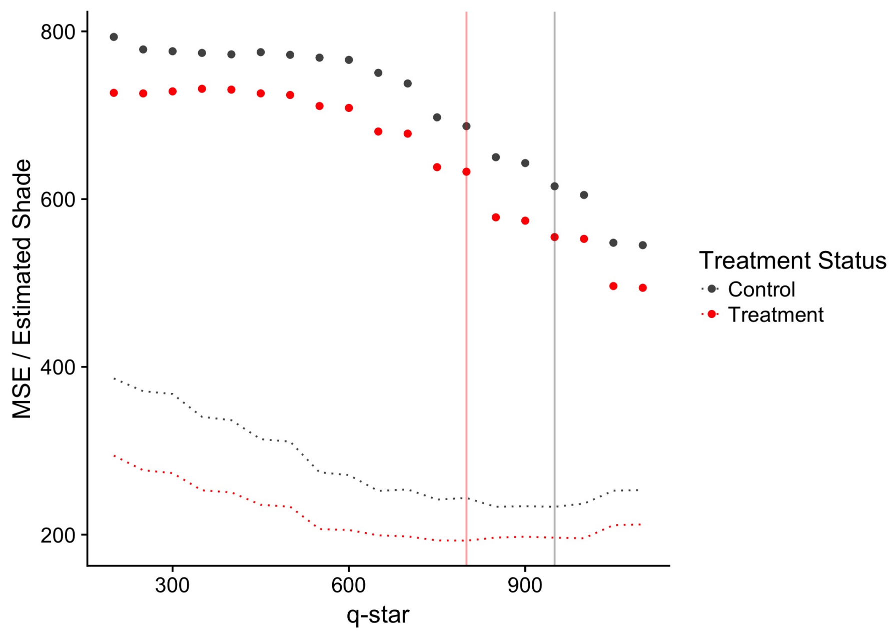




Notes: The outcome in these figures is whether or not the driver made the target (owner income = target). The data is collected from owner daily surveys. In the first panel, each point on the graph is the average difference in the probability of making the target between treatment and control in a particular month. Note that because installations were rolled out over time, Month 1 (through 6) represent the first month (through the sixth month) since *installation*. In the second panel we pool the data from months 2 onwards and look at whether the probability of making the target differs between treatment and control for a particular amount of revenue. The x-axis here depicts revenue relative to an average revenue day on that route (normalized by the target). Note we use gross revenue for this outcome instead of net revenue like we did for the under-reporting amount - because the variable only depends on drivers reporting, which means we have more data to work with.

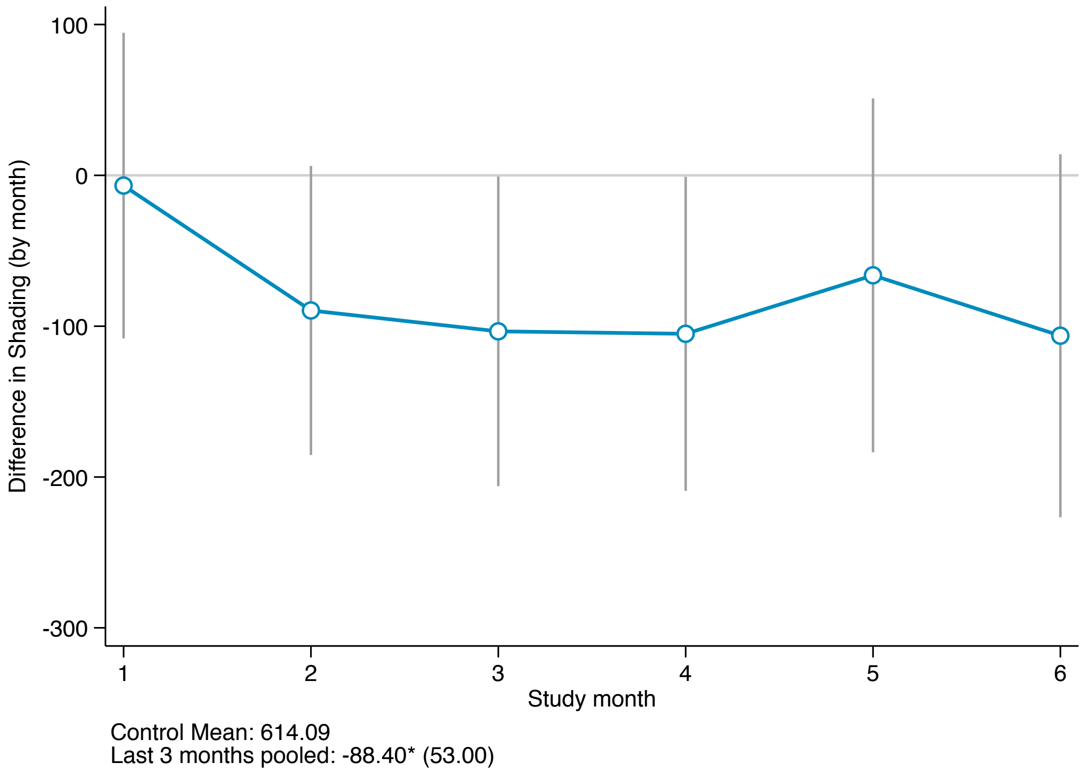
Figure 2.14: Prediction 4 → Less Under-reporting





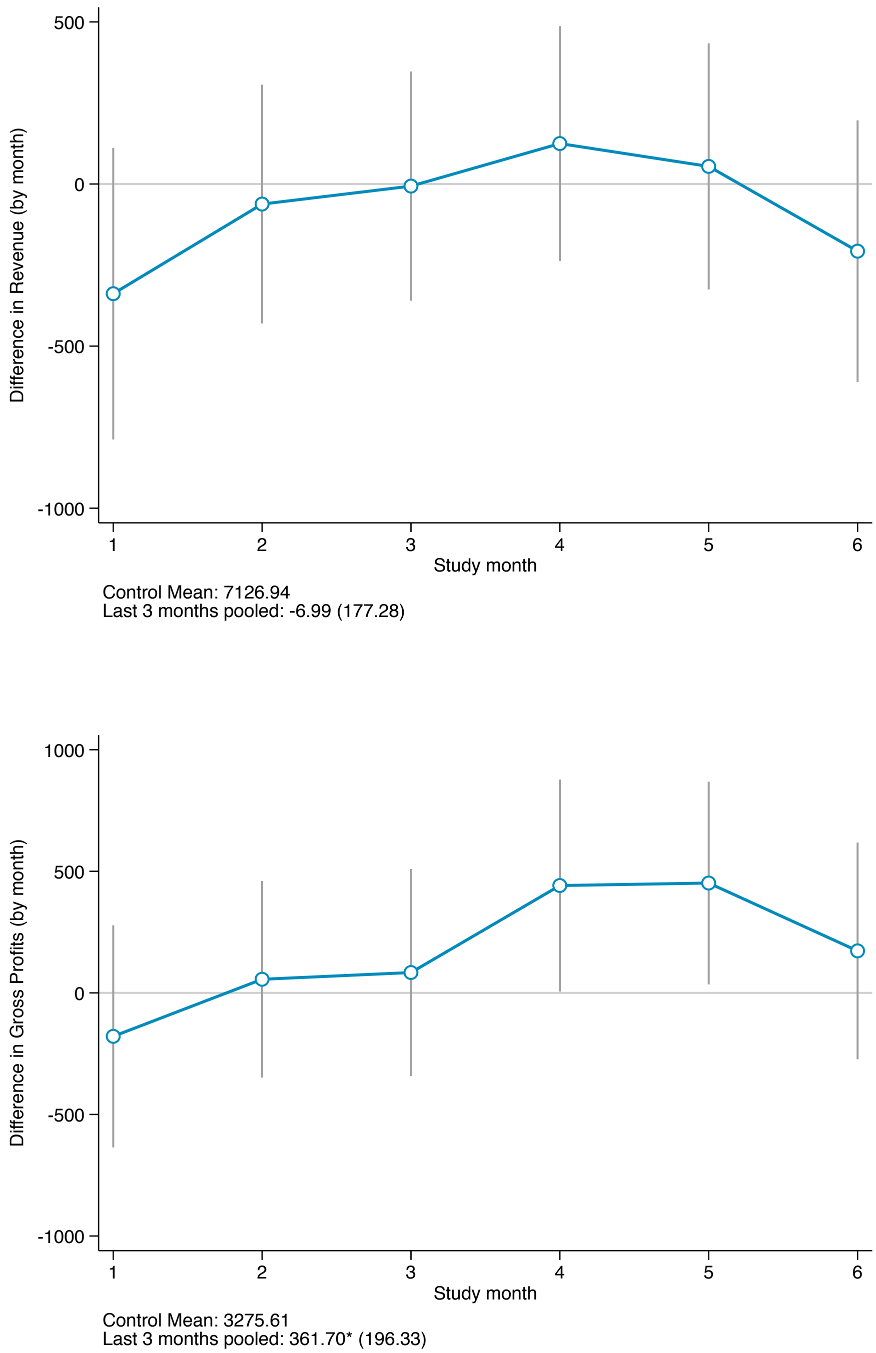
Notes: The outcome in these figures is the amount of revenue drivers under-report. The first panel reproduces Figure 2.5, but separates treatment drivers from control drivers. We also overlay a non-parametric smoothing function. The second panel imposes the model's step function and computes the average under-reported amount for different levels of q^{*} for treatment (red dots) and control (black dots). The dotted lines on the bottom represent the MSE from each regression.

Figure 2.15: Under-report



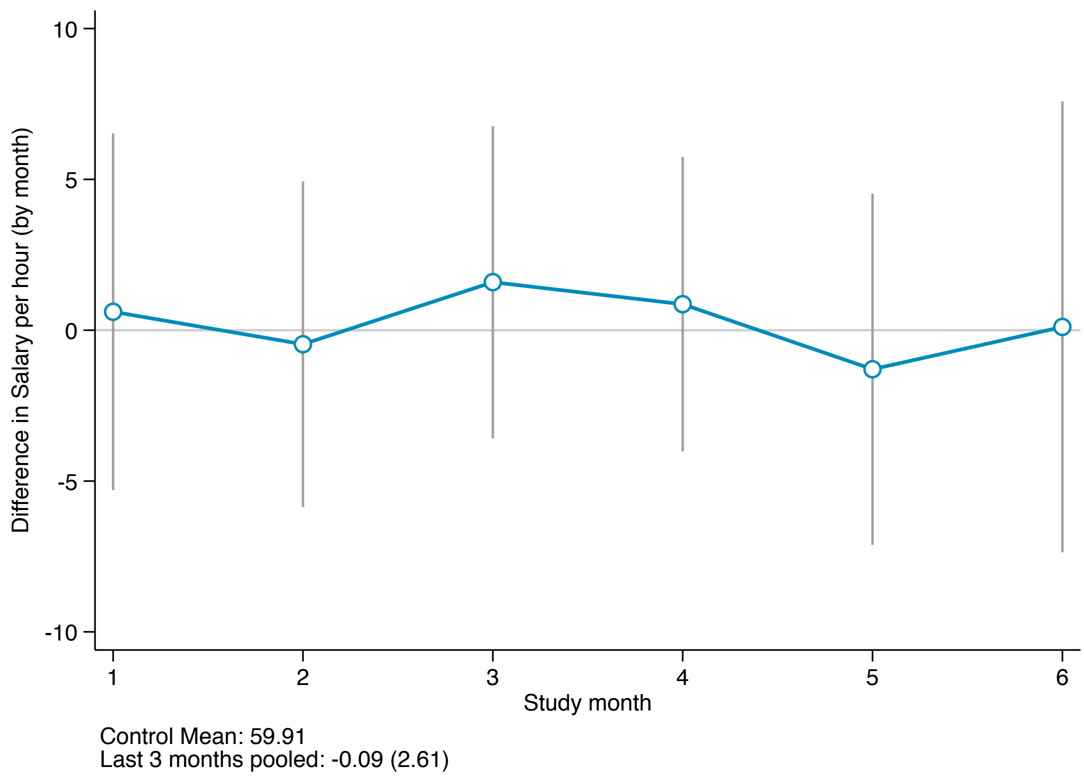
Notes: Each point on the graph is the difference between the average amount of revenue treatment and control drivers under-report in a particular month. The data was collected from owner and driver surveys - see Figure 2.5 for a description of how under-reported revenue was computed. Note that because installations were rolled out over time, Month 1 (through 6) represent the first month (through the sixth month) since *installation*.

Figure 2.16: Company Outcomes



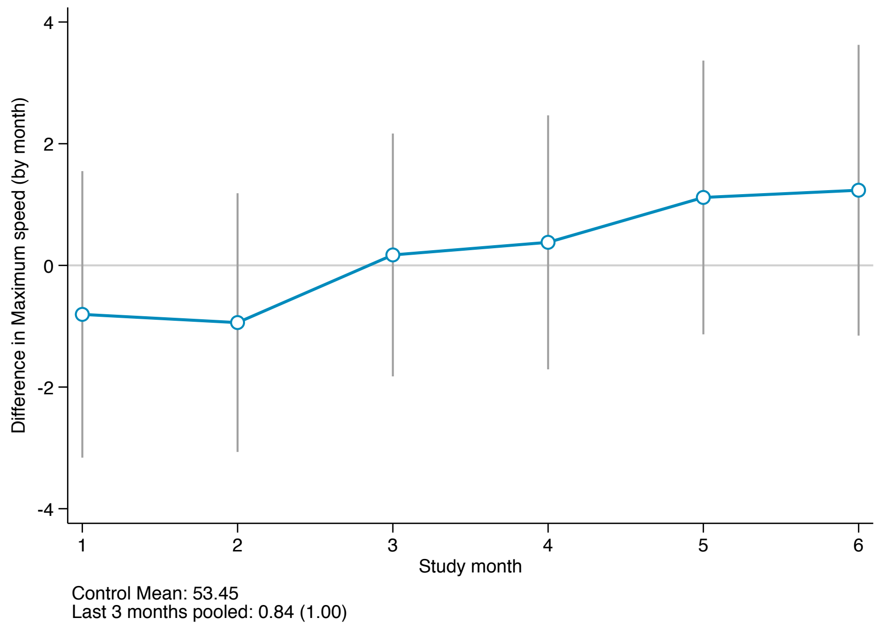
Notes: Each point on the graphs is the average difference between treatment and control in a particular month for a particular outcome. Note that because installations were rolled out over time, Month 1 (through 6) represent the first month (through the sixth month) since *installation*. The outcome in the first panel is the amount of revenue collected by the driver (captured through daily driver surveys). The outcome in the second panel is amount of profits the business generates, where $profits = revenue - costs - driversalary$ - which are captured from owner and driver daily reports.

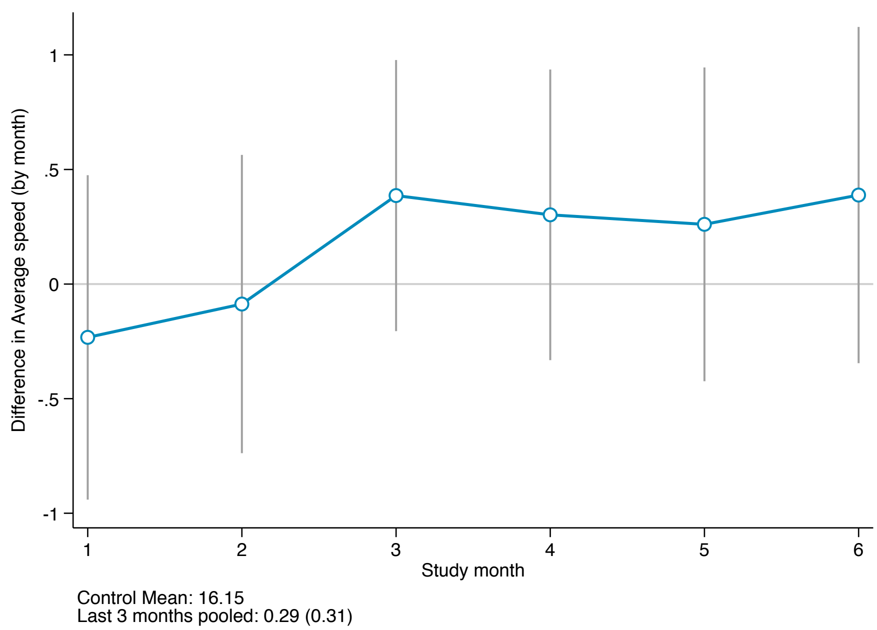
Figure 2.17: Salary per hour



Notes: Each point on the graph is the difference in driver salary per hour between treatment and control drivers. The data was collected from daily driver surveys. Note that because installations were rolled out over time, Month 1 (through 6) represent the first month (through the sixth month) since *installation*.

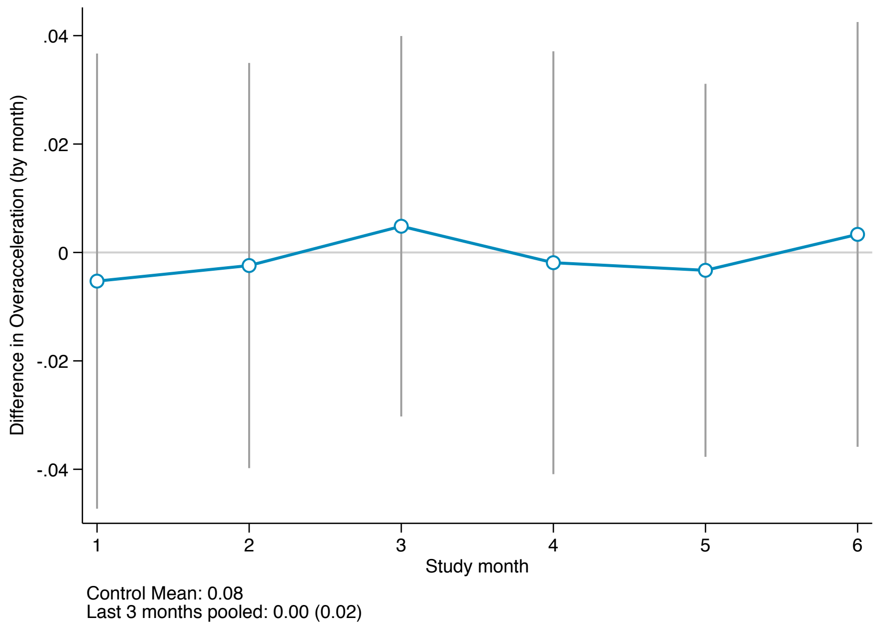
Figure 2.18: Speeding

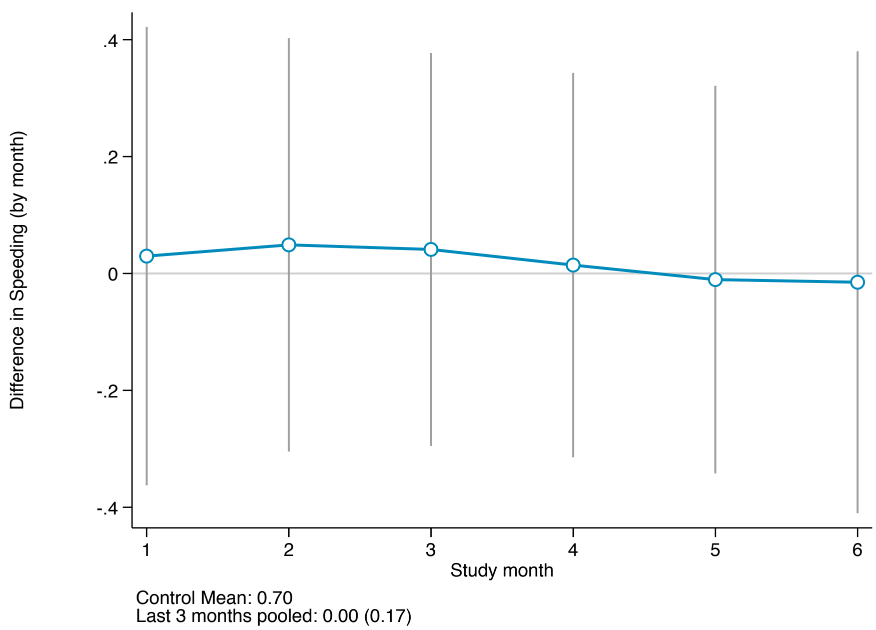




Notes: Each point on the graphs is the average difference between treatment and control in a particular month for a particular outcome. Note that because installations were rolled out over time, Month 1 (through 6) represent the first month (through the sixth month) since *installation*. The outcome in the first panel is maxspeed, while the outcome in the second panel is average speed. Both are measured directly by the tracker.

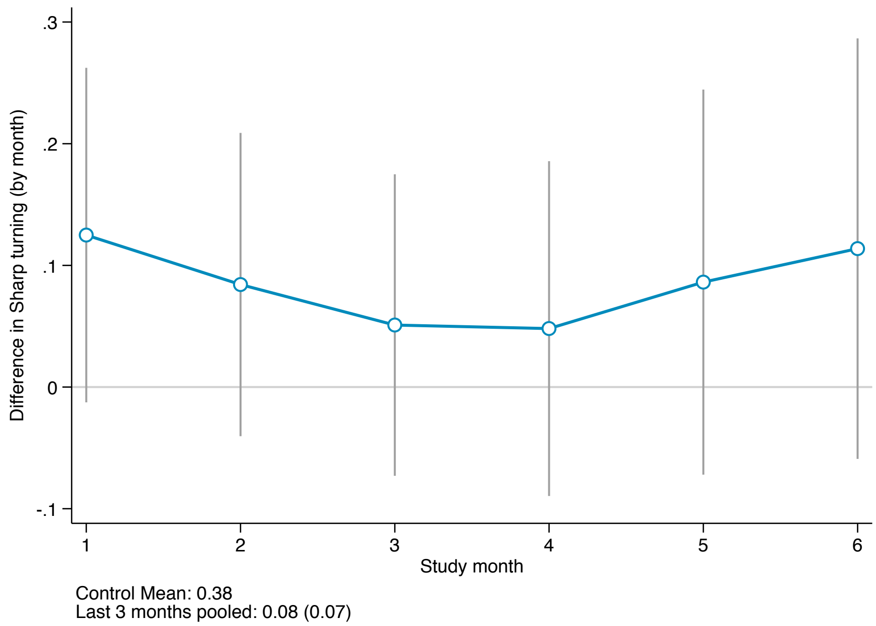
Figure 2.19: Over-acceleration and Over-speeding

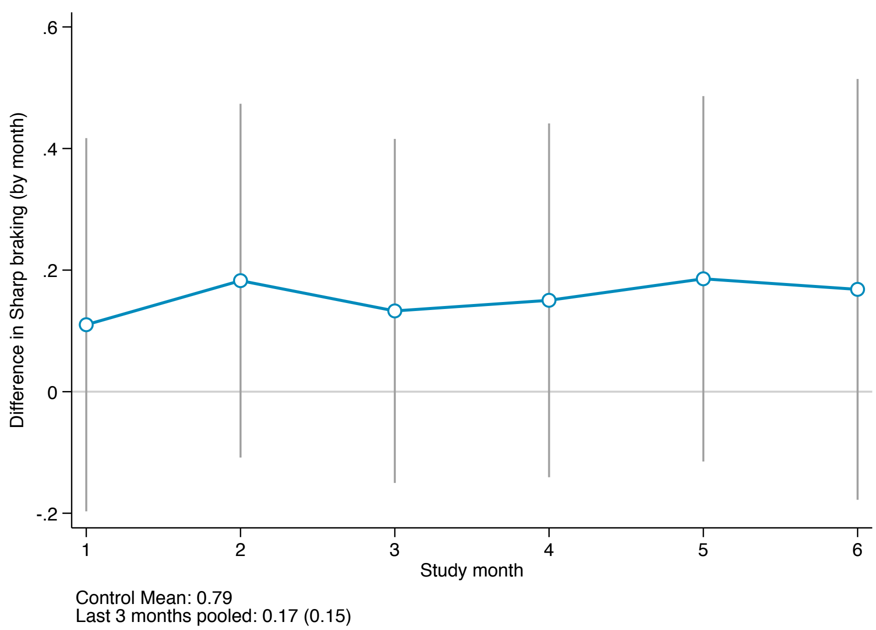




Notes: Each point on the graphs is the average difference between treatment and control in a particular month for a particular outcome. Note that because installations were rolled out over time, Month 1 (through 6) represent the first month (through the sixth month) since *installation*. The outcome in the first panel is over-acceleration alerts, while the outcome in the second panel is over-speeding alerts. Both are measured directly by the tracker

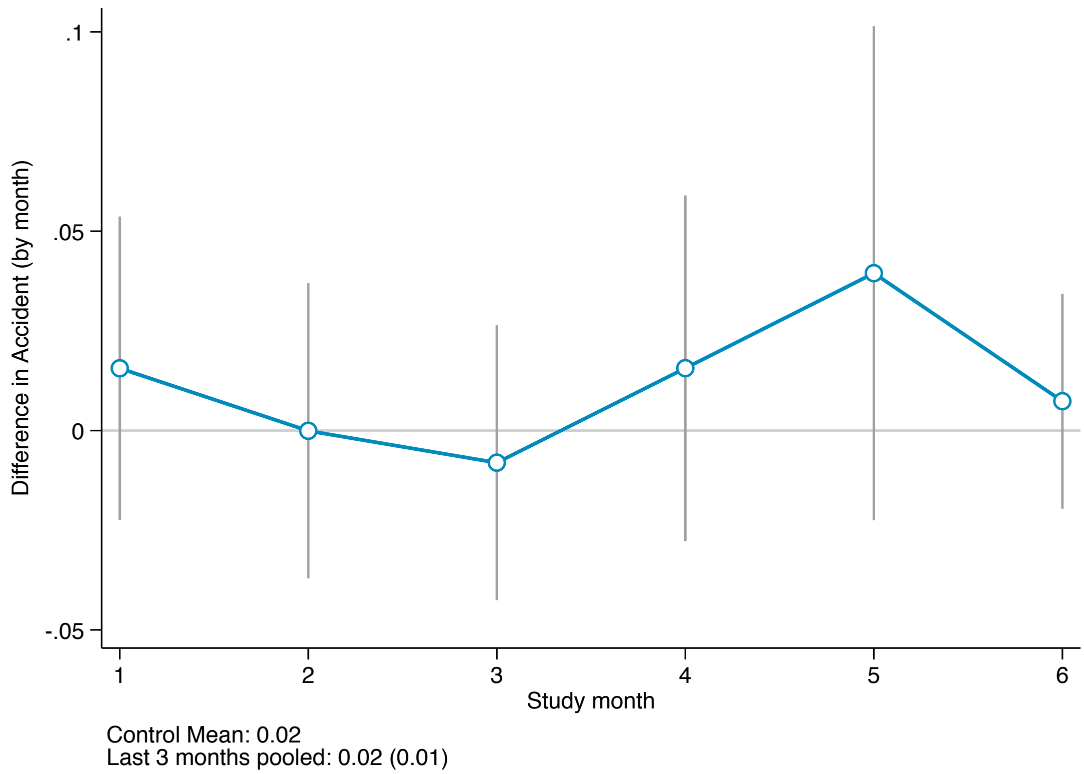
Figure 2.20: Sharp-turning and Sharp-braking





Notes: Each point on the graphs is the average difference between treatment and control in a particular month for a particular outcome. Note that because installations were rolled out over time, Month 1 (through 6) represent the first month (through the sixth month) since *installation*. The outcome in the first panel is sharp-turning alerts, while the outcome in the second panel is sharp braking alerts. Both are measured directly by the tracker.

Figure 2.21: Accidents



Notes: Each point on the graph is the difference in the number of accidents treatment and control drivers get into. The data was collected from daily owner/driver daily surveys, and then validated by an enumerator who called the owner directly. Note that because installations were rolled out over time, Month 1 (through 6) represent the first month (through the sixth month) since *installation*.

3 | Public Transit Safety and Consumer Choice

Chapter abstract: Road traffic accidents represent the eighth leading cause of death worldwide, claiming more than 1.35 million lives annually and causing up to 50 million injuries. Crashes disproportionately occur in low and middle income countries, where government capacity to effectively implement regulations is limited. Despite the severity of this issue, we know very little about the efficacy of different strategies designed to improve road safety in these contexts. To address this problem, we use a randomized control trial to examine whether disseminating safety information to passengers can help reduce their exposure to unsafe driving in Kenya's public transit industry. We fit six companies operating from Nairobi to Kisumu with GPS trackers, and intercept passengers as they enter the terminal. Some passengers receive a pamphlet about safety in general, while others receive a pamphlet that reveals which bus on the route is safest according to the GPS data. We cross randomize a second treatment, which provides incentives to certain passengers and not others to take a particular bus. We find three main behavioral responses. Passengers are willing to switch their bus choice in response to a monetary incentive. Rather surprisingly, however, passengers propensity to switch buses is reduced when the subsidy is combined with an explanation that the bus we are asking them to take is ranked safest on the route. Finally, there is some evidence that among passengers who we identify as caring about safety, the probability they choose a bus they *already* believed was safe ex-ante is increased when we make the issue of safety salient via the pamphlet. This suggests that passengers' failure to respond to the subsidy when we tell them which bus is safest may be due to their skepticism about the quality of the safety information we provided. This emphasizes the importance of disseminating information through channels that every-day passengers are familiar with and trust.

Road traffic injuries represent the eight leading cause of death worldwide, and the first cause of death among children and young adults aged 5-29 (WHO, 2018). The statistics have not improved in recent years despite the prevalence of new technologies designed to monitor unsafe driving. Most accidents are reported in middle and low income countries, where regulatory environments are weak. Low income countries feature 1% of the world's vehicles but 13% of all deaths (WHO, 2018). By way of comparison, high income countries feature 40% of the world's vehicles but

7% of all deaths. The situation is most dire in Africa, where road traffic fatalities are almost three times what they are in Europe. One of the main constraints in low and middle income countries is that legislation and compliance are lacking. Fewer than 15% of African countries have good laws governing speeding, drunk-driving or helmet use. Police are notoriously corrupt and do not enforce the limited laws that do exist. They collect bribes to supplement their income, thereby allowing unsafe driving to proliferate with impunity.

Small public transportation vehicles are amongst the leading perpetrators of unsafe driving in these countries. Bus drivers face incentives to drive recklessly in order to increase their take-home pay, and reach their destination more quickly. This endangers passengers on board, other vehicles on the road, and pedestrians. These dangerous driving conditions disproportionately affect the poor, who rely heavily on public transportation to get to work. Kenya's public transportation is dominated by minibuses or 'matatus' that are notorious for their unsafe driving: drivers often over-accelerate, speed, stop suddenly, and turn sharply in order to collect more passengers. According to one study in Kenya, matatus account for 11% of registered vehicles but 70.2% of passenger casualties (Macharia et al., 2009). As a point of comparison, buses in the US account for 1% of registered vehicles and 0.4% of casualties (BTS, 2016).

In response to these trends, private institutions have directed additional funding, knowledge and technical assistance towards building new systems that reduce the number of traffic injuries and deaths worldwide. This research builds on these endeavors by investigating the role that passengers can play in the industry to avoid unsafe driving. Globally, we have seen that commuters have the power to affect dynamics within the transportation industry. Companies like Uber have risen to prominence because of passenger demand - passengers value the convenience and services that Uber provides, and reward the company by taking their vehicles instead of standard taxis. These changes in the dynamics of the transportation industry rely on the fact that passengers have information about the various alternatives that they can take. If commuters are not informed, or misinformed, about the vehicles they take, the power of commuters to avoid unsafe driving and improve public transport services will be limited. Unfortunately, most commuters in developing countries have no way of systematically learning about how safe their public transportation choices are, and in many cases, they may underestimate the power of taking their business elsewhere, or find it too burdensome to change their habits.

In light of the role that passengers can play in shaping the transit industry, this paper investigates the effectiveness of an intervention designed to encourage passengers to take safer buses. The intervention is appealing because it bypasses the need for strong regulatory bodies and empowers private stakeholders, who may be in a better position to affect changes in road safety. To this end, we developed our own GPS tracking device that we designed specifically for the Kenyan minibus industry, which we fit to 60 medium-range minibuses traveling from Nairobi to Kisumu. This is one of the most dangerous routes in Kenya, connecting the capital to the western regional hub. This required forming close connections with the six companies operating on this route, and coordinating with managers to fit 15 vehicles in their fleet with our tracker. We then monitored the safety performance of the buses on a monthly basis and awarded a Top Safety Performer title to the bus company with the best safety performance that particular month (relative to other companies). The assessment of a bus' safety performance was primarily based on their incidents of speeding in the countryside, which is where most of the accidents take place. Finally, we spent six months intercepting thousands of passengers as they entered the terminal and randomly providing them with pamphlets about a) safe driving in general, or b) pamphlets that displayed which company was awarded the top safety performer title we generated. We then cross randomized this first "information treatment" with a second "subsidy treatment" where we offered small incentives for taking a particular bus. Our research design allows us to determine the extent to which passengers value safety, whether their decisions about which bus to take are affected by information about safe driving, and whether companies ultimately respond to passenger preferences.

In preliminary results, we have generated a number of novel insights that shed light on the role that passengers can play in shaping safer driving in the public transportation industry. First, we show that travelers react strongly to a small (10%) subsidy for the safest bus company: the share of travelers taking the safest company triples, nudging hundreds of travelers towards a safer public transportation experience. The size of this effect is significant, especially given the low cost of the intervention (the subsidy only cost 1 USD per traveler). Second, we show that while money can move travelers towards safer choices, the effect of information is much more complex. When we provide the subsidy, and explain that the subsidy is for the safest bus, we see fewer people switching to the safer bus relative to when we provided the subsidy with no information. Instead of taking the information at face value, travelers may have thought that our enumerators were making a sales pitch to them by offering a subsidy for a particular bus they claimed was safest. This finding sheds light on the importance of creating a trustworthy information environment for public information campaigns to succeed. This can often be achieved by ensuring that information campaigns are highly visible (i.e. online and easily accessible to all), and invite user feedback. Third, our preliminary findings suggest that highlighting safety as an important dimension of public transport services nudges travelers to fall back on their prior beliefs about which company is safest. This result underlines the role of designing public information campaigns on the basis of travelers' (potentially uninformed) prior beliefs about the safety of their public transport options.

There is relatively little previous work that has addressed the issue of road safety in low and middle income countries. The most closely related research is a paper by Habyarimana and Jack (2015) who investigate the impact of encouraging commuters to speak up when minibus drivers engage in dangerous driving through the use of stickers in minibus'. They find that encouraging passengers to express themselves lowers the average speed of minibuses operating on long distance routes in Kenya. This result suggests that the industry can be responsive to the demands of their customers. Our paper builds on this work by informing passengers about minibus safety ratings *before* they board the bus, which enables them to reduce the probability of encountering bad driving in the first place. Furthermore, information provision to customers has the potential to incentivize bus companies to improve their safety performance to increase demand.

There also exists some research on the role of monitoring devices in the public transit industry more broadly in both the developing and developed world. This body of work focuses primarily on the efficiency of the industry, and the relationship between bus owners and drivers. In a companion paper we focus on the impact that monitoring technologies have on the contractual relationship between owners and drivers, and how this subsequently affects the productivity of the business as a whole. In work focusing on Liberia's trucking industry, de Rochambeau (2018) finds that monitoring technologies improve performance of poorly motivated drivers but crowds out high performing drivers' intrinsic motivation. In the developed world, Baker and Hubbard (2003, 2004) explore how contracts and firm organization change as the use onboard diagnostic computers increases. We contribute to this literature by investigating the role of monitoring technologies in affecting the *quality* of transport services that are delivered to customers.

Overall, this project provides some of the first rigorous evidence on the role that information and subsidies can play on commuters' bus choices. These findings are policy relevant as they suggest new avenues that policymakers and private institutions can take to address the growing safety concern generated by reckless driving in low and middle income countries. First, small price subsidies turn out to be very effective in shifting traveler choices towards safer buses - provided passengers have some way of validating the way safety is being measured. If replicated on a large scale, this type of intervention has the potential to shift behavior on the supply side as well by providing strong incentives for competitors to catch up to the safest bus company. Second, large-scale information campaigns need to be built in such a way that passengers have faith in the information provided. We under-estimated the importance of this channel by providing information through pamphlets rather than a more trusted medium such as a mobile app or USSD short-code (a short digit sequence that individuals can text to receive short message about mobile operators' services). In future work we are placing large posters at the entrance of the terminal, to inform passengers that the companies on this particular route are being tracked for safety, and to encourage them to find out which company is the top safety performer. We will continue providing the same pamphlets to determine whether passengers responses to the information and subsidy treatments change as a result of these additional signs, which are aimed to build commuters trust in the tracking methodology we use.

The rest of the paper is organized as follows. Section 3.1 provides context on the matatu industry and the technology we used to measure safety. Section 3.2 outlines the experimental design and 3.3 gives an overview of the data used in the analysis. Section 3.4 reports the results from the experiment and finally, Section 3.5 concludes.

3.1 Context

3.1.1 Road Safety

According to the WHO's Global Status on Road Safety an estimated 1.35 million people are killed annually in road accidents and as many as 50 million individuals are injured worldwide (WHO, 2018). In Kenya alone, approximately 3,000 to 13,000 people die each year as a result of reckless driving (WHO, 2015). These road accidents disproportionally affect disadvantaged youths. According to the WHO's World report on road traffic injury prevention (2004), more than 75% of road traffic casualties involve economically productive young adults. Moreover, poor households are more at risk than others because they are constrained to take public transportation vehicles, which are notoriously unsafe in these environments.

Many international organizations are trying to implement solutions that reverse these trends. This is particularly difficult to achieve in these contexts because government regulations are non-existent, and corruption is rampant. Only 28 countries, representing 449 million people (and 7% of the world's population), have implemented laws that address all five risk factors (speed, drunk driving, helmets, seat-belts and child restraints). Moreover, less than 35% of low- and middle-income countries have policies in place to protect road users, despite experiencing the highest fatality rates in the world. In parallel, the few laws that do exist are poorly enforced. The Kenyan police service, for example, is one of the most corrupt institutions in the country, regularly extorting bribes from Kenyans, both on and off the road. A survey conducted by Transparency International among various stakeholders in the industry, revealed that most minibus drivers (approximately 75% of respondents) regularly pay bribes to the police (Transparency International, 2018).

3.1.2 Road Safety and the Semi-Formal Public Transit Industry

A significant share of road traffic accidents in Kenya involve public transportation providers. Kenya's transportation system grew out of the need for mobility in and around the major urban centers, which small-scale entrepreneurs responded to by retrofitting old vehicles and transporting passengers. These private businesses were legalized in 1973, but remained largely unregulated until 2003 when the government passed the Michuki rules, requiring that buses install speed limiters, safety belts, and ensure that all drivers exhibit valid licenses (Michuki, 2003). To date, these regulations are rarely enforced. In 2010, the Ministry of Transport issued a new directive to further formalize the industry and eliminate the presence of gangs that were becoming increasingly active in the sector. This required that all minibus owners form or join transport Savings and Credit Cooperatives (SACCOs) or transport companies licensed to a particular route (McCormick et al., 2013). At present, industry newcomers must first register with a SACCO or transport company before they can put their vehicle on the road. Transport companies and SACCOs both operate on routes between Kenya's major cities. Transport companies typically operate larger, safer, and more expensive buses while SACCOs operate smaller, cheaper matatus.

The industry is notoriously unsafe because drivers face incentives to operate recklessly on the road. In the case of long range buses, which we work with for this study, speeding on open stretches of road is common as drivers want to get to their destinations more quickly so they can complete more trips throughout the week. While these maneuvers are extremely dangerous in their own right, they are exacerbated in the presence of bumpy roads, non-existent sidewalks and other reckless drivers. According to one study in Kenya, minibuses account for 11% of registered vehicles but 70.2% of casualties (Macharia et al., 2009). As a point of comparison, buses in the US account for 1% of registered vehicles but only 0.4% of casualties (BTS, 2016). The United States Department of State recently cautioned travelers that road safety (along with crime) remains the biggest threat to traveling within the country, highlighting that accidents are frequent, and often fatal, when matatus are involved (Council, 2019).

3.1.3 Tracking Devices as a solution

While transit systems in low and middle income countries are widely viewed as unsafe, they are often difficult for most passengers to avoid as alternative forms of transportation are too costly. Furthermore, information about which buses are safest is difficult (if not impossible) to come by, which makes avoiding unsafe driving, and rewarding companies that implement safe driving practices, extremely difficult to achieve. One way to fight these dynamics is to work with governments to ensure regulations are implemented and properly enforced. As detailed in the previous sections, however, this is extremely difficult to accomplish in many low and middle income countries. An alternative is to empower passengers, which are the industry's key stakeholders. There have been some attempts to do just that in Kenya. Ma3 route is a mobile/web/SMS platform in Nairobi that crowd-sources for up to date transportation data, and provides users with information on traffic, matatu directions and driving reports.

We built on these grassroots initiatives by installing monitoring devices in minibuses, and conveying the information they captured about unsafe driving to passengers. This is an appealing solution because monitoring technologies are becoming widely available in many low-income countries, including in Kenya. In 2016, the Kenyan government mandated that long-range bus companies that travel between the country's main cities be equipped with tracking devices. Moreover, some banks in Nairobi recently announced that they would only issue loans for minibuses whose location could be tracked with a device. All of the companies that we worked with for the purposes of this research already had trackers installed in their vehicles. However, they did not maintain the systems, nor did they use the data the trackers collected because it was not stored in a user friendly way. It was important for our research team to be able to collect reliable data, which we did by installing our own tracker in the minibuses we worked with. This device was attractive to managers of long range minibuses because we also developed software that they could easily access on their phone at the end of the study, that provided the information they required in a user-friendly way.

The new system we designed was considerably cheaper, more flexible and more powerful than traditional tracking devices. The physical tracking units were procured for 125$ from a company in the United States (CalAmp). They feature GPS, internal back-up battery packs, 3-axis accelerometer for motion sense, tilt and impact detection. The device was designed to capture and transmit the information we required to evaluate the safety of each vehicle, including the 95th percentile and average forward/backward/lateral/vertical acceleration, as well as the 95th percentile and average forward/backward jerk. The device was also calibrated to generate alerts for every instance of vehicle speeding, over-acceleration, sharp braking and sharp turning. These safety alerts were calculated by an internal algorithm built into the CalAmp device with threshold parameters as inputs, using the full sequence of acceleration and speed data to identify unsafe driving actions. Further processing of the CalAmp system data on the server provided additional measures of interest including the total number of kilometers traveled that day, the total time the matatu was running, and a safety index (from aggregating the day's safety alerts).

The data we captured from each device was analyzed to determine the relative safety performance of each company we were working with. We first calculated the rate of safety violations for each fitted bus by dividing the number of alerts by the number of hours they operated. We took the average of this measure for all buses operating under the same SACCO to construct a company-wide rating. This measure formed the basis for the safety comparison between all SACCOs, which we then communicated to certain passengers through a pamphlet we provided. Upon completion of the study, managers received disaggregated data on each of their fitted buses via a mobile application that was specifically designed to present information simply. In addition to the safety measures, the app also provided information on the bus' real-time location, the exact route traveled, and summary statistics of the bus's activities to assist in the management of the minibus. The results of providing this type of information to matatu managers is not the focus of this paper, but appears in a companion paper.

3.2 Experimental Design

3.2.1 Bus Sample Recruitment

We began an extensive recruitment campaign in the summer of 2018. We canvased bus stations in Nairobi that provided medium and long-range bus services to other major cities throughout Kenya. Bus stations (or terminals) in Nairobi are located in the central business district, and generally refer to a city block where multiple SACCOs operate their buses from. We had several criteria in selecting the location for the experiments. First, the bus terminal area needed to be well defined so that potential passengers could be reliably distinguished from other pedestrians.

On a related point, we discarded any bus terminals that were too crowded and therefore too difficult to operate in. Second, we needed a sufficient number of different SACCOs to be operating from the bus station. Third, we required that the SACCOs operate their buses to the same destinations so that they were in direct competition with one another. Finally, we needed to secure the support of the majority of SACCO managers to fit their vehicles with our tracking device and to provide the information collected to potential passengers. After four months of scouting, we selected one bus station located on Mafangano Lane, which carries passengers to Kisumu. Five SACCOs operate from the bus station and all of them agreed to participate in our study, which meant fitting a minimum of six of their matatus with our tracking device. Each SACCO was compensated with a one-time payment of 5000 KES (50 USD) for the time their vehicles spent off-road to perform the installations of the devices. After the installations were completed, the manager of each SACCO received training on how to use of the phone app that displayed real-time information from the devices (which they would receive upon completion of the study).

After we completed the installations of the tracking devices, we collected one month of tracking data for each bus. We used this information to reliably compare the safety performance of one SACCO relative to another based on the average safety scores of their tracked matatus. The analysis of the safety data occurred every two months thereafter so that we could update the relative safety ranking of each SACCO. This process properly captured any improvements in SACCO safety performance that occurred over the study period.

3.2.2 Passenger Recruitment and Treatment Assignment

Our field team intercepted passengers as they entered the bus terminal area, but before they had an opportunity to purchase a ticket for a specific SACCO. We successfully completed 2016 passenger surveys using this approach. This represents a 30% interception rate on average across most of the days in our sample. For the most part passengers who declined to participate in the survey mentioned they did not have time to complete the survey, or they were not interested in participating. Study participants were compensated with 50 KES (50 cents USD). Enumerators would confirm before starting the survey that passengers were indeed traveling along the route serviced by the SACCOs participating in the experiment, and were willing to respond to the full set of questions (approximately 4 minutes of their time). It was made clear to passengers that the enumerators were not representing any SACCO and they were instead conducting independent research with a local NGO and their researcher partners in the US.

Once passengers agreed to participate and they completed the baseline survey, they were randomly assigned to one of three arms in the ‘information treatment": control, safety salience, and safety information. We created a pre-randomized list that matched a treatment assignment and a specific unique identifier (a “key"). The list was provided to EchoMobile’s field manager (our implementing partner), who then printed the three types of pamphlets with the unique identifier on each one.

The pamphlets were organized in the same order as the randomization lists and provided to the enumerators for distribution. The enumerators were unaware of the process.

In the control group, passengers received a pamphlet with a key printed on it, but nothing more. They were asked to respond to the baseline survey and report which bus they chose to a second enumerator stationed where the buses were leaving from. In the salience group, passengers were given a pamphlet that contained a message about the safety of the matatu industry, a picture of a matatu after an accident, and a list of the five SACCOs participating in the study (see Figure **??**). After handing over the pamphlet to the passengers, the enumerators would carefully read through the statements with them. Finally, passengers assigned to the safety information group, received the same pamphlet as the salience group with one notable exception: one of the five SACCOs appeared prominently with the message "Top Safety performer" alongside it. The enumerator carefully explained what this meant and how the title was awarded.

The three information treatment arms are designed to identify two effects. First, comparing the control group to the safety salience group identifies whether priming passengers about matatu safety affects their choice of vehicle. Without additional information about which of the five SACCOs operating on the route is safest, however, passengers have to rely only on their own priors about the safety (and other attributes) of each bus. Therefore, we would expect that being exposed to the salience treatment will increase the probability that passengers choose the SACCO that they previously believed to be the safest. Second, the comparison of the information group to the salience group will identify the impact of revealing new information about which SACCO is in fact the safest option on the route. To the extent that passengers use this new information to update their priors about safety, we would expect this treatment to increase the probability that passengers choose the bus marked as "Certified Safe."

Upon completing the baseline survey and receiving the associated pamphlet, each passenger was cross randomized into an additional "subsidy treatment". We selected half of the respondents to receive an additional 100 KES ($1 USD) subsidy should they purchase a ticket from the Top Safety performer. The other half of the respondents did not receive any additional incentives beyond the 50 KES that they were awarded for completing our surveys. When the subsidies were provided to passengers in the control and safety salience group, enumerators were careful not to tell passengers why this particular bus was being subsidized. When subsidies were provided to passengers in the safety information group, the enumerators pointed out that the subsidy was indeed for the bus that received the Top Safety performer title, as indicated on the pamphlet.

The introduction of the subsidy treatment serves two purposes. First, in the information control group, the effect of the subsidy captures how many passengers are willing to change their choice of bus in response to the subsidy. This provides a benchmark measure for how strong passengers preferences are, and how difficult it is to convince them to move away from their preferred bus. Second, the interaction between the subsidy and the salience and safety information treatment arms indicates how many more (or fewer) passengers switch their choice of bus once they have been primed to care about safety, or received information about which SACCO is safest. We would expect the subsidy to have *less* impact in the salience arm if passengers ex-ante believe that another SACCO is safest, while we expect the subsidy to have a *larger* impact on passengers in the information treatment.

3.3 Data Collection

We collected data from three different sources. First, we have passengers' survey responses and bus choices collected at the bus terminal area. Second, we collected travel experience surveys by asking passengers to rate their perceptions of how safe their journey was, which we collected via SMS once they completed their trip. Finally, we have GPS tracker data, which we use to measure driving behavior, and generate the safety ratings detailed above.

3.3.1 Passenger Surveys

We administered a baseline survey to passengers before handing them their assigned pamphlet. This survey collected information about their demographics, their experience riding matatus on this particular route; their preferences for various matatu characteristics including speed, comfort, safety, and style; and their beliefs about which of the five SACCOS was the best along each of these dimensions. After completing the survey, we explained to passengers that the five digit key at the bottom of their pamphlet should be presented to a second enumerator after they purchased their bus ticket in order to claim their 50 KES incentive (note that in in the case of control group the key constituted the entire pamphlet).

Passengers would then purchase their ticket and approach the second enumerator with their key and their ticket receipt. The second enumerator entered both these data points into an "endline survey" and issued the 50 KES reward to all passengers, and an additional 100 KES to those who were in the subsidy treatment and purchased a ticket with the Top Safety performer. The key was recorded to facilitate matching with baseline, and the ticket receipt was entered to record passengers' bus choice. Finally, we asked each passenger whether we could contact them later that day to conduct a travel experience survey. We recorded the phone number of those who agreed, and contacted them via SMS that evening. The survey asked them to rank the overall safety of their journey on a scale from 1 to 5, where 1 was very unsafe and 5 was very safe. We also asked them to flag any specific instances of unsafe driving such as speeding or near accidents.

3.3.2 Tracking data

The CalAmp tracking device transmitted high frequency data on forward/backward/lateral/vertical acceleration, jerk, location and a timestamp. The tracker subsequently fed the raw data into an algorithm that computed the number of safety events that occurred in a 30 second time frame. An “event" was recorded if these raw measures exceeded a certain threshold. We needed to calibrate the thresholds for the Kenyan roads because default thresholds captured an unreasonably high number of safety violations, thereby diluting the effectiveness of these measures of unsafe driving. These events included instances of speeding, over-acceleration, sharp braking, and sharp turns. The data was then further aggregated on the backend to produce reports on the number of safety violations, which is what we used to determine who the Top Safety performer was in a given month.

3.4 Results

3.4.1 Baseline Characteristics

We successfully intercepted 2,176 passengers during the 6 months we were in the field. We only worked on Friday’s and Saturday’s because footfall was higher and we needed to minimize costs. The passengers we surveyed are approximately 30 years old and have 14 years of education, which corresponds to two years of college (Table 3.1. This population is slightly older and more educated than average in Kenya, likely because we were working in the capital. Passengers use the SACCOs on this route approximately 13 times a year, which averages to about once a month. Anecdotally we know that many passengers travel home on a monthly basis. Table 3.1 tests for balance across the three information treatment arms and we find no statistical difference across groups in any of these baseline characteristics.

3.4.2 Subsidy, Salience and Information Treatment Arms

We run the following regression model to understand the impact of our treatments on passenger bus choices.

$$y_{ids} = \beta D_{ids} + \alpha_d + \tau_s + X_i\gamma + \varepsilon_{ids}$$

where y_{ird} is an indicator equal to 1 if passenger i on day d selected safety certified bus s; α_d is a day fixed effect; τ_s is a fixed effect for which bus was safety certified; and X_i are controls for the time of day and day of week when the passenger interview took place, and ε_{ids} is an error term. D is a treatment indicator. Specifically, we run one regression with two separate treatment indicators equal to 1 if passengers are in the salience and information arms, respectively (as the two did not interact). We run a second regression where the treatment indicator is equal to 1 if passengers received a subsidy for taking the Top Safety performer.

We are also interested in understanding how the subsidy and information treatments interact with one another. In order to test whether the salience or safety information arms increase the probability of opting for the safe SACCO with the subsidy we run the following regression model:

$$y_{ids} = \eta Subsidy + \beta salience + \psi Info + \pi Salience * Subsidy + \rho Info * Subsidy + \alpha_d + \tau_s + X_i\gamma + \varepsilon_{ids}$$

where y_{ird} is an indicator equal to 1 if passenger i on day d selected safety certified bus s; α_d is a day fixed effect; τ_s is a fixed effect for which bus was safety certified; and X_i are controls for the time of day and day of week when the passenger interview took place, and ε_{ids} is an error term. $Subsidy$ is an indicator equal to 1 if passengers were offered a subsidy to take the safe bus in the "subsidy treatment". $Salience$ and $Info$ are indicators for the salience and safety information arms of the "information treatment". $Salience*Subsidy$ and $Info*Subsidy$ represent the interactions between the two treatments, the results of which we focus on to test our main hypothesis.

Table 3.2 presents the effects of the price subsidy (Column 1), the two information treatments (Column 2), and their interactions (Column 3) on the probability that a passenger buys a ticket with the Top Safety Performer. Column 1 shows that the probability that passengers choose the safe bus increases by 31 percentage points with the subsidy. This represents a 200% increase in the probability of choosing a safe bus. This is a very large effect and demonstrates that many passengers have low switching costs and are willing to change their bus choice for a relatively modest 10% price subsidy ($1). Column 2 examines the effects of being in the salience and safety arms of the "information treatment". Making safety salient to passengers does not have any impact on the probability of choosing the safest bus. We anticipated this result, as passengers have only been primed about the importance of safety in general and are not aware of which bus is safest. Rather surprisingly, however, we see a zero (and if anything slightly negative) effect of the safety information treatment on the probability that passengers choose the safe bus. This result does not align with our predictions, and if taken at face value, suggests that passengers have a negative willingness to pay for bus safety.

Moving onto the interactions, Column 3 explores whether the effects of the salience and safety arms of the information treatment vary in the presence of the subsidy. The salience treatment arm had a negative impact on the probability of choosing the safest bus without the subsidy. This may be driven by the fact that without new information, passengers are more likely to choose a bus that they thought was safe previously. With the subsidy, however, the salience treatment arm has no effect on the probability that passengers choose the safe bus, thereby removing the negative effect of the salience treatment on it's own. The reversal of the base negative effect may come from the fact that passengers inferred that we were offering a subsidy for the bus that was safest (even though they were not told this explicitly). Next we consider the impact of revealing which company was the Top Safety performer on the route. We find that this treatment arm had no effect on the probability of choosing the safe bus without the subsidy. When coupled with the subsidy, however, it has a negative 6.6 percentage point effect (p-value of 0.054). This is a surprising result and suggests that passengers where unresponsive to the safety information on its own and reacted even more negatively when we coupled the information with the subsidy (a "backfire" effect). While the overall effect of the subsidy treatment is still highly positive, it is nearly 25% lower when coupled with the safety information. This first set of results is surprising, and on their own suggests that disseminating quantitative safety data is not sufficient to change behaviors. It may even deter the additional caution we were hoping to inspire under certain circumstances. We investigate some theories about what might be driving this negative effect in the next section.

Table 3.3 shows that effect of the interventions on the probability that passengers choose the bus they previously believed to be the safest at baseline. Column 1 shows that providing a subsidy to passengers for the Top Safety performer reduces the probability that they select the bus they believed to be the safest in the past by 10 percentage point, a 25% reduction. This is not surprising, as most passengers are receiving a subsidy to take a bus that they did not previously believe to be the safest. Column 2 reports on the impact of the salience and safety information treatment arms. In contrast to our previous results, we might expect that the salience treatment would encourage passengers to take the bus they initially believed to be safest by priming them to care about this issue. While we do find a positive point estimate, the salience treatment does not significantly encourage more passengers to take the bus they initially perceived to be the safest. While the safety information treatment had a similar null effects - this is a result we would have expected. Finally, Column 3 reports on whether these treatments varied with the presence of the subsidy. We find no evidence that there is any change in these null effects when the subsidy is also offered to passengers.

3.4.3 Heterogeneity

There are a few surprising results that warrant further investigation. First, the safety information on it's own did not induce passengers to take buses from the Top Safety performing SACCO. Moreover, when combined with a subsidy it appears to move passengers further *away* from the safe SACCO. Second, when primed about safety, we do not see passengers moving towards buses they previously thought were safe at baseline. It is important to try to understand what might be driving these unexpected results. The first hypothesis that we investigate is that passengers simply do not value safety, and hence we should not expect that safety information will influence their choices. We have already seen in Table 3.1 that 55% of passengers reported safety as their first priority when choosing a bus. However, it is possible that this response reflects social conformity, or interviewer demand effects. and not a true underlying preference for safety, or any significant willingness to pay for safety.

To test this hypothesis, we examine whether the responses to the salience and safety treatment arms differ among the group that reported safety as their first priority and those that did not. Table 3.4 highlights the effects of the two treatments on the probability passengers choose the Top Safety performer. Column 1 and 2 demonstrate the effect of the salience and safety information treatments for the sample that reported safety as most important, and those who did not, respectively. In neither sub-sample do we see passengers responding significantly to either treatment. Column 3 tests whether the probability of taking a Top Safety performer's bus in response to the treatments is the same across each sub-group. While we cannot reject equality of their responses, if anything we see that the passengers that reported caring most about safety are most responsible for driving the negative “backfire" effect.

While this result may suggest that passengers simply do not care about bus safety, it is also possible that passengers may not believe the information provided to them. They may be concerned about enumerators’ ulterior motives if they believe them to be making a sales pitch, and therefore distrust the information provided. In an attempt to disentangle these different possibilities, Table 3.5 examines whether the treatments have differential impacts on the likelihood that the two groups of passengers (those who value safety and those who don’t) take the bus they believed to be safe at baseline. If passengers as a whole do not significantly value safety, then no one will respond to the salience treatment. However, if some passengers do not trust the information provided to them, then we should see passengers who care about safety responding to the salience treatment, and increasing the probability they take the bus they already perceived to be safe.

Table 3.5, Column 1 reports the results for the group that rated safety as their top priority, and Column 2 shows the results for those who did not. We see that among passengers who care about safety, the salience treatment increases the likelihood they choose their perceived safe bus by 7.6 percentage points, which is an 18% increase. In contrast, there is no effect of the salience treatment among those who did not rate safety as their top priority. Interestingly, the safety information treatment has no effect in both groups, suggesting that the safety information is being heavily discounted by all passengers. Finally, Column 3 formally tests whether the two groups respond differently to the treatments, and shows that the passengers who care about safety are 9.5 percentage points more likely to take their perceived safe bus. This movement is approximately a third of the movement we observed from the 100 KES subsidy, suggesting that the salience treatment is equivalent (assuming a linear response) to a 33 KES subsidy for the bus passengers perceive to be safest.

The results above suggest that at least a subset of passengers care about safety and are likely to change which bus they choose in response to being primed about the issue. However, we find that passengers either do not respond or even respond *negatively* to the safety information treatment. This gives more weight to the hypothesis that passengers may not have trusted the enumerators as a credible source of information, or may have been suspicious of a scam and therefore responded by avoiding the bus suggested by the pamphlet. This may also be the reason why we observed a significantly negative effect of the information when it was provided with a subsidy in Table 3.2. To further explore this hypothesis, we examine whether passengers’ response to the treatment differs by whether they exhibit certain characteristics that may make them more (less) likely to believe (distrust) the safety information we provided. First, we examine differential responses by whether the passenger attended college or not, where we assume that higher educated passengers would be more skeptical of the new information. Second, we examine differential responses by whether passengers reported a strong ex-ante preference for one SACCO over another. We say that a passenger has a strong preference for a particular SACCO if they report the SACCO as being the best performer across all the attributes we asked about at baseline.

We focus first on education as the important source of heterogeneity. Table 3.6 demonstrates how passengers' choice of the Top Safety performer change in response to the salience and safety information treatment both with and without the subsidy. Column 1 reports the results for college educated passengers. We see that among this group, the subsidy is highly effective (37 percentage points) in moving passengers towards the safe bus. In contrast, neither the salience nor safety information have any effect on their own. Interestingly, when an educated passenger is presented with *both* the safety information and the subsidy, they are 16.5 percentage points *less* likely to take the safety certified bus, as 45% percent reduction in the subsidy's base treatment effect. This large reduction in the effectiveness of the subsidy when coupled with the safety information lends credibility to the hypothesis that they were a more skeptical group to begin with. They may have avoided taking the bus suggested by the enumerators because they found it suspicious to couple information with money. Scams are relatively common in Kenya while research is not, which could affect the assumptions people fall back on as we intercept them in the streets.

Column 2 in Table 3.6 shows results for the same regression among non-college educated passengers. Interestingly, the subsidy has a much more muted effect among this group at a 19.8 percentage increase in the probability of taking the certified bus, a 54% reduction from the college educated group. This is surprising, as we might have expected the non-college educated to be relatively less wealthy and therefore more price sensitive. Furthermore, we also see that the non-college educated react quite differently treatments when offered with a subsidy as compared to the college educated group. First, we see that when the salience treatment is given with a subsidy, passengers are 18.5 percentage points more likely to take the certified bus (cancelling out a nearly 10 percentage point decrease with salience alone). This is in contrast to a zero effect observed in the college group regardless of the presence of the subsidy. Second, we see that the non-college educated passengers are 12.8 percentage points more likely to take the certified bus when the safety information treatment is also given with a subsidy. While, this positive effect is not statically significant (p-value of 0.105), we can reject that the non-college educated group has a same negative response seen in the college group. This suggests that non-college educated passengers did not see the presence of the subsidy as potentially signaling a scam. Therefore, rather than creating the backfire effect seen among the college group, the subsidy had the intended outcome of magnifying the effectiveness of the salience and information treatments.

Next, we move on to examine whether we observe similar patterns between passengers who reported a strong ex-ante preference for one SACCO (approximately 17% of the sample) as compared to those who did not. Table 3.7 reports the results on the effects of the information and subsidy treatments across these two groups. Column 1 reports results for passengers with a strong preference for a particular SACCO. First, we observe that the base effect of the subsidy for this group is similar to the result for passengers overall. This suggests that passengers are equally willing to change their choice of bus in response to a price subsidy, despite their stronger preference for a particular SACCO. We see no base effect of the salience or safety information treatment on their own. However, as before, we see very large negative impacts of the safety information treatment when combined with a subsidy. The safety information lowers the probability that passengers choose the certified bus by 24 percentage points, which constitutes 75% of the base effect of the subsidy. This suggests that passengers with strong beliefs about the quality of a particular SACCO exhibit greater mistrust when money and information are combined. Column 2 reports the same results for passengers that do not have a strong preference for one particular SACCO. Here, we see that passengers have similar responses to the subsidy, but do not respond to the safety information treatment at all regardless of the presence of the subsidy. While our treatments do not nudge these people as we might have expected, we do not observe the backlash behavior that we observe among the group with strong preferences.

3.5 Conclusion

In recent years, international institutions have invested funding, knowledge and technical assistance into building systems aimed at reducing the number of traffic injuries and deaths worldwide (World Bank, 2014). These efforts are typically difficult to evaluate because the investments are multi-faceted and rolled out across an entire city. One exception is a program that was launched in Kenya, which placed stickers inside Nairobi's matatus to encourage passengers to complain to their drivers about unsafe driving (Habyarimana and Jack, 2015). Our intervention complements their approach by asking whether providing information to passengers about safe driving can change their choice of minibus, which could eventually put pressure on companies to reform. To this end, we implement a research design that allows us to rigorously test this hypothesis. This required forming close connections with SACCOs operating on the Nairobi - Kisumu route and fitting GPS trackers into their vehicles. We then tracked the companies' safety performance over the course of a month to understand which company on the route was safest relative to the others. Finally, we provided this information to passengers via pamphlets, where the exact information that was displayed on the pamphlet differed based on the treatment arm the passenger was assigned to. A certain subset of passengers also received a subsidy for taking a particular bus.

We find three main behavioral responses to our treatments. First, we show that passengers' choice of bus responds to a 10% subsidy for the safest bus company, causing the share of travelers taking the safest SACCO to triple. This is a significant effect, especially in light of the cost of the intervention. Second, our findings suggest that increasing the salience of bus safety causes *certain* passengers to rely on their prior beliefs about which company is safest and choose this SACCO. Finally, we show that while money can induce passengers towards safer choices, the effect of providing them alongside information about safety is much more complicated. Instead of taking the information at face value and updating their prior beliefs as expected, there is evidence that passengers perceive this intervention as part of a scam. This finding emphasizes the need to create a credible mechanism or institution to transmit information to the public to build passengers trust and encourage them to make better informed choices.

These results are important for a number of key stakeholders, including policy makers working to improve road safety conditions in developing countries. Policy makers have struggled to find solutions to improving road safety in environments where regulation is weak and the police face incentives to collect bribes. In these contexts, empowering customers of public transit vehicles is discussed as an alternative way to improve safety outcomes. Our results highlight that the method of delivery of this information is vital in order for it to have its intended effects. In the next phase of this research we are experimenting with alternative ways of providing safety information to passengers. Namely, we have created large posters that indicate to passengers that the whole route is being tracked for safety, and encourages them to find out who the top safety performer is. We will continue with the distribution of pamphlets and see whether the results of our intervention change accordingly. Future work might think about developing a mobile application or a website that lends even more credibility to the information source

Tables

Table 3.1: Summary Statistics and Balance Test

	(1)	(2)	(3)	(4)	(5)	(6)
	Full Sample	Control	Salience	Information	Diff Control-Salience	Diff Control-Info
Age	29.65	30.04	29.36	29.43	0.69	0.62
	(7.30)	(7.46)	(7.17)	(7.29)	(0.36)	(0.41)
Years Education	13.60	13.56	13.66	13.57	-0.11	-0.01
	(3.10)	(3.09)	(3.03)	(3.20)	(0.15)	(0.18)
Female	0.56	0.55	0.57	0.54	-0.02	0.01
	(0.50)	(0.50)	(0.50)	(0.50)	(0.02)	(0.03)
Yearly Bus Use	13.64	13.78	13.23	14.06	0.54	-0.28
	(17.95)	(18.17)	(17.49)	(18.35)	(0.87)	(1.02)
Safety Most Important	0.55	0.55	0.56	0.54	-0.01	0.01
	(0.50)	(0.50)	(0.50)	(0.50)	(0.02)	(0.03)
Observations	2176	832	811	509	1667	1365

Notes: Column 1 shows summary statistics for the full sample. Columns 2-4 show summary statistics for the three treatment arms (control, safety salience, and safety information). Columns 5 and 6 the difference between the control group and the salience and information group respectively. Standard deviations and standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.2: Choice of Safety Certified Bus

	(1) Chose Safe Bus	(2) Chose Safe Bus	(3) Chose Safe Bus
Subsidy	0.312***		0.295***
	(0.018)		(0.030)
Salience		-0.017	-0.058*
		(0.022)	(0.030)
Safety		-0.041	-0.020
		(0.027)	(0.035)
Salience X Subsidy			0.074*
			(0.042)
Safety X Subsidy			-0.046
			(0.048)
Controls	Yes	Yes	Yes
Mean in Control	0.16	0.16	0.16
p-value Salience + Salience X Subsidy			0.595
p-value Safety + Safety X Subsidy			0.054
Observations	2150	2174	2150

Notes: The outcomes is an indicator for the respondent buying a ticket for the safest bus company as measured by the tracking devices. “Subsidy" is the subsidy treatment where passengers were given a 100 Ksh discount to take the safest bus. “Salience" is an indicator for passengers receiving a pamphlet that increases the salience of safety on matatus. “Safety" is an indicator for passengers receiving a pamphlet that indicates which bus has been “safety certified". * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.3: Choice of Perceived Safe Bus

	(1) Perceived Safe Bus	(2) Perceived Safe Bus	(3) Perceived Safe Bus
Subsidy	-0.105***		-0.116***
	(0.024)		(0.038)
Salience		0.022	-0.007
		(0.027)	(0.039)
Safety		-0.010	0.005
		(0.032)	(0.045)
Salience X Subsidy			0.053
			(0.054)
Safety X Subsidy			-0.038
			(0.061)
Controls	Yes	Yes	Yes
Mean in Control	0.44	0.44	0.44
p-value Salience + Salience X Subsidy			0.219
p-value Safety + Safety X Subsidy			0.443
Observations	1728	1749	1728

Notes: The outcomes is an indicator for the respondent buying a ticket for the bus company they previously believe was safest. ‘Subsidy" is the subsidy treatment where passengers were given a 100 Ksh discount to take the safest bus. “Salience" is an indicator for passengers receiving a pamphlet that increases the salience of safety on matatus. “Safety" is an indicator for passengers receiving a pamphlet that indicates which bus has been “safety certified". * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.4: Choice of Safety Certified Bus by Safety Importance

	(1) Chose Safe Bus Safety Most Important	(2) Chose Safe Bus Safety Not Most Important	(3) Chose Safe Bus Full Sample
Salience	-0.007	-0.026	-0.034
	(0.031)	(0.035)	(0.033)
Safety	-0.055	-0.016	-0.018
	(0.037)	(0.040)	(0.038)
Salience X Care Safety			0.031
			(0.045)
Safety X Care Safety			-0.044
			(0.051)
Safety Most Important			-0.007
			(0.031)
Controls	Yes	Yes	Yes
Mean in Control	0.17	0.15	0.16
p-value Salience + Salience X Care			0.920
p-value Safety + Safety X Care			0.086
Observations	1191	983	2174

Notes: The outcomes is an indicator for the respondent buying a ticket for the safest bus company as measured by the tracking devices. “Care Safety" is an indicator for passengers who rated safety as their top criteria for taking a bus. “Salience" is an indicator for passengers receiving a pamphlet that increases the salience of safety on matatus. “Safety" is an indicator for passengers receiving a pamphlet that indicates which bus has been “safety certified". Column 1 limits the sample only to passengers who rated safety as their highest priority. Column 2 limits the sample to passengers who listed some other attribute as their highest priority. Column 3 includes the entire sample. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.5: Choice of Perceived Safe Bus by Safety Importance

	(1) Perceived Safe Bus Safety Most Important	(2) Perceived Safe Bus Safety Not Most Important	(3) Perceived Safe Bus Full Sample
Salience	0.076*	-0.031	-0.027
	(0.039)	(0.040)	(0.039)
Safety	0.023	-0.040	-0.037
	(0.047)	(0.046)	(0.044)
Salience X Care Safety			0.095*
			(0.054)
Safety X Care Safety			0.056
			(0.061)
Safety Most Important			-0.035
			(0.038)
Controls	Yes	Yes	Yes
Mean in Control	0.42	0.46	0.44
p-value Salience + Salience X Care			0.072
p-value Safety + Safety X Care			0.670
Observations	901	848	1749

Notes: The outcomes is an indicator for the respondent buying a ticket for the bus company they previously believe was safest. “Care Safety" is an indicator for passengers who rated safety as their top criteria for taking a bus. “Salience" is an indicator for passengers receiving a pamphlet that increases the salience of safety on matatus. “Safety" is an indicator for passengers receiving a pamphlet that indicates which bus has been “safety certified". Column 1 limits the sample only to passengers who rated safety as their highest priority. Column 2 limits the sample to passengers who listed some other attribute as their highest priority. Column 3 includes the entire sample. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.6: Choice of Safety Certified Bus by Higher Education

	(1) Chose Safe Bus - College	(2) Chose Safe Bus - No College
Subsidy	0.367***	0.198***
	(0.038)	(0.047)
Salience	-0.035	-0.097*
	(0.039)	(0.049)
Safety	0.003	-0.039
	(0.046)	(0.058)
Salience X Subsidy	0.008	0.185***
	(0.054)	(0.070)
Safety X Subsidy	-0.165***	0.128
	(0.062)	(0.079)
Controls	Yes	Yes
Mean in Control	0.16	0.17
p-value Salience + Salience X Subsidy	0.468	0.078
p-value Safety + Safety X Subsidy	0.000	0.113
Observations	1343	807

Notes: The outcomes is an indicator for the respondent buying a ticket for the safest bus company as measured by the tracking devices. "Salience" is an indicator for passengers receiving a pamphlet that increases the salience of safety on matatus. "Safety" is an indicator for passengers receiving a pamphlet that indicates which bus has been "safety certified". Column 1 limits the sample only to passengers who have some college education. Column 2 limits the sample to passengers who have no college education. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.7: Choice of Safety Certified Bus by Strong Bus Preference

	(1) Chose Safe Bus - Strong Pref.	(2) Chose Safe Bus - No Pref.
Subsidy	0.309***	0.287***
	(0.071)	(0.033)
Salience	0.003	-0.065**
	(0.076)	(0.033)
Safety	0.013	-0.026
	(0.092)	(0.039)
Salience X Subsidy	0.115	0.072
	(0.108)	(0.047)
Safety X Subsidy	-0.238**	-0.008
	(0.121)	(0.053)
Controls	No	No
Mean in Control	0.10	0.17
p-value Salience + Salience X Subsidy	0.120	0.843
p-value Safety + Safety X Subsidy	0.008	0.369
Observations	344	1801

Notes: The outcomes is an indicator for the respondent buying a ticket for the safest bus company as measured by the tracking devices. “Salience" is an indicator for passengers receiving a pamphlet that increases the salience of safety on matatus. “Safety" is an indicator for passengers receiving a pamphlet that indicates which bus has been “safety certified". Column 1 limits the sample only to passengers who reported the same bus as having being the best in all bus attributes. Column 2 limits the sample to passengers listed different buses as being the best in different attributes. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figures

Figure 3.1: Safety Saliency Pamphlet

Notes: The pamphlet was distributed to passengers in the safety saliency groоup after completion of the baseline survey. Enumerators read the text out loud to each passenger to ensure understanding. The "KEY: #####" was a randomly generated number that allowed passenger baselines to be matched to their ticket purchase collected by a second enumerator.

Figure 3.2: Safety Information Pamphlet

Notes: The pamphlet was distributed to passengers in the safety information group after completion of the baseline survey. Enumerators read the text out loud to each passenger and explained the meaning of the "Safety Certified" designation to ensure understanding. The "KEY: #####" was a randomly generated number that allowed passenger baselines to be matched to their ticket purchase collected by a second enumerator.

Bibliography

Angelucci, Manuela, Dean Karlan, and Jonathan Zinman. 2015. "Microcredit Impacts: Evidence from a Randomized Microcredit Program Placement Experiment by Compartamos Banco." *American Economic Journal: Applied Economics* 7 (1):151–182.

Baker, George P. and Thomas N. Hubbard. 2003. "Make Versus Buy in Trucking: Asset Ownership, Job Design, and Information." *American Economic Review* 93 (3):551–572.

———. 2004. "Contractibility and Asset Ownership: On-Board Computers and Governance in U. S. Trucking." *The Quarterly Journal of Economics* 119 (4):1443–1479.

Banerjee, Abhijit, Esther Duflo, Rachel Glennerster, and Cynthia Kinnan. 2015. "The Miracle of Microfinance? Evidence from a Randomized Evaluation." *American Economic Journal: Applied Economics* 7 (1):22–53.

Banerjee, Abhijit, Dean Karlan, and Jonathan Zinman. 2015. "Six Randomized Evaluations of Microcredit: Introduction and Further Steps." *American Economic Journal: Applied Economics* 7 (1):1–21.

Barboni, Giorgia. 2017. "Repayment Flexibility in Microfinance Contracts: Theory and Experimental Evidence on Take up and Selection." *Journal of Economic Behavior & Organization* 142:425–450.

Beaman, Lori, Dean Karlan, Bram Thuysbaert, and Christopher Udry. 2014. "Self-Selection into Credit Markets: Evidence from Agriculture in Mali." Working Paper 20387, National Bureau of Economic Research.

Berge, Lars Ivar Oppedal, Kjetil Bjorvatn, and Bertil Tungodden. 2014. "Human and Financial Capital for Microenterprise Development: Evidence from a Field and Lab Experiment." *Management Science* 61 (4):707–722.

Björkman, Martina and Jakob Svensson. 2009. "Power to the People: Evidence from a Randomized Field Experiment on Community-Based Monitoring in Uganda." *The Quarterly Journal of Economics* 124 (2):735–769.

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

Bloom, Nicholas, Gregory Fischer, Imran Rasul, Andreas Rodriguez-Clare, Tavneet Suri, Christopher Udry, Eric Verhoogen, Christopher Woodruff, and Gulia Zane. 2014. "Firm Capabilities and Economic Growth." Evidence Paper, IGC.

Bloom, Nicholas, Raffaella Sadun, and John Van Reenen. 2017. "Management as Technology?" Working Paper, NBER.

Brammer, H. 1990. "Floods in Bangladesh: Geographical Background to the 1987 and 1988 Floods." *The Geographical Journal* 156 (1):12–22.

Casaburi, Lorenzo and Jack Willis. 2018. "Time vs. State in Insurance: Experimental Evidence from Contract Farming in Kenya." *American Economic Review* .

Clarke, Daniel J. 2016. "A Theory of Rational Demand for Index Insurance." *American Economic Journal: Microeconomics* 8 (1):283–306.

Cole, Shawn, Xavier Giné, Jeremy Tobacman, Petia Topalova, Robert Townsend, and James Vickery. 2013. "Barriers to Household Risk Management: Evidence from India." *American Economic Journal: Applied Economics* 5 (1):104–135.

Cole, Shawn A. and Wentao Xiong. 2017. "Agricultural Insurance and Economic Development." *Annual Review of Economics* 9 (1):235–262.

Council, Overseas Security Advisory. 2019. "Kenya 2019 Crime and Safety Report." Tech. rep., US Department of State.

de Mel, Suresh, David McKenzie, and Christopher Woodruff. 2014. "Business Training and Female Enterprise Start-up, Growth, and Dynamics: Experimental Evidence from Sri Lanka." *Journal of Development Economics* 106:199–210.

de Rochambeau, Golvine. 2018. "Monitoring and Intrinsic Motivation: Evidence from Liberia's Trucking Firms." Working Paper, Columbia University.

Deaton, Angus. 1991. "Saving and Liquidity Constraints." *Econometrica* 59 (5):1221–1248.

———. 1992. "Household Saving in LDCs: Credit Markets, Insurance and Welfare." *Scandinavian Journal of Economics* 94 (2):253–73.

Demont, Timothée. 2014. "Microcredit as Insurance: Evidence from Indian Self-Help Groups." Tech. Rep. 1410, University of Namur, Department of Economics.

Dercon, Stefan. 2002. "Income Risk, Coping Strategies, and Safety Nets." *The World Bank Research Observer* 17 (2).

Donovan, Kevin. 2016. "Agricultural Risk, Intermediate Inputs, and Cross-Country Productivity Differences." *Unpublished Paper, University of Notre Dame* .

Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan. 2013. "Truth-Telling by Third-Party Auditors and the Response of Polluting Firms: Experimental Evidence from India." SSRN Scholarly Paper ID 2294736, Social Science Research Network, Rochester, NY.

Duflo, Esther, Rema Hanna, and Stephen P. Ryan. 2012. "Incentives Work: Getting Teachers to Come to School." *American Economic Review* 102 (4):1241–1278.

Field, Erica and Rohini Pande. 2010. "Repayment Frequency and Default in Microfinance: Evidence from India." *Journal of the European Economic Association* 6 (2-3):501–509.

Field, Erica, Rohini Pande, John Papp, and Natalia Rigol. 2013. "Does the Classic Microfinance Model Discourage Entrepreneurship among the Poor? Experimental Evidence from India." *American Economic Review* 103 (6):2196–2226.

Ghatak, Maitreesh and Priyanka Pandey. 2000. "Contract Choice in Agriculture with Joint Moral Hazard in Effort and Risk." *Journal of Development Economics* 63:303–326.

Gross, David B. and Nicholas S. Souleles. 2002. "Do Liquidity Constraints and Interest Rates Matter for Consumer Behavior? Evidence from Credit Card Data." *The Quarterly Journal of Economics* 117 (1):149–185.

Grossman, Sanford J. and Oliver D. Hart. 1983. "An Analysis of the Principal-Agent Problem." *Econometrica* 51 (1):7–45.

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 of the United States of America* 112 (34):E4661–4670.

Harris, Milton and Artur Raviv. 1979. "Optimal Incentive Contracts with Imperfect Information." *Journal of Economic Theory* 20 (2):231–259.

Hart, Oliver and Bengt Holmstrom. 1987. *The Theory of Contracts*. Cambridge: Cambridge University Press, advances in economic theory ed.

Hoddinott, John. 2006. "Shocks and Their Consequences Across and Within Households in Rural Zimbabwe." *The Journal of Development Studies* 42:301–321.

Hölmstrom, Bengt. 1979. "Moral Hazard and Observability." *The Bell Journal of Economics* 10 (1):74–91.

Hubbard, Thomas N. 2003. "Information, Decisions, and Productivity: On-Board Computers and Capacity Utilization in Trucking." *American Economic Review* 93 (4):1328–1353.

Janzen, Sarah A and Michael R Carter. 2018. "After the Drought: The Impact of Microinsurance on Consumption Smoothing and Asset Protection." Working Paper 19702, National Bureau of Economic Research.

Jensen, Nathaniel and Christopher Barrett. 2017. "Agricultural Index Insurance for Development." *Applied Economic Perspectives and Policy* 39 (2):199–219.

Karlan, Dean. 2014. "Innovation, Inclusion and Trust: The Role of Non-Profit Organizations in Microfinance." Tech. rep., Innovations for Poverty Action.

Karlan, Dean, Robert Osei, Isaac Osei-Akoto, and Christopher Udry. 2014. "Agricultural Decisions after Relaxing Credit and Risk Constraints." *The Quarterly Journal of Economics* 129 (2):597–652.

Karlan, Dean and Jonathan Zinman. 2011. "Microcredit in Theory and Practice: Using Randomized Credit Scoring for Impact Evaluation." *Science* 332 (6035):1278–1284.

Kenya Roads Board, Roads. 2007. "Kenyan Transport Sector Details Annex 3.1." Tech. rep., Kenya Roads Board.

Macharia, WM, EK Njeru, F Muli-Musiime, and V Nantulya. 2009. "Severe Road Traffic Injuries in Kenya, Quality of Care and Access." *African Health Sciences* 9 (2):118–124.

Maes, J. and L. Reed. 2011. "State of the Microcredit Summit Campaign Report 2012." Paper, Microcredit Summit Campaign.

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

McIntosh, Craig, Alexander Sarris, and Fotis Papadopoulos. 2013. "Productivity, Credit, Risk, and the Demand for Weather Index Insurance in Smallholder Agriculture in Ethiopia." *Agricultural Economics* 44 (4-5):399–417.

McKenzie, David and Christopher Woodruff. 2016. "Business Practices in Small Firms in Developing Countries." *Management Science* 63 (9):2967–2981.

Michuki, John. 2003. "The Traffic Act."

Mutongi, Kenda. 2017. *Matatu, A History of Popular Transportation in Nairobi*. The University of Chicago Press.

Payet, Charmel. 2018. "Green Light for Taxi Incentive Programme." *eThekwini Municipality* .

Pierce, Lamar, Daniel C. Snow, and Andrew McAfee. 2015. "Cleaning House: The Impact of Information Technology Monitoring on Employee Theft and Productivity." *Management Science* 61 (10):2299–2319.

Reed, Rachel C. 2018. "Transportation Turned Performance Art: Nairobi's Matatu Crews." *New York Times* .

Serfilippi, Elena, Michael Carter, and Catherine Guirkinger. 2018. "Insurance Contracts When Individuals "Greatly Value" Certainty: Results from a Field Experiment in Burkina Faso." Working Paper 25026, National Bureau of Economic Research.

Shapiro, Carl and Joseph E. Stiglitz. 1984. "Equilibrium Unemployment as a Worker Discipline Device." *The American Economic Review* 74 (3):433–444.

Transparency International, Kenya. 2018. "Traffic Legislation Gaps and Drivers of Corruption in Traffic Matters." Tech. rep., Transparency International, Kenya.

UNDP. 2015. "MDG Bangladesh Progress Report." Tech. rep., UNDP.

Valvidia, Martin. 2012. "Training or Technical Assistance for Female Entrepreneurship? Evidence from a Field Experiment in Peru." Tech. rep., GRADE.

WHO, Geneva. 2015. "Global Status Report on Road Safety." Tech. rep., World Health Organization, Geneva.

———. 2018. "Global Status Report on Road Safety." Tech. rep., World Health Organization, Geneva.

World Bank. 2016. "Bangladesh: Growing the Economy through Advances in Agriculture." http://www.worldbank.org/en/results/2016/10/07/bangladesh-growing-economy-through-advances-in-agriculture.

World Bank, DC. 2014. "Global Road Safety Facility (GRSF) Strategic Plan 2013-2020." Tech. rep., World Bank.

A | Credit Lines as Insurance: Evidence from Bangladesh

A.1 Comparative Statics

In this section we will more formally derive the comparative statics for input choice x and first period borrowing b^1 with respect to the increase in second period borrowing b_B^2. Starting with the maximization problem defined in equation 9:

$$\max_{x,b^1,b_B^2} \mathscr{L} = u(Y - x + b^1) + q\beta u\left(-Rb^1 + b_B^2\right) + (1-q)\beta u\left(m_G f(x) - Rb^1\right) +$$
$$q\beta^2 u(I - Rb_B^2) + (1-q)\beta^2 u(I) + \lambda_1[\bar{B}_1 - b^1] + \lambda_2[\bar{B}_2 - b_B^2]$$

Where the FOCs are given by:

$$\begin{aligned}
\frac{\partial \mathscr{L}}{\partial x} =& -u'(c_1) + (1-q)\beta u'(c_G^2) m_G f' \\
\frac{\partial \mathscr{L}}{\partial b^1} =& u'(c_1) - q\beta R u'(c_B^2) - (1-q)\beta R u'(c_g^2) - \lambda_1 \\
\frac{\partial \mathscr{L}}{\partial b_B^2} =& q\beta u'(c_B^2) - qR\beta^2 u'(c_B^3) - \lambda_2
\end{aligned}$$

Note, we assume the constraints do not bind ($\lambda_t = 0$) so that the choice of x and b^1 can adjust. We also know from the implicit function theory that we can calculate $\frac{\partial x}{\partial b_B^2}$ and $\frac{\partial b^1}{\partial b_B^2}$ by:

$$\begin{bmatrix} \frac{\partial x}{\partial b_B^2} \\ \frac{\partial b^1}{\partial b_B^2} \end{bmatrix} = -\begin{bmatrix} \frac{\partial \mathscr{L}}{\partial x \partial x} & \frac{\partial \mathscr{L}}{\partial x \partial b^1} \\ \frac{\partial \mathscr{L}}{\partial b^1 \partial x} & \frac{\partial \mathscr{L}}{\partial b^1 \partial b^1} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial \mathscr{L}}{\partial x \partial b_B^2} \\ \frac{\partial \mathscr{L}}{\partial b^1 \partial b_B^2} \end{bmatrix}$$

Calculating each term separately:

$$\begin{aligned}
\frac{\partial \mathscr{L}}{\partial x \partial x} &= u''(c_1) + (1-q)\beta m_G \left[(f')^2 u''(c_G^2) + f'' u'(c_G^2)\right] < 0 \\
\frac{\partial \mathscr{L}}{\partial x \partial b^1} &= -u''(c_1) - q\beta R m_G f' u''(c_G^2) > 0 \\
\frac{\partial \mathscr{L}}{\partial b^1 \partial x} &= -u''(c_1) - q\beta R m_G f' u''(c_G^2) > 0 \\
\frac{\partial \mathscr{L}}{\partial b^1 \partial b^1} &= u''(c_1) + \beta R^2 \left[q u''(c_B^2) + (1-q) u''(c_G^2)\right] < 0 \\
\frac{\partial \mathscr{L}}{\partial x \partial b_B^2} &= 0 \\
\frac{\partial \mathscr{L}}{\partial b^1 \partial b_B^2} &= -q\beta R u''(c_B^2) > 0
\end{aligned}$$

Inverting the matrix

$$\begin{bmatrix} \frac{\partial x}{\partial b_B^2} \\ \frac{\partial b^1}{\partial b_B^2} \end{bmatrix} = -\frac{1}{\frac{\partial \mathscr{L}}{\partial x \partial x}\frac{\partial \mathscr{L}}{\partial b^1 \partial b^1} - \frac{\partial \mathscr{L}}{\partial x \partial b^1}\frac{\partial \mathscr{L}}{\partial b^1 \partial x}} \begin{bmatrix} \frac{\partial \mathscr{L}}{\partial b^1 \partial b^1} & -\frac{\partial \mathscr{L}}{\partial x \partial b^1} \\ -\frac{\partial \mathscr{L}}{\partial b^1 \partial x} & \frac{\partial \mathscr{L}}{\partial x \partial x} \end{bmatrix} \begin{bmatrix} \frac{\partial \mathscr{L}}{\partial x \partial b_B^2} \\ \frac{\partial \mathscr{L}}{\partial b^1 \partial b_B^2} \end{bmatrix}$$

The denominator of the fraction is the determinate of a 2x2 hessian from a maximization problem, and is therefore positive. Then, the matrices are pre-multiplied by a negative value, which we will replace with $-\frac{1}{Det}$. Multiplying out the matrices we find

$$\frac{\partial x}{\partial b_B^2} = \underbrace{-\frac{1}{Det}}_{-} \underbrace{\left[\frac{\partial \mathscr{L}}{\partial b^1 \partial b^1} \cdot 0 - \frac{\partial \mathscr{L}}{\partial x \partial b^1}\frac{\partial \mathscr{L}}{\partial b^1 \partial b_B^2}\right]}_{-} > 0$$

$$\frac{\partial b^1}{\partial b_B^2} = \underbrace{-\frac{1}{Det}}_{-} \underbrace{\left[-\frac{\partial \mathscr{L}}{\partial b^1 \partial x} \cdot 0 + \frac{\partial \mathscr{L}}{\partial x \partial x}\frac{\partial \mathscr{L}}{\partial b^1 \partial b_B^2}\right]}_{-} > 0$$

Therefore, we conclude that the choice of inputs x and first period borrowing b^1 will both increase with the offer of the Emergency Loan.

A.2 Selection

In this appendix we examine whether selection into eligibility in 2017 matters for the results. First, we simply examine whether there was differential Emergency Loan eligibility in 2017 across treatment and control branches. We see in Table A.1 shows that there is no statistically significant difference in the probability that households are Emergency Loan eligible between treatment and control branches. Ignoring statistical significance, the point estimate suggests that treatment branches were three percentage points *less* likely to be Emergency Loan eligible in 2017. This is the opposite effect as what might be expected ex-ante, that households in treatment branches improve repayment rates and are therefore more likely to become eligible. Finally, I also report ex-post outcomes without controlling for flooding.

Table A.1: 2017 Eligiblity

	(1) EL Eligible
Treatment Branch	-0.030
	(0.029)
Flood Last Year	Yes
District FE	Yes
Observations	3939

Notes: Sample includes all surveyed households in 2017. The outcome variable is a binary indicator for the household being Emergency Loan eligible in 2017. Flood last year is an indicator for being flooded in 2016.

As a robustness check, I reproduce the results on household investment and ex-post outcomes with two different specifications. First, I limit the analysis to only 2016 when there are no selection concerns. Second, I instrument for eligibility using branch treatment status. With the exception of non-agriculture investment, the results are consistent with those found with the other specifications.

Table A.2: Land Farmed 2016

	(1) Own land	(2) Rented land	(3) Sharecrop land	(4) Total land	(5) Any Cult.
Treatment	0.001	0.067***	-0.006	0.059*	0.034
	(0.014)	(0.020)	(0.004)	(0.030)	(0.027)
Controls	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var	0.15	0.22	0.02	0.39	0.50
Observations	2986	2986	2986	2986	2986

Notes: Sample includes only eligible BRAC members from both treatment and control groups. Data is from only the 2016 Aman season. Controls are included for precision, and are comprised of baseline measures of total land owned, household size, and the age and education of the head of household. Standard errors clustered at the branch level. Land measured in acres. Total land is the sum of own land, rented land, and sharecropped land. Any Cult. is an indicator for whether or not a household planted any crops during the season.

Table A.3: Ex-Ante Investments 2016

	(1) Fert. Applied	(2) Pest. Applied	(3) Cost Seeds per acre	(4) Input Cost per Acre	(5) Non-Ag Invest
Treatment	6.15	0.36*	1.05	1.20	1.09
	(5.62)	(0.18)	(0.89)	(2.49)	(3.35)
Controls	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var	129.93	1.34	14.45	60.53	7.84
Observations	1479	1479	1375	1375	2986

Notes: Sample includes only eligible BRAC members from both treatment and control groups. Data is only from the 2016 Aman season. Controls are included for precision, and are comprised of baseline measures of total land owned, household size, and the age and education of the head of household. Standard errors clustered at the branch level. Fertilizer and pesticide measured in kg/L per acre. Input cost per acre is the sum of the cost of fertilizer, pesticide, and seeds. Cost and investment measured in dollars.

Table A.4: IV Land Farmed

	(1) Own land	(2) Rented land	(3) Sharecrop land	(4) Total land	(5) Any Cult.
Treatment	-0.004	0.071^{***}	-0.007^{*}	0.057^{*}	0.034
	(0.015)	(0.019)	(0.004)	(0.029)	(0.028)
Controls	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var	0.13	0.18	0.02	0.33	0.44
Observations	5981	5977	5980	5976	5982

Notes: Sample includes all observations from both treatment and control groups. Treatment is instrumented using first year eligibility interacted by year. Data is pooled from both the 2016 and 2017 Aman season. Controls are included for precision, and are comprised of baseline measures of total land owned, household size, and the age and education of the head of household. Standard errors clustered at the branch level. Land measured in acres. Total land is the sum of own land, rented land, and sharecropped land. Any Cult. is an indicator for whether or not a household planted any crops during the season.

Table A.5: IV Inputs

	(1) Fert. Applied	(2) Pest. Applied	(3) Cost Seeds per acre	(4) Input Cost per Acre	(5) Non-Ag Invest
Treatment	5.71	0.28	0.39	1.79	1.15
	(5.41)	(0.18)	(0.83)	(2.38)	(7.51)
Controls	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var	141.48	1.60	16.88	66.87	56.02
Observations	2638	2559	2504	2431	5982

Notes: Sample includes all observations from both treatment and control groups. Treatment is instrumented using first year eligibility interacted by year. Data is pooled from both the 2016 and 2017 Aman season. Controls are included for precision, and are comprised of baseline measures of total land owned, household size, and the age and education of the head of household. Standard errors clustered at the branch level. Land measured in acres. Total land is the sum of own land, rented land, and sharecropped land. Any Cult. is an indicator for whether or not a household planted any crops during the season.

Table A.6: Ex-Ante Land by Risk Aversion: 2016

	(1) Own land	(2) Rented land	(3) Sharecrop land	(4) Total land	(5) Any Cult.
Treatment	-0.02	0.04	-0.01*	-0.00	0.03
	(0.02)	(0.03)	(0.01)	(0.04)	(0.03)
Risk Aversion X Treatment	0.03	0.05	0.01	0.11*	0.00
	(0.03)	(0.05)	(0.01)	(0.06)	(0.05)
Risk Aversion	0.14*	0.09	-0.02	0.19*	0.12
	(0.08)	(0.06)	(0.01)	(0.10)	(0.09)
Controls	Yes	Yes	Yes	Yes	Yes
Controls X Risk Aversion	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var	0.15	0.22	0.02	0.39	0.50
Observations	2986	2986	2986	2986	2986
p-value Treat + Risk X Treat	0.654	0.001	0.900	0.008	0.352

Notes: Sample includes only eligible BRAC members from both treatment and control groups. Data is only from the 2016 Aman season. Controls are included for precision, and are comprised of baseline measures of total land owned, household size, and the age and education of the head of household. Standard errors clustered at branch level. Land is measured in acres. Total land is the sum of own land, rented land, and sharecropped land. Any Cult. is an indicator for whether or not a household planted any crops during the season. Risk aversion was measured at baseline and ranges 0 to 1, where 0=most risk loving and 1=most risk averse.

Table A.7: Ex-Ante Inputs by Risk Aversion: 2016

	(1) Fert. Applied	(2) Pest. Applied	(3) Cost Seeds per acre	(4) Input Cost per Acre	(5) Non-Ag Invest
Treatment	7.40	0.25	1.85	2.02	-0.77
	(9.98)	(0.30)	(1.44)	(4.43)	(5.33)
Risk Aversion X	-3.29	0.20	-1.57	-1.89	3.58
Treatment	(16.87)	(0.44)	(1.88)	(6.52)	(8.29)
Risk Aversion	9.33	0.31	-4.80	-5.65	-8.93
	(28.41)	(0.81)	(3.56)	(12.56)	(13.77)
Controls	Yes	Yes	Yes	Yes	Yes
Controls X Risk Aversion	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var	129.93	1.34	14.45	60.53	7.84
Observations	1479	1479	1375	1375	2986
p-value Treat + Risk X Treat	0.689	0.101	0.808	0.971	0.600

Notes: Sample includes only eligible BRAC members from both treatment and control groups. Data is only from the 2016 Aman season. Controls are included for precision, and are comprised of baseline measures of total land owned, household size, and the age and education of the head of household. Standard errors clustered at branch level. Fertilizer and pesticide are measured in kg / L per acre. Input cost per acre is the sum of the cost of fertilizer, pesticide, and seeds. Cost and investment measured in dollars. Risk aversion ranges 0 to 1, where 0=most risk loving and 1=most risk averse.

Table A.8: IV Ex-Ante Land by Risk Aversion

	(1) Own land	(2) Rented land	(3) Sharecrop land	(4) Total land	(5) Any Cult.
Treatment	-0.032	0.044	-0.014*	-0.010	-0.007
	(0.026)	(0.036)	(0.007)	(0.051)	(0.053)
Risk Aversion X	0.033	0.055	0.015	0.115**	0.066
Treatment	(0.031)	(0.044)	(0.010)	(0.057)	(0.055)
Risk Aversion	0.094	0.040	-0.001	0.126	0.081
	(0.065)	(0.051)	(0.010)	(0.082)	(0.073)
Controls	Yes	Yes	Yes	Yes	Yes
Controls X Risk Aversion	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var	0.13	0.18	0.02	0.33	0.44
Observations	5736	5732	5735	5731	5737
p-value Treat + Risk X Treat	0.949	0.000	0.964	0.001	0.031

Notes: Sample includes all observations from both treatment and control groups. Treatment is instrumented using first year eligibility interacted by year. Data is pooled from both the 2016 and 2017 Aman season. Controls are included for precision, and are comprised of baseline measures of total land owned, household size, and the age and education of the head of household. Standard errors clustered at branch level. Land is measured in acres. Total land is the sum of own land, rented land, and sharecropped land. Any Cult. is an indicator for whether or not a household planted any crops during the season. Risk aversion was measured at baseline and ranges 0 to 1, where 0=most risk loving and 1=most risk averse.

Table A.9: IV Ex-Ante Inputs by Risk Aversion

	(1) Fert. Applied	(2) Pest. Applied	(3) Cost Seeds per acre	(4) Input Cost per Acre	(5) Non-Ag Invest
Treatment	-2.62	-0.05	0.62	-0.55	-13.80
	(9.41)	(0.40)	(1.75)	(4.98)	(25.64)
Risk Aversion X	17.52	0.55	-1.01	4.76	29.34
Treatment	(13.18)	(0.56)	(2.30)	(6.55)	(33.27)
Risk Aversion	-12.62	-0.95	-4.22	-20.20**	11.39
	(22.71)	(0.70)	(3.74)	(10.04)	(36.66)
Controls	Yes	Yes	Yes	Yes	Yes
Controls X Risk Aversion	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var	140.08	1.56	16.83	66.61	52.67
Observations	2550	2473	2423	2352	5737
p-value Treat + Risk X Treat	0.051	0.054	0.722	0.157	0.172

Notes: Sample includes all observations from both treatment and control groups. Treatment is instrumented using first year eligibility interacted by year. Data is pooled from both the 2016 and 2017 Aman season. Controls are included for precision, and are comprised of baseline measures of total land owned, household size, and the age and education of the head of household. Standard errors clustered at branch level. Fertilizer and pesticide are measured in kg / L per acre. Input cost per acre is the sum of the cost of fertilizer, pesticide, and seeds. Cost and investment measured in dollars. Risk aversion ranges 0 to 1, where 0=most risk loving and 1=most risk averse.

Table A.10: Ex-Post Outcomes 2016

	(1) Log Cons PerCap	(2) Log Income	(3) Crop Prod. (Kg)	(4) Livestock
Treatment	0.013	-0.004	128.861**	-0.118
	(0.048)	(0.050)	(55.976)	(0.114)
Flood X Treatment	0.144*	-0.090	-142.596*	0.310*
	(0.074)	(0.077)	(80.684)	(0.170)
Flood	-0.094	0.058	-20.675	0.037
	(0.076)	(0.077)	(60.773)	(0.142)
Controls	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Week Interviewed FE	Yes	No	No	No
Mean Dep. Var	5.86	10.73	327.80	1.53
Observations	2969	2826	2971	2971
p-value Treat + Flood X Treat	0.005	0.120	0.797	0.130

Notes: Sample includes only eligible BRAC members from both treatment and control groups. Data is from only the 2016 Aman season. Controls are included for precision, and are comprised of baseline measures of total land owned, household size, and the age and education of the head of household. Week interviewed fixed effects are included for the log consumption regression due to the presence of holidays over the course of the survey period that changed standard consumption patterns. Standard errors clustered at branch level. Income is measured in dollars. Flood is an indicator that equals one if flooding occurred and the Emergency Loan was activated.

Table A.11: IV Ex-Post Outcomes

	(1) Log Cons PerCap	(2) Log Income	(3) Crop Prod. (Kg)	(4) Livestock
Treatment	0.040	-0.027	110.893**	-0.155
	(0.053)	(0.053)	(46.853)	(0.121)
Flood X Treatment	0.066	0.014	-100.623*	0.513***
	(0.062)	(0.065)	(51.432)	(0.144)
Flood	-0.019	0.029	-3.086	-0.130
	(0.047)	(0.049)	(31.068)	(0.104)
Controls	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Week Interviewed FE	Yes	No	No	No
Mean Dep. Var	5.94	10.78	258.54	1.47
Observations	5980	5726	5982	5982
p-value Treat + Flood X Treat	0.004	0.738	0.747	0.000

Notes: Sample includes only eligible BRAC members from both treatment and control groups. Treatment is instrumented using first year eligibility interacted by year. Data is pooled from both the 2016 and 2017 Aman season. Controls are included for precision, and are comprised of baseline measures of total land owned, household size, and the age and education of the head of household. Week interviewed fixed effects are included for the log consumption regression due to the presence of holidays over the course of the survey period that changed standard consumption patterns. Standard errors clustered at branch level. Income is measured in dollars. Flood is an indicator that equals one if flooding occurred and the Emergency Loan was activated.

Table A.12: Ex-Post Outcomes with out Flood Controls

	(1) Log Cons PerCap	(2) Log Income	(3) Crop Prod. (Kg)	(4) Livestock
Treatment	0.080**	-0.019	47.896*	0.118
	(0.031)	(0.029)	(28.093)	(0.076)
Controls	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Week Interviewed FE	Yes	No	No	No
Mean Dep. Var	5.93	10.77	275.22	1.51
Observations	4743	4531	4745	4745

Notes: Sample includes only eligible BRAC members from both treatment and control groups. Data is pooled from both the 2016 and 2017 Aman season. Controls are included for precision, and are comprised of baseline measures of total land owned, household size, and the age and education of the head of household. Week interviewed fixed effects are included for the log consumption regression due to the presence of holidays over the course of the survey period that changed standard consumption patterns. Standard errors clustered at branch level. Income is measured in dollars.

A.3 Spillovers

In this appendix I report the spillovers on the ineligible households for the main ex-ante and ex-post outcomes. In general, I find no evidence of significant spillovers onto the ineligible population.

Table A.13: Spillovers: Ineligible Land Farmed

	(1) Own land	(2) Rented land	(3) Sharecrop land	(4) Total land	(5) Any Cult.
Treatment branch	0.000	-0.010	-0.005	-0.013	-0.035
	(0.015)	(0.014)	(0.003)	(0.022)	(0.022)
Controls	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var	0.12	0.14	0.01	0.28	0.40
Observations	3193	3193	3193	3193	3193

Notes: Sample includes only ineligible BRAC members both treatment and control groups. Data is pooled from both the 2016 and 2017 Aman season. Controls are included for precision, and are comprised of baseline measures of total land owned, household size, and the age and education of the head of household. Standard errors clustered at the branch level. Land measured in acres. Total land is the sum of own land, rented land, and sharecropped land. Any Cult. is an indicator for whether or not a household planted any crops during the season.

Table A.14: Spillovers: Ineligible Inputs

	(1) Fert. Applied	(2) Pest. Applied	(3) Cost Seeds per acre	(4) Input Cost per Acre	(5) Non-Ag Invest
Treatment branch	-0.78	-0.02	-0.88	1.24	-4.24
	(6.26)	(0.16)	(1.11)	(2.65)	(12.76)
Controls	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var	140.31	1.46	18.45	68.69	71.63
Observations	1271	1208	1204	1146	3193

Notes: Sample includes only ineligible BRAC members from both treatment and control groups. Data is pooled from both the 2016 and 2017 Aman season. Controls are included for precision, and are comprised of baseline measures of total land owned, household size, and the age and education of the head of household. Standard errors clustered at the branch level. Land measured in acres. Total land is the sum of own land, rented land, and sharecropped land. Any Cult. is an indicator for whether or not a household planted any crops during the season.

Table A.15: Spillovers: Ineligible Ex-Post Outcomes

	(1) Log Cons PerCap	(2) Log Income	(3) Crop Prod. (Kg)	(4) Livestock
Treatment branch	0.083*	-0.028	-7.616	-0.135
	(0.048)	(0.046)	(31.866)	(0.122)
Flood X Treatment	-0.024	-0.016	-5.095	0.234
	(0.061)	(0.064)	(39.846)	(0.152)
Flood	0.082	-0.023	-28.800	-0.269**
	(0.054)	(0.059)	(33.960)	(0.135)
Controls	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Week Interviewed FE	Yes	No	No	No
Mean Dep. Var	6.01	10.82	210.93	1.27
Observations	3176	3057	3177	3177
p-value Treat + Flood X Treat	0.120	0.284	0.633	0.330

Notes: Sample includes only ineligible BRAC members from both treatment and control groups. Data is pooled from both the 2016 and 2017 Aman season. Controls are included for precision, and are comprised of baseline measures of total land owned, household size, and the age and education of the head of household. Week interviewed fixed effects are included for the log consumption regression due to the presence of holidays over the course of the survey period that changed standard consumption patterns. Standard errors clustered at branch level. Income is measured in dollars. Flood is an indicator that equals one if flooding occurred and the Emergency Loan was activated.

B | The Impact of Monitoring Technologies on Contracts and Employee Behavior: Experimental Evidence from Kenya's Transit Industry

B.1 Tables

Table B.1: Submitting daily survey report

	(1) Owner report submitted	(2) Driver report submitted
Info treatment group	0.042	0.029
	(0.040)	(0.037)
Control Mean of DV	0.45	0.55
Day FE	Y	Y
Route FE	Y	Y
Matatu N	255	255
Matatu-Day N	46,920	46,920

The data are from the owner and drivers that submitted data throughout the study period. The dependent variable is a binary indicator for whether the owner (Column 1) or driver (Column 2) submitted a report that day. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table B.2: Differential Revenue Reporting

	(1) Revenue	(2) Revenue
Mileage	8.848***	
	(1.089)	
Info treatment group	-259.461	
	(267.894)	
Mileage X Treat	1.241	
	(1.491)	
Mileage quartile 1		-1489.213***
		(209.834)
Mileage quartile 2		-1248.806***
		(203.465)
Mileage quartile 3		-552.104***
		(180.865)
Mileage quartile 1 X Treat		-331.924
		(266.860)
Mileage quartile 2 X Treat		30.419
		(181.920)
Mileage quartile 3 X Treat		-119.329
		(204.766)
Mileage quartile 4 X Treat		38.374
		(235.454)
Control Mean of DV	7126.94	7126.94
Day FE	Y	Y
Route FE	Y	Y
Matatu N	241	241
Matatu-Day N	22,436	23,514

The data are from days where we have both the driver reports (reported revenue), and the tracking data (number of miles). The dependent variable is the amount of revenue the driver reports. In the first column we regress revenue on the number of miles the vehicle travelled, and indicator for treatment, and an interaction between the two terms. The interaction captures the differential relationship between mileage and reported revenue between treatment and control. We might be concerned that drivers are reporting differently in the treatment than the control group. The coefficient on the interaction term is not significant however. Column 2 further investigates whether there is differential reporting across different quartiles of the mileage distribution. Again, the interaction terms are all insignificant. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels. Note there are 14 drivers that did not report revenue at all during the 6 months of the study.

Table B.3: Contract

	(1) Daily Target	(2) Daily Target	(3) Daily Target
Treat X Month1	-35.141		
	(68.842)		
Treat X Month2	-46.502		
	(91.193)		
Treat X Month3	-63.407		
	(86.703)		
Treat X Month4	-94.940		
	(83.842)		
Treat X Month5	-120.861		
	(92.118)		
Treat X Month6	-134.693		
	(97.085)		
Treat X First 3 months		-51.322	
		(79.284)	
Treat X Last 3 months		-112.564	
		(86.747)	
Treat X Trend			-9.332
			(17.225)
Control Mean of DV	3057.38	3057.38	3057.38
Controls	Y	Y	Y
Day FE	Y	Y	N
Route FE	Y	Y	Y
Matatu N	237	237	237
Matatu-Day N	15,884	15,884	15,542

This table shows the results of three different regression specifications. The first column shows the treatment effects broken out by treatment month, as specified in the main body of the paper (through figures). The second column pools the first three months and the last three months - to show the effects in the “early" versus “later" months of the experiment. The final column demonstrates the results of a regression of the outcome of interest on a linear time trend, treatment, and the interaction of the two. The interaction term reveals how treatment trended differently over time relative to the control group. The outcome of interest in this table is the target. Note there are 237 matatus included in these regressions because 18 owners failed to report the target throughout the study (balanced between treatment and control). Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels. Standard errors are clustered at the matatu level.

Table B.4: Deviceon

	(1) Hours tracking device on	(2) Hours tracking device on	(3) Hours tracking device on
Treat X Month1	-0.842		
	(0.658)		
Treat X Month2	0.623		
	(0.590)		
Treat X Month3	0.986*		
	(0.554)		
Treat X Month4	0.691		
	(0.640)		
Treat X Month5	1.468**		
	(0.715)		
Treat X Month6	1.474*		
	(0.757)		
Treat X First 3 months		0.388	
		(0.503)	
Treat X Last 3 months		1.119*	
		(0.637)	
Treat X Trend			0.177
			(0.207)
Control Mean of DV	14.79	14.79	14.79
Controls	Y	Y	Y
Day FE	Y	Y	N
Route FE	Y	Y	Y
Matatu N	254	254	254
Matatu-Day N	45,654	45,654	44,444

This table shows the results of three different regression specifications. The first column shows the treatment effects broken out by treatment month, as specified in the main body of the paper (through figures). The second column pools the first three months and the last three months - to show the effects in the “early" versus “later" months of the experiment. The final column demonstrates the results of a regression of the outcome of interest on a linear time trend, treatment, and the interaction of the two. The interaction term reveals how treatment trended differently over time relative to the control group. The outcome of interest in this table is the number of hours the device was on (as a proxy for effort). Note there are 254 matatus included in these regressions because 1 device was faulty (the matatu was in an accident). Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels. Standard errors are clustered at the matatu level.

Table B.5: Mileage

	(1) Kilometers	(2) Kilometers	(3) Kilometers
Treat X Month1	-4.302		
	(5.966)		
Treat X Month2	2.053		
	(5.601)		
Treat X Month3	7.802		
	(5.318)		
Treat X Month4	4.411		
	(5.700)		
Treat X Month5	9.640		
	(6.386)		
Treat X Month6	13.132*		
	(6.917)		
Treat X First 3 months		2.911	
		(5.095)	
Treat X Last 3 months		8.365	
		(5.793)	
Treat X Trend			1.533
			(1.650)
Control Mean of DV	96.64	96.64	96.64
Controls	Y	Y	Y
Day FE	Y	Y	N
Route FE	Y	Y	Y
Matatu N	254	254	254
Matatu-Day N	45,654	45,654	44,444

This table shows the results of three different regression specifications. The first column shows the treatment effects broken out by treatment month, as specified in the main body of the paper (through figures). The second column pools the first three months and the last three months - to show the effects in the “early" versus “later" months of the experiment. The final column demonstrates the results of a regression of the outcome of interest on a linear time trend, treatment, and the interaction of the two. The interaction term reveals how treatment trended differently over time relative to the control group. The outcome of interest in this table is the number of hours the device was on (as a proxy for effort). Note there are 254 matatus included in these regressions because 1 device was faulty (the matatu was in an accident). Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels. Standard errors are clustered at the matatu level.

Table B.6: Repair Cost

	(1) Repair Cost	(2) Repair Cost	(3) Repair Cost
Treat X Month1	67.838		
	(70.046)		
Treat X Month2	-50.802		
	(73.520)		
Treat X Month3	-125.119		
	(79.387)		
Treat X Month4	-186.449**		
	(89.291)		
Treat X Month5	-187.876**		
	(94.119)		
Treat X Month6	-226.720**		
	(104.752)		
Treat X First 3 months		-46.459	
		(70.294)	
Treat X Last 3 months		-187.616**	
		(89.675)	
Treat X Trend			-40.116*
			(20.406)
Control Mean of DV	483.48	483.48	483.48
Controls	Y	Y	Y
Day FE	Y	Y	N
Route FE	Y	Y	Y
Matatu N	238	238	238
Matatu-Day N	15,881	15,881	15,539

This table shows the results of three different regression specifications. The first column shows the treatment effects broken out by treatment month, as specified in the main body of the paper (through figures). The second column pools the first three months and the last three months - to show the effects in the "early" versus "later" months of the experiment. The final column demonstrates the results of a regression of the outcome of interest on a linear time trend, treatment, and the interaction of the two. The interaction term reveals how treatment trended differently over time relative to the control group. The outcome of interest in this table is the amount of repairs the owner incurred. Note there are 238 matatus included in these regressions because 17 owners failed to report the target throughout the study (balanced between treatment and control).Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels. Standard errors are clustered at the matatu level.

Table B.7: Repair Cost (Binary)

	(1) Large Repairs	(2) Large Repairs	(3) Large Repairs
Treat X Month1	0.033		
	(0.030)		
Treat X Month2	-0.019		
	(0.030)		
Treat X Month3	-0.030		
	(0.032)		
Treat X Month4	-0.048		
	(0.033)		
Treat X Month5	-0.066**		
	(0.033)		
Treat X Month6	-0.081**		
	(0.038)		
Treat X First 3 months		-0.010	
		(0.029)	
Treat X Last 3 months		-0.060*	
		(0.032)	
Treat X Trend			-0.009
			(0.006)
Control Mean of DV	0.17	0.17	0.17
Controls	Y	Y	Y
Day FE	Y	Y	N
Route FE	Y	Y	Y
Matatu N	238	238	238
Matatu-Day N	15,881	15,881	15,539

This table shows the results of three different regression specifications. The first column shows the treatment effects broken out by treatment month, as specified in the main body of the paper (through figures). The second column pools the first three months and the last three months - to show the effects in the “early" versus “later" months of the experiment. The final column demonstrates the results of a regression of the outcome of interest on a linear time trend, treatment, and the interaction of the two. The interaction term reveals how treatment trended differently over time relative to the control group. The outcome of interest in this table is the number of large repairs the owner incurred. Note there are 238 matatus included in these regressions because 18 owners failed to report the target throughout the study (balanced between treatment and control).Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels. Standard errors are clustered at the matatu level.

Table B.8: Made Target

	(1) Made Target	(2) Made Target	(3) Made Target
Treat X Month1	-0.099**		
	(0.045)		
Treat X Month2	-0.013		
	(0.046)		
Treat X Month3	0.068		
	(0.048)		
Treat X Month4	0.125**		
	(0.051)		
Treat X Month5	0.083		
	(0.053)		
Treat X Month6	0.059		
	(0.053)		
Treat X First 3 months		-0.010	
		(0.043)	
Treat X Last 3 months		0.084*	
		(0.048)	
Treat X Trend			0.028**
			(0.011)
Control Mean of DV	0.43	0.43	0.43
Controls	Y	Y	Y
Day FE	Y	Y	N
Route FE	Y	Y	Y
Matatu N	237	237	237
Matatu-Day N	15,888	15,888	15,546

This table shows the results of three different regression specifications. The first column shows the treatment effects broken out by treatment month, as specified in the main body of the paper (through figures). The second column pools the first three months and the last three months - to show the effects in the "early" versus "later" months of the experiment. The final column demonstrates the results of a regression of the outcome of interest on a linear time trend, treatment, and the interaction of the two. The interaction term reveals how treatment trended differently over time relative to the control group. The outcome of interest in this table is whether or not the driver made the target. Note there are 237 matatus included in these regressions because 17 owners failed to report the target throughout the study (balanced between treatment and control).Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels. Standard errors are clustered at the matatu level.

Table B.9: Under-report

	(1) shade	(2) shade	(3) shade
Treat X Month1	-6.782		
	(51.685)		
Treat X Month2	-89.596*		
	(48.896)		
Treat X Month3	-103.442**		
	(52.358)		
Treat X Month4	-105.062**		
	(53.112)		
Treat X Month5	-66.293		
	(59.854)		
Treat X Month6	-106.336*		
	(61.400)		
Treat X First 3 months		-69.041	
		(47.008)	
Treat X Last 3 months		-88.403*	
		(52.955)	
Treat X Trend			-15.213
			(12.967)
Control Mean of DV	521.20	521.20	521.20
Controls	Y	Y	Y
Day FE	Y	Y	N
Route FE	Y	Y	Y
Matatu N	215	215	215
Matatu-Day N	7,426	7,426	7,320

This table shows the results of three different regression specifications. The first column shows the treatment effects broken out by treatment month, as specified in the main body of the paper (through figures). The second column pools the first three months and the last three months - to show the effects in the "early" versus "later" months of the experiment. The final column demonstrates the results of a regression of the outcome of interest on a linear time trend, treatment, and the interaction of the two. The interaction term reveals how treatment trended differently over time relative to the control group. The outcome of interest in this table is the amount revenue drivers under-report. Note there are 215 matatus included in these regressions because 17 owners failed to report the target throughout the study (balanced between treatment and control); and 14 drivers didn't report their revenue/salary (balanced between treatment and control) - both of which are required to compute this measure - as detailed in the main figures. We also needed both owners and drivers to report in a particular day to compute this measure. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels. Standard errors are clustered at the matatu level.

Table B.10: Revenue

	(1) Revenue	(2) Revenue	(3) Revenue
Treat X Month1	-338.338		
	(229.414)		
Treat X Month2	-61.938		
	(187.910)		
Treat X Month3	-6.492		
	(180.413)		
Treat X Month4	124.687		
	(184.769)		
Treat X Month5	54.217		
	(193.679)		
Treat X Month6	-207.192		
	(205.937)		
Treat X First 3 months		-131.349	
		(176.217)	
Treat X Last 3 months		-6.992	
		(177.277)	
Treat X Trend			-26.140
			(58.137)
Control Mean of DV	7126.94	7126.94	7126.94
Controls	Y	Y	Y
Day FE	Y	Y	N
Route FE	Y	Y	Y
Matatu N	241	241	241
Matatu-Day N	22,436	22,436	22,107

This table shows the results of three different regression specifications. The first column shows the treatment effects broken out by treatment month, as specified in the main body of the paper (through figures). The second column pools the first three months and the last three months - to show the effects in the “early" versus “later" months of the experiment. The final column demonstrates the results of a regression of the outcome of interest on a linear time trend, treatment, and the interaction of the two. The interaction term reveals how treatment trended differently over time relative to the control group. The outcome of interest in this table is reported revenue. Note there are 241 matatus included in these regressions because 14 owners failed to report the revenue throughout the study (balanced between treatment and control).Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels. Standard errors are clustered at the matatu level.

Table B.11: Profits

	(1) Gross Profits	(2) Gross Profits	(3) Gross Profits
Treat X Month1	-178.812		
	(233.078)		
Treat X Month2	55.929		
	(206.280)		
Treat X Month3	83.617		
	(217.394)		
Treat X Month4	441.369**		
	(222.362)		
Treat X Month5	451.597**		
	(212.750)		
Treat X Month6	172.690		
	(227.400)		
Treat X First 3 months		-13.173	
		(191.659)	
Treat X Last 3 months		361.696*	
		(196.326)	
Treat X Trend			51.548
			(75.291)
Control Mean of DV	3275.61	3275.61	3275.61
Controls	Y	Y	Y
Day FE	Y	Y	N
Route FE	Y	Y	Y
Matatu N	216	216	216
Matatu-Day N	10,406	10,406	10,277

This table shows the results of three different regression specifications. The first column shows the treatment effects broken out by treatment month, as specified in the main body of the paper (through figures). The second column pools the first three months and the last three months - to show the effects in the “early" versus “later" months of the experiment. The final column demonstrates the results of a regression of the outcome of interest on a linear time trend, treatment, and the interaction of the two. The interaction term reveals how treatment trended differently over time relative to the control group. The outcome of interest in this table is profits. Note there are 216 matatus included in these regressions because 18 owners failed to report the target throughout the study (balanced between treatment and control); and 14 drivers didn’t report their revenue/salary (balanced between treatment and control) - both of which are required to compute this measure - as detailed in the main figures. We also needed both owners and drivers to report in a particular day to compute this measure. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels. Standard errors are clustered at the matatu level.

B.2 Model derivation

Step 1: Owner's punishment

We assume the owner's signal is noisy but unbiased, so that $\hat{q}$ is defined as follows:

$$\hat{q} = q - \sigma$$

$$\sigma \sim U\left(-\frac{1}{\alpha}, \frac{1}{\alpha}\right) \qquad f(\hat{q}) = \frac{1}{q + \frac{1}{\alpha} - \left(q - \frac{1}{\alpha}\right)} = \frac{\alpha}{2}$$

where α is the precision of the owners' signal about true revenue. Any monitoring technology we introduce will provide more information to the owner about driver behavior. This will increase the precision of the owner's signal about what revenue should be, which gives the driver less leeway to significantly under-report revenue on a particular day. In the case of our monitoring technology specifically, the owner can observe the number of kilometers driven, and where the driver was operating at any point in time. Owners can use this information to estimate the number of trips to and from the city center, which provides a more accurate measure of total daily revenue.[1] With $\hat{q}$ defined in this way, the owners can be sure that real revenue q falls in the interval:

$$\left[\hat{q} - \frac{1}{\alpha}\ ,\ \hat{q} + \frac{1}{\alpha}\right]$$

On days when the driver reports making the target, the owner receives the target amount and does not punish the driver. On days when the driver does not report making the target, the owner will punish them if they can be sure the driver is lying to them. In other words, they will punish if the reported revenue comes in below the possible range for q. This assumes the owner is really unwilling to punish a driver incorrectly, which makes sense because firing costs are high in this setting. The actual punishment applied is some function of the difference between this lower bound, $\hat{q} - \frac{1}{\alpha}$, and the reported amount $\tilde{q}$ (owners are less upset on days where the driver reports below the target and they know for a fact that conditions were difficult). We assume for simplicity that this function is linear.

$$\begin{aligned}
E[punishment] &= E\left[\left(\hat{q} - \frac{1}{\alpha}\right) - \tilde{q} \mid \hat{q} - \frac{1}{\alpha} > \tilde{q}\right] \cdot \Pr\left(\hat{q} - \frac{1}{\alpha} > \tilde{q}\right) \\
&= E\left[\hat{q} - \tilde{q} - \frac{1}{\alpha} \mid \hat{q} > \tilde{q} + \frac{1}{\alpha}\right] \cdot \Pr\left(\hat{q} > \tilde{q} + \frac{1}{\alpha}\right) \\
&= \int_{\tilde{q}+\frac{1}{\alpha}}^{q+\frac{1}{\alpha}} \left(\hat{q} - \tilde{q} - \frac{1}{\alpha}\right) \cdot f(\hat{q}) d\hat{q}
\end{aligned}$$

[1] While they do not know the exact number of passengers that board, they know that the vehicle generally waits at the terminal until it fills up.

$$
\begin{aligned}
&= \frac{\alpha}{2} \int_{\tilde{q}+\frac{1}{\alpha}}^{q+\frac{1}{\alpha}} \left(\hat{q} - \tilde{q} - \frac{1}{\alpha}\right) d\hat{q} \\
&= \frac{\alpha}{2} \left[\frac{\hat{q}^2}{2} - \left(\tilde{q} + \frac{1}{\alpha}\right) \hat{q} \Big|_{\tilde{q}+\frac{1}{\alpha}}^{q+\frac{1}{\alpha}} \right] \\
&= \frac{\alpha}{2} \left[\frac{1}{2}\left(q + \frac{1}{\alpha}\right)^2 - \left(\tilde{q} + \frac{1}{\alpha}\right)\left(q + \frac{1}{\alpha}\right) - \frac{1}{2}\left(\tilde{q} + \frac{1}{\alpha}\right)^2 + \left(\tilde{q} + \frac{1}{\alpha}\right)^2 \right] \\
&= \frac{\alpha}{2} \left[\frac{1}{2}\left(q + \frac{1}{\alpha}\right)^2 - \left(\tilde{q} + \frac{1}{\alpha}\right)\left(q + \frac{1}{\alpha}\right) + \frac{1}{2}\left(\tilde{q} + \frac{1}{\alpha}\right)^2 \right] \\
&= \frac{\alpha}{4} \left[\left(q + \frac{1}{\alpha}\right)^2 - 2\left(\tilde{q} + \frac{1}{\alpha}\right)\left(q + \frac{1}{\alpha}\right) + \left(\tilde{q} + \frac{1}{\alpha}\right)^2 \right] \\
&= \frac{\alpha}{4} \left[\left(q + \frac{1}{\alpha}\right)^2 - \left(\tilde{q} + \frac{1}{\alpha}\right)^2 \right] \\
&= \frac{\alpha}{4} (q - \tilde{q})^2
\end{aligned}
$$

Step 2: Solve the agent's optimal shading amount

Below we provide the full derivation of the optimal shading amount

$$
\begin{aligned}
\frac{\partial U^D}{\partial \tilde{q}} = -1 + \frac{\alpha}{2}(q - \tilde{q}) &= 0 \\
\frac{\alpha}{2}(q - \tilde{q}) &= \frac{1}{2} \\
q - \tilde{q} &= \frac{2}{\alpha} \\
-\tilde{q} &= -q + \frac{2}{\alpha} \\
\tilde{q} - q &= \frac{2}{\alpha}
\end{aligned}
$$

Step 3: Switch point

Below we provide the full derivation of the switch point

$$
\begin{aligned}
q - T - \beta r &= (q - \tilde{q}) - \frac{\alpha}{4}(q - \tilde{q})^2 - \beta r \\
q - T &= \left(q - q + \frac{2}{\alpha}\right) - \frac{\alpha}{4}\left(q - q + \frac{2}{\alpha}\right)^2 \\
q - T &= \frac{2}{\alpha} - \frac{\alpha}{4}\left(\frac{4}{\alpha^2}\right) \\
q - T &= \frac{1}{\alpha}
\end{aligned}
$$

$$q^* = T + \frac{1}{\alpha}$$

Step 4: Driver's optimal choice of effort and risk

The driver chooses effort to maximize his utility

$$\max_{e,r} \quad E[(q - T - \beta r) \mid q \geq q^*] \cdot Pr(q \geq q^*) + E[(q - \tilde{q}) - \frac{\alpha}{4}(q - \tilde{q})^2 - \beta r \mid q < q^*] \cdot Pr(q < q^*) - h(e, r)$$

Simplifying

$$\max_{e,r} \quad E\left[(q - T - \beta r) \mid q \geq q^*\right](1 - F(q^*)) + E\left[\frac{1}{\alpha} - \beta r \mid q < q^*\right] \cdot F(q^*) - h(e, r)$$

Expressing in terms of exogenous variable

$$\max_{e,r} \quad E\left[(e + r\varepsilon - T - \beta r) \mid \varepsilon \geq \frac{q^* - e}{r}\right] \cdot \left(1 - F_\varepsilon\left(\frac{q^* - e}{r}\right)\right) + E\left[\frac{1}{\alpha} - \beta r \mid \varepsilon < \frac{q^* - e}{r}\right] \cdot F_\varepsilon\left(\frac{q^* - e}{r}\right) - h(e, r)$$

Using integral notation:

$$L = \int_{\frac{q^*-e}{r}}^{\infty} (e + r\varepsilon - T - \beta r) f_\varepsilon(\varepsilon) d\varepsilon + \int_0^{\frac{q^*-e}{r}} \left(\frac{1}{\alpha} - \beta r\right) f_\varepsilon(\varepsilon) d\varepsilon - h(e, r)$$

Taking the derivative with respect to e

$$\begin{aligned}
\frac{\partial L}{\partial e} &= \int_{\frac{q^*-e}{r}}^{\infty} 1 \cdot f_\varepsilon(\varepsilon) d\varepsilon + \cancel{\frac{1}{r}(q^* - T - \beta r) f_\varepsilon\left(\frac{q^* - e}{r}\right)} + 0 - \cancel{\frac{1}{r}\left(\frac{1}{\alpha} - \beta r\right) f_\varepsilon\left(\frac{q^* - e}{r}\right)} - h'_e \\
&= \int_{\frac{q^*-e}{r}}^{\infty} f_\varepsilon(\varepsilon) d\varepsilon - h'_e \\
&\rightarrow \underbrace{1 - F_\varepsilon\left(\frac{q^* - e}{r}\right) - h'_e}_{F.O.C} = 0
\end{aligned}$$

Taking the derivative with respect to r

$$\frac{\partial L}{\partial r} = \int_{\frac{q^*-e}{r}}^{\infty} (\varepsilon - \beta) \cdot f_\varepsilon(\varepsilon) d\varepsilon + \cancel{\left(\frac{q^* - e}{r^2}\right)(q^* - T - \beta r) f_\varepsilon\left(\frac{q^* - e}{r}\right)} + \int_0^{\frac{q^*-e}{r}} (-\beta) \cdot f_\varepsilon(\varepsilon) d\varepsilon - \cancel{\left(\frac{q^* - e}{r^2}\right)\left(\frac{1}{\alpha} - \beta r\right) f_\varepsilon\left(\frac{q^* - e}{r}\right)} - h'_r$$

$$= \int_{\frac{q^*-e}{r}}^{\infty} \varepsilon f_\varepsilon(\varepsilon)d\varepsilon - h'_r - \beta\left(\int_{\frac{q^*-e}{r}}^{\infty} f_\varepsilon(\varepsilon)d\varepsilon + \int_0^{\frac{q^*-e}{r}} f_\varepsilon(\varepsilon)d\varepsilon\right)$$

$$\rightarrow \underbrace{\int_{\frac{q^*-e}{r}}^{\infty} \varepsilon f_\varepsilon(\varepsilon)d\varepsilon - h'_r - \beta}_{F.O.C} = 0$$

Next we investigate how a change in T affects effort and risk:

$$\begin{bmatrix} \frac{\partial e}{\partial T} \\ \frac{\partial r}{\partial T} \end{bmatrix} = -\begin{bmatrix} \frac{\partial^2 L}{\partial e^2} & \frac{\partial L}{\partial r \partial e} \\ \frac{\partial L}{\partial e \partial r} & \frac{\partial^2 L}{\partial r^2} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial L}{\partial T \partial e} \\ \frac{\partial L}{\partial T \partial r} \end{bmatrix}$$

$$= -\frac{1}{\underbrace{\text{Determinant}}_{S.O.C \text{ for Hessian } > 0}} \begin{bmatrix} \frac{\partial^2 L}{\partial r^2} & -\frac{\partial L}{\partial r \partial e} \\ -\frac{\partial L}{\partial e \partial r} & \frac{\partial^2 L}{\partial e^2} \end{bmatrix} \begin{bmatrix} \frac{\partial L}{\partial T \partial e} \\ \frac{\partial L}{\partial T \partial r} \end{bmatrix}$$

Taking each term in turn:

$$\frac{\partial^2 L}{\partial r^2} = 0 - \frac{\partial}{\partial r}\left(\frac{q^*-e}{r}\right)\left(\frac{q^*-e}{r}\right) f_\varepsilon\left(\frac{q^*-e}{r}\right) - h''_{rr} - 2\beta$$

$$= \left(\frac{q^*-e}{r^2}\right)\left(\frac{q^*-e}{r}\right) f_\varepsilon\left(\frac{q^*-e}{r}\right) - h''_{rr} - 2\beta$$

$$= \underbrace{\frac{1}{r}\left(\frac{q^*-e}{r}\right)^2 f_\varepsilon(\cdot) - h''_{rr}}_{S.O.C < 0}$$

$$\frac{\partial^2 L}{\partial e^2} = f_\varepsilon(\cdot)\left(\frac{1}{r}\right) - h''_{ee}$$

$$= \underbrace{\frac{1}{r} f_\varepsilon(\cdot) - h''_{ee}}_{S.O.C < 0}$$

$$\frac{\partial L}{\partial e \partial r} = 0 - \left[-\frac{1}{r}\left(\frac{q^*-e}{r}\right) f_\varepsilon\left(\frac{q^*-e}{r}\right)\right] - h''_{er}$$

$$= \underbrace{\left(\frac{q^*-e}{r^2}\right) f_\varepsilon(\cdot) - h''_{er}}_{< 0}$$

$$\frac{\partial L}{\partial r \partial e} = \underbrace{-f_\varepsilon(\cdot)\left(\frac{q^*-e}{r^2}\right) - h''_{er}}_{< 0}$$

$$\frac{\partial L}{\partial T \partial e} = -\frac{1}{r} f_\varepsilon\left(\frac{q^*-e}{r}\right)$$

$$= \underbrace{-\frac{1}{r} f_\varepsilon(\cdot)}_{< 0}$$

$$\frac{\partial L}{\partial T \partial r} = 0 - \frac{1}{r}\left(\frac{q^* - e}{r}\right) f_\varepsilon\left(\frac{q^* - e}{r}\right)$$

$$= \underbrace{-\left(\frac{q^* - e}{r^2}\right) f_\varepsilon(\cdot)}_{>\,0}$$

We can sign most of these terms because of 1) second order conditions and 2) the shape of the distribution of revenue (q), which is skewed to the left, and the fact that drivers make the target 44% of the time. This means ($q^* - e < 0$). Note, we would expect the cross partial to be negative because as the driver increases risk (fatter tails), they are less likely to make the target, which means the returns to making the target decrease and effort will be reduced. Putting it altogether:

$$\frac{\partial e}{\partial T} = -\underbrace{\frac{1}{\text{Determinant}}}_{+}\left[\overbrace{\frac{\partial^2 L}{\partial r^2}}^{-} \cdot \overbrace{\frac{\partial L}{\partial T \partial e}}^{-} - \overbrace{\frac{\partial L}{\partial r \partial e}}^{-} \cdot \overbrace{\frac{\partial L}{\partial T \partial r}}^{+}\right]$$

$$< 0$$

$$\frac{\partial r}{\partial T} = -\underbrace{\frac{1}{\text{Determinant}}}_{+}\left[-\overbrace{\frac{\partial L}{\partial e \partial r}}^{-} \cdot \overbrace{\frac{\partial L}{\partial T \partial e}}^{-} + \overbrace{\frac{\partial^2 L}{\partial e^2}}^{-} \cdot \overbrace{\frac{\partial L}{\partial T \partial r}}^{+}\right]$$

$$> 0$$

Next we investigate how a change in α affects effort and risk:

$$\begin{bmatrix} \frac{\partial e}{\partial \alpha} \\ \frac{\partial r}{\partial \alpha} \end{bmatrix} = -\begin{bmatrix} \frac{\partial^2 L}{\partial e^2} & \frac{\partial L}{\partial r \partial e} \\ \frac{\partial L}{\partial e \partial r} & \frac{\partial^2 L}{\partial r^2} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial L}{\partial \alpha \partial e} \\ \frac{\partial L}{\partial \alpha \partial r} \end{bmatrix}$$

$$= -\underbrace{\frac{1}{\text{Determinant}}}_{S.O.C \text{ for Hessian} \,>\, 0} \begin{bmatrix} \frac{\partial^2 L}{\partial r^2} & -\frac{\partial L}{\partial r \partial e} \\ -\frac{\partial L}{\partial e \partial r} & \frac{\partial^2 L}{\partial e^2} \end{bmatrix} \begin{bmatrix} \frac{\partial L}{\partial \alpha \partial e} \\ \frac{\partial L}{\partial \alpha \partial r} \end{bmatrix}$$

Computing the additional terms

$$\frac{\partial L}{\partial \alpha \partial e} = f_\varepsilon\left(\frac{q^* - e}{r}\right)\left(\frac{1}{r \cdot \alpha^2}\right)$$

$$= \frac{1}{\alpha^2 \cdot r} f_\varepsilon(\cdot)$$

$$\frac{\partial L}{\partial \alpha \partial r} = 0 + \frac{1}{r \cdot 4\alpha^2}\left(\frac{q^* - e}{r}\right) f_\varepsilon\left(\frac{q^* - e}{r}\right)$$

$$= \frac{1}{\alpha^2 \cdot r}\left(\frac{q^* - e}{r}\right) f_\varepsilon(\cdot)$$

Putting it altogether:

$$\frac{\partial e}{\partial \alpha} = -\underbrace{\frac{1}{\text{Determinant}}}_{+}\left[\overbrace{\frac{\partial^2 L}{\partial r^2}}^{-} \cdot \overbrace{\frac{\partial L}{\partial \alpha \partial e}}^{-} - \overbrace{\frac{\partial L}{\partial r \partial e}}^{-} \cdot \overbrace{\frac{\partial L}{\partial \alpha \partial r}}^{+}\right]$$

$$> 0$$

$$\frac{\partial r}{\partial \alpha} = -\underbrace{\frac{1}{\text{Determinant}}}_{+}\left[-\overbrace{\frac{\partial L}{\partial e \partial r}}^{-} \cdot \overbrace{\frac{\partial L}{\partial \alpha \partial e}}^{-} + \overbrace{\frac{\partial^2 L}{\partial e^2}}^{-} \cdot \overbrace{\frac{\partial L}{\partial \alpha \partial r}}^{+}\right]$$

$$< 0$$

Finally we investigate how a change in β affects effort and risk:

$$\begin{bmatrix} \frac{\partial e}{\partial \beta} \\ \frac{\partial r}{\partial \beta} \end{bmatrix} = -\begin{bmatrix} \frac{\partial^2 L}{\partial e^2} & \frac{\partial L}{\partial r \partial e} \\ \frac{\partial L}{\partial e \partial r} & \frac{\partial^2 L}{\partial r^2} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial L}{\partial \beta \partial e} \\ \frac{\partial L}{\partial \beta \partial r} \end{bmatrix}$$

$$= -\underbrace{\frac{1}{\text{Determinant}}}_{S.O.C \text{ for Hessian } > 0} \begin{bmatrix} \frac{\partial^2 L}{\partial r^2} & -\frac{\partial L}{\partial r \partial e} \\ -\frac{\partial L}{\partial e \partial r} & \frac{\partial^2 L}{\partial e^2} \end{bmatrix} \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

Putting it altogether:

$$\frac{\partial e}{\partial \beta} = -\underbrace{\frac{1}{\text{Determinant}}}_{+}\left[\overbrace{\frac{\partial L}{\partial r \partial e}}^{-} \cdot \overbrace{1}^{+}\right]$$

$$> 0$$

$$\frac{\partial r}{\partial \beta} = -\underbrace{\frac{1}{\text{Determinant}}}_{+}\left[\overbrace{\frac{\partial^2 L}{\partial e^2}}^{-} \cdot \overbrace{(-1)}^{-}\right]$$

$$< 0$$

Step 5: Owner's optimal reporting choice

Constrained case

The owner chooses T to maximize his utility:

$$\max_{T} \quad T \cdot Pr(q \geq q^*) + E[\tilde{q} \mid q < q^*] \cdot Pr(q < q^*) - \gamma(r) \qquad \text{s.t}$$

$$E\Big[q - T - \beta r \mid q \geq q^*\Big] \cdot Pr(q \geq q^*) + E\Big[q - \tilde{q} - \frac{\alpha}{4}(q - \tilde{q})^2 - \beta r \mid q < q^*\Big]$$
$$\cdot Pr(q < q^*) - h(e^*, r^*) > R$$

Expressing in terms of exogenous variables:

$$\max_{T} \quad T \cdot Pr\left(\varepsilon \geq \frac{q^* - e^*}{r}\right) + E\left[e + r\varepsilon - \frac{1}{2\alpha} \mid \varepsilon < \frac{q^* - e^*}{r}\right]$$

$$\cdot Pr\left(\varepsilon < \frac{q^* - e^*}{r}\right) - \gamma(r) \qquad \text{s.t}$$

$$E\left[(e + r\varepsilon - T - \beta r) \mid \varepsilon \geq \frac{q^* - e}{r}\right] \cdot Pr(\varepsilon \geq \left(\frac{q^* - e}{r}\right) + E\left[\frac{1}{\alpha} - \beta r \mid \varepsilon < \frac{q^* - e}{r}\right]$$

$$\cdot Pr\left(\varepsilon < \frac{q^* - e}{r}\right) - h(e, r) \geq 0$$

Translating into integral notation:

$$L = \underbrace{T \int_{\frac{q^*-e^*}{r*}}^{\infty} f_\varepsilon(\varepsilon) d\varepsilon}_{A} + \underbrace{\int_0^{\frac{q^*-e^*}{r*}} \left(e + r\varepsilon - \frac{1}{2\alpha}\right) f_\varepsilon(\varepsilon) d\varepsilon}_{B} - \gamma(r) +$$

$$\lambda \left[\underbrace{\int_{\frac{q^*-e^*}{r*}}^{\infty} (e + r\varepsilon - T - \beta r)\, f_\varepsilon(\varepsilon) d\varepsilon}_{C} + \underbrace{\int_0^{\frac{q^*-e^*}{r*}} \left(\frac{1}{\alpha} - \beta r\right) f_\varepsilon(\varepsilon) d\varepsilon - h(e, r)}_{D} \right]$$

Taking the derivative with respect to T

$$= \underbrace{\int_{\frac{q^*-e^*}{r*}}^{\infty} f_\varepsilon(\varepsilon) d\varepsilon + T\left[0 - \frac{\partial}{\partial T}\left(\frac{q^* - e^*}{r^*}\right) f_\varepsilon(\cdot)\right]}_{A'} +$$

$$\underbrace{\int_0^{\frac{q^*-e^*}{r*}} \left(\frac{\partial e}{\partial T} + \frac{\partial r}{\partial T}\varepsilon\right) f_\varepsilon(\varepsilon) d\varepsilon + \frac{\partial}{\partial T}\left(\frac{q^* - e^*}{r^*}\right)\left(T - \frac{1}{\alpha}\right) f_\varepsilon(\cdot)}_{B'} - \gamma'(r)\frac{\partial r}{\partial T}$$

$$+ \lambda\Bigg[\underbrace{\int_{\frac{q^*-e^*}{r*}}^{\infty} \left(\frac{\partial e}{\partial T} + \frac{\partial r}{\partial T}\varepsilon - 1 - \left(\beta \cdot \frac{\partial r}{\partial T}\right)\right) f_\varepsilon(\varepsilon) d\varepsilon - \frac{\partial}{\partial T}\left(\frac{q^* - e^*}{r^*}\right)\frac{1}{\alpha} f_\varepsilon(\cdot)}_{C'}$$

$$+ \underbrace{\int_0^{\frac{q^*-e}{r}} -\left(\beta \cdot \frac{\partial r}{\partial T}\right) f_\varepsilon(\varepsilon) d\varepsilon + \frac{\partial}{\partial T}\left(\frac{q^* - e^*}{r^*}\right)\frac{1}{\alpha} f_\varepsilon(\cdot) - h'\left(\frac{\partial e}{\partial T} + \frac{\partial r}{\partial T}\right)}_{D'} \Bigg]$$

$$= \underbrace{1 - F_\varepsilon(\cdot) - \cancel{T\frac{\partial}{\partial T}\left(\frac{q^* - e^*}{r^*}\right) f_\varepsilon(\cdot)}}_{A'} +$$

$$\underbrace{\frac{\partial e}{\partial T}F_\varepsilon(\cdot)+\frac{\partial r}{\partial T}\int_0^{\frac{q^*-e^*}{r^*}}\varepsilon f_\varepsilon(\varepsilon)d\varepsilon+\cancel{\frac{\partial}{\partial T}\left(\frac{q^*-e^*}{r^*}\right)\left(T-\frac{1}{\alpha}\right)}f_\varepsilon(\cdot)}_{B'}-\gamma'(r)\frac{\partial r}{\partial T}$$

$$+\lambda\Bigg[\underbrace{-(1-F_\varepsilon(\cdot))+\frac{\partial e}{\partial T}(1-F_\varepsilon(\cdot))+\frac{\partial r}{\partial T}\int_{\frac{q^*-e^*}{r^*}}^{\infty}\varepsilon f_\varepsilon(\varepsilon)d\varepsilon-\cancel{\frac{\partial}{\partial T}\left(\frac{q^*-e^*}{r^*}\right)\frac{1}{\alpha}}f_\varepsilon(\cdot)}_{C'}$$

$$+\underbrace{\cancel{\frac{\partial}{\partial T}\left(\frac{q^*-e^*}{r^*}\right)\frac{1}{\alpha}}f_\varepsilon(\cdot)-h'\left(\frac{\partial e}{\partial T}+\frac{\partial r}{\partial T}\right)}_{D'}-\beta\frac{\partial r}{\partial T}\Bigg]$$

$$=\underbrace{1-F_\varepsilon(\cdot)}_{A'}+\underbrace{\frac{\partial e}{\partial T}F_\varepsilon(\cdot)+\frac{\partial r}{\partial T}\int_0^{\frac{q^*-e^*}{r^*}}\varepsilon f_\varepsilon(\varepsilon)d\varepsilon-\frac{\partial}{\partial T}\left(\frac{q^*-e^*}{r^*}\right)\left(\frac{1}{\alpha}\right)f_\varepsilon(\cdot)}_{B'}-\gamma'(r)\frac{\partial r}{\partial T}$$

$$+\lambda\left[\underbrace{-(1-F_\varepsilon(\cdot))+\frac{\partial e}{\partial T}(1-F_\varepsilon(\cdot))+\frac{\partial r}{\partial T}\int_{\frac{q^*-e^*}{r^*}}^{\infty}\varepsilon f_\varepsilon(\varepsilon)d\varepsilon}_{C'}+\frac{\partial r}{\partial T}(-\beta)-\underbrace{h'\left(\frac{\partial e}{\partial T}+\frac{\partial r}{\partial T}\right)}_{D'}\right]=0$$

Taking the derivative with respect to λ

$$\int_{\frac{q^*-e^*}{r^*}}^{\infty}(e+r\varepsilon-T-\beta r)\,f_\varepsilon(\varepsilon)d\varepsilon+\int_0^{\frac{q^*-e^*}{r^*}}\left(\frac{1}{\alpha}-\beta r\right)f_\varepsilon(\varepsilon)d\varepsilon-h(e,r)=0$$

Next we apply the IFT to understand how T changes with α.

$$\begin{aligned}\begin{bmatrix}\frac{\partial T}{\partial\alpha}\\ \frac{\partial\lambda}{\partial\alpha}\end{bmatrix}&=-\begin{bmatrix}\frac{\partial^2L}{\partial T^2}&\frac{\partial L}{\partial\lambda\partial T}\\ \frac{\partial L}{\partial T\partial\lambda}&\frac{\partial^2L}{\partial\lambda^2}\end{bmatrix}^{-1}\begin{bmatrix}\frac{\partial L}{\partial\alpha\partial T}\\ \frac{\partial L}{\partial\alpha\partial\lambda}\end{bmatrix}\\
&=-\frac{1}{\text{Determinant}}\begin{bmatrix}\frac{\partial^2L}{\partial\lambda^2}&-\frac{\partial L}{\partial\lambda\partial T}\\ -\frac{\partial L}{\partial T\partial\lambda}&\frac{\partial^2L}{\partial T^2}\end{bmatrix}^{-1}\begin{bmatrix}\frac{\partial L}{\partial\alpha\partial T}\\ \frac{\partial L}{\partial\alpha\partial\lambda}\end{bmatrix}\\
&=-\frac{1}{\text{Determinant}}\begin{bmatrix}0&-\frac{\partial L}{\partial\lambda\partial T}\\ -\frac{\partial L}{\partial T\partial\lambda}&\frac{\partial^2L}{\partial T^2}\end{bmatrix}^{-1}\begin{bmatrix}\frac{\partial L}{\partial\alpha\partial T}\\ \frac{\partial L}{\partial\alpha\partial\lambda}\end{bmatrix}\end{aligned}$$

Taking each term in turn:

$$\begin{aligned}\frac{\partial L}{\partial\lambda\partial T}&=-(1-F_\varepsilon(\cdot))+\frac{\partial e}{\partial T}(1-F_\varepsilon(\cdot))+\frac{\partial r}{\partial T}\int_{\frac{q^*-e^*}{r^*}}^{\infty}\varepsilon f_\varepsilon(\varepsilon)d\varepsilon+\frac{\partial r}{\partial T}(-\beta)-h'\left(\frac{\partial e}{\partial T}+\frac{\partial r}{\partial T}\right)\\
&=-(1-F_\varepsilon(\cdot))+\underbrace{\frac{\partial e}{\partial T}\left(1-F_\varepsilon(\cdot)-h'_e\right)}_{F.O.C\;=\;0}+\underbrace{\frac{\partial r}{\partial T}\left(\int_{\frac{q^*-e^*}{r^*}}^{\infty}\varepsilon f_\varepsilon(\varepsilon)d\varepsilon-h'_r-\beta\right)}_{F.O.C\;=\;0}\end{aligned}$$

$$= \underbrace{-(1 - F_\varepsilon(\cdot))}_{<\,0}$$

$$\frac{\partial L}{\partial\alpha\partial\lambda} = \int_{\frac{q^*-e^*}{r*}}^{\infty}\left(\frac{\partial e}{\partial\alpha} + \frac{\partial r}{\partial\alpha}\varepsilon - \beta\frac{\partial r}{\partial\alpha}\right) f_\varepsilon(\varepsilon)d\varepsilon - \cancel{\frac{\partial}{\partial\alpha}\left(\frac{q^*-e^*}{r^*}\right)\left(\frac{1}{\alpha} - \beta r\right) f_\varepsilon(\cdot)}$$

$$- \int_0^{\frac{q^*-e^*}{r*}}\left(\frac{1}{\alpha^2} + \beta\frac{\partial r}{\partial\alpha}\right) f_\varepsilon(\varepsilon)d\varepsilon + \cancel{\frac{\partial}{\partial\alpha}\left(\frac{q^*-e^*}{r^*}\right)\left(\frac{1}{\alpha} - \beta r\right) f_\varepsilon(\cdot)} - h'\left(\frac{\partial e}{\partial\alpha} + \frac{\partial r}{\partial\alpha}\right)$$

$$= \int_{\frac{q^*-e^*}{r*}}^{\infty}\left(\frac{\partial e}{\partial\alpha} + \frac{\partial r}{\partial\alpha}\varepsilon - \beta\frac{\partial r}{\partial\alpha}\right) f_\varepsilon(\varepsilon)d\varepsilon - \int_0^{\frac{q^*-e^*}{r*}}\left(\frac{1}{\alpha^2} + \beta\frac{\partial r}{\partial\alpha}\right) f_\varepsilon(\varepsilon)d\varepsilon - h'\left(\frac{\partial e}{\partial\alpha} + \frac{\partial r}{\partial\alpha}\right)$$

$$= \frac{\partial e}{\partial\alpha}(1 - F_\varepsilon(\cdot)) + \frac{\partial r}{\partial\alpha}\int_{\frac{q^*-e^*}{r*}}^{\infty}\varepsilon f_\varepsilon(\varepsilon)d\varepsilon - \int_0^{\frac{q^*-e^*}{r*}}\left(\frac{1}{\alpha^2}\right) f_\varepsilon(\varepsilon)d\varepsilon - h'\left(\frac{\partial e}{\partial\alpha} + \frac{\partial r}{\partial\alpha}\right) - \beta\frac{\partial r}{\partial\alpha}$$

$$= \underbrace{\frac{\partial e}{\partial\alpha}(1 - F_\varepsilon(\cdot) - h'_e)}_{F.O.C\,=\,0} + \underbrace{\frac{\partial r}{\partial\alpha}\left(\int_{\frac{q^*-e^*}{r*}}^{\infty}\varepsilon f_\varepsilon(\varepsilon)d\varepsilon - h'_r - \beta\right)}_{F.O.C\,=\,0} - \int_0^{\frac{q^*-e^*}{r*}}\left(\frac{1}{\alpha^2}\right) f_\varepsilon(\varepsilon)d\varepsilon$$

$$= \underbrace{-\frac{1}{\alpha^2}F_\varepsilon(\cdot)}_{<0}$$

Putting it all together:

$$\frac{\partial T}{\partial\alpha} = -\frac{1}{0 - \left(\frac{\partial L}{\partial T\partial\lambda}\right)^2}\left[-\frac{\partial L}{\partial\lambda\partial T}\cdot\frac{\partial L}{\partial\alpha\partial\lambda}\right]$$

$$= \underbrace{\frac{1}{\left(\frac{\partial L}{\partial T\partial\lambda}\right)^2}}_{+}\left[-\underbrace{\frac{\partial L}{\partial\lambda\partial T}}_{-}\cdot\underbrace{\frac{\partial L}{\partial\alpha\partial\lambda}}_{-}\right]$$

$$< 0$$

We can also apply the IFT to understand how T changes with β.

$$\begin{bmatrix}\frac{\partial T}{\partial\beta}\\ \frac{\partial\lambda}{\partial\beta}\end{bmatrix} = -\begin{bmatrix}\frac{\partial^2 L}{\partial T^2} & \frac{\partial L}{\partial\lambda\partial T}\\ \frac{\partial L}{\partial T\partial\lambda} & \frac{\partial^2 L}{\partial\lambda^2}\end{bmatrix}^{-1}\begin{bmatrix}\frac{\partial L}{\partial\beta\partial T}\\ \frac{\partial L}{\partial\beta\partial\lambda}\end{bmatrix}$$

$$= -\frac{1}{\text{Determinant}}\begin{bmatrix}\frac{\partial^2 L}{\partial\lambda^2} & -\frac{\partial L}{\partial\lambda\partial T}\\ -\frac{\partial L}{\partial T\partial\lambda} & \frac{\partial^2 L}{\partial T^2}\end{bmatrix}^{-1}\begin{bmatrix}\frac{\partial L}{\partial\beta\partial T}\\ \frac{\partial L}{\partial\beta\partial\lambda}\end{bmatrix}$$

$$= -\frac{1}{\text{Determinant}}\begin{bmatrix}0 & -\frac{\partial L}{\partial\lambda\partial T}\\ -\frac{\partial L}{\partial T\partial\lambda} & \frac{\partial^2 L}{\partial T^2}\end{bmatrix}^{-1}\begin{bmatrix}\frac{\partial L}{\partial\beta\partial T}\\ \frac{\partial L}{\partial\beta\partial\lambda}\end{bmatrix}$$

We compute the only new term required:

$$
\begin{aligned}
\frac{\partial L}{\partial \beta \partial \lambda} &= \int_{\frac{q^*-e^*}{r^*}}^{\infty} -r f_\varepsilon(\varepsilon) d\varepsilon - \cancel{\frac{\partial}{\beta}\left(\frac{q^*-e^*}{r^*}\right)(q^* - T - \beta r)}\, f_\varepsilon(\cdot) \\
&\quad + \int_0^{\frac{q^*-e^*}{r^*}} -r f_\varepsilon(\varepsilon) d\varepsilon + \cancel{\frac{\partial}{\beta}\left(\frac{q^*-e^*}{r^*}\right)\left(\frac{1}{\alpha} - \beta r\right)} f_\varepsilon(\cdot) \\
&= -r
\end{aligned}
$$

Putting it all together:

$$
\begin{aligned}
\frac{\partial T}{\partial \beta} &= \underbrace{\frac{1}{\left(\frac{\partial L}{\partial T \partial \lambda}\right)^2}}_{+} \left[- \underbrace{\frac{\partial L}{\partial \lambda \partial T}}_{-} \cdot \underbrace{\frac{\partial L}{\partial \beta \partial \lambda}}_{-} \right] \\
&< 0
\end{aligned}
$$