Title
Essays in Development Economics

Author
Kelley, Erin Munro

Publication Date
2019

Peer reviewed|Thesis/dissertation

Essays in Development Economics

By

Erin Munro Kelley

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 Jeremy Magruder, Chair
Professor Elisabeth Sadoulet
Professor Edward Miguel

Summer 2019

Abstract

Essays in Development Economics

by

Erin M Kelley

Doctor of Philosophy in Agricultural and Resource Economics

University of California, Berkeley

Professor Jeremy Magruder, Chair

This dissertation presents a three-part study in development economics. Each chapter investigates how technological innovations enable firms and individuals to overcome problems of asymmetric information - whether it be monitoring devices to help firms overcome problems of moral hazard; online job-boards to overcome search and matching frictions for job-seekers; or new agricultural techniques to increase information diffusion within social networks.

The first 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.

The second chapter investigates how agents in a social network can be encouraged to obtain information from outside their peer groups. Using a field experiment in rural Bangladesh, we show that demonstration plots in agriculture — a technique where the first users of a new variety cultivate it in a side-by-side comparison with an existing variety — facilitate social learning by inducing conversations and information sharing outside of existing social networks. We compare these im-

provements in learning with those from seeding new technology with more central farmers in village social networks. The demonstration plots — when cultivated by randomly selected farmers — improve knowledge by just as much as seeding with more central farmers. Moreover, the demonstration plots only induce conversations and facilitate learning for farmers that were unconnected to entry points at baseline. Finally, we combine this diffusion experiment with an impact experiment to show that both demonstration plots and improved seeding transmit information to farmers that are less likely to benefit from the new innovation.

Finally, the third chapter explores the impact of online job platforms on labor market frictions in India. In recent years innovative job market information systems have emerged in order to match recent graduates with employers via integrated websites and call center based platforms. We use a randomized control trial to evaluate the ability of such job-portals to ease search frictions in India. We partner with Job Shikari, an online platform operating primarily in Northern India, and upload a randomly selected sample of recent vocational training graduates onto their platform. A second subset of graduates are uploaded to the platform, and receive “priority status", which means they subsequently receive many more text messages than their peers about potentially job opportunities. We find that being uploaded to the job portal has a negative impact on the probability of being employed, but no significant impacts on job-search. The treatment priority group experiences a less strong disemployment effect, but is much more likely to respond to the treatment by migrating to urban centers. These results differ by job-seekers' observable characteristics such as age and martial status. We interpret our results as evidence that the impact of job-portals depends significantly on job-seekers beliefs about what the job-portal can do for them, rather than just on how well the platform can match the job-seeker with a particular job.

C | Contents

F | Figures

T | Tables

Acknowledgements

Coauthor, data, and funding Chapter 1 of this dissertation is coauthored with Gregory Lane and David Schönholzer. Chapter 2 of this dissertation is coauthored with Manzoor H. Dar, Alain de Janvry, Kyle Emerick, and Elisabeth Sadoulet. Chapter 3 of this dissertation is coauthored with Christopher Ksoll and Jeremy Magruder. I thank Echo Mobile (Jeremy, Boris, James, Langat, Wambururu), and the amazing team that made this research possible. I also thank and acknowledge CEGA, DIL, PEDL, the Weiss Foundation, and the IGC for funding these research projects. All errors are my own.

Personal thanks I feel very fortunate to have received the support of a village over the last 6 years. My life would not be where it is today without Greg. You've made everything better and possible during these 6 years. Ninette, Tim and Dave, thank you for providing direction and guidance throughout. There is no way I would have a thesis today without David, Jeremy, Betty and Ted. The opportunities, patience, and support they provided was staggering. I also want to thank the community of my fellow PhD students that helped me through the highs and lows: Kyle, Fiona, Ed, Shelley, Matt W., Deirdre, Matt. P, Manaswini, Dave M., Eleanor, Kate, Megan, Caroline, Alessandra, Juliana.

1 | 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.

1.1 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 3.2 discusses Kenya's transportation system, the prevalence of moral hazard, and the scope for monitoring. Section 3.3 details the field experiment, and Section 1.4 reviews the data. We present a simple theoretical framework in Section 2.4. Section 3.4 discusses each of our results. We then discuss the implications of the findings and conclude in the final section.

1.2 Context

1.2.1 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.

1.2.2 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.

1.2.3 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 1.1). 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).

1.3 Experimental Design

1.3.1 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 1.2). 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).

1.3.2 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.

1.3.3 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).

1.4 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.

1.4.1 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 C.1 and A.2).

1.4.2 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.

1.4.3 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 1.4 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.

1.5 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.

1.5.1 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 1.5 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:

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

Where

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

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

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

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):

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

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

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

1.5.2 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.

1.6 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).

1.6.1 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 1.1 and Table 1.2).

1.6.2 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.

1.6.3 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 1.6 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 1.3 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 A.3 though A.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 1.4).[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 1.7 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 1.7 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 1.8 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 1.4 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 1.9 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 1.10 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 1.11). 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 1.9 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 1.12 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 1.12).

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 1.13 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 1.13. 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 1.5). Figure 1.13 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 1.14 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 1.5 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 1.15 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 1.14. 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 1.17 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 1.16 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 1.16 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 1.7, 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 1.7, 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 1.6 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 1.12), and the amount they under-report decreases (Figure 1.14). While their salary per hour remains unchanged (Figure 1.17), 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 1.19). We see no effect on sharp turns or sharp braking (Figure 1.20). 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 1.21). 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.

1.6.4 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 1.9 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.

1.7 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 1.1: 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 1.2: 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 1.3: 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 1.4: 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 1.5: 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 1.6: 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 1.7: 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 1.8: 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.

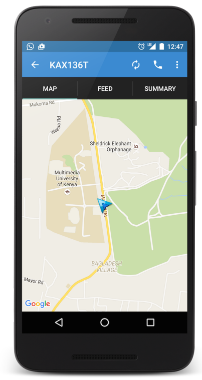
Table 1.9: Effect of no cash (ongoing)

	(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

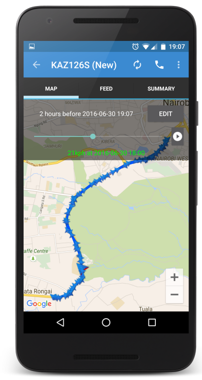
The data are from the tracking device throughout the study period. Column (1) and (2) capture average and maximum speeds throughout the day, respectively. Column (3)-(6) capture daily alerts for speeding over 80 km/hour, sharp breaking, over-acceleration, and sharp turning, respectively. We control for 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

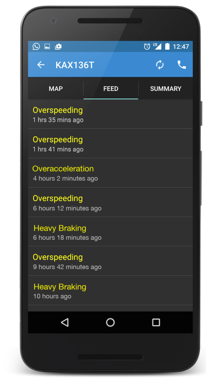
Figure 1.1: Mobile app “SmartMatatu"



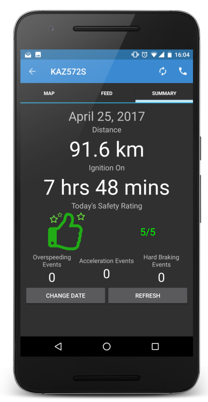
(a) Map Viewer



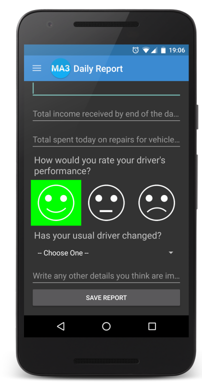
(b) Historical Map Viewer



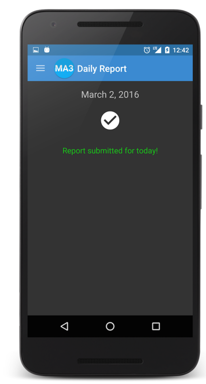
(c) Safety Feed



(d) Productivity Summary



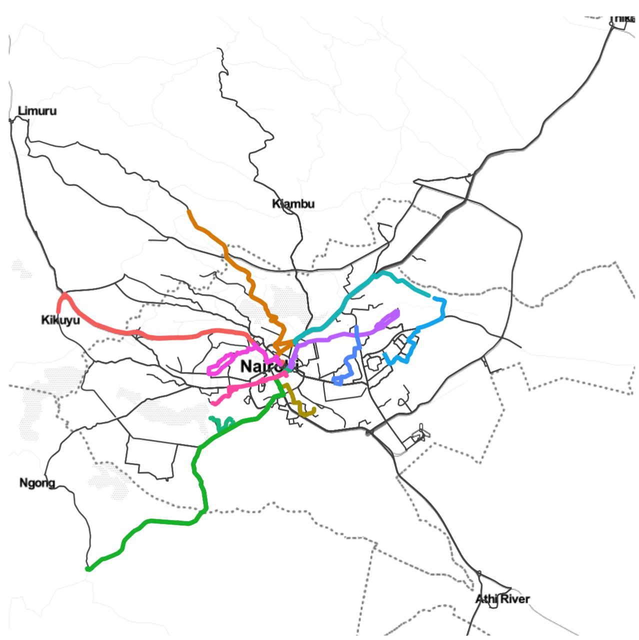
(e) Report Submit



(f) Report Complete

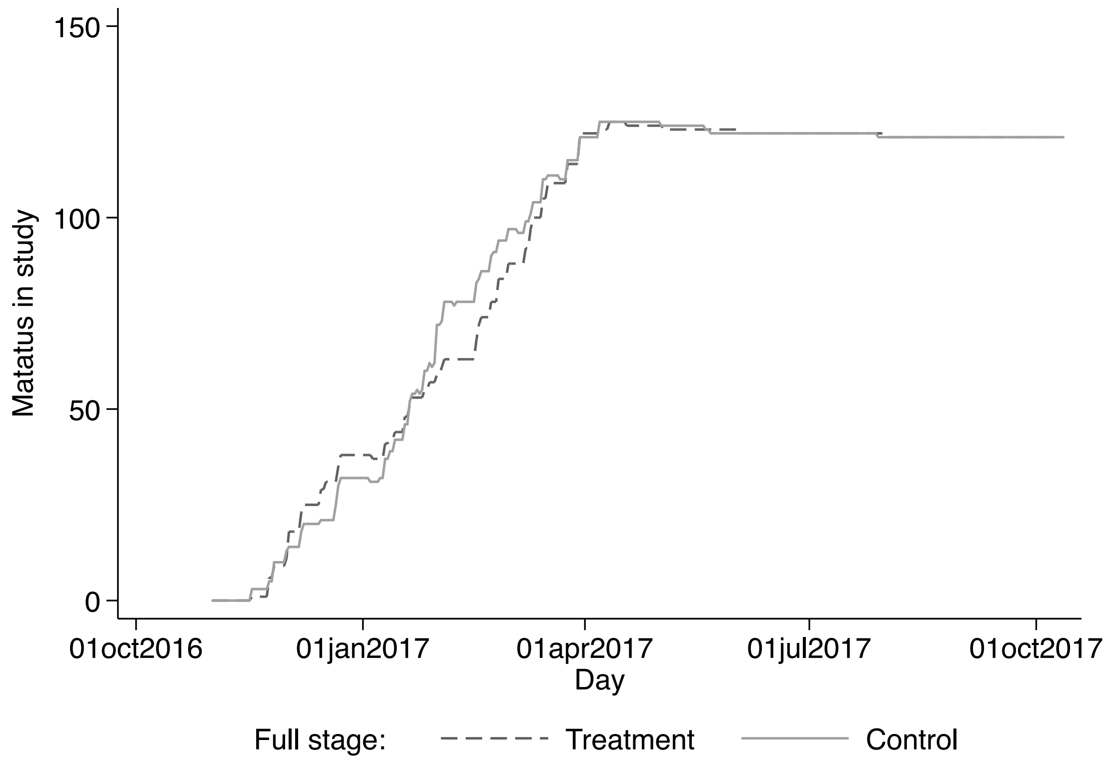
Figure 1.2: Device Location

(a) Designated bus routes in Nairobi (black)



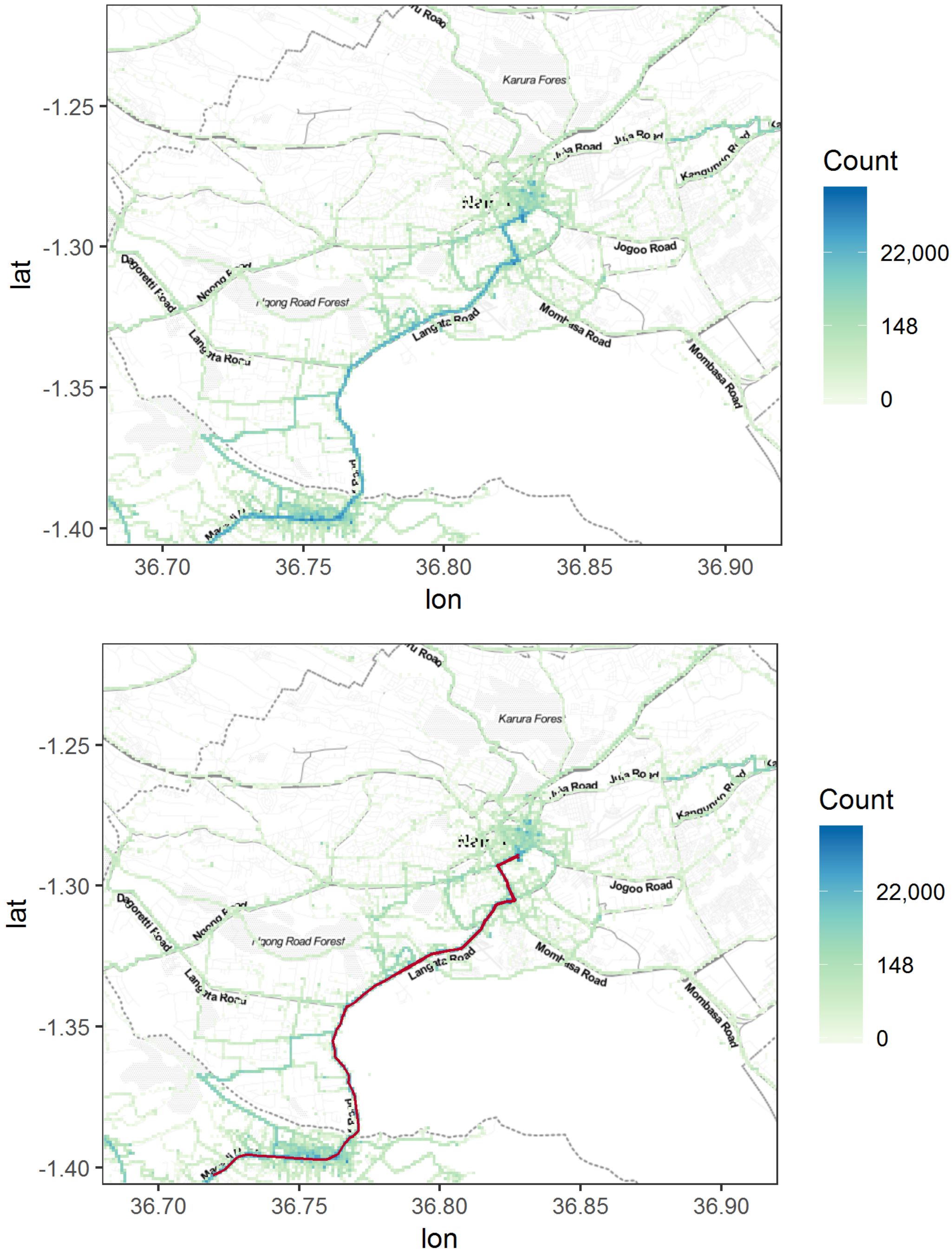
(b) Designated bus routes in Nairobi (black) and routes in our sample (colored)

Figure 1.3: Study Timeline



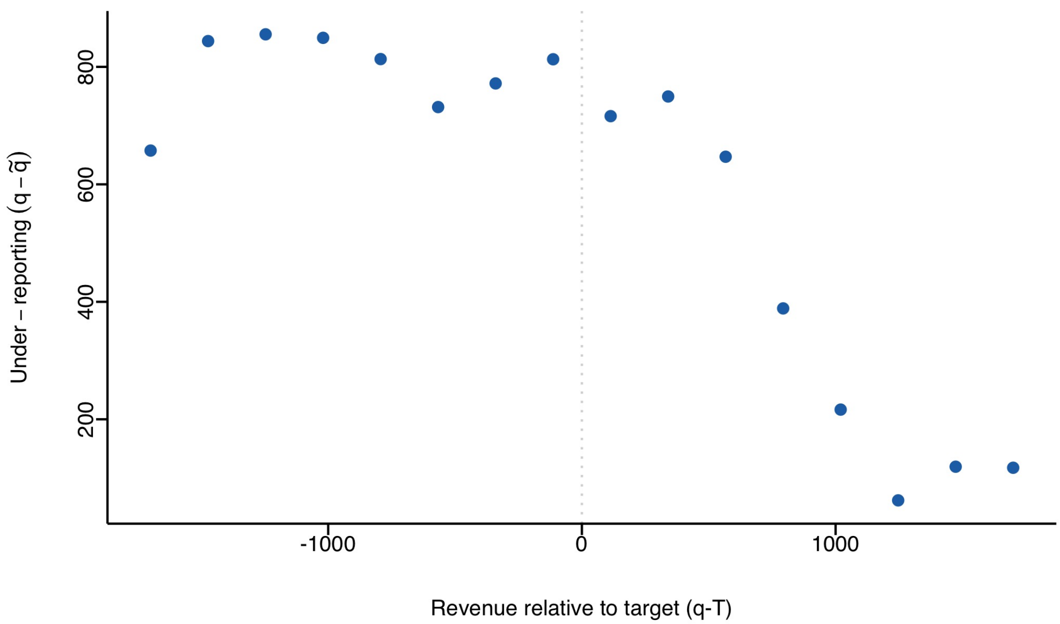
Notes: The figure depicts the number of matatus that were fitted with GPS trackers (and hence were added to the study) per week. The first installation took place in November 2016, and continued until April 2017. On average, the field team was able to fit GPS trackers to 15 matatus per week. As a result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

Figure 1.4: Device Location



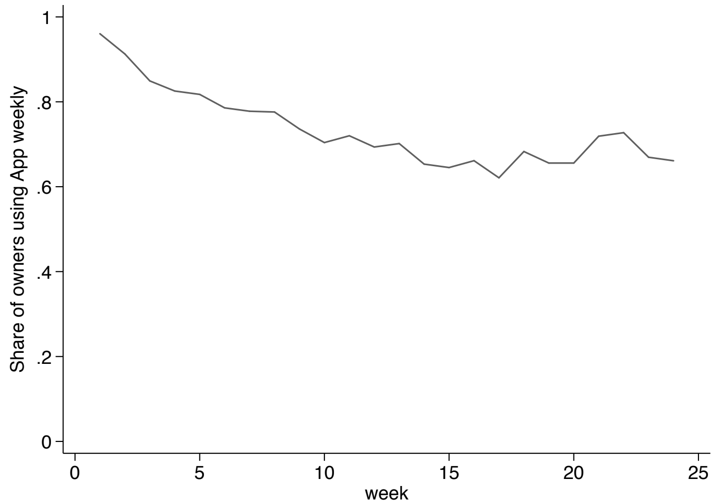
Notes: These maps use data from the trackers that were installed in vehicles licensed to operate on Route 126 (Ongata-Rongai line). Specifically, we count the number of times that vehicles passed through particular longitudinal and latitudinal cells on the map. A deeper shade of blue demonstrates that 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 1.5: Model Validation (Constant Shading)

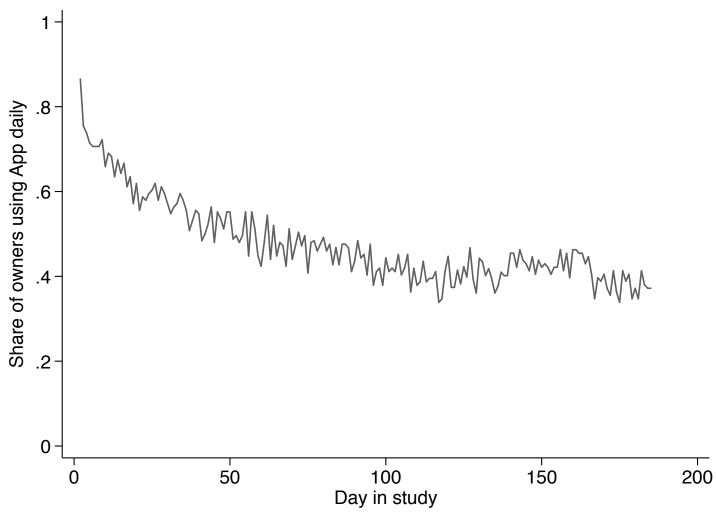


Notes: This figure depicts the amount of under-reporting on the y-axis, and the amount of revenue relative to the target on the x-axis. The data are from the control group. Under-reporting is the amount of revenue the driver withholds from the owner. According to the contract, the driver must deliver 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 1.6: Device Usage (API Calls)



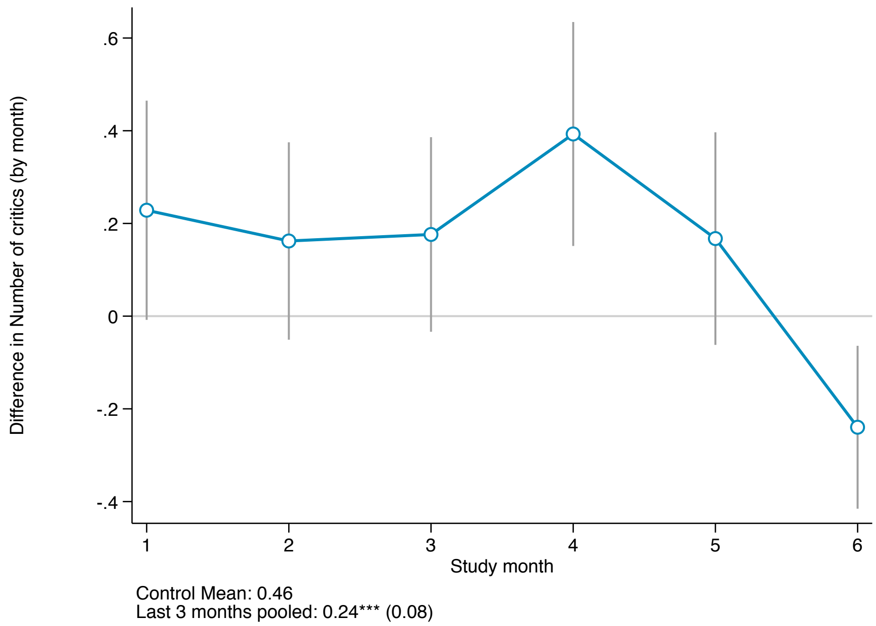
(a) Weekly Usage

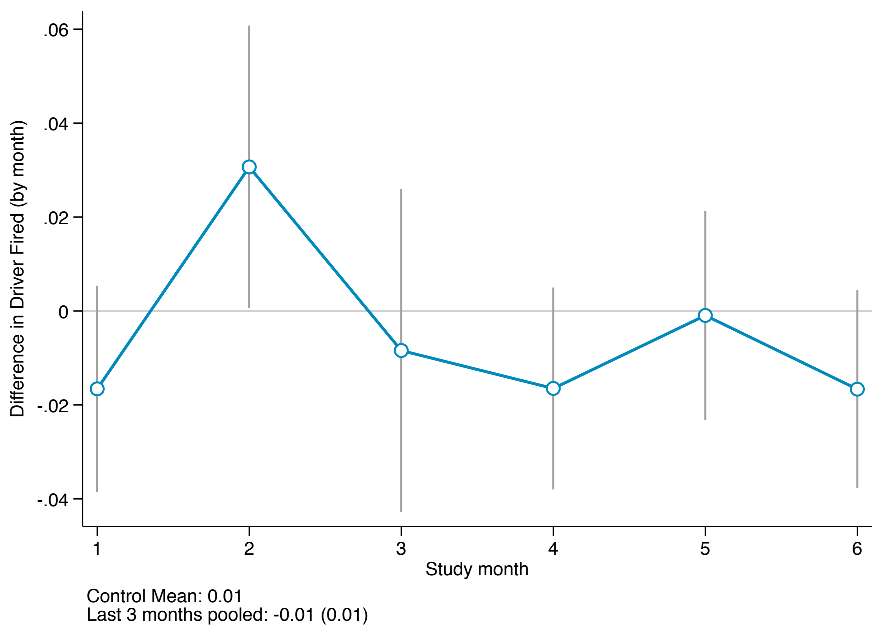


(b) Daily Usage

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.

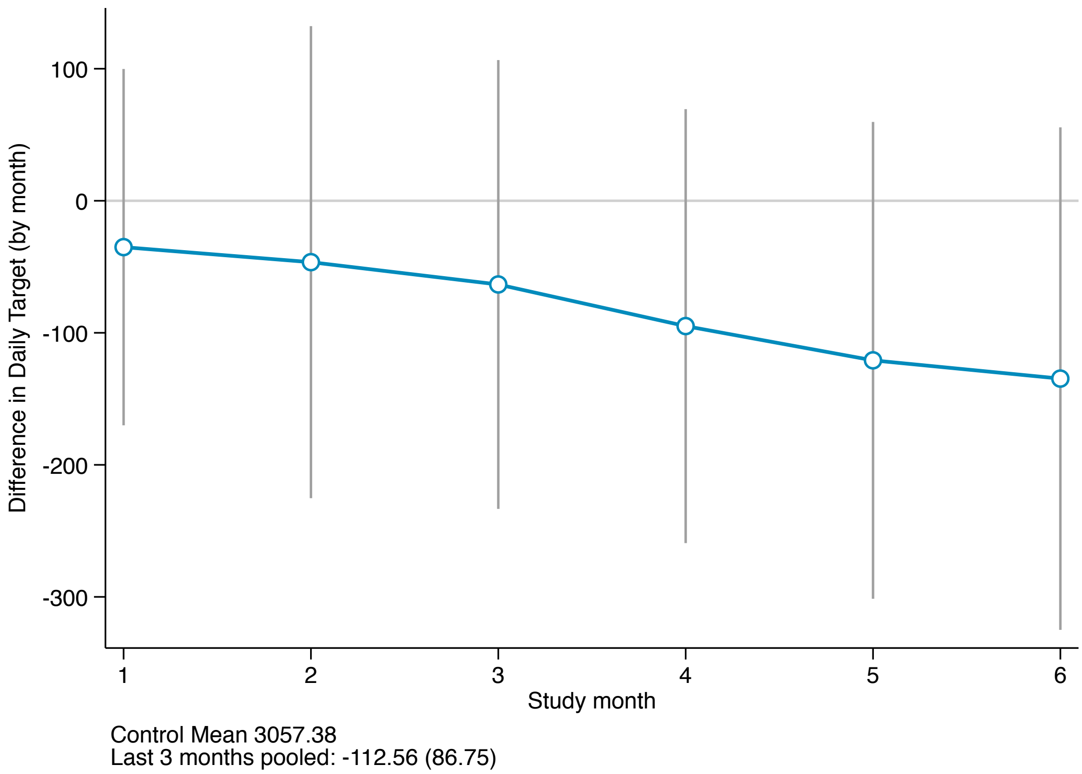
Figure 1.7: Reprimands and Firing





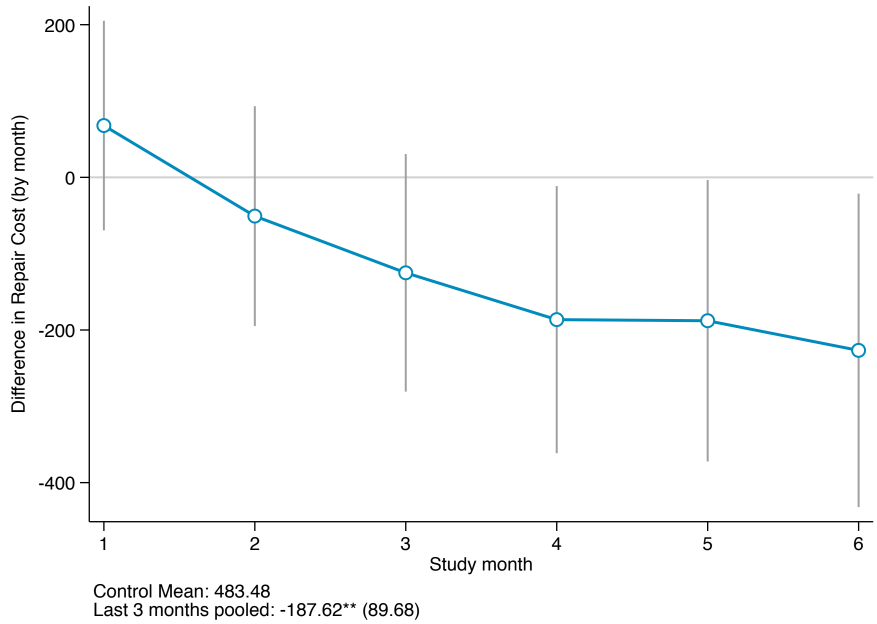
Notes: Each point on the graphs is the average difference between treatment and control in a particular month for a particular outcome. Note that because installations were rolled out over time, Month 1 (through 6) represent the first month (through the sixth month) since *installation*. The outcome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

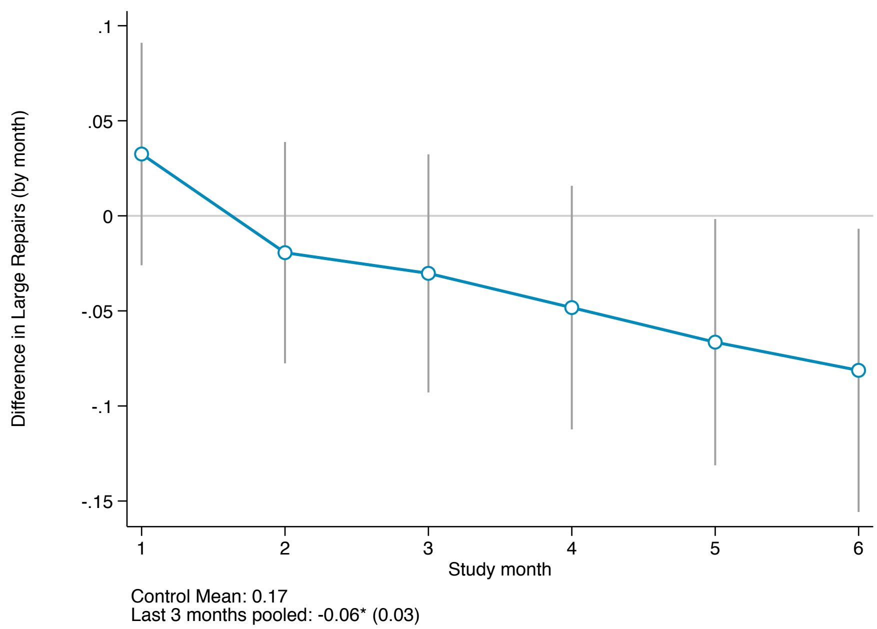
Figure 1.8: Prediction 1 → Target



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*.

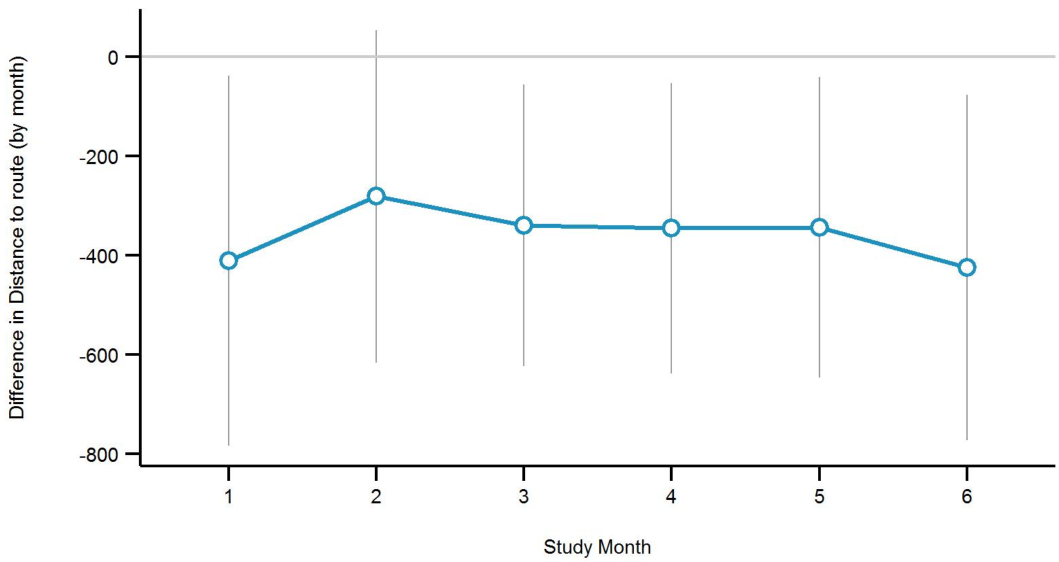
Figure 1.9: Prediction 2 → Damaging (Risky) Driving





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).

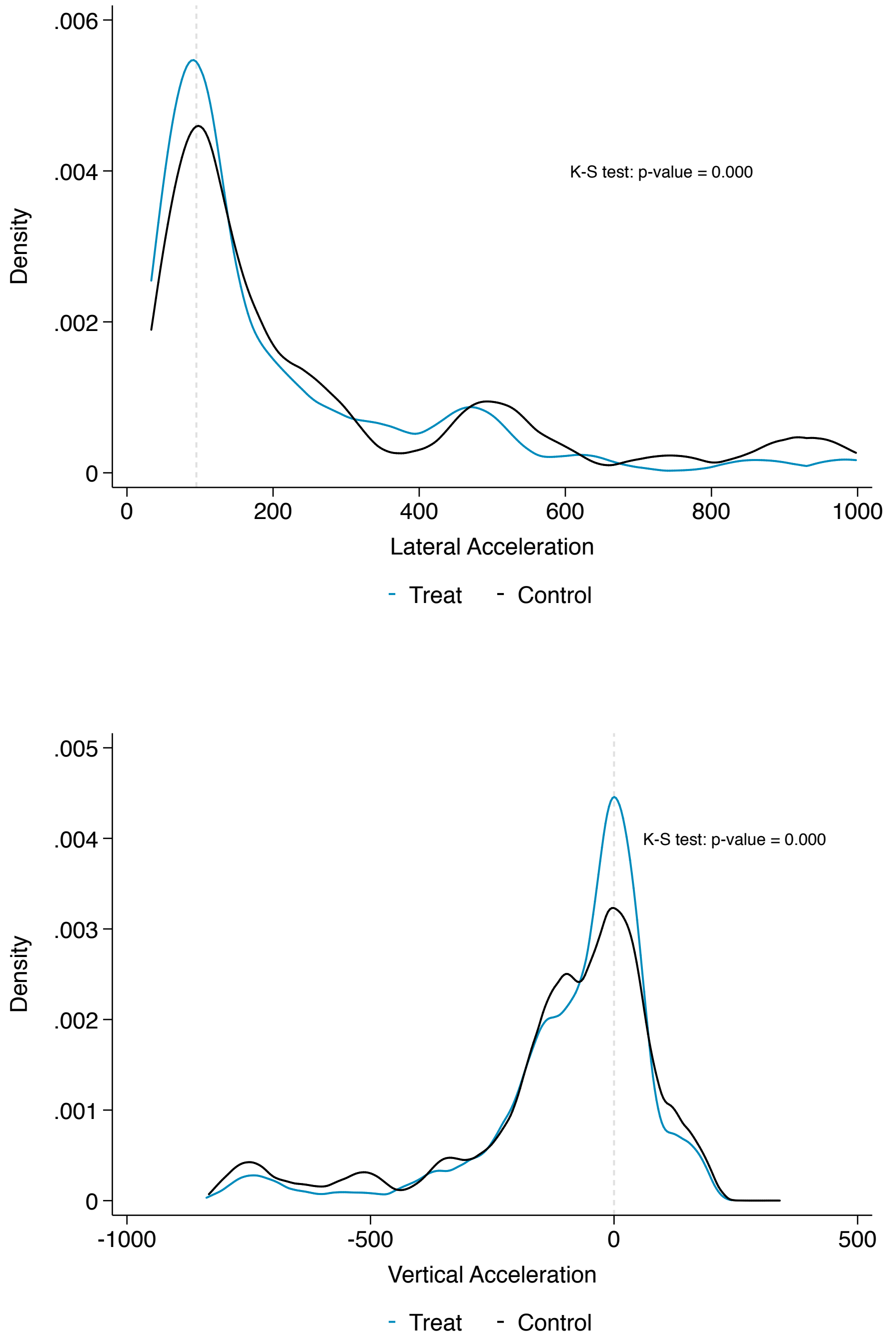
Figure 1.10: Prediction 2 → Damaging (Risky) Driving



Control Mean: 815.475
Last 3 months pooled: -371.62** (154.251)

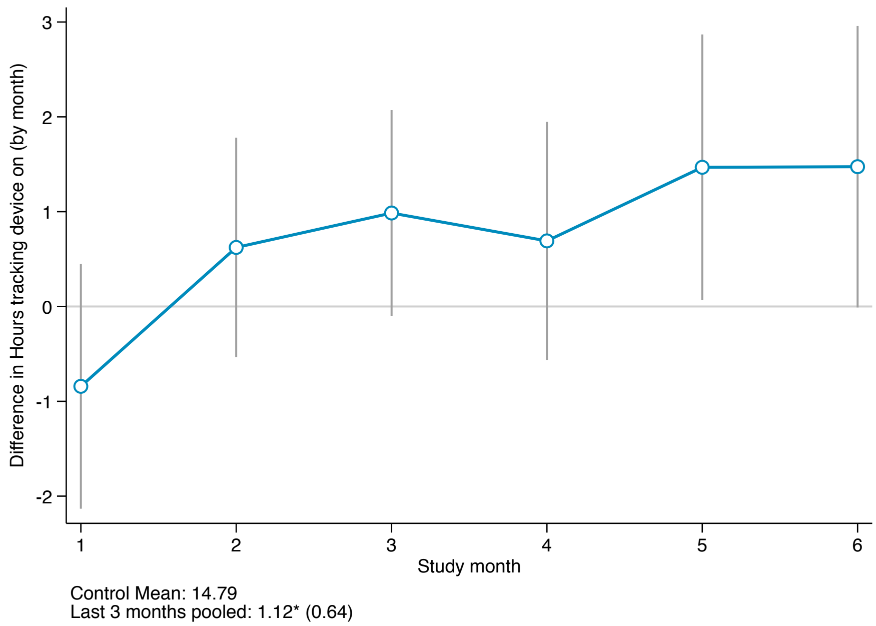
Notes: Each point on the graph is the difference between the average distance to designated route calculated for treatment and control owners in a particular month. “Distance to designated route" is the shortest distance between the GPS tracker data point and the line corresponding to the route the vehicle 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*.

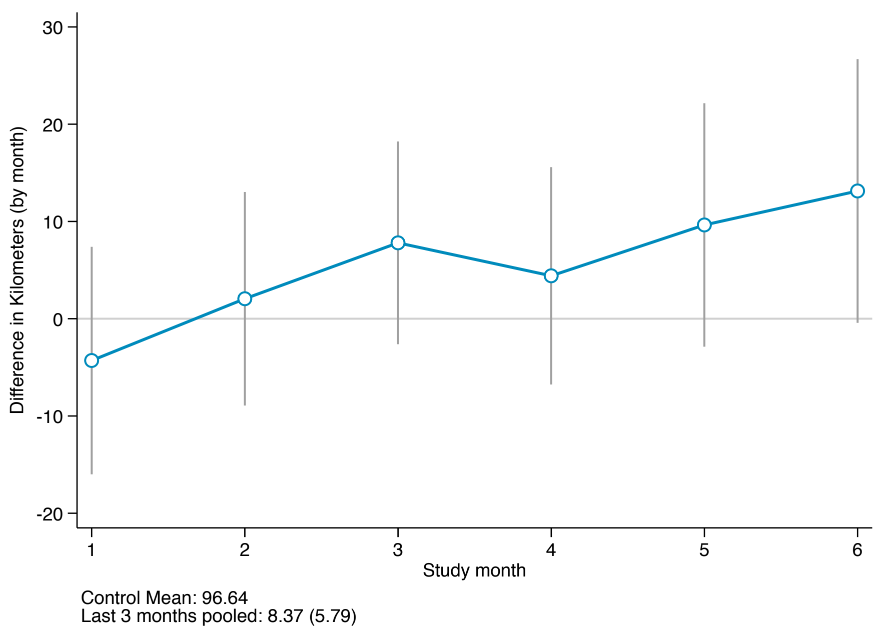
Figure 1.11: Prediction 2 → Damaging (Risky) Driving



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.

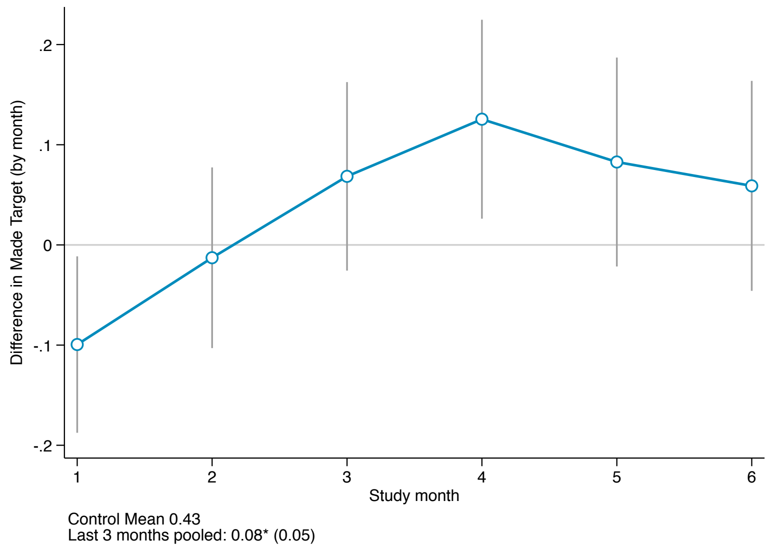
Figure 1.12: Prediction 3 → Effort

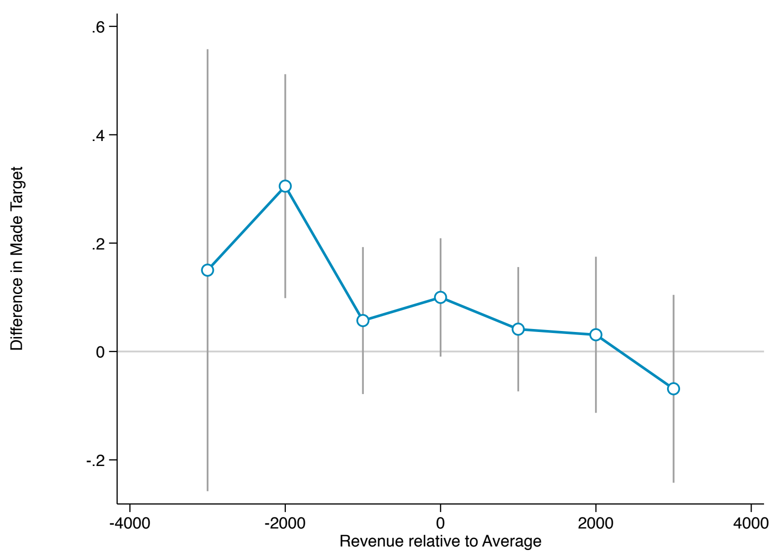




Notes: Each point on the graphs is the average difference between treatment and control in a particular month for a particular outcome. Note that because installations were rolled out over time, Month 1 (through 6) represent the first month (through the sixth month) since *installation*. The outcome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

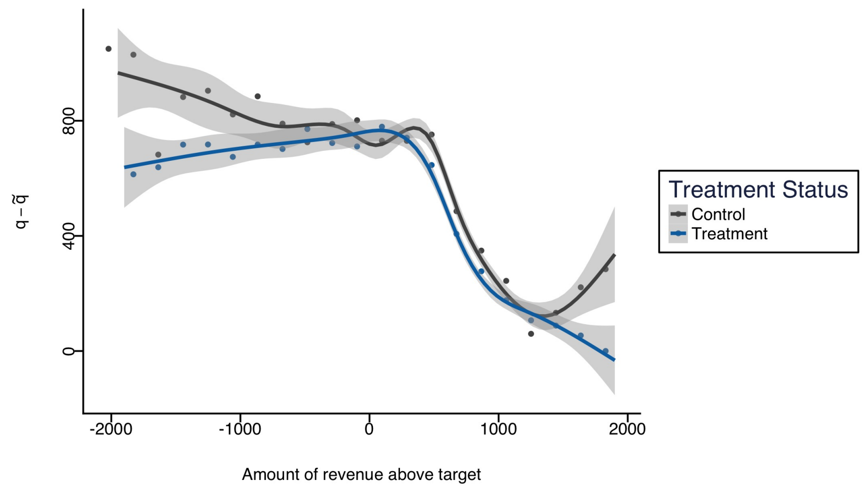
Figure 1.13: Prediction 5 → Achieving Target

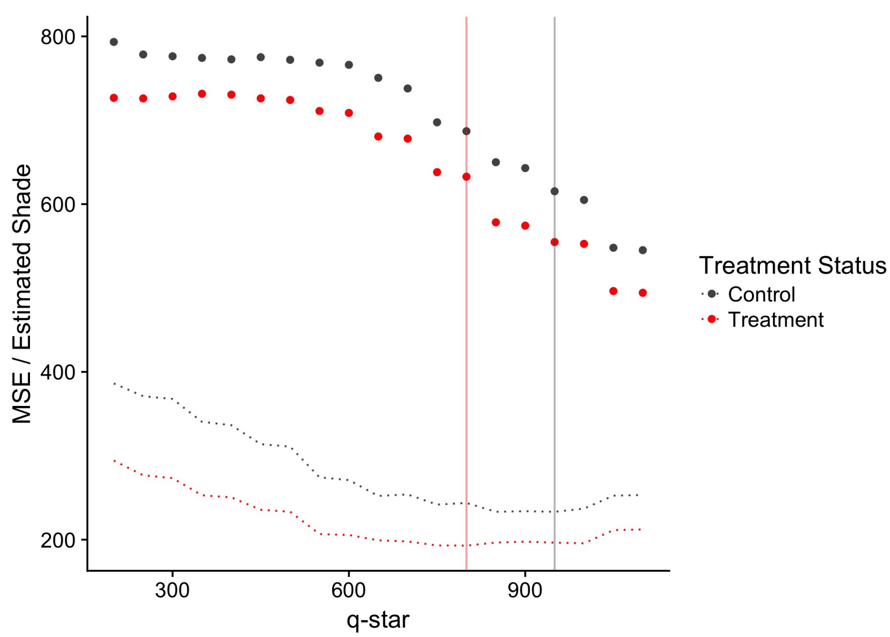




Notes: The outcome in these figures is whether or not the driver made the target (owner income = target). The data is collected from owner daily surveys. In the first panel, each point on the graph is the average difference in the probability of making the target between treatment and control in a particular 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.

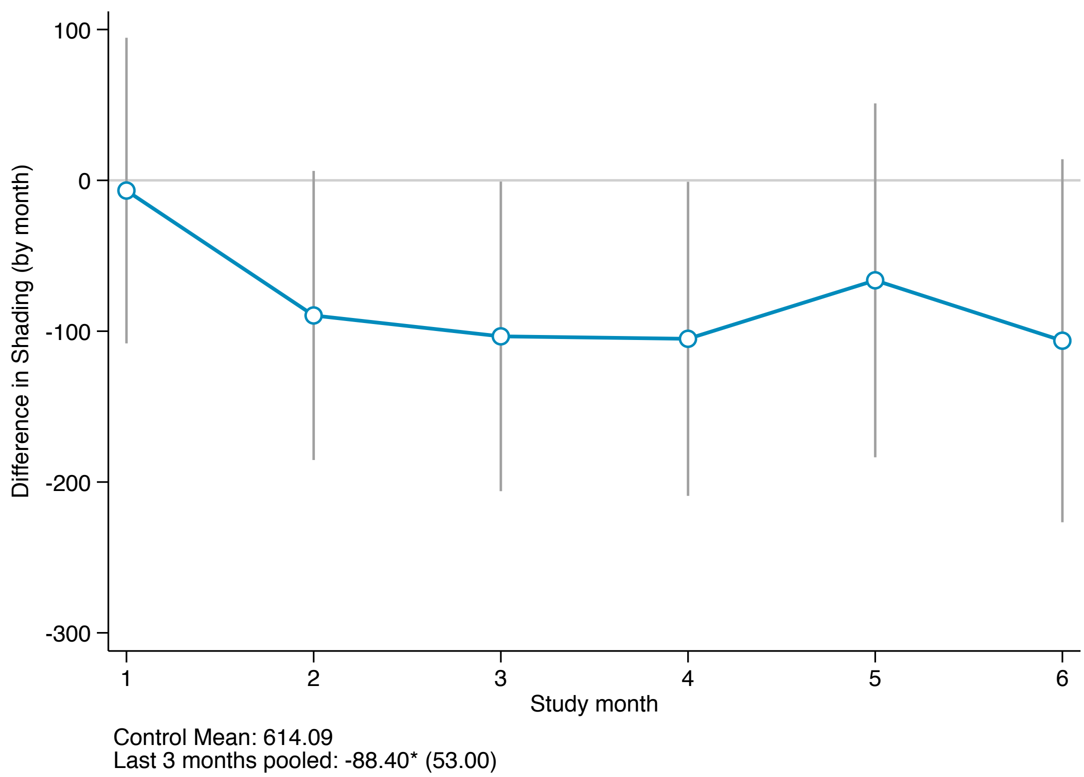
Figure 1.14: Prediction 4 → Less Under-reporting





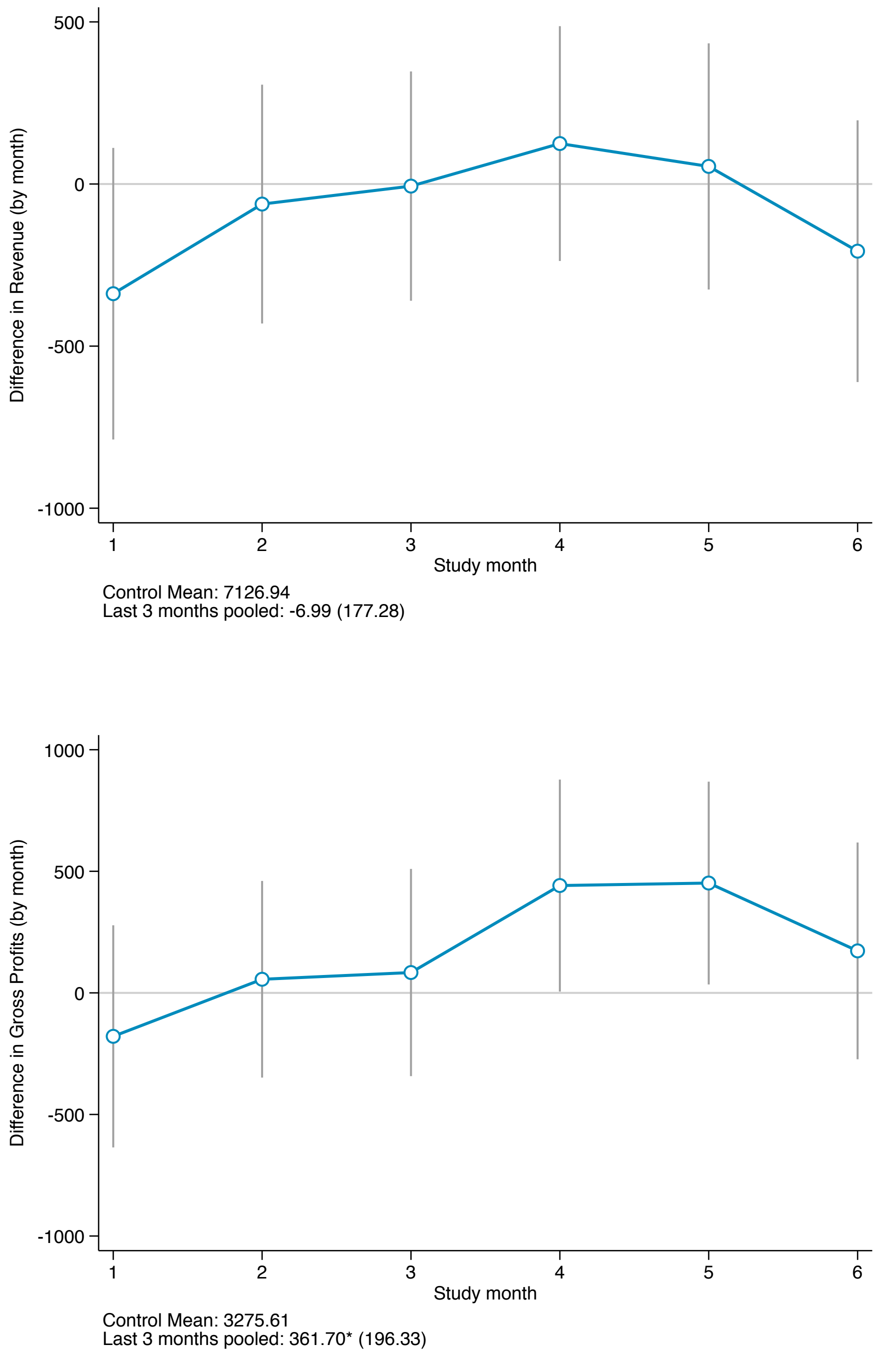
Notes: The outcome in these figures is the amount of revenue drivers under-report. The first panel reproduces Figure 1.5, but separates treatment drivers from control drivers. We also overlay a non-parametric smoothing function. The second panel imposes the model's step function and computes the average 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 1.15: Under-report



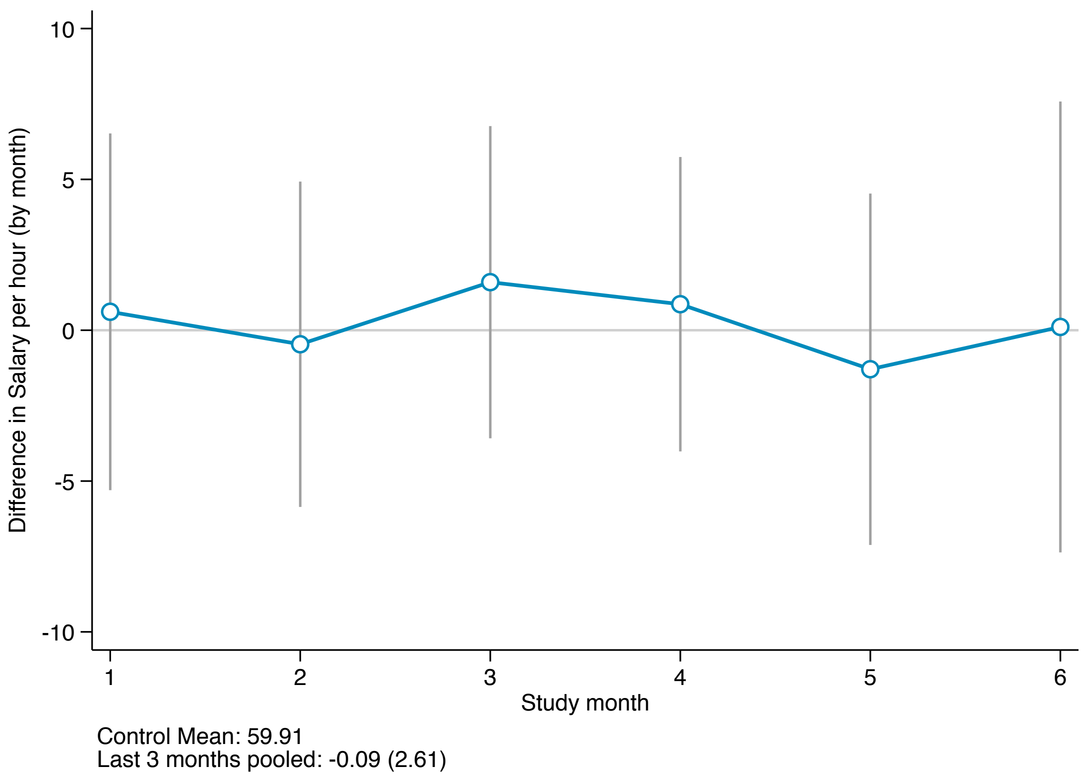
Notes: Each point on the graph is the difference between the average amount of revenue treatment and control drivers under-report in a particular month. The data was collected from owner and driver surveys - see Figure 1.5 for a description of how under-reported revenue was computed. Note that because installations were rolled out over time, Month 1 (through 6) represent the first month (through the sixth month) since *installation*.

Figure 1.16: Company Outcomes



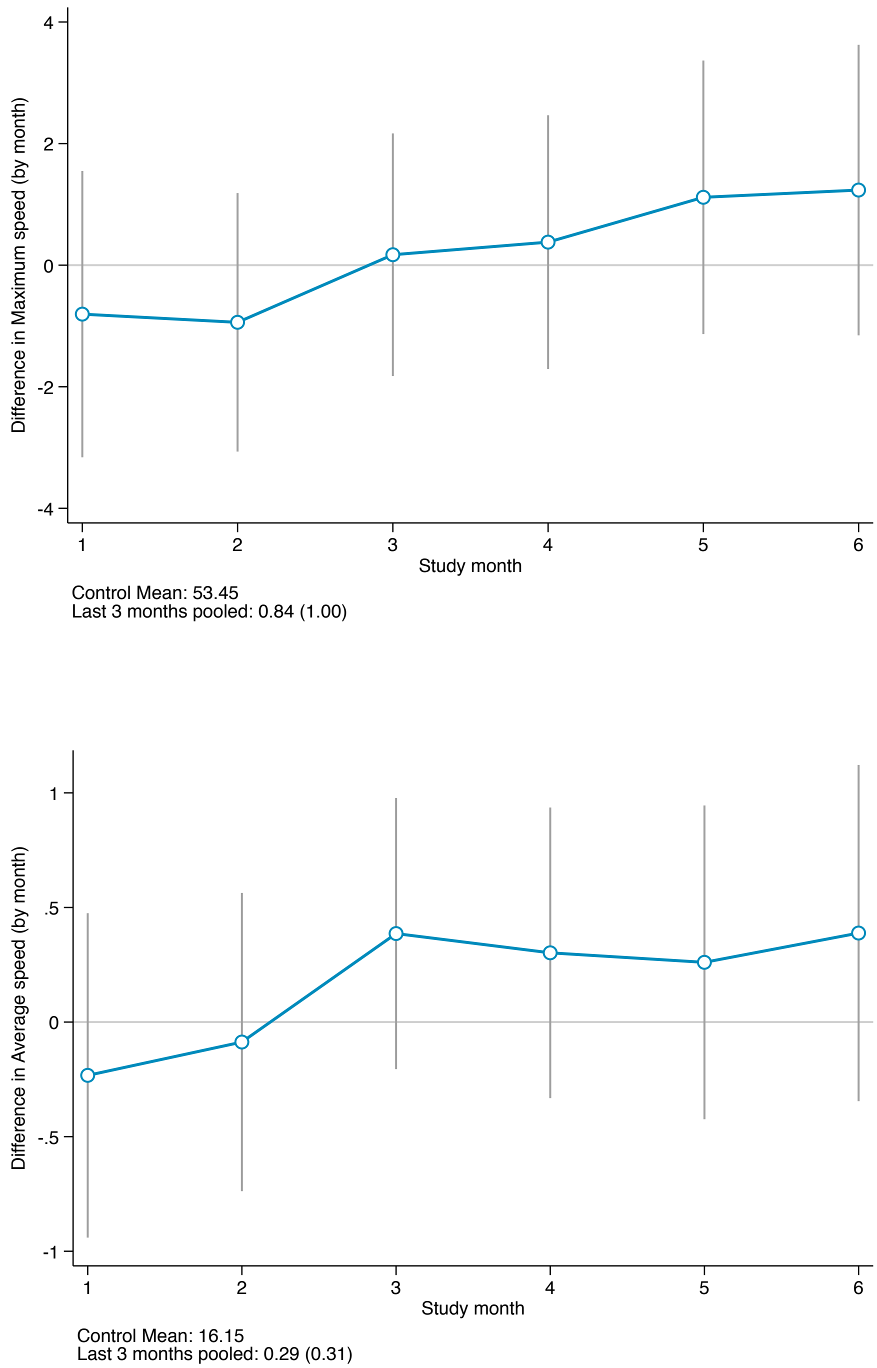
Notes: Each point on the graphs is the average difference between treatment and control in a particular month for a particular outcome. Note that because installations were rolled out over time, Month 1 (through 6) represent the first month (through the sixth month) since *installation*. The outcome 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 1.17: Salary per hour



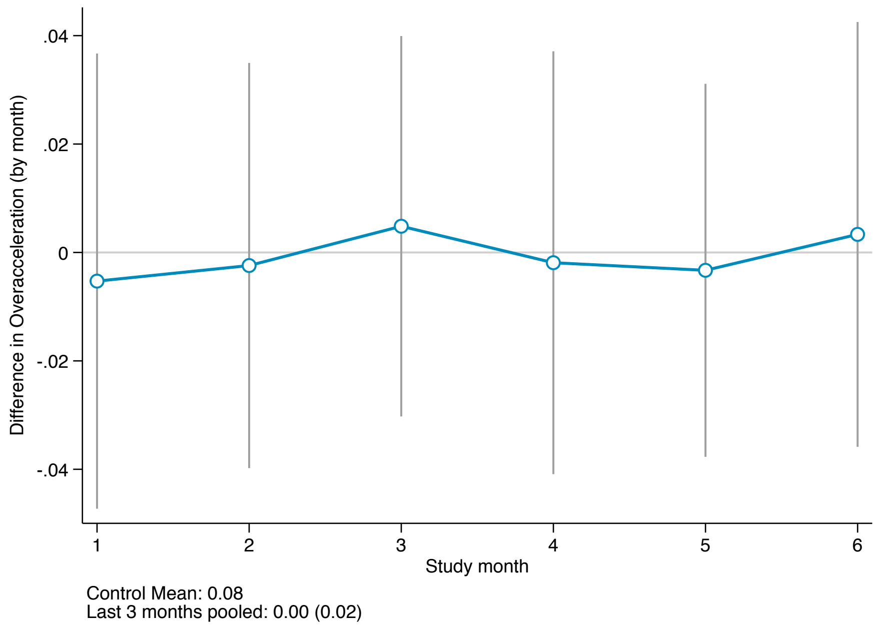
Notes: Each point on the graph is the difference in driver salary per hour between treatment and control drivers. The data was collected from daily driver surveys. Note that because installations were rolled out over time, Month 1 (through 6) represent the first month (through the sixth month) since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

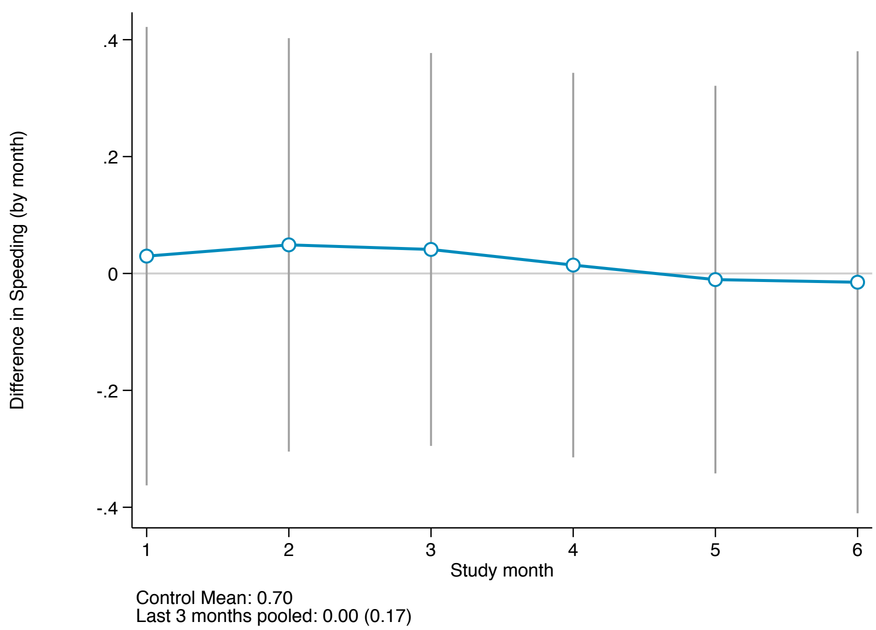
Figure 1.18: Speeding



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.

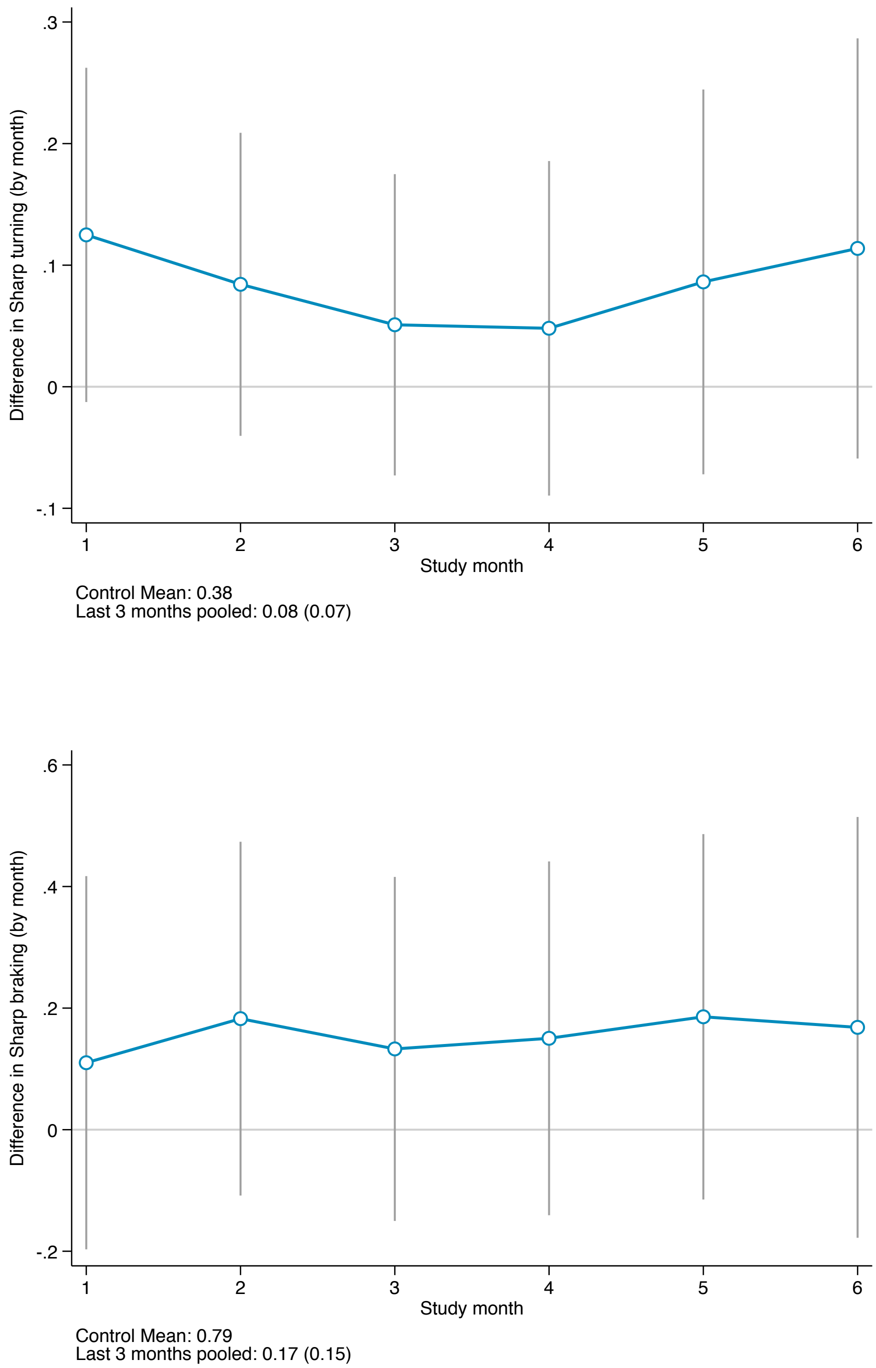
Figure 1.19: Over-acceleration and Over-speeding





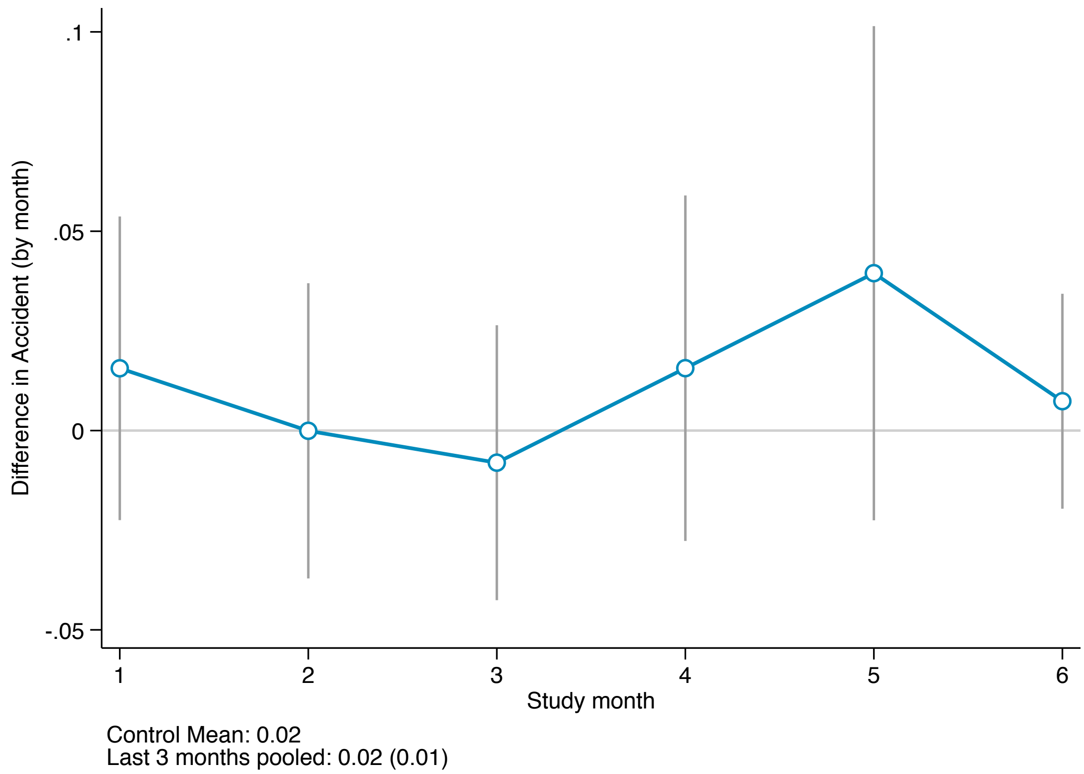
Notes: Each point on the graphs is the average difference between treatment and control in a particular month for a particular outcome. Note that because installations were rolled out over time, Month 1 (through 6) represent the first month (through the sixth month) since *installation*. The outcome 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 1.20: 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 1.21: 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*.

2 | Endogenous Information Sharing and the Gains from Using Network Information to Maximize Technology Adoption

Chapter abstract: Can agents in a social network be induced to obtain information from outside their peer groups? Using a field experiment in rural Bangladesh, we show that demonstration plots in agriculture — a technique where the first users of a new variety cultivate it in a side-by-side comparison with an existing variety — facilitate social learning by inducing conversations and information sharing outside of existing social networks. We compare these improvements in learning with those from seeding new technology with more central farmers in village social networks. The demonstration plots — when cultivated by randomly selected farmers — improve knowledge by just as much as seeding with more central farmers. Moreover, the demonstration plots only induce conversations and facilitate learning for farmers that were unconnected to entry points at baseline. Finally, we combine this diffusion experiment with an impact experiment to show that both demonstration plots and improved seeding transmit information to farmers that are less likely to benefit from the new innovation.

2.1 Introduction

People commonly rely on their peers for information. Research has established the existence of such peer effects across a variety of domains, ranging from learning in school to adoption of new innovations in poor countries.[1] Building on the importance of networks for transmitting knowledge, recent work has sought to answer

[1]A non-exhaustive subset of research in this area includes peer effects on academic performance (Sacerdote, 2001), purchases of financial assets (Bursztyn et al., 2014), adoption of improved sanitation in developing countries (Guiteras, Levinsohn, and Mobarak, 2015), the decision of whether to purchase crop insurance (Cai, de Janvry, and Sadoulet, 2015), and the adoption of agricultural technology (Foster and Rosenzweig, 1995; Munshi, 2004; Bandiera and Rasul, 2006; Conley and Udry, 2010).

the question as to who should be chosen as the initial recipients of information in order to make that information proliferate more extensively throughout the network? The proven methods for selecting these optimal entry points or "seeds" include collecting information on the full network in a community (Beaman et al., 2015) or asking a smaller number of individuals for the right people to seed information (Banerjee et al., 2018b). These approaches rely on the structure of the social network being roughly fixed and stable over time. Typically, agents pass information only amongst their connected peers.

Another strategy — and one that has received less attention in the literature — is to consider the exchange of information between individuals without existing social ties, and ask what can be done to encourage individuals to seek out information from outside their networks? Is it possible for the policymaker to intervene in a way that encourages people to seek information from off-network sources? Or, are information-sharing relationships sufficiently rigid to only respond to the selection of optimal seeds within the network? In this paper we focus on the efficacy of a commonly used technique in agricultural extension, namely the use of demonstration plots that showcase the features of a new seed variety relative to traditional varieties. We find evidence suggesting that this low-cost approach to encouraging learning can effectively substitute for the typically difficult task of finding optimal entry points in social networks.

We arrive at this conclusion using two related experiments spread across 256 villages in rural Bangladesh. The first experiment contrasts the two approaches detailed above to spreading information about a new rice variety called BRRI Dhan 56 (or BD56 for short).[2] We introduced the new variety to five farmers, referred to as "entry points" throughout the remainder of the paper, in a random subset of 192 villages. We then cross randomized villages across two treatments: (1) the selection criteria for entry-point farmers and (2) the demonstration method, aimed at encouraging other farmers from the village to seek information. The different treatment cells are described visually in Figure 2.1.

In terms of selection of entry points, we randomized villages across three different selection methods. In the first 64 villages we randomly selected farmers (as a benchmark). In the next 64 villages we relied on the local knowledge of agricultural extension officers (known as "sub-agricultural officers" or SAO's) to identify farmers who would be most effective at demonstrating the features of a new rice variety. In the remaining 64 villages, we ranked farmers according to farm size and selected the five largest farmers.[3] We opted for these selection mechanisms because they can

[2]Effective use of BD56 necessitates a large change in the production process. In short, the variety is ready for harvest a full month before the common variety being grown in the rainy season. As a result, the technology provides enough time for a third crop to be grown in between the rainy and dry-season rice crops. This large change in the cropping system — going from producing two rice crops to producing a third crop in between — enhances the potential for social learning about BD56 and its attributes.

[3]We found during piloting that large landholders often act as opinion leaders during focus group discussions about agriculture. Moreover, other farmers seemed to look at large farmers to learn about new technology. These casual observations led to the inclusion of the arm where farm size was used to select entry points.

be implemented at relatively low cost, which is attractive from a policy perspective. Indeed, neither approach requires an expensive survey nor detailed data collection efforts to characterize social networks.

In the second treatment, we randomized whether farmers were asked to set up demonstration plots or not. This treatment was cross-randomized with the selection of entry points detailed above. This means that 32 villages out of the 64 assigned to a particular "entry-point" selection arm were also assigned to demonstration plots. There the team assisted farmers in setting up "head-to-head demonstration plots", which involved cultivating BD56 alongside a counterfactual seed variety of the farmer's choosing. We provided two markers to make comparison across the two varieties more visible — one reading "BD56" and the other listing the name of the chosen alternative variety — to be placed in the two fields (see Figure 2.2 for an example).[4] We refer to this treatment as "demonstration plots" throughout the remainder of the paper. The demonstration plots indicate to other farmers that the entry point is comparing the attributes of the new variety to a known one. This serves as a mechanism to focus attention of other farmers on existence of a new learning opportunity.[5] In the remaining 32 villages that were not assigned to demonstration plots, the entry points used BD56 on one of their plots and were provided with a single "BD56" marker. It follows that the demonstration plots must do more than broadcast information about the existence of BD56 because farmers in these comparison villages also labelled their field with a marker.

These two approaches to increasing knowledge transmission rely on very different assumptions about how information diffuses in networks. The improved selection of entry points, using either large farmers or those hand picked by SAO's, seeks to test scalable approaches to finding the right influential farmers in the network. Importantly, the notion of optimal seeds is designed to exploit the network as it exists at baseline, and does not consider that the intervention itself may change the network. In contrast, the demonstration plots are meant to spark interest and potentially induce communication beyond existing network links. The demonstration plots can do so by either 1) capturing attention and thus increasing the demand for information or 2) providing a more precise signal to entry points and therefore increasing their supply of information to others.

This experiment delivers four main results. First, and using only baseline information, we verify that the large and SAO-selected entry points are far more central in the village networks than random seeding, and are thus well positioned to spread information to a larger number of farmers. At baseline our survey teams visited all farming households in each village — making a total of almost 22,000 visits — and posed the question "which farmers in this village do you regularly speak to about rice farming?"[6] Using these data, we observe that the average entry point in the

[4]This method of demonstration plots is commonly used, especially by private sector seed companies seeking to promote and demonstrate the attributes of their new seed varieties.

[5]We didn't provide any further assistance (such as inputs or advice) with the actual cultivation of the two plots. This was purposeful to lessen the cost of the demonstration plots treatment, and make the approach easily scalable.

[6]Chandrasekhar and Lewis (2016) show that measures of network centrality are misleading when

random villages is connected to 4.6 other farmers. This increases sharply to 8.2 and 9.1 connections for entry points in SAO and large farmer villages, respectively. Similarly, the eigenvector centrality of entry points increases by 47 percent under SAO selection and 80 percent with large-farmer selection. Based solely on the network measured at baseline, both treatments would therefore be expected to increase the spread of knowledge.

Despite these noticeable differences in network centrality, our second result is that demonstration plots with random farmers create just as much additional knowledge as entering with large and SAO-selected farmers. After the harvest of BD56 — and sowing of the next crop — we conducted a survey with 10 random farmers in each village. In addition to awareness about existence of BD56 and its basic attributes, we also collected information on reported conversations about BD56. Using large farmers as entry points increases awareness by 7.4 percentage points or 12.3 percent in villages without demonstrations. Similarly, entry points selected by extension agents increase awareness by 6.7 percentage points (11.2 percent) in villages without demonstrations. However, providing seeds to more central farmers provides no additional benefits for knowledge diffusion when we introduce demonstration plots.

At the same time, setting up demonstration plots increases awareness by 7.2 percentage points (12 percent) with random entry points. Noticeably, the impact of demonstration plots under random selection is of the same magnitude as the impacts of improved entry-point selection in the absence of demonstration plots. The results remain similar when looking at the number of reported conversations about BD56: demonstration plots with random farmers induce conversation by about the same amount as introducing seeds with more central farmers. Given the ease of setting up demonstrations, i.e. simply placing two signs in adjacent fields, our experiment offers insight into how interventions that attract attention to the fact that there is something to be learned can substitute for seeding with more central entry points in networks.

Third, we then move on to investigate why demonstration plots are effective. The results are consistent with the idea that demonstration plots cause farmers to learn from people outside their network. Focusing on the 64 villages with random entry-point selection, we observe strong peer effects on knowledge. Farmers that were randomly connected to an additional entry point are 7.7 percentage points more likely to know about BD56. This average effect differs meaningfully between villages with and without demonstration plots: an additional connection with an entry point has no effect in demonstration villages, but it increases awareness by 13.5 percentage points (22 percent) in non-demonstration villages. In other words, demonstration plots lead to information exchange outside of baseline networks and therefore greater transmission of knowledge. In doing so, demonstration plots entirely eliminate peer effects. The same pattern of results again appears when considering conversations about the technology.

estimated using only a sample of nodes within the network. Our approach of fully characterizing the network by sampling each household in the village alleviates this concern.

Also consistent with this network interpretation, we find that demonstration plots were most effective for farmers that are least connected in the baseline information network — where connectivity is measured by eigenvector centrality. A plausible explanation of this result is that the demonstration plots induced these less connected farmers to endogenously seek information about BD56.

Finally, we show that their network centralities at least partly explain why large and SAO-selected entry points lead to better information diffusion. The effects of large and SAO selection on knowledge decrease by 43 and 31 percent, respectively, when conditioning on the average degree of entry points. In addition, average degree centrality of entry points is itself positively correlated with BD56 awareness. Conditioning on average degree is not a perfect test — since the particular network measure that should "knock out" the treatment effects depends on the specific model of diffusion.[7] Nonetheless, the result is consistent with the idea that diffusion via network links partly explains why the entry-point treatments were successful.

While we mainly consider effects on knowledge transmission, we also observe uptake when BD56 was made available for sale at subsidized prices. The results on seed adoption are noisier, but qualitatively consistent with our observations on knowledge.

Our second experiment allows us to further test whether demonstration plots (or alternative seeding strategies) are more likely to deliver information *to the farmers most likely to benefit from the new technology*. Returning to Figure 2.1, we also randomly selected 64 control villages where we provided a new long-duration rice seed to a set of farmers identified using the same criteria as in the 192 BD56 villages. This experiment allows us to characterize the impact of short-duration rice on agricultural practice and profits. The main benefit of the short-duration seed is the ability to grow a third crop in between the rainy and dry-season rice crops. Using the machine learning methods developed in Chernozhukov et al. (2018), we estimate a mapping between baseline covariates and the treatment effect of BD56 on the number of crops grown. This heterogeneity index serves as a prediction of which farmers are most likely to benefit from BD56 by increasing cropping intensity.

Using this heterogeneity index with the 192 villages in our first diffusion experiment, we find suggestive evidence that both demonstration plots and our selection treatments increase knowledge and conversations only for farmers that have *below-median* expected treatment effects of BD56 on the number of crops grown. In other words, intervening to increase knowledge diffusion in networks may only be effective for farmers with lower expected benefits from adopting an innovation.[8] These findings point to an important consideration for research that studies alternative mechanisms for increasing diffusion of products that have heterogeneous benefits.

[7]Degree is a suitable measure if we think of diffusion models with few periods, the probability of passing information to connected friends is high, and the farmers with the largest degrees are sufficiently spread out in the network.

[8]An alternative explanation is that the diffusion treatments succeeded in informing farmers that were the least likely to benefit from BD56, and therefore prevented their adoption. This explanation is less consistent with the positive (but noisy) point estimates we observe when estimating the effects of the diffusion treatments on seed adoption.

Combining diffusion experiments with standard impact evaluations allows the researcher to estimate treatment effect heterogeneity and use that to measure which diffusion strategies reach the people most likely to benefit from an innovation, even without a strong prior on which observables drive the heterogeneity.[9]

We then go on to show how a simple diffusion model, when amended to allow for formation of new links with entry points, explains our pattern of results. In the model, farmers can either become informed by receiving information flowing through the structure of existing links, or by actively communicating with entry points. Farmers in the worst position to learn from the network, i.e. those that are the least connected to entry points, benefit from having the opportunity to engage directly with entry points. The demonstration plots treatment appears to deliver these benefits by inducing unconnected and more isolated farmers to learn directly from entry points.

Turning to literature, theoretical work has considered information transmission when network communication is endogenous (Acemoglu, Bimpikis, and Ozdaglar, 2014; Calvó-Armengol, Martí, and Prat, 2015), but empirical work in this area is scarce. Mobius, Phan, and Szeidl (2015) show that having conversations is correlated with possessing the correct information in a field experiment amongst college students, but that frictions still exist in the diffusion of information. Chandrasekhar, Golub, and Yang (2016) make an important contribution by considering one particular friction arising with endogenous social learning: the need for information can reveal low skill, therefore creating a stigma effect that represents part of the costs of information seeking. Banerjee et al. (2018a) show evidence consistent with this same idea during India's recent demonetization. In particular, villagers were less willing to seek information about demonetization rules when everybody knew that information was provided widely and thus seeking information signals an inability to process one's own information. Our experiment delivers insights on the potential to intervene to induce communication, overcome some of these frictions, and therefore facilitate social learning.

Looking for ways to induce communication and learning builds on an active literature that treats networks as fixed and asks how to find the optimal entry points within these networks. These studies include demonstration of agricultural inputs (Beaman et al., 2015; Beaman and Dillon, 2017), diffusion of information about microfinance (Banerjee et al., 2013), diffusion of health products (Kim et al., 2015), and information on how to capitalize on a financial opportunity or the uptake of vaccines (Banerjee et al., 2018b).[10] There are a couple of limitations with this approach. First, it can be difficult to identify the most central entry points in a social

[9]Rigol, Hussam, and Roth (2017) is the closest example where the authors use machine learning methods to estimate treatment effect heterogeneity for microfinance in India. They then show that using community information on the returns to microfinance is more effective than the machine learning algorithm when applied to the set of observables in their baseline data.

[10]In addition to selection, other work has looked at different ways of making entry points communicate more, including compensation (BenYishay and Mobarak, 2015) and training (Kondylis, Mueller, and Zhu, 2017).

network without fully characterizing the network, which may be cost prohibitive.[11] Second, the efficacy of finding better entry points for seeding information likely depends on the underlying structure of the social network or the specific model of diffusion (Centola, 2010; Valente, 2012; Golub and Jackson, 2012). Akbarpour, Malladi, and Saberi (2018) show that in many network structures the benefits from seeding information with a slightly larger number of agents outweigh the benefits of identifying the most central individuals. Our experiment considers scalable methods of entry-point selection as a benchmark and shows that their benefits can also be obtained by interventions that trigger social learning.

Focusing on agricultural development and policy, learning frictions is one of the frequently proposed reason why farmers do not adopt new technology.[12] Agricultural extension is expected to serve as the policy tool to improve learning. In practice, the standard method of agricultural extension attempts to leverage social learning by seeding information with contact farmers (entry points) and relying on information diffusion through social networks (Birkhaeuser, Evenson, and Feder, 1991; Anderson and Feder, 2007; Kondylis, Mueller, and Zhu, 2017). Our experiment provides evidence on how demonstration plots can offer a substitute for the policymaker when improved selection of these contact farmers is difficult to implement.

The remainder of this paper is organized as follows. Section 3.3 discusses the design and implementation of the experiment, data collection, and basic characteristics of the sample. Section 3.4 presents each of our individual results, focusing on how selection and demonstration plots influence learning, and on understanding what drives the effectiveness of the two approaches. Section 2.4 outlines a simple theoretical framework that is consistent with our results. We then provide an overview and discuss implications of the findings in the final section.

2.2 Overview of the Experiment

In this section we review the details of the randomized control trial: the sampling strategy, the experimental design, and the data collection activities. We conducted the study in 11 sub-districts (upazilas) scattered across 3 districts of Rajshahi division, consulting with the Department of Agricultural Extension to identify upazilas that were suitable for the rice variety being introduced.[13] We randomly selected 20% of the villages with not more than 150 households, resulting in a final sample of 256 villages. This includes the 192 villages that received the new BD56 rice variety in the diffusion experiment, as well as 64 control villages that received the longer duration rice variety for the impact evaluation experiment. This village-level

[11] Banerjee et al. (2018b) show how to overcome this difficulty by asking a sample of villagers who are the important people for diffusing information.

[12] In addition to learning, numerous studies highlight a wide range of explanations, including behavioral biases, profitability, and risk (Duflo, Kremer, and Robinson, 2011; Suri, 2011; Karlan et al., 2014; Emerick et al., 2016; Cole, Giné, and Vickery, 2017).

[13] See Figure B.1 for the location of the 11 upazilas included in the study.

randomization was stratified by upazila.

BD56 has two key features. First, it requires less water, allowing farmers to save on supplemental irrigation fees and preserving groundwater resources. Second, it matures approximately 25 days earlier than other varieties commonly planted in the area, providing farmers with the option of harvesting and selling an additional crop between the two rice seasons (Aman and Boro rice seasons last from June/July to October/November and December/January to April/May, respectively). The 64 control villages received a long duration rice variety called BRRI Dhan 51 (BD51) which was chosen due to its similarity to the most popular variety at baseline.[14]

2.2.1 Experimental Design

Figure 2.1 summarizes the experimental design for the diffusion and impact evaluation experiments detailed below. The design for the diffusion experiment consists in the cross-randomization of two dimensions of treatment across 192 villages: the method used to select the five seed recipients, and the implementation of "demonstration plots" by these entry-point farmers.

For the first dimension, the villages were subdivided into three groups of 64 villages each. The types of farmers we selected to receive the seeds (the "entry points") differed based on the specific treatment arm that village was assigned to. In the first group of 64 villages, the seeds were distributed to five farmers selected at random. In the second group, we ranked farmers by landholding sizes and distributed 5kg minikits of BD56 seeds to the top five farmers. In the third group of 64 villages, we asked the Sub-Agricultural Officer (SAO) to identify five farmers in the village that would be effective at demonstrating the new variety and provided BD56 seeds to them.

Within each group of 64 villages, we then selected 32 villages to receive additional assistance setting up demonstration plots. We included this additional dimension in the experiment to determine how this approach to boosting diffusion rates compared with the alternative of selecting entry points that rely on the underlying social networks to spread information. To this end, we asked farmers to select a counterfactual variety that they would like to plant beside the new BD56 seeds we distributed. We then provided two sticks to farmers: one with the name of the new variety (BD56), and another with the name of the variety they had selected to plant beside it. We asked that farmers keep these two signs in their fields throughout the cropping season to showcase the performance of BD56 relative to the counterfactual they selected. Farmers in the remaining 32 "non-demonstration" villages received a single sign for their BD56 plot. The provision of the single sign ensures that any effect we detect in the demonstration plot villages with two signs

[14] BD51 is released as "Swarna-Sub1" in India and several other countries. Emerick et al. (2016) show that Swarna-Sub1 is similar to Swarna, besides Swarna-Sub1 being more flood tolerant. However, our sample is not a flood prone area. We introduced Swarna-Sub1 in the control villages because Swarna is not officially released in Bangladesh (and thus not available for sale) despite being the most popular variety in our sample at baseline. More precisely, 77 percent of farmers in the village census reported growing Swarna at baseline.

goes beyond the attention effect of placing one sign in the field.

For the impact evaluation experiment, we selected up to 15 “counterfactual" farmers to receive equal sized amounts of BD51 seed in each of the 64 control villages. This included five farmers with the largest landholdings, five farmers selected by the SAO, and five farmers selected at random. These sets overlapped in some cases and therefore the number of farmers per control village is occasionally less than 15. Identifying these particular farmers in control villages was necessary in order to compare how the entry points we identified in the treatment groups cultivated BD56 relative to the longer duration counterfactual we distributed in the control villages. In the rest of the paper we refer to these counterfactual farmers as "entry-points" as well.

2.2.2 Timeline and Data Collection

Census with network information of each village
Figure 2.3 presents a complete timeline of the study. We began by performing a complete census of each village in March 2016. Our field team surveyed 21,926 households over the period of three months. Villages have between 14 and 184 households, 86 on average. We administered a short questionnaire to each member of the village, asking about their agricultural production (landholdings size, fertilizer use, production, and varieties sown), and their social networks (the name of the person they considered to be the best farmer, and the names of up to 10 farmers they turn to for advice on rice cultivation). We use these data to identify the largest farmers in each village, select the random entry-points, compute network statistics, and to forecast heterogeneous impacts of BD56 as a function of observable covariates.

Seed distribution
The distribution of 5kg minikits to the farmers selected to be entry-points in all 256 villages took place in early June 2016, in time for the Aman season. A total of 1,795 farmers were reached, achieving a response rate of 99 percent.[15] During these visits, the field team carefully explained the features of the seeds being distributed, and provided farmers with calendars to record the dates crops were sown, harvested, irrigated and applied with inputs. We also supplied farmers in the treatment villages with sticks and cards to place in their fields as a way of demonstrating to other farmers the variety they were planting. We briefly visited farmers 6 weeks later to make sure that we answered any remaining questions they had about the seeds, and to verify that the sticks were properly displayed in the fields.

Information diffusion survey
In April 2017 we visited all treatment villages, and randomly selected 10 additional farmers to conduct a small survey to determine if they had any knowledge of the

[15]The total number of entry points is five per BD56 village, or 960, and up to 15 counterfactual farmers in the BD51 control villages. All 960 BD56 entry-points were reached, as well as 835 counterfactual farmers.

BD56 seeds that were distributed to the entry points 9 months earlier. Specifically we asked whether they had heard about the new variety, which farmers they spoke to, and whether they could articulate some of the variety's key features. We use this information to assess how the diffusion of BD56 knowledge differs based on the village's treatment status.

Seed sale

In early June 2017, we visited each treatment village in our sample to sell 5kg and 2kg bags of BD56 seeds at subsidized prices. The field team called a select sample of farmers in each village to inform them about the date and time of the seed sale. The sample included the original minikit recipients, and the ten randomly selected farmers who were surveyed about their BD56 knowledge 2 months prior. The field team travelled to each village on the pre-determined date, and set-up their truck in the middle of the village (often at the local market) with a large sign indicating that the new BD56 seeds were available for purchase and that the main benefit of the seed is that its shorter duration allows for an additional crop to be grown. They recorded each sale that was made in a tablet. While we did not record the identity of the buyer, the survey provides a measure of BD56's diffusion within the village. Unfortunately, we ran out of seeds before traveling to all of the villages, and hence this information is only available for 168 of the 192 villages.

Agricultural surveys of entry-points.

In addition to the main sources of data described above, we conducted a series of agricultural surveys with the entry-points in all 256 villages. The goal of these surveys was to rigorously establish the impact of BD56 on agronomic practices, cropping intensity, and annual income. A baseline survey was administered at the time of the minikit distribution in June 2016. Important outcomes of interest included area cultivated, plot-level information on crops sown, inputs, and production volumes. We asked farmers to provide this information for 3 of their plots, which we selected randomly when farmers had more than 3.

This survey was followed by three other rounds in order to fully characterize annual production. In January-February 2017, we collected detailed information about the recently harvested Aman rice crop, and recorded whether farmers were planting a second post-Aman (known as the Rabi season) crop (midline 1). The survey asked specifically about seed variety choice, planting methods, and production at the plot level. In April 2017 we collected additional information about the Rabi crop, and the final Boro rice crop (midline 2). Finally, in August 2017, we asked farmers about their crop production levels during the Boro rice season, as well as their crop choice for the 2017 Aman rice season in order to gauge their propensity to re-plant the new rice variety (BD56) that we had offered them the previous year (endline). All surveys successfully reached the initial 1,795 farmers, except for two farmers missing in the April survey.

2.2.3 Baseline Characteristics from Census Data

Table B.1 presents summary statistics from the household census and verifies randomization balance.[16] Our sample consists primarily of farmers cultivating long-duration rice varieties (only 1.17% of treatment farmers and 2.4% of control farmers planted short-duration varieties in the 2015 crop cycle). While approximately 35% of farmers only grow rice throughout the season, a non-negligible share grow a Rabi crop (including wheat, potato, pulses, onion, and garlic). Finally, farms in the sample are small: average area sown with Aman rice (the main crop) is approximately 1.33 acres.

We asked farmers to provide the names of up to 10 farmers they talked to about rice farming during last Aman season. We define two farmers as being connected or being peers if either name the other among the farmers they talk to. We use this information to create various centrality statistics including degree centrality (the number of connections a person has), eigenvector centrality and betweenness centrality (the number of times a node acts as a bridge along the shortest path between two other nodes). Figure 2.4 displays the distribution of these three measures for the entire sample of households interviewed during the census. Farmers have on average 4 connections with whom they talk to about rice cultivation, though we see a strong right tail with some farmers having up to 26 connections. The distributions of eigenvector and betweenness centrality display similar patterns with long right tails: while most farmers have a few connections some have a disproportionately high number. The long right tails in the distributions provides some initial evidence that the agricultural information networks do have some highly central farmers that in theory would serve as more effective entry points.

2.3 Results

We now provide our main results — starting with differences between entry points in their social network status. We then seek to understand how entry points demonstrated the attributes of BD56, by comparing them to the farmers growing long-duration rice in the control villages. Building on these results, we focus on how the demonstration plots affect awareness and how these effects compare to the selection treatments. We argue that the efficacy of the demonstration plots is being driven by their ability to induce new communication links. In contrast, additional evidence suggests that the standard diffusion model in networks explains much of the effectiveness of our selection treatments. Finally, we combine our different data sources to test whether the treatments deliver information to farmers most likely to capture BD56's main benefit by increasing cropping intensity.

[16] Table B.2 further shows randomization balance for the sample of entry points, for the impact evaluation experiment.

2.3.1 The network centrality of entry points

Recent literature has tested various mechanisms for identifying the most central nodes in social networks. These mechanisms include both direct elicitation of the entire social network (Beaman et al., 2015; Kim et al., 2015) or trying to infer centrality by asking a sample of individuals who is suitable for diffusing information (Banerjee et al., 2018b). Compared to eliciting the entire network, our two methods of selecting entry points are generally less demanding in terms of data; entering with large farmers requires only administrative data on farm sizes and SAO-based selection requires only a short interview with an agricultural extension agent. Yet, despite their ease of implementation, there is no guarantee that either of these two methods will deliver the entry points that are theoretically optimal for diffusion. We first check this using our social network survey.

The main measures of network centrality all differ noticeably between random entry points, those identified by SAO's, and the largest five farmers in the village. Table 2.1 displays average characteristics for the 960 entry points across the 192 BD56 treatment villages. The average farmer that was selected randomly has about 4.6 connections with other farmers in the village. The entry points selected by SAO's have an average of 8.14 connections and the five largest farmers in each village have an average degree of 9.04. Eigenvector centrality of entry points increases by 47 percent when selected by SAO's and 80 percent when chosen as the five largest farmers in the village. Banerjee et al. (2013) introduce diffusion centrality of farmer i as a measure of the expected number of times that farmers obtain a piece of information that was introduced with i.[17] Comparing to random entry points, average diffusion centrality of SAO-selected entry points is higher by 0.64 standard deviations and the diffusion centrality of large entry points increases by 0.87 standard deviations. Figure B.2 in the appendix shows the cumulative distribution functions of the different network centrality measures across the selection treatments. Most importantly, the increases in centrality for large and SAO farmers occur throughout the distributions.

These findings deliver an important verification that our experiment compares demonstration plots to meaningful methods of selecting entry points. Put differently, the SAO and large farmers in theory would be suitable for demonstrating technology in order to spread awareness. The high centrality of large and SAO farmers — relative to random entry points — is comparable to other mechanisms to selecting entry points that have been tested in the literature. For instance, Banerjee et al. (2018b) find that the median diffusion centrality of people identified by other villagers as suitable for spreading information is larger than that of other villagers by around 0.5 to 1 standard deviations. We find that relative to random targeting, larger farmer targeting would increase median diffusion centrality by around 0.51 standard deviations (Figure B.2).

[17]This measure requires as parameters the number of periods for the diffusion process and the probability that an informed agent passes information to their social connections. We set the number of periods to 5 and the information passing probability to 0.5. We also normalize the measure by subtracting the village-specific mean and dividing by the village-specific standard deviation.

Besides network centrality, Table 2.1 also shows how several other observable characteristics vary across the different types of entry points. There are two notable observations. First, SAO-selected entry points tend to be larger farmers: farm size increases by about 5.9 bigha (3 bigha = 1 acre) for SAO entry points relative to random farmers. We show in Table B.3 that controlling for farm size reduces substantially the gaps in network centralities between SAO-selected and random farmers. Put differently, extension officers could use knowledge of farm size in selecting entry points, and this explains part of the reason why SAO-selected entry points are more influential in networks. The ability of extension agents to select influential entry points contrasts with Beaman et al. (2015) who find that Malawian extension agents possess little information on the optimal entry points within social networks. The sharp correlation between farm size — an easily observable characteristic — and network centrality offers one possible explanation for the greater ability of extension agents in our sample. This phenomenon is visually evident in Figure 2.5 where we show the network structures for 6 randomly selected villages with either SAO or large-farmer selection. All three of the large-farmer villages in the top panel of the figure have at least one relatively larger farmer that is well connected in the network. Focusing on the SAO-selection villages in the bottom panel, both village 99 and 224 have medium or larger size farmers that are central in the network and were selected as entry points by the SAO.

Second, our door-to-door census asked each household to list the "best farmer" in the village. A random farmer is only named about 0.79 times while SAO entry points are named 5.27 times, and the five largest farmers in each village 6.4 times. These numbers further suggest that our selection treatments identify entry points that are both better networked and that other farmers consider to be knowledgeable about agriculture.[18]

2.3.2 How do entry points demonstrate the technology?

In this section, we use the impact evaluation experiment to explore how BD56 affect farmers' cultivation practices and profits.

Take up

During the first midline survey we asked farmers whether they planted the seeds provided to them. We do not find any evidence of differential adoption rates (Table B.4). Focusing exclusively on treatment villages, we further investigate whether take up varies across different types of entry points. We find that adoption rates remain fairly consistent across treatment arms, albeit slightly higher among large farmers assigned to the demonstration plots (Column 3). This last finding would suggest that the potential impacts of demonstration plots should be greater among

[18] The number of nominations as the best farmer is correlated with network centrality. It explains 44 percent of the variation of degree centrality, 32 percent for eigenvector centrality, and 41 percent for betweenness centrality.

large farmers because of the higher adoption rates, a result working against our main findings presented later in the paper.

Cultivation practices and profitability

The intervention affected cropping systems. Farmers planting BD56 harvested those fields 25 days earlier than farmers sowing BD51 (in late October rather than mid November) (Table B.5 and Figure B.3). Treatment farmers used this additional month between their two rice crops to increase the likelihood of planting a post-Aman (Rabi) crop. On average, BD56 plots were 27.8 percentage points more likely to be sown with the Rabi crop than BD51 plots (Table B.5). Mustard, pulses, and potatoes were the most frequent short-season Rabi crop induced by the treatment.

Importantly for knowledge diffusion, this change in cropping systems is heterogeneous by type of entry point. The treatment effect for growing the Rabi crop is 17 and 11 percentage points higher for large and SAO farmers, respectively (Table B.6). Because growing additional crops is such a visible activity, this offers a potential mechanism as to why knowledge diffuses faster with large and SAO selection — a possibility we investigate in Section 2.3.4.

While the BD56 treatment led to a sharp increase in cropping intensity, we still observe that 46 percent of the BD56 plots were left fallow in between the two rice crops.[19] In addition, BD56 naturally leads to lower yields given its shorter duration: the yield of BD56 plots was 31 percent lower than that of the longer duration BD51 plots. We also discovered that BD56 fetched a slightly lower market price, which farmers attributed to less familiarity by millers. In combination, profits during the Aman season were lower by 4,576 taka for BD56 plots, or around 44 percent (Table B.7).

The average gain in profit from Rabi cultivation is 1,436 taka (60 percent). Assuming all of these benefits come from the extensive margin of growing the crop, BD56 led to an increase in Rabi profits of 5,241 taka for farmers that complied by growing the additional crop afforded by the treatment. This suggests that the technology was profitable among the subset of farmers who fully complied by planting the Rabi crop, but was not profitable on average since not all farmers capitalized on this main benefit of the technology.

2.3.3 How does knowledge diffuse across treatments?

Do demonstration plots increase awareness about new technology? If so, how do these effects compare to those generated by improved selection of entry points? The follow up information survey in the 192 BD56 villages allows us to answer these questions. The farmer-level specification compares awareness across the six

[19] There are a number of explanations including that farmers were not prepared to grow an additional crop when the rice matured much earlier than anticipated (despite being told of the duration when receiving the seeds), an inability to access land with plows when it is surrounded by maturing rice, and lack of access to capital for planting an additional crop.

different arms of the diffusion experiment. The corresponding regression is

$$\begin{aligned} aware_{ivs} = \beta_0 &+ \beta_1 RandomDemo_{vs} + \beta_2 SAONoDemo_{vs} + \beta_3 SAODemo_{vs} \\ &+ \beta_4 LargeNoDemo_{vs} + \beta_5 LargeDemo_{vs} + \alpha_s + \varepsilon_{ivs}, \end{aligned} \tag{2.1}$$

where the dependent variable is an indicator for whether farmer i in village v and upazila s is aware of BD56, $SAONoDemo_{vs}$ is an indicator for villages with SAO selection and no demonstration plots, and the remainder of the variables are defined analogously. As in all of the analysis, we include strata (upazila) fixed effects and cluster standard errors at the village level.

The results in Table 2.2 deliver three insights. First, the demonstration plots increase knowledge when cultivated by randomly selected farmers. Specifically, the rate of awareness increases by 7.2 percentage points when random farmers grow BD56 side by side with a chosen comparison variety. Sixty percent of farmers were knowledgeable of BD56 in control villages, meaning that the treatment effect of demonstration plots amounts to a 12 percent effect. The large rate of awareness in *RandomNoDemo* villages (the omitted category) shows that information diffuses, even under the benchmark where entry points are random and demonstration plots are absent.

Second, the demonstration plots have no effects when cultivated by the better-connected large and SAO farmers. The estimates of β_2 and β_3 are nearly identical, meaning that the demonstration plots didn't spread knowledge with SAO selection. Similarly, the estimates of β_4 and β_5 indicate that adding demonstration plots failed to increase awareness with large-farmer entry points.

Third, the effect of demonstration plots with random farmers is roughly the same magnitude as the effects of entering with large and SAO farmers. As we would expect based on their network connections, entering with large and SAO-selected farmers increases awareness. Amongst non-demo villages, SAO selection increases awareness by 6.7 percentage points (11.2 percent) and entering with the largest farmers increases awareness by 7.4 percentage points (12.3 percent). These effects are quite similar and statistically indistinguishable from the effect of demonstration plots with random entry points. Moreover, the demonstration plots eliminate the effects of targeting more central farmers. Specifically, the estimates of β_1, β_3, and β_5 are indistinguishable.

Columns 2 and 3 of Table 2.2 show effects on the number of conversations farmers reported having about BD56. While somewhat noisier, these data are also consistent with the demonstration plots creating just as much conversation as the improved selection of entry points. The demonstration plots led to 0.12 more conversations per farmer about BD56 when entry points were selected randomly (col. 2), almost all of which with entry points (col. 3). This 14 percent effect is similar to the effect on knowledge reported in column 1. Focusing on the mean outcomes, the average respondent in the control villages reported 0.84 conversations about BD56, 0.72 of which were with the entry points, and by difference 0.12 were with any of ten other randomly selected farmers that each farmer was asked about. Importantly, the reported conversations should not be interpreted as the total number of conversations about BD56, but rather the number of conversations with the 15 farmers asked about in our survey.[20]

The data on conversations help to rule out an alternative explanation where the side-by-side comparison — and two markers in the field — was more effective at simply broadcasting information on the existence of BD56. Instead, the demonstration plots caused farmers to engage in social learning, rather than just learn about the new technology from the sign in the field.

2.3.4 What mechanism explains the effects?

We highlight one mechanism which makes the demonstration plots work. Once allowing for endogenous communication across network links, the demonstration plots make people pay attention to farmers outside of their immediate network. This mechanism is consistent with the entry-point selection effects being eliminated by demonstration plots. The 64 villages with random entry-point selection allow us to further consider this mechanism. Within these villages, the number of entry points in a farmer's network is as good as randomly assigned when conditioning on the total number of connections of that farmer.[21] As a result, the average effect of being connected to an additional entry point can be estimated with

$$aware_{ivs} = \beta_0 + \beta_1 Entry\ Point\ Peers_{ivs} + \beta_2 Total\ Peers_{ivs} + \alpha_s + \varepsilon_{ivs}, \tag{2.2}$$

where $Entry\ Point\ Peers_{ivs}$ is the variable measuring how many of the five entry points farmer i is connected to and $Total\ Peers_{ivs}$ is the network degree of farmer i. Under our framework the demonstration plots should make baseline relationships less important for awareness, i.e. cause a decrease in β_1.

The average peer effects are indeed consistent with social learning. Column 1 in Table 2.3 shows that being connected to an additional entry point increases awareness by 7.7 percentage points, i.e. about 12 percent. More interestingly, the second column shows that this social learning only exists in villages without demonstration plots. An additional connection to an entry point increases knowledge by 13.5 percentage points without demonstration plots, but when entry points set up head-to-head demonstrations this effect goes down significantly to only 1.9 percentage points. In addition, the demonstration plots only increase learning for farmers that were unconnected to entry points at baseline. Demonstration plots increase awareness by 11.3 percentage points for farmers having no baseline connections to entry points. The effect disappears with just one connection to an entry point as the coefficient on the interaction term is nearly identical to the coefficient on the demonstration villages indicator. The similarity of the effects points to how the demonstration plots substitute for social connections to entry points: the increased awareness generated by demonstration plots is nearly the same as the peer effects

[20] We collected information on conversations between the respondent and the five entry points as well as 10 randomly selected other farmers for each village.

[21] Miguel and Kremer (2004) use a similar strategy when estimating spillover effects from deworming in Kenya.

in non-demo villages. Lastly, measuring connections with a binary variable for being connected to at least one entry point does not change the results (columns 3 and 4).

Table B.8 shows the analogous results where the dependent variable is instead the number of reported conversations between respondents and entry points. The results are consistent with those on knowledge. The demonstration plots induced conversations between entry points and farmers that were outside of their baseline information networks.

Finally, the demonstration plots were only effective for the least networked farmers. We investigate this by limiting to the random-entry-point villages and estimating

$$aware_{ivs} = \beta_0 + \beta_1 Demo_{vs} + \beta_2 Eigenvector_{ivs} + \beta_3 Eigenvector_{ivs} * Demo_{vs} + \alpha_s + \varepsilon_{ivs}, \quad (2.3)$$

where $Eigenvector_{ivs}$ is the baseline eigenvector centrality of farmer i. The estimated coefficient on the interaction term β_3 measures whether the demonstration plots had a differential effect for farmers that were more central at baseline. Intuitively, the more connected farmers have a number of ways (both direct and indirect) to find out about BD56. In contrast, the least central farmers are the most likely to benefit from making new connections to entry points.

Table 2.4 shows that the effect of demonstrations varies significantly according to the farmer's eigenvector centrality. The coefficient on the interaction term (β_3) is negative and precisely estimated. The magnitude of the coefficient is large. Going from the 10th to the 90th percentile of the eigenvector centrality distribution causes the effect of demonstrations to go from 0.165 to -0.022. Put another way, demonstration plots become ineffective for the most central farmers in the network.

This evidence favors the explanation that demonstration plots substitute for social learning in networks. This substitution also offers an explanation for why the identities of entry points become irrelevant when they cultivate demonstration plots. Demonstration plots effectively cause farmers to pay attention, thus "turning off" peer effects and eliminating the need to rely on information being transmitted from entry points to other farmers. Rather, the demonstration plots induce information to flow to people that would have otherwise been excluded from peer-to-peer social learning.

In contrast to inducing conversations in endogenous social networks, the baseline network structure explains part of the effects of SAO and large-farmer selection. Any network-based model of diffusion with an exogenous network would favor targeting more central entry points. As a result, the effects of SAO and large-farmer selection would be predicted to decrease when conditioning on the average centrality of entry points. The data show exactly this. Table 2.5 shows that conditioning on the average degree centrality of entry points (moving from column 1 to 2) causes the effects of large and SAO selection to decrease by 43 and 31 percent, respectively.[22] In addition, the average degree of entry points is strongly correlated with

[22]We limit the data to the non-demo villages for this analysis since the selection treatments are only effective in these villages.

knowledge diffusion. We of course can not provide a causal interpretation of this parameter. The exercise instead offers evidence that the ability of large and SAO entry points to increase knowledge is due to their more central network positions.

Beyond network centrality, the act of growing the additional crop — something more likely done by large and SAO farmers — appears to have captured attention and resulted in increased knowledge of BD56. Column 3 shows that the effects of large and SAO selection also decrease when conditioning on the number of entry points that planted a Rabi crop on their BD56 plot. The effects of large farmer and SAO selection are smaller by 63 percent and 51 percent, respectively, when conditioning on both degree centrality and the choice of growing a Rabi crop.

The villages with random and SAO-based selection also include large farmers as entry points. At least one of the five largest farmers was selected as an entry point in 19 of the 64 random villages and 38 of the 64 SAO selection villages. Table 2.6 shows analysis where we compare villages where at least one large farmer was targeted with those having no large-farmer entry points.[23] We find similar results when exploiting this variation. In column 1, having at least one large-farmer entry point increases awareness by 7.9 percentage points in non-demonstration villages. Similar to the previous analysis, conditioning on the average network degree of entry points and the number growing the Rabi crop absorbs much of this effect (columns 2 and 3). Introducing demonstration plots makes these effects disappear. Columns 4-6 show that hitting a large farmer has no effect in demonstration villages and that the correlation between the degree centrality of entry points and awareness is much weaker with demonstration plots.[24] Consistent with the other findings, the demonstration plots decrease the relevance of the baseline social network for knowledge diffusion.

In sum, two attributes explain much of the reason why the selection treatments were effective. First, large and SAO farmers are more central in the network and therefore share information with more farmers. Second, these farmers do a better job of showcasing the benefits of new technology. Nonetheless, despite these two mechanisms, demonstration plots with random farmers are equally effective at spreading awareness about new technology.

2.3.5 Effects on seed purchases

Awareness about new technology is our main outcome variable. In addition to information diffusion, we also obtained data on purchases of BD56 seeds. A local NGO visited each village prior to the 2017 rainy season — a year after BD56 had been introduced in the village. The NGO set up a small stand and made BD56 seeds available to any farmer wishing to purchase. Importantly, the seeds were subsidized at a rate of 60 percent, and farmers visiting the shop were told that BD56 shortens

[23]We only worked with SAO's to select entry points in the 64 SAO villages and the 64 BD51 villages. We therefore can't repeat this analysis using SAO-identified farmers rather than large farmers.

[24]Table B.9 gives the statistical test showing that across all villages, the impact of hitting at least one large farmer is significantly smaller in demonstration villages.

the season, gives lower yield relative to longer duration varieties, but allows for an additional crop to be grown during the year.[25] The NGO representative explained to farmers that a third crop is needed to make BD56 profitable on an annual basis.

Table 2.7 shows regression results akin to Equation (2.1), but at the village level where the dependent variable is either the number of farmers purchasing or the adoption rate (number of buyers divided by village size). The point estimates are much noisier, but the coefficients are sizable and the directions line up with what we observe on knowledge diffusion. About 1.7 farmers purchased seeds per village in control villages, and this increased by around 0.67 farmers (40 percent) when adding demonstration plots. The number of farmers purchasing seeds also increases in large and SAO villages. Turning to column 2, the degree centrality of entry points and the number growing the third crop are positively correlated with the number of purchasing farmers and absorb some of the selection effects. Columns 3 and 4 show that the pattern remains when considering the share of farmers purchasing, rather than the absolute number. Overall, the coefficients in Table 2.7 show a similar pattern to what we observe on awareness creation, despite being less precisely estimated.

2.3.6 Who becomes informed from the different treatments?

Combining our two experiments (diffusion and impact evaluation experiments) give a unique opportunity to test whether alternative mechanisms for encouraging information diffusion deliver information to the people expected to benefit the most from new technology. Capturing the short-duration benefits of BD56 requires growing an additional short-season crop — an action that was not universally taken by farmers in our BD56 treatment group (Table B.5). We next ask two questions (1) do observable characteristics explain variability across farmers in the treatment effect of BD56 on the number of crops grown across the year and (2) do the demonstration plots or the entry-points selection strategies differentially inform the farmers who, based on these same observable characteristics, are the most likely to increase their number of crops grown when adopting BD56?

We start by using the impact experiment to estimate a function measuring the treatment-effect heterogeneity of BD56, which we refer to as the heterogeneity index. The first step is to implement the method suggested in Chernozhukov et al. (2018) to generate a linear prediction of the ATE conditional on the observed covariates from our door-to-door census, denoted as z_i. Our sample of treatment and control farmers is first divided into two samples: one "training" sample where we seek to estimate the heterogeneity index, denoted as $s_0(z_i)$, and another "validation" sample where we seek to measure whether this estimate $\hat{s}_0(z_i)$ is a significant determinant of the heterogeneous treatment effect of BD56 on the number of crops grown. First for the training sample, we estimate separate LASSO regressions for the treatment and control groups to pick which of the covariates in z predict

[25] That year the government rate for BD56 seed ranged between 34 and 40 tk/kg. The NGO sold the seeds for 15tk/kg.

the number of crops grown y. Using these covariates, we generate estimates of the conditional expectations of the number of crops grown as $E(y_i|D_i = 1, z_i)$ and $E(y_i|D_i = 0, z_i)$.[26] The difference between these two conditional expectations serves as the heterogeneity index, $\hat{s}_0(z_i)$.

Turning to the validation sample, we want to verify that the treatment effect varies according to this measure $\hat{s}_0(z_i)$. We do this in two ways. First, we add an interaction between the treatment indicator and $\hat{s}_0(z_i) - \bar{s}$ in a regression where the dependent variable is the number of crops grown.[27] Second, we estimate separate treatment effects for the four quartiles of the distribution of $\hat{s}_0(z_i)$. Finally, this process is iterated 100 times, delivering 100 separate sample divisions and 100 estimates of the heterogeneity index.

The observed covariates predict treatment-effect heterogeneity in the validation sample. Figure 2.6 shows the 100 estimates of the ATE and the linear heterogeneity term. The heterogeneous effect is almost always larger than zero, suggesting that the heterogeneity index $\hat{s}_0(z_i)$ does proxy for the true heterogenous effect of BD56 on the number of crops grown. In other words, farmers with larger values of $\hat{s}_0(z_i)$ appear more likely to increase cropping intensity if adopting short-duration rice. Figure 2.7 shows the separate treatment effects by quartile of the heterogeneity index. Treatment effects increase with the heterogeneity index and are largest in the top two quartiles of the distribution of $\hat{s}_0(z_i)$.

We then return to the diffusion experiment in 192 villages to test whether the treatment effects differ based on values of the heterogeneity index. We possess 100 estimates of $\hat{s}_0(z_i)$ for each of the 1,920 farmers for which we elicited knowledge of BD56. We take the median of $\hat{s}_0(z_i)$ across these 100 sample divisions and estimate whether the treatment effects of demonstration plots or entry-points selection strategies depend on this predicted heterogeneity index.

Table 2.8 shows the regression results. Columns 1 and 2 interact the treatment variables directly with $\hat{s}_0(z_i)$, while columns 3 and 4 use an indicator for observations with above-median values of this heterogeneity index. The estimates are noisy, but the heterogeneity index is positively associated with learning and having conversations, indicating that information is more likely to flow to farmers with higher returns when entry points are selected randomly and there are no demonstrations. For instance, farmers with an above-median heterogeneity index are 10.1 percentage points more likely to learn about BD56 in random and no demo villages (column 3). From column 4, these same farmers are expected to have .26 more conversations (about 31 percent). However, the point estimates on the interaction terms between the heterogeneity index and the treatment indicators are generally negative and the coefficients on the five treatment indicators are positive and of similar magnitudes. These findings indicate that while our treatments increased knowledge by either exploiting existing network structure or triggering conversa-

[26] We use the OLS regressions with the covariates selected by the two LASSO procedures. The selected covariates can be different in the treatment and control groups.

[27] The coefficient on the treatment indicator in this regression measures the average treatment effect, while the coefficient on the interaction between treatment and $\hat{s}_0(z_i) - \bar{s}$ measures whether the heterogeneity index predicts actual treatment-effect heterogeneity.

tions, the gains in knowledge were concentrated amongst farmers that would be less likely to capitalize on the main benefit of BD56 if adopting.

As examples in column 4, demonstration plots increased conversations about BD56 for farmers with *below-median* predicted effects on cropping intensity by around 0.29, but had no effects for farmers who appeared more likely to increase cropping intensity if adopting BD56 (0.29 + -0.32). Similarly, seeding with SAO-selected farmers and with the largest farmers increases conversations for farmers with below-median values of the heterogeneity index, but had no effect for farmers that are above the median.

This finding sheds light on who is induced to have conversations and learn when a policymaker intervenes with either an alternative seeding strategy or demonstration plots. Put simply, the people impacted by these treatments do not appear to be those that would be the most likely to enjoy the technology's main benefit if it was randomly introduced to them. One reasonable interpretation is that conversations take place and information is obtained endogenously. Therefore, farmers with the highest returns from obtaining information are more likely to engage in social learning regardless of the dissemination strategy. Intervening to trigger the spread of information only affects those with lower returns, i.e. those who are less likely to endogenously seek information regardless of the dissemination strategy.

2.4 A model that rationalizes the experimental results

This section presents a basic diffusion model that allows for endogenous interaction between unlinked agents and can rationalize our experimental findings. Our goal is not to capture all of the strategic elements of network formation.[28] Instead, we opt for the simplest formulation that captures the tradeoffs introduced by our treatments. First, the policymaker can introduce information to central entry points therefore taking advantage of the existing network structure. Or, the policymaker can seek to induce more communication in a world where communication networks are endogenous. We show how the two approaches act as substitutes in a way that is consistent with our empirical findings.

2.4.1 Model Environment

We consider a village with N farmers indexed by $i \in \{1, 2, ..., N\}$. The village social network is described by the $N \times N$ adjacency matrix G, where $g_{ij} = 1$ indicates that

[28]There are several papers looking at different aspects of how information links are formed. These range from differences between actively sharing information and passively listening (Calvó-Armengol, Martí, and Prat, 2015), the types of initial network structures that allow for efficient information aggregation as societies grow large (Acemoglu, Bimpikis, and Ozdaglar, 2014), and the stigma from seeking information when it might signal low ability or understanding (Chandrasekhar, Golub, and Yang, 2016; Banerjee et al., 2018a).

farmers i and j have an information-sharing link. In terms of our data, this feature of the model corresponds to the baseline social network module.

The policymaker first seeds technology with five entry points, indexed by $j = 1, \ldots 5$. Each entry point is now "informed" and can then spread the information to others, as in the standard information cascade. We assume, for now, that the network is fixed and a farmer can only become informed if they receive information that emanated from an entry point. The parameter q represents the exogenous probability that any informed farmer passes information to a connected peer. The probability that a farmer becomes informed, denoted as h_i, is a function of the information-passing probability q and the length of the possible paths between i and each entry point j. Intuitively, farmers having the greatest number of shortest paths to entry points become the most likely to be informed. In contrast, a farmer that is completely isolated from entry points is not informed.

We build on this standard framework by adding the ability to form new links. In addition to passively waiting for information to arrive from entry points, a farmer can increase the probability of becoming informed by forming a link with an entry point. Suppose the cost of forming a new link is c and the probability that an entry point will pass information is p. A new link between farmer i and entry point j is denoted as $l_{ij} = 1$ if i chooses to form the link and 0 otherwise. Overall, the probability of gaining information from one of the two channels, denoted as μ_i, is

$$\mu_i = 1 - (1 - h_i) * \prod_j (1 - p)^{l_{ij}} . \tag{2.4}$$

Finally, we write v as the utility of being informed and normalize the utility of not being informed to 0.

2.4.2 The link-formation decision

The simple problem of the farmer is to choose whether to link with each of the entry points. More formally, the farmer's optimization problem is written as

$$\max_{l_{ij}} v \left(1 - (1 - h_i) * \prod_j (1 - p)^{l_{ij}} \right) - \sum_j c * l_{ij}. \tag{2.5}$$

The problem can be simplified to choosing the number of new contacts with entry points, denoted as m, since each entry point adds the same probability of learning p. The exact decision rule is that the farmer will link with m entry points if and only if:

$$v(1 - h_i)p(1 - p)^{m-1} > c. \tag{2.6}$$

The farmer seeks to connect directly to entry points for information if the costs of doing so are low (c is low), or if he is in a poor position to obtain information via diffusion in the network (h_i is low). Increasing the probability that an entry point shares useful information (p) increases the likelihood of connecting with one entry point ($m = 1$). If p is sufficiently large, then increasing p causes the marginal benefit of linking with further entry points to decrease because the farmer is likely to obtain the information from the first entry point newly added to her network.

2.4.3 Consistency between the model and experimental findings

This small modification to a standard diffusion framework predicts treatment effects that are in line with our RCT results. Put differently, the two mechanisms for being informed — receiving information via the existing social network or communicating with an entry point outside the network — can explain the main impacts as well as the heterogeneity across the sample. The discussion that follows links this simple theory to our results.

Effectiveness of entering with Large/SAO farmers: By being more central in networks, large and SAO-selected farmers facilitate diffusion. Holding networks fixed, the probability h_i in equation (2.6) increases, causing the probability of receiving information to increase. This prediction is the obvious one that supports network-based approaches to identifying entry points when network structure remains fixed. The main effects of large farmer and SAO-based selection on knowledge, and the sensitivity of these effects to conditioning on centrality of entry points (Tables 2.2, 2.5, and 2.6) are all consistent with this standard mechanism.

Effectiveness of demonstration plots: We argue that the demonstration plots convey active experimentation and signal a farmer that is paying attention to how the new technology performs against its' relative alternative. This is nontrivial as Hanna, Mullainathan, and Schwartzstein (2014) show that inattention bias can hinder what farmers learn from experimentation. As a result, fellow villagers perceive that farmer to be more likely to pass useful information, i.e. the parameter p increases with the introduction of demonstration plots. The marginal benefit of forming a link with a single entry point then increases. Correspondingly, the new information link increases the likelihood of becoming informed. The main results on knowledge and reported conversations in Table 2.2 appear consistent with this reasoning.

Interactions between network-based selection and demonstration plots: The endogenous network mechanism posits substitutability between seeding with more central farmers and demonstration plots. Returning to equation (2.6), seeding with more central farmers increases the probability of learning via existing network links and therefore reduces the marginal benefit of endogenously seeking information from entry points. The findings show this exactly: demonstration plots have no effect when cultivated by more central farmers and the effect of entering with more central farmers is eliminated when the policymaker introduces demonstration plots.

Network effects: Farmers directly connected to entry points gain less from demonstration plots because the existing network connection increases the likelihood of learning through the information-diffusion mechanism. We found that baseline connections with entry points do increase knowledge, however, demonstration plots eliminate this advantage by giving a channel for unconnected farmers to learn. This is compatible with equation (2.6) where the benefit of making new links with entry points declines with h_i.

In sum, our experimental findings, along with the simple model, emphasize the two mechanisms that serve to boost learning in networks: optimizing the selection of entry points or inducing communication to facilitate learning. The latter mechanism has received less attention in the literature. However, our results suggest that it gives the policymaker an important alternative for making information diffuse faster in networks.

2.5 Concluding Remarks

We have shown experimental evidence on a new mechanism for spreading awareness about technology. Demonstration plots — where partnering farmers cultivate a new technology side by side with an existing one — increase awareness relative to a control where new technology is demonstrated on its own. We found that this relatively straightforward method of agricultural extension made an additional seven percent of farmers aware about a new seed variety. Demonstration plots raised awareness only for farmers that lacked social connections with adopters. In addition, the demonstration plots were the most effective for the farmers that were the most isolated — in terms of their eigenvector centralities — in the baseline information network. All of these results, when taken together, are consistent with a model where the effectiveness of demonstration plots is explained by their ability to facilitate learning by triggering communication in information-sharing networks.

The experiment benchmarked these demonstration plots against scalable and policy-relevant alternatives where entry points were selected strategically to increase their centralities in the social network. These were entering with the largest farmers and those hand picked by government extension agents. Indeed, these methods do increase knowledge compared to random selection. But the gains in awareness are about the same as the gains from the demonstration plots. And in contrast to endogenous communication between people without links, the selection treatments seem to be effective because of how they exploit the structure of the baseline social network and because the more central entry points did a better job of demonstrating a key benefit of new technology.

We also showed evidence on who becomes informed. The seed variety we introduced has heterogeneous benefits. Specifically, only some farmers took advantage of the early maturation by increasing cropping intensity. Applying machine-learning methods to identify the characteristics associated with taking this action, we found farmers who are the most likely to grow an additional crop when adopting BD56 are not those learning from demonstration plots. We found the same result for improved seeding strategies. This finding highlights a tradeoff: intervening to increase knowledge transmission may be ineffective for the highest return individuals who learn even in the absence of additional intervention by the policymaker.

Overall, our analysis highlights the potential for alternate mechanisms of agricultural extension, outside of those that rely on information cascades through social networks. Despite the ample evidence that networks are important for knowledge transmission, there have been few studies that compare information cascades with alternative methods of spreading knowledge.[29] Focusing on policy, it is important to consider such alternatives because policymakers may face difficulty in identifying the entry points that are theoretically positioned for the best spread of information. Either it could be prohibitively expensive to do a full network survey, or there may not be observable characteristics (such as farm size) that correlate strongly with less observable measures of network centrality. Our results show that in these contexts, improving learning can be achieved by taking small steps to capture people's attention, and get them to communicate and engage in learning from new people.

[29] Outside of agricultural technology, Banerjee et al. (2018a) is the only paper we are aware of that compares seeding information about rules for India's 2016 demonetization with broadcasting that information more widely.

Tables

Table 2.1: Differences in baseline characteristics for different entry points

	Coefficients and SE:			
	(1) Constant	(2) SAO	(3) Large farmers	(4) p-value (2)-(3)
Network Variables:				
Degree	4.562*** (0.355)	3.582*** (1.042)	4.481*** (0.853)	0.473
Eigenvector centrality	0.089*** (0.006)	0.042*** (0.012)	0.071*** (0.011)	0.030
Diffusion centrality	-0.010 (0.054)	0.643*** (0.141)	0.872*** (0.110)	0.157
Betweenness centrality	164.186*** (27.926)	394.084*** (103.540)	315.640*** (69.762)	0.509
Household Characteristics:				
Area cultivated all seasons (bigah)	9.013*** (0.658)	5.865*** (1.368)	21.396*** (2.689)	0.000
Times named best farmer	0.790*** (0.206)	4.477*** (0.785)	5.589*** (0.707)	0.275
Log revenue per bigah	10.061*** (0.057)	-0.016 (0.077)	-0.014 (0.075)	0.970
Number livestock owned	3.950*** (0.217)	-0.008 (0.284)	1.968*** (0.512)	0.000
Number of overseas migrants	0.138*** (0.031)	-0.021 (0.039)	-0.026 (0.037)	0.881
Education	4.647*** (0.304)	1.247*** (0.464)	0.925* (0.488)	0.536
Age	42.222*** (0.739)	0.712 (1.026)	3.594*** (1.078)	0.007
Tubewell owner	0.097*** (0.022)	0.094*** (0.036)	0.181*** (0.051)	0.108

The data are limited to the 960 selected entry points in the 192 BD56 villages. Each row is the result from a separate regression where the characteristic is regressed on a constant and indicators for SAO and large farmer villages. The omitted group is the villages where demonstrators were selected randomly (meaning the first column is the mean value for random entry points). The standard errors in each regression are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table 2.2: Treatment effects on knowledge

	(1) Heard About	(2) Conversations	(3) Conversations w/ Entry Points
Random w/ demo	0.072*	0.116	0.106
	(0.041)	(0.086)	(0.088)
SAO no demo	0.067*	0.046	0.042
	(0.039)	(0.067)	(0.071)
SAO w/ demo	0.065	0.074	0.096
	(0.040)	(0.089)	(0.086)
Large no demo	0.074**	0.123*	0.114*
	(0.036)	(0.068)	(0.066)
Large w/ demo	0.049	0.108	0.113
	(0.044)	(0.075)	(0.079)
Strata fixed effects	Yes	Yes	Yes
Mean in Control	0.60	0.84	0.72
Number of Observations	1919	1920	1920
R squared	0.171	0.212	0.250

The data are for the 10 random farmers per village that were selected for the information survey. The dependent variable in column 1 is an indicator for having knowledge of BD56. The dependent variable in column 2 is the number of conversations the farmer had with 15 other farmers about BD56 (the five entry points and 10 randomly selected farmers). The dependent variable in column 3 is the number of conversations specifically with entry points. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels. Note, the rates of awareness are similar amongst the different selection arms in demonstration plot villages, i.e. the estimates on SAO w/ demo, Large w/demo, and Random w/demo are indistinguishable.

Table 2.3: Peer effects on knowledge, separate for villages with and without demonstration plots

	(1)	(2)	(3)	(4)
Peer connections w/ entry points	0.077**	0.135***		
	(0.034)	(0.046)		
Peer connections w/ entry points * Demonstration Village		-0.116*		
		(0.061)		
Connected to at least 1 entry point			0.092*	0.156***
			(0.053)	(0.059)
Connected to at least 1 entry point * Demonstration Village				-0.130
				(0.097)
Number of connections	-0.004	-0.005	-0.002	0.002
	(0.003)	(0.006)	(0.003)	(0.005)
Number of connections * Demonstration Village		0.000		-0.005
		(0.006)		(0.005)
Demonstration Village		0.113**		0.132***
		(0.045)		(0.045)
Strata fixed effects	Yes	Yes	Yes	Yes
Mean in Control	0.60	0.60	0.60	0.60
Number of Observations	635	635	635	635
R squared	0.186	0.199	0.185	0.197

The dependent variable in all regressions is an indicator for having heard of BD56 amongst the 10 randomly surveyed farmers per village. The data are limited to the 64 villages where entry points where chosen randomly and peer effects can therefore be causally identified. The variable *Peer connections w/ entry points* is the number of entry points (from 0 to 5) that the farmer is connected with while *Connected to at least 1 entry point* is an indicator variable for being connected to at least one of the entry points. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table 2.4: Effects of demonstration plots as a function of baseline network centrality

	(1)	(2)
Demonstration Village	0.072*	0.179***
	(0.041)	(0.057)
Eigenvector Centrality * Demo		-1.067***
		(0.381)
Eigenvector Centrality		0.935***
		(0.325)
Strata fixed effects	Yes	Yes
Mean in Control	0.60	0.61
Number of Observations	639	517
R squared	0.183	0.202

The dependent variable in both regressions is an indicator for having heard of BD56. The data are limited to the 64 villages where entry points were chosen randomly. *Eigenvector Centrality* is the baseline network centrality of the respondent. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table 2.5: Effects of entry-point treatments when conditioning on observable attributes of entry points

	(1)	(2)	(3)	(4)
SAO no demo	0.068*	0.047	0.056	0.033
	(0.038)	(0.039)	(0.039)	(0.040)
Large no demo	0.076**	0.043	0.064*	0.028
	(0.035)	(0.035)	(0.035)	(0.036)
Average degree of entry points		0.006***		0.006***
		(0.002)		(0.002)
Number entry points growing rabi crop			0.020*	0.022*
			(0.011)	(0.011)
Strata fixed effects	Yes	Yes	Yes	Yes
Mean in Control	0.60	0.60	0.60	0.60
Number of Observations	960	960	960	960
R squared	0.173	0.179	0.176	0.182

The data are for the 10 random farmers per village that were selected for the information survey and are limited to the 96 villages without demonstration plots. The dependent variable in all regressions is an indicator for being aware of BD56. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table 2.6: Effects of having at least one large-farmer entry point on BD56 knowledge

	Non-Demo Villages			Demo Villages		
	(1)	(2)	(3)	(4)	(5)	(6)
At least 1 large entry point	0.079**	0.048	0.035	-0.017	-0.024	-0.025
	(0.032)	(0.034)	(0.035)	(0.037)	(0.040)	(0.038)
Average degree of entry points		0.005**	0.006***		0.003	0.003
		(0.002)	(0.002)		(0.003)	(0.003)
Number entry points growing rabi crop			0.022*			0.020
			(0.012)			(0.013)
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Mean in Control	0.62	0.62	0.62	0.68	0.68	0.68
Number of Observations	960	960	960	959	959	959
R squared	0.175	0.179	0.182	0.175	0.176	0.180

The data are for the 10 random farmers per village that were selected for the information survey. The dependent variable is an indicator for being aware of BD56. Column 1-3 are for the villages without demonstration plots and columns 4-6 are for the demonstration villages. *At least 1 large entry point* is an indicator for villages where one of the five largest farmers was selected as an entry point. The mean in the control group is defined as the mean awareness rate in villages where none of the entry points were large farmers. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table 2.7: Treatment effects on seed purchasing behavior

	Number of farmers		Share of village	
	(1)	(2)	(3)	(4)
Random w/ demo	0.673	0.613	0.00622	0.00493
	(0.813)	(0.807)	(0.00916)	(0.00917)
SAO no demo	0.697	0.392	0.0116	0.00809
	(0.694)	(0.690)	(0.0107)	(0.0104)
SAO w/ demo	0.116	-0.154	0.00411	0.00135
	(0.615)	(0.640)	(0.00744)	(0.00774)
Large no demo	0.272	-0.171	0.00817	0.00325
	(0.535)	(0.621)	(0.00782)	(0.00868)
Large w/ demo	0.866	0.515	0.0195**	0.0147
	(0.641)	(0.654)	(0.00946)	(0.00914)
Average degree of entry points		0.0642*		0.000567
		(0.0382)		(0.000507)
Number entry points growing rabi crop		0.166		0.00291
		(0.156)		(0.00214)
Strata fixed effects	Yes	Yes	Yes	Yes
Mean in Control	1.679	1.679	0.019	0.019
Number of Observations	168	168	168	168
R squared	0.085	0.106	0.091	0.106

The data are from seed sales that were carried out for each village prior to the 2017 rainy season. We are missing data for 24 of the 192 villages because the seed supply ran out before those villages could be completed. The dependent variables are the number of farmers purchasing BD56 seeds (columns 1-2) and the share of farmers purchasing (columns 3-4). Robust standard errors are in parentheses. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

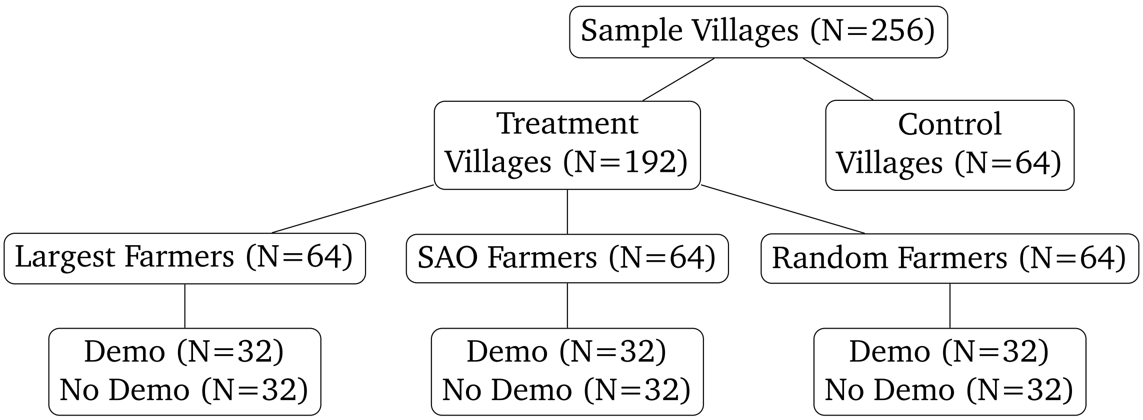
Table 2.8: Heterogeneous effects on knowledge and conversations by predicted impact of BD56 on number of crops grown

	Linear Heterogeneity		Effect Above Median	
	(1) Heard About	(2) Conversation	(3) Heard About	(4) Conversations
Random w/ demo	0.082	0.247*	0.121**	0.294**
	(0.069)	(0.131)	(0.060)	(0.128)
SAO no demo	0.073	0.096	0.089	0.125
	(0.072)	(0.115)	(0.060)	(0.094)
SAO w/ demo	0.096	0.184*	0.111*	0.253**
	(0.064)	(0.110)	(0.057)	(0.099)
Large no demo	0.148**	0.247**	0.148***	0.284***
	(0.061)	(0.101)	(0.053)	(0.101)
Large w/ demo	0.094	0.190	0.123*	0.265***
	(0.083)	(0.116)	(0.068)	(0.096)
Heterogeneity	0.034	0.236	0.101*	0.261**
	(0.123)	(0.195)	(0.057)	(0.103)
SAO no demo * Heterogeneity	0.001	-0.145	-0.014	-0.111
	(0.164)	(0.282)	(0.072)	(0.115)
SAO w/ demo * Heterogeneity	-0.096	-0.369	-0.072	-0.337**
	(0.143)	(0.227)	(0.077)	(0.148)
Large no demo * Heterogeneity	-0.230	-0.386*	-0.123*	-0.276**
	(0.152)	(0.227)	(0.072)	(0.135)
Large w/ demo * Heterogeneity	-0.136	-0.252	-0.124	-0.274**
	(0.192)	(0.240)	(0.081)	(0.119)
Random w/ demo * Heterogeneity	-0.020	-0.430*	-0.078	-0.318**
	(0.161)	(0.241)	(0.077)	(0.141)
Strata fixed effects	Yes	Yes	Yes	Yes
Mean in Benchmark	0.60	0.85	0.60	0.85
Number of Observations	1910	1911	1910	1911
R squared	0.174	0.216	0.175	0.219

These regressions test whether the different treatments increase knowledge and spark conversations more (or less) for farmers that are predicted to have the largest impact of of BD56 on the number of crops grown. Columns 1 and 2 show linear heterogeneity where the treatment indicators are interacted with $\hat{s}_0(z_i) = E(y_i|D_i = 1, z_i) - E(y_i|D_i = 0, z_i)$ and columns 3 and 4 partition the sample into farmers that are above and below the median in the distribution of $\hat{s}_0(z_i)$. For each farmer we calculate $\hat{s}_0(z_i)$ as the median value across the 100 sample divisions in Figures 2.6 and 2.7. The dependent variable in columns 1 and 3 is an indicator for having knowledge of BD56. The dependent variable in columns 2 and 4 is the number of conversations the farmer had with 15 other farmers about BD56. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Figures

Figure 2.1: Experimental Design



Notes: Figure shows a schematic representation of the experimental design. The 192 BD56 treatment villages were divided into three groups for entry-point selection: random selection, relying on the five largest farmers, and selecting those indicated by the ag. extension officer (SAO). Demonstration plots were set up on half of the 64 villages within each of these arms.

Figure 2.2: A visualization of the demonstration plot in comparison to the control group

A: Demonstration Plot

B: Control Plot

Notes: Panel A on the left shows an example demonstration plot. The plot on the left side is the BD56 plot while the plot on the right is the popular longer duration variety Swarna. Panel B on the right shows an example from the comparison villages where farmers were only given one marker to denote the BD56 plot.

Figure 2.3: A timeline of the experiment and data collection

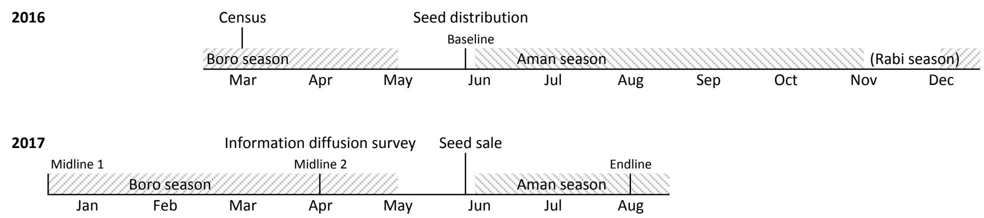
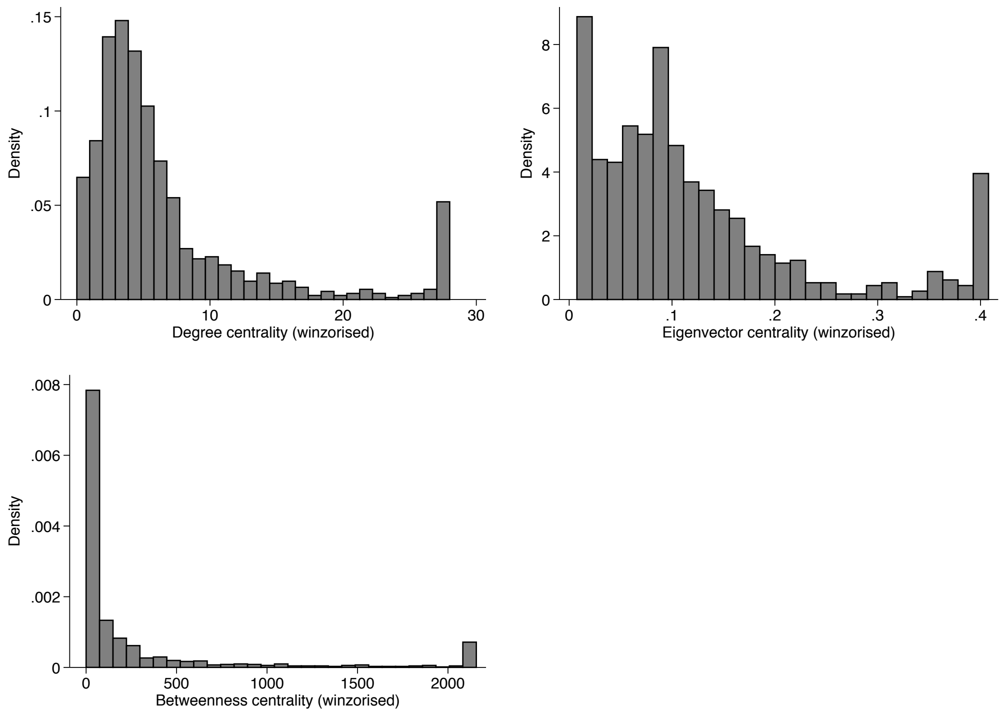
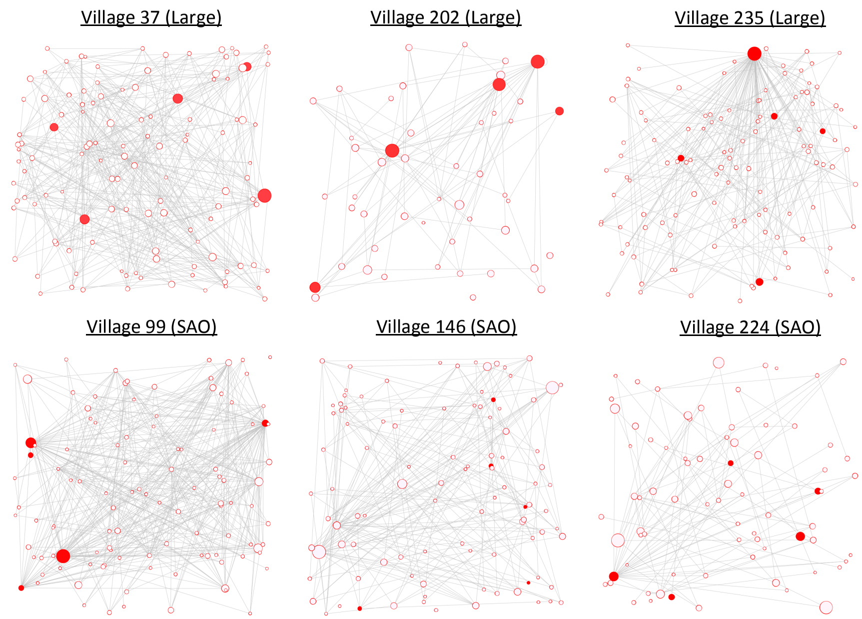


Figure 2.4: Distribution of centrality measures from social network survey



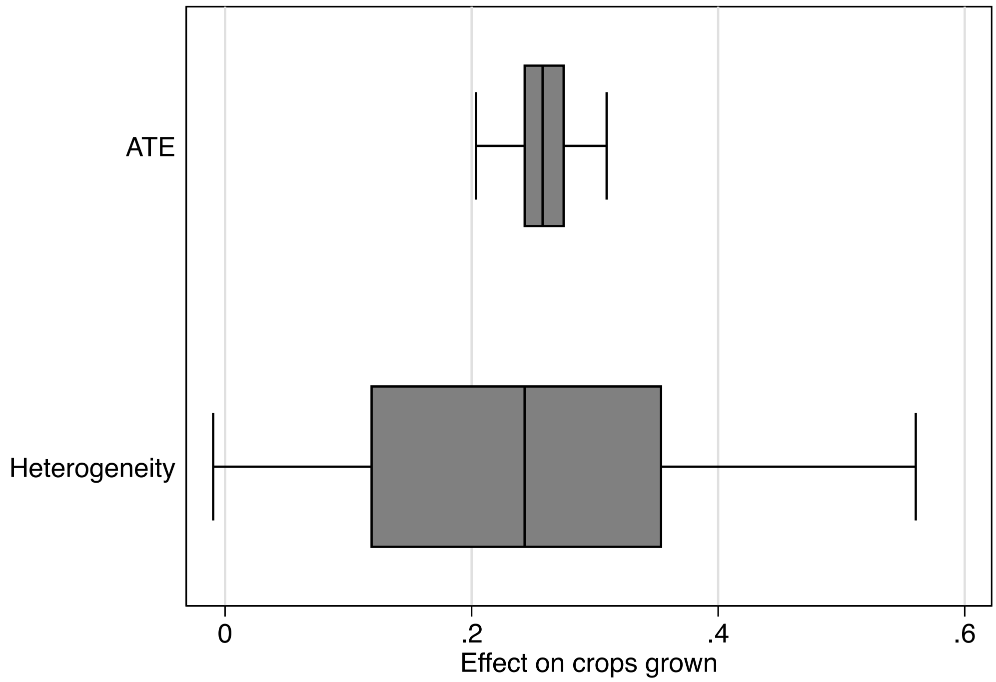
Notes: Figure shows the histograms for the 3 centrality measures from the baseline social network survey with all households (N=21,926).

Figure 2.5: Example network diagrams for 6 villages in the sample



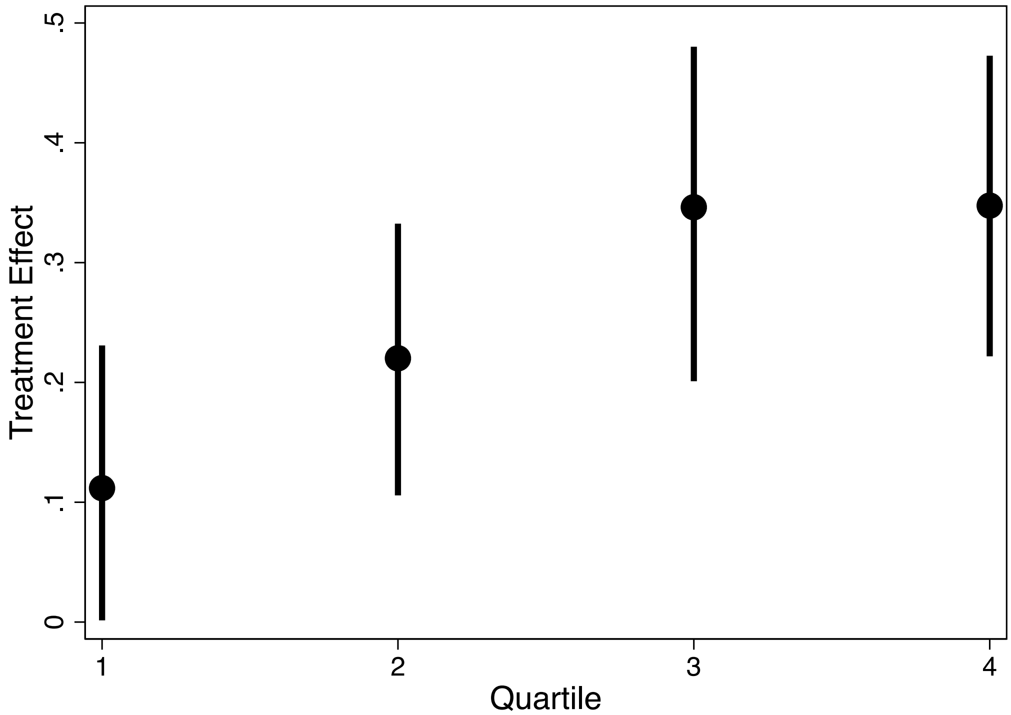
Notes: The figure maps the social network for 6 BD56 villages. The top 3 villages are large-farmer villages and the bottom 3 are villages with SAO selection. The nodes (dots) represent farmers and the size of nodes is proportional to farm size, where larger dots indicate larger farmers. The shaded red dots indicate the 5 farmers chosen as entry points while the hollow red dots denote the remaining farmers.

Figure 2.6: ATE and heterogeneous effect on number of crops grown



Notes: The figure shows the average treatment effects and the heterogeneous effect on the number of crops grown across across 100 equal-sized splits into training and validation datasets datasets. For each split, we estimate separate LASSO regressions for treatment (BD56) and control (BD51) farmers in the training dataset. In each case the number of crops grown is regressed on a set of 24 covariates, z_i. Using the selected covariates for each group, we calculate the estimated heterogeneity index for each farmer in the validation dataset as $\hat{s}_0(z_i) = E(y_i|D_i = 1, z_i) - E(y_i|D_i = 0, z_i)$. Using the validation dataset, we then regress the observed number of crops on the treatment, $\hat{s}_0(z_i) - \bar{s}_0$, the interaction between treatment and $\hat{s}_0(z_i) - \bar{s}_0$, and upazila fixed effects. The top bar in the figure shows the distribution of the 100 estimates of the ATE (the coefficients on the treatment indicator). The bottom bar shows the 100 estimates of the heterogeneity effect (the coefficient on the interaction between treatment and $\hat{s}_0(z_i) - \bar{s}_0$). The vertical line represents the average across the 100 splits, the box the interquartile range, and the whiskers give the min and max.

Figure 2.7: Effects on number of crops grown by quartiles of the predicted effect



Notes: The figure shows the estimated treatment effects by quartile of the heterogeneity index for 100 equal-sized splits into training and validation datasets. For each split, we estimate separate LASSO regressions for treatment (BD56) and control (BD51) farmers in the training dataset. In each case the number of crops grown is regressed on a set of 24 covariates, z_i. Using the selected covariates for each group, we calculate the estimated heterogeneity index for each farmer in the validation dataset as $\hat{s}_0(z_i) = E(y_i|D_i = 1, z_i) - E(y_i|D_i = 0, z_i)$. Using the validation dataset, we then regress the observed number of crops on the treatment and upazila fixed effects *separately for the four quartiles of* $\hat{s}_0(z_i)$. The heavy dots show the averages across the 100 sample divisions while the bands display the range from the 5th to 95th percentiles.

3 How do Online Job Portals affect Employment and Search outcomes? Evidence from India

Chapter abstract: In recent years innovative job market information systems have emerged in order to match recent graduates with employers via integrated websites and call center based platforms. We use a randomized control trial to evaluate the ability of such job-portals to ease search frictions in India. We partner with Job Shikari, an online platform operating primarily in Northern India, and upload a randomly selected sample of recent vocational training graduates onto their platform. A second subset of graduates are uploaded to the platform, and receive “priority status", which means they subsequently receive many more text messages than their peers about potentially job opportunities. We find that being uploaded to the job portal has a negative impact on the probability of being employed, but no significant impacts on job-search. The treatment priority group experiences a less strong disemployment effect, but is much more likely to respond to the treatment by migrating to urban centers. These results differ by job-seekers’ observable characteristics such as age and martial status. We interpret our results as evidence that the impact of job-portals depends significantly on job-seekers beliefs about what the job-portal can do for them, rather than just on how well the platform can match the job-seeker with a particular job.

3.1 Introduction

Youth unemployment is a policy priority throughout the developing world. In India, the importance of solving the youth unemployment problem was buoyed by 2017-18 job numbers which revealed that youth joblessness in both urban and rural areas had spiked to approximately 18% (Slater, 2019). These very high joblessness rates appeared in a context where youth unemployment and the skills gap were already a governmental priority highlighted in the country’s eleventh five year plan (Planning Comission, 2012). The policy response sought by the government has been to expand skills training, with a goal of reaching 500 million youth by 2022. However, it is unclear whether expanding skills training will be associated with new jobs: recent data from Orissa, Andhra Pradesh, and Maharashtra suggests that only 16%, 41%, and 35% of new graduates are wage or self-employed (National Knowledge Comission, 2009).

Search frictions may explain part of this gap. They also may disproportionately affect the young (who do not have well-established work networks or detailed CVs) and the poor (who may lack the resources to pay the costs of undirected search). A number of recent papers (e.g. Abebe et al. (2018a,b); Beam (2016); Abel et al. (2019); Crepon et al. (2019); Carranza et al. (2018)) test for the effects of reducing search frictions on employment in developing countries through interventions aimed at increasing matches (Abebe et al., 2018b; Beam, 2016); providing information about applicants (Abebe et al., 2018a,b; Carranza et al., 2018), or paying search costs (Abebe et al., 2018a; Crepon et al., 2019). These researcher-led innovations often generate changes in the types of jobs acquired by at least some workers, but in general have had more muted impacts on employment rates (McKenzie, 2017; Crépon and van den Berg, 2016).

At the same time, modern communication technology has provided a market solution to the problem of incomplete information in job search. In the last few years internet access in India has almost tripled (World Bank Data, 2017), which has facilitated the expansion of online job portals such as Naukri, Monster, Indeed, Quikr, and Shine, many of which can push SMS messages to users who may not be active on the internet. Evidence from the US on the role of online job search has been somewhat mixed. In the early years of internet job search (1998-2000), Kuhn and Skuterud (2004) find that search durations were if anything longer for internet users once observable characteristics were controlled for. More recently, Kroft and Pope (2014) find no evidence that the expansion of Craigslist.com reduced the unemployment rate. By contrast, Kuhn and Skuterud (2004) find that by 2005-2008 online job search was associated with large (25%) reduction in unemployment durations. They conclude that this change in the relationships between internet, search and unemployment durations may be attributable to an improvement in the technological capabilities of online portals. Of course, absent exogenous variation in the use of internet search, it is difficult to distinguish secular trends in portal capacity from a variety of other explanations, including trends in the sophistication of portal users and changes in unobservable selection.

This paper investigates whether accessing an online job portal impacts search and employment outcomes in India using a randomized controlled trial. To this end, we partner with a new online job portal, and vocational training institutes that supplied a sample of 2,662 recent graduates. The portal we work with, Job Shikari.com[1], provides SMS information on low-skilled jobs to candidates registered on the portal. Firms who contract with Job Shikari pay for a certain number of SMSs to be pushed to the specific registrants that match the skill and locational requirements of the job. We enroll a randomly selected subset of graduates on the Job Shikari platform, and send them a brief message indicating that the portal will

[1]Job Shikari is no longer active as it was ultimately purchased by another job portal after study completion.

be contacting them with information about new jobs unless they opt out. This is our first treatment group. For a second randomly-selected subset of new graduates, we provide access to the portal *and* grant them a priority ranking within Job Shikari's algorithm. This priority ranking guarantees that their names will be among the first given to firms who are matched. We refer to this second sample as the treatment priority group. Our study ultimately resulted in 1 additional SMS message about job opportunities being sent to the treatment group, and an additional 16 messages being sent to the treatment priority group for a truly intensive information intervention. We can compare these two groups to control respondents who are not registered (by us) on Job Shikari in order to estimate 1) the causal impact of portal access on search and employment outcomes, and 2) how these employment responses change as respondents receive more information from the portal.

We find a strong, but unexpected response to enrollment on the portal: new graduates are 9 percentage points less likely to be working 6 and 12 months later when they have access to Job Shikari. This result is consistent with the results in (Kuhn and Skuterud, 2004), though larger and statistically precise. We also show that a steady stream of information for the treatment priority job seekers results in these graduates "catching up" to the control group. Treatment priority job seekers are only 4 percentage points less likely to be employed than control. This reversal in employment rates relative to the treatment group is statistically significant. We also find that this stream of information is productive in one particular - and important - dimension of search: treatment priority responses are 6 percentage points more likely to live in a city, suggesting that one piece of useful information provided by the portal may be the spatial location of jobs (Bryan, Chowdhury, and Mobarak, 2014; McKenzie, 2017).

While these results were initially surprising to the research team, they are consistent with seminal models of job search (McCall, 1970; Jovanovic, 1979) if these graduates have inaccurate expectations about the effectiveness of the portal. Indeed, perceptions of access to new sources of job opportunities should boost reservation wages, and reduce employment if those job opportunities fail to materialize. Interestingly, there is a growing body of work documenting that job-seekers misjudge their prospects on the labor market. Krueger and Mueller (2016) find evidence that workers in New Jersey do not accurately predict their probability of employment, while Abebe et al. (2018b) find that job-seekers have reservation wages well above what firms in the area can offer. Finally, Rasul et al. (2019) find evidence of inaccurate expectations about job arrival rates in Uganda. However, the treatment priority group may have been able to overcome this impact by updating their perceptions of the new job arrival rate more effectively as they received more information about the characteristics of the jobs offered by the portal, which in general were low paid and located in urban areas. Alternatively, treatment priority job-seekers may also have accepted more job-offers and formed new matches. Examining our data, we find that the positive employment effects are strongest for treatment priority job seekers who are more likely to be impatient with the new technology: most significantly, older job seekers, but also married and lower caste job seekers. This plus descriptive data on the geographic distribution of jobs and respondents leads us to conclude that the primary impact of the priority treatment is a quicker reconciliation of beliefs with the actual effectiveness of these portals.

The remainder of this paper is organized as follows: Section 3.2 discusses the proliferation and use of job portals in India, as well as the expansion of vocational training programs. Section 3.3 details the field experiment and the data we collected. Section 3.4 discusses each of our results. We then discuss the implications of the findings and conclude in the final section.

3.2 Context

3.2.1 The Internet and Job Portals in India

India has one of the largest and fastest growing populations of internet users in the world. An estimated 500 million Indians use the internet, up from 6 million in 2001 (World Bank Data, 2017). Internet usage is growing by almost 20% annually as millions of Indians go online every day, using the Internet to make purchases, access financial services and education, and interact with friends and family (Kantar IMRB, 2018). The proliferation of the internet has also made it an increasingly popular tool for millions of job seekers searching for work throughout the country. At the forefront of this surge, are job portals connecting prospective employees with potential employers. The popularity of these portals has grown significantly in the last 10 years, evidently responding to the growing need to better link these two groups. In India, there are over 10 job portals currently operating nation-wide that are trying to meet job-seekers' demand for alternatives solutions to establishing connections with prospective employers (including LinkedIn, Naukri, Shine, Monster, Indeed, Timesjob, Quikr).

While the use of the internet and the prevalence of job portals has surged, the existing platforms are primarily focused on servicing white collar jobs including marketing/sales roles, IT (software-engineer) positions, and banking/finance jobs. When we launched our experiment, Job-Shikari was one of the only job-portals that was advertising employment opportunities for job-seekers from lower socio-economic backgrounds. It is also important to note that despite being easily accessible, job-seekers' efficiency at using these platforms likely varies dramatically by socio-economic status. As we will see below, the majority of job-seekers in our sample are from lower socio-economic strata and may not have had substantial experience with other job-portals.

3.2.2 Vocational Training Institutes in India

Our sample of job-seekers is drawn from vocational training institutes. These vocational institutes are part of a program that the Ministry of Skill Development and Entrepreneurship (MSDE) launched in July 2015. The Skill India Campaign was designed to dramatically improve access to high quality skills training as quickly and efficiently as possible. As part of these efforts, they created the National Skill Development Corporation - a public-private partnership that supports NGO's, business associations, skill development organizations and private companies in their efforts to bolster vocational training initiatives (National Skill Development Corporation, 2019). One of the NSDC's largest programs is the Pradhan Mantri Kaushal Vikas Yojana (PMKVY) scheme, which encourages youth to sign up for training programs by offering them monetary rewards upon successful completion of the program. Trainees are instructed to sign up for a course of their choosing that is offered by one of the NSDC's training partners. Students are required to complete the program and take an assessment in order to receive the funds from the NSDC, which are deposited directly into their bank accounts. The assessments are carried out by appointed Assessment Agencies, which operate independently of the training provider, to ensure impartiality. Nevertheless, placement rates for the 1 million graduates per year from existing government affiliated institutes are low. Available data from three states, Orissa, Andhra Pradesh and Maharashtra, show that only 16%, 41% and 35%, respectively, of new graduates were wage or self-employed as of 2009 (National Knowledge Comission, 2009). Recent work by Banerjee and Chiplunkar (2018) suggests that part of this phenomenon may be explained by the inefficient placement services that graduates receive upon completion of the training.

3.3 Experimental Design and Data Collection

3.3.1 Design

Our experiment was run in partnership with a job-matching platform, Job Shikari, where employers payed the company to post job opportunities, and the portal sent SMS messages to qualified candidates. Employees then followed up directly with the employer to set up interviews. All SMSs were constructed in the following way: *JOB: Data Entry Operator in Delhi, Salary 11500 Rupees, Please call +91******** - www.Job Shikari.in*. Interested job-seekers contacted the phone number listed in the SMS to proceed with the next stages of the interview process. Job Shikari was not involved in any of these later stage discussions.

Our sample of job-seekers was comprised of recent graduates from vocational training institutes that offer courses through the NSDC's PMKVY program. We randomly assigned these graduates to one of three groups, stratifying by location and their registered trade. Job-seekers assigned to the treatment group were uploaded to the platform, and received a welcoming SMS introducing them to the platform. While they were then eligible to receive many SMSs about various job opportunities, in practice they did not receive more than 1. Job-seekers assigned to the treatment priority group were provided with priority access on the portal by appearing first on the list of job-seekers who matched the portal's query for a given job opportunity. Job Shikari would only send text messages to the first 100 job-seekers that matched their query for a particular job. As a result, appearing first on these lists meant receiving many more SMSs about new job opportunities. Finally, job-seekers in the control group were not enrolled on the platform at all.

Table 3.1 presents the number of SMSs that job-seekers in our sample received by treatment status. The impact for job-seekers in the treatment group relative to the control group is captured by the "treatment" indicator, while the impact for job-seekers in the treatment priority group relative to control group is recovered by adding the "treatment" to the "treatment priority" group indicator. While job-seekers in the treatment group were eligible to receive more text messages, the portal placed a premium on sending job advertisements to the treatment priority group. Column 1 demonstrates that job-seekers in the treatment group received 1 text message on average, while job-seekers in the treatment priority group received an average of 17 text messages over the course of the year. This translates into a 32 percentage point increase in the probability of receiving a text message in the case of the treatment group, and a 58 percentage point increase in the probability of receiving a text message in the case of the treatment priority group (Column 2). Finally, we also surveyed job-seekers about the number of text messages they received. Treatment job-seekers were 4 percentage points more likely to report receiving a text message about job opportunities than the control group. This number increases to 8 percentage points in the treatment priority group. This last column reveals that some of the text messages we were sending to job-seekers were likely being overlooked. This is not altogether surprising as most marketing initiatives in India overload subscribers with text messages, and job-seekers may have missed some of the specific ones sent by Job Shikari.

3.3.2 Data

The NSDC agreed to provide the names and contact information of recent PMKVY graduates from 98 training institutes spanning the entire country, which meant that we received information for over 829,812 recent graduates from October 2014 to March 2015. We only selected training centers that belonged to one of 4 pre-selected trades (Telecom, Logistics, Sales and Security). These trades had the most employers and the highest rate of job offers on Job Shikari's portal. Next, we restricted the sample to job seekers located in India's (broadly) Northern States, in order to avoid any language barriers between the enumerators and the respondents. This includes: Delhi, Haryana, Punjab, Uttar Pradesh, Rajasthan, Uttarakhand, Chandigarh, Maharashtra, Madhya Pradesh, Bihar and Himachal Pradesh. We categorized these states into four broad geographic zones: Delhi-NCR, North, South-West and East India. Finally, within each training center, we randomly selected 30 graduates to call. This resulted in a sample of 15,268 observations. Over 80% of the calls did not lead to a completed interview: the phone numbers either did not exist, no one would pick up, or the number did not belong to the respondent we had on file. We ultimately completed 2,662 surveys between April and July 2015.

Once the surveys were completed for all respondents, we randomly assigned them to one of the two treatment groups, and the control group. We stratified by geographic zone and trade. Job-seekers assigned to the treatment group were uploaded to Job Shikari and received a text message introducing them to the platform. Job seekers in the treatment priority group were also uploaded to the platform and received the introductory text message. They were, however, more prominently featured on the platform and received more SMSs about job opportunities. Job-seekers in the control group were not uploaded to the platform.

We proceeded with a midline survey 9 months later, between December 2015 and April 2016, and we managed to reach 83% of respondents. The endline was conducted between June and September 30, 2016, and we managed to reach 71% of respondents. We do not see differential attrition (Table C.1). All of the surveys were completed over the phone. Enumerators first sent an SMS to each respondent. The message included a greeting, a brief sentence about the research, and details of the mobile recharge they would receive upon completing our surveys. The enumerators then followed up by calling the respondents directly. An integrated digital data collection software system was custom made for the purposes of this research project. The software used a cloud telephonic service, Exotel, to make calls to the respondents. To minimize enumerator errors, the software produced a list of numbers for each enumerator to call, which updated periodically. It included the phone numbers of people we had not reached yet, as well as the numbers of people who had stopped the survey mid-way through and needed to be contacted again to complete the survey. Finally, the software processed the mobile recharges upon successful completion of the surveys. Conducting surveys over the phone made it possible to reach job-seekers that were geographically spread out. It was, however, a challenge to reach our sample in a timely manner. Job-seekers were often busy, and did not always have the time to take the survey. The field team took special care to call job-seekers back multiple times and assure respondents that they would be immediately compensated for their time.

The surveys collected basic demographic information, including gender, education, language fluency, parents' education and occupation, caste and religion, and place of birth/current residence. We also collected detailed information on formal skills trainings, and employment history (including details about the trade, location, wage, and benefits of all jobs held in the last two years). Finally, we asked questions about job search strategies, the usefulness of respondent's occupational network, and knowledge/use of job portals. We also elicited their reservation wages.

Finally, we also rely on a dataset shared by Job Shikari that has every text message that was ever sent to job-seekers (both in and outside of our sample) over the course of the study period. This dataset presents the time and date the SMS was sent, the job-role that was being advertised, the location of the job, the educational requirements of the job, the wage offered, and the exact content of the text message that was sent.

3.3.3 Sample characteristics

Table 3.3 provides information about the basic demographic characteristics of the graduates in our sample. Most of our sample is male - only 11% of the respondents are female. The job-seekers we are working with have recently graduated from their vocational training institutes after completing a college degree. They are relatively young, approximately 23 years of age, and only a third of the sample is married. These vocational training programs typically cater to households from disadvantaged backgrounds, and over 60% of our sample comes from STs, SCs, or OBCs. Only 30% of the sample is employed, and 65% say they are actively looking for work. While the vast majority of graduates have access to the internet (approximately 76-80%), and many say they use the internet to find job opportunities, fewer than 35% were able to provide the name of an active job portal when we asked. We try to capture job-seekers' reservations wages, which suggest that 12,000 rupees (172 USD) is the lowest wage rate job-seekers in our sample would be willing to accept for a particular type of job. Note that while most of the differences between each treatment group are small and insignificant, there remains some imbalance across caste, reservation wage, and access to the Internet. Table C.2 in the Appendix presents the same balance table for the NSDC sample after controlling for strata, training institute, and education. The imbalance we see across these few observables largely disappears, confirming the importance of including individual fixed effects in our main specifications.

Table 3.4 provides some basic statistics about the jobs in our sample. Taken together the evidence demonstrates that the jobs on the portal are typically low wage, urban jobs. Over 90% of the jobs only require a high-school education, and the average salary being offered is 9863 rupees per month (141 USD). As can be seen in Figured 3.1 the job advertisements were primarily for positions such as data entry operators, telecallers, and field executives (who perform a variety of administrative roles related to sales). The figure is subdivided into four panels, where each panel demonstrates the number of SMSs about a particular job role going out to job-seekers in our sample based on the geographic zone they were drawn from at baseline (North, South-West, East and Delhi-NCR). The job-seekers in Delhi-NCR are receiving relatively more data entry positions, while job-seekers in the other three zones are seeing just as many, if not more, field executive roles. Finally, Figure 3.2 displays the locations of the jobs that Job-Shikari was advertising throughout the study period. The figure is subdivided into the same four panels, where each panel demonstrates the location of jobs and job-seekers in our sample based on the geographic zone they were drawn from at baseline. Job-seekers in the Delhi-NCR region (panel a) are mostly concentrated in Delhi, as are the jobs they are receiving text messages about. Job-seekers in the North (panel b) are primarily seeing jobs in Delhi as well. Job-seekers in the South (panel c) and the East (panel d) are seeing jobs in Delhi, but also in Mumbai and a few in Kolkata.

Figure 3.3 presents the average wage offers from the Job Shikari platform relative to the baseline wages that we collected from job-seekers during our phone surveys. The figure has the same four panels we presented above. The jobs on the platform were advertising wages that were below the average accepted wages for job-seekers in Delhi. In the North, the distributions of wage offers and baseline wages look fairly similar. Finally, in the South and East the wage offers appear slightly higher. These figures suggest that job-seekers perceptions of what the job-portal can do for them may depend on the types of jobs they are seeing, as well as their own wage offers/reservation wages.

3.3.4 Estimation

We estimate the effects of our intervention by pooling the two follow-up survey rounds and running the following regression:

$$y_{it} = \beta_0 + \beta_1 T_{it} + \beta_2 TP_{it} + \gamma_i + \delta_t + u_{it}$$

where y_{it} is an outcome of interest for job-seeker i in time period t (baseline, midline or endline). T is a dummy equal to 1 if the job-seeker was assigned to either of our treatments; TP is an indicator equal to 1 for being in the treatment priority group specifically; γ represents individual fixed effects; and δ_t represents an time fixed effect (survey round). The coefficients of interest are β_1, which represents the average effect of being uploaded to the portal across all survey rounds, and β_2, which represents the average effect of receiving additional text messages as a result of being in the treatment priority group (relative to treatment) across all survey rounds. We cluster all regressions at the individual level (our unit of randomization).

3.4 Results

3.4.1 Main effects on Employment

We first investigate the impacts of joining a job portal on job-seekers' employment probabilities. A priori, the impact of job-portals on employment rates is ambiguous. On the one hand, job-portals are designed to ease search frictions by increasing the arrival rate of jobs. Job-seekers should respond to the notification that they were uploaded to a job portal by increasing their reservation wage, which will increase the probability that they are employed at a higher wage rate if these jobs materialize over the course of the study period. To the best of our knowledge only one paper in the US documents this type of effect (Kuhn and Mansour, 2014). On the other hand, if job seekers have unrealistic expectations about what job portals can do for them, and over-estimate the arrival rate of jobs, they may remain unemployed for longer durations as they wait for jobs that fail to materialize. Previous work by Kuhn and Skuterud (2004); Kroft and Pope (2014) suggests that job-portals have negative or no impacts on unemployment durations, respectively. The notion that job-seekers have incorrect beliefs about the arrival rate of jobs, or their employment prospects more broadly, has been increasingly documented in the literature. Krueger and Mueller (2016); Abebe et al. (2018b); Rasul et al. (2019), find evidence that workers do not accurately predict their probability of being employed in the short-run, their likelihood of receiving a favorable wage offer, and the arrival rate of jobs, respectively.

Table 3.5 presents our first results. Treatment job-seekers are informed about their enrollment on the job portal, and can only adjust their *expectations* about the new arrival rate of jobs. As such, we hypothesize that job-seekers will react to Job Shikari's enrollment notification by increasing their reservation wage as they anticipate gaining better access to higher paying jobs through the portal. However, the treatment group only received 1 text message on average. With so few SMSs, it is unlikely that the jobs they were expecting, materialized. This intuition is confirmed in our data. We we see that treatment job-seekers' labor supply decreases in response to the notification that they were uploaded to the Job Shikari portal: the probability that a job-seeker in the treatment group is employed falls by 9.2 percentage points (Column 1). This disemployment effect is muted in the treatment priority group, which only experiences a 4.8 percentage point decrease relative to control. This could be because the probability of receiving a wage offer above their reservation wage increases with the number of text messages they receive. Alternatively, as they receive more information about the types of jobs that the platform provides access to, they may be re-adjusting their expectations downwards.

We have suggestive evidence that reservation wages respond in the way we would predict. Table 3.6 presents the results on four different measures of beliefs that we gather from job-seekers. We find that being uploaded to the job portal leads to a small (1.3%) increase in reservation wages for the treatment group relative to control (Column 1). This is consistent with our hypothesis that job-seekers respond to the notification that they were uploaded to the portal by forming positive expectations about what the portal will do for them. We also hypothesize that beliefs about the effectiveness of the intervention can be reversed once job-seekers have more information about the types of jobs that the portal is providing access to, and the arrival rate of these jobs. We see that reservation wages decrease significantly in the treatment priority group, by approximately 4.1%. Similar trends, albeit not statistically significant, are visible in Column 2, which displays the treatment and treatment priority's assessment of the wages they can expect in their current location. In the last three columns of table 3.6 we ask job seekers to estimate the probability they can get a job in their current location that pays 10,000, 16,000 and 20,000 rupees respectively. The stronger negative effects on the treatment priority group seem to suggests that job-seekers who receive more text messages have a more negative assessment of their probability of actually getting a job at these wage rates.

Next, we look at the impacts of the portal on actual wages for employed job-seekers, and overall earnings for the full sample. Table 3.5 presents the results on log wages (column 2), wages (column 3) and earnings (column 4) - where we impute a zero for the earnings of unemployed job-seekers. The results suggest that being in the treatment group is associated with a negative impact on wages. The 11% decrease is difficult to interpret, however, because of selection: access to the portal changes who is employed, and only the employed report wages. It is conceivable that job-seekers in the treatment group experience higher reservation wages on average, and only the lower ability types decide to accept lower-paying jobs. This would reduce the average wages we see in the treatment group. The effect of treatment priority relatively to treatment is positive, but is not statistically significant. As treatment priority job-seekers receive more text messages about jobs, they adjust their expectations about what the job portal can do, thereby lowering their reservation wages and increasing the probability they accept a job. If more job-seekers (not just low-ability types) are accepting jobs, we might expect to see this positive result. The results on earnings (Column 4) follow the same trends as those on wages.

3.4.2 Main effects on Job-Search

We consider two main dimensions of job-search. First, we look at where people are living. Specifically, we are interested in whether job-seekers have moved to urban centers in response to the interventions. We find that job-seekers in the treatment group are slightly less likely to be located in a city, but this result is small and imprecise (Column 1, Table 3.7). These particular job-seekers were not receiving any information about where the jobs were located. It is unlikely that the notification of being uploaded to the portal could have induced a strong migratory response from this sample. In the treatment-priority group, however, we see many more job-seekers moving to the city. Treatment priority job-seekers are 6 percentage points more likely to be in cities relative to treatment, and 4 percentage points more likely than the control group. This sample was seeing many more jobs, advertised exclusively in urban hubs including Delhi, Mumbai and in some cases Kolkata. This may have prompted a subset of job-seekers to migrate to areas they felt more confident would have significantly more employment opportunities.

We also look at whether job-seekers spend more or less time searching and applying for jobs when they have access to the job portal. We find that being uploaded to the portal has no significant impacts on whether job-seekers are searching for a job, the number of hours that they search, or the number of applications that they submit (Table 3.8). This is true for both employed and unemployed job-seekers. While these results suggest that job-seekers are not spending more or less time searching for work, we interpret these results with some caution. We are not able to measure job-seekers' intensity of search on different platforms, and it could be that Job Shikari changes the amount of effort and emphasis job-seekers place on certain job-search tools over others.

3.4.3 Heterogeneity effects on Employment

We have shown that job-seekers have biased beliefs in our sample: everyone responds to being added to the portal by increasing their reservation wage, which reduces the probability that they are employed as jobs fail to materialize. This highlights a fundamental challenge with these matching platforms as a whole because it suggests that their effectiveness depends on job-seekers' beliefs, which will then affect their reservation wages. If job-seekers' beliefs about what the portal can do are wrong, the portal may lead to lower match rates, as they do in our data. Nevertheless, job-seekers' beliefs about the effectiveness of the intervention can be updated if they have more information about what the intervention is actually doing. We see evidence of this among our treatment priority group graduates. This process of 'updating' is going to differ, however, based on how job-seekers are processing the incoming information. This heterogeneity in job-seekers' responses is what we investigate next.

We begin by examining whether the treatment *priority* effects differ by the different geographical zones that workers come from. We stratified on geographic zone, as we anticipated that the appeal of the Job Shikari jobs might differ based on job-seekers' location of origin. A job-seekers' location affects how far they have to move for the jobs they see on the portal, and how attractive these jobs are because of spatial differences in the wage setting. This is confirmed by our data, which shows that the distribution of wage offers is differentially attractive across geo-zones because of the prevailing baseline wage in the given location (Figure 3.3). Wage offers from Job Shikari are more appealing in the SouthWest and East where the baseline wages are lower. Table 3.9 presents the main results on employment broken out by geographic zone. We see that the negative employment effects are visible in the full sample of treatment graduates, albeit strongest in the SouthWest. The effect of being in the treatment priority group, however, is only detected in the SouthWest and the East. Job-seekers in the SouthWest experience a 8.9 percentage point increase in the probability of being employed relative to treatment job-seekers from the same zone. Similarly, job-seekers in the East experience a 7.5 percentage point increase in the probability of being employed relative to treatment. The treatment priority effects in the North and DelhiNCR are approximately 0 and imprecisely estimated.

These results are surprising at first glance because many more SMSs were sent to job-seekers in DelhiNCR and the North (Table 3.2). On average job-seekers in the SouthWest received 4 SMSs while job-seekers in the East received 13 SMS. This is relatively low compared to the sample in Delhi who received upwards of 50 SMS and, to a lesser extent, the sample in the North that received 19 SMS. To validate that the results we see on employment are indeed concentrated among those who received SMSs in these regions, we include an indicator for whether the treatment priority job-seekers received 0 SMS in our standard specifications. If the treatment priority job-seeker did not receive any SMSs, we would expect them to look similar to the treatment group. As such, adding the estimates of being in the treatment priority group and being in the treatment priority group but receiving no SMS, should yield a coefficient close to zero. Table 3.9 displays the treatment effects broken up by geo-zone, including an indicator for whether the treatment priority job-seeker received 0 SMS ("TP0"). The coefficient on TP0 for the SouthWest is negative and of similar magnitude to the treatment priority group, which is consistent with what we would expect. The estimate on the East is negative as well but imprecisely estimated. This confirms that the employment effects are concentrated among those in the treatment priority group who are receiving SMSs.

The job-seekers' geo-zone matters because it determines how far they have to move to find the advertised jobs, and how attractive those jobs look based on their previous wages. Nevertheless, geography also matters because job-seekers demographic characteristics differ significantly across regions. Table C.3 confirms that job-seekers in the SouthWest for example are much older than those in DelhiNCR. Similarly, there are many more general caste job-seekers in DelhiNCR than anywhere else. Given the paucity of jobs being sent to job-seekers, it is likely that their personal characteristics more heavily influenced their beliefs about what they could expect on the job-portal, rather than the features of the jobs themselves (e.g distance and wage offers). Furthermore, we can show that the differences between geographical zones are explained by a set demographic characteristics. Table 3.13 presents regressions of our main outcomes of interest (employment status, living in a city, and hours spent searching) broken out by geo-zone. The treatment priority effects for employment are particularly pronounced in the South and East, while the migration results are stronger for treatment priority job-seekers in the North. The next three columns take these same regressions and control for a predicted geo-zone measure, which we constructed by regressing indicators for being in these zones on a set of demographic characteristics (including age, gender, caste, education, martial status, religion and whether they lived in a village at baseline). The treatment priority effects across zones are no longer significant when we do so. This confirms that differences across geo-zone are explained by these demographic characteristics, the importance of which we investigate next.

It is difficult to draw broader policy conclusions from the fact that people in the South-West react more or less strongly than their counterparts in the North for example. Rather, it is more informative to learn about whether different socio-demographic characteristics affect job-seekers' responses to these types of matching interventions. Individual characteristics (age, marriage, gender, education, rural/urban) should affect how responsive job-seekers' beliefs and reservation wages are to the information from the job portal. We anticipate that this, in turn, will affect how likely certain types of job-seekers are to be employed throughout the study. In our sample we have two measures in particular that are relevant, namely job-seeker's age, and their marital status. Indeed, we anticipate that older and married job-seekers cannot be as patient, and may revise their beliefs about the portal more quickly than their younger, single counterparts. Table 3.14 shows that age does indeed matter. Column 1 presents the results for older job-seekers who are above 23 years of age (the mean age in the sample). Column 2 presents the results for younger job-seekers who are below 23 years of age. We see that the older treatment priority group is much more likely to react to the text messages than the younger job-seekers. The negative employment effect disappears altogether from this older treatment priority group (adding the treatment and treatment priority coefficients yields a point estimate of 0.014). Moreover, the difference between younger and older treatment priority job-seekers is large (7 percentage points) and statistically significant (Column 3). The impact of age becomes even more stark as we focus on older cohorts (Figure 3.4). The coefficient on treatment priority increases from 0.09 to 0.20 as we move from 23 year olds to 28 year olds in the sample.

Next, we consider another important characteristic, namely marital status. Table 3.15 presents the results on employment status for job-seekers who are married (column 1) versus single (column 2). We see similar patterns as above, where married job-seekers are more likely to respond to treatment priority status. This cohort experience a 9.6 percentage point increase in employment relative to treatment. This number falls to 3 percentage points for single treatment priority job-seekers relative to treatment. While the impact for married job-seekers is less strong than for older ones, it remains statistically significant. This suggests that married job-seekers are more likely to revise their beliefs downwards, and accept a job, relative to single job-seekers. This is likely because they have dependents and cannot afford to wait as long. The difference between married and single treatment priority job-seekers is not statistically significant (column 3), but we take this as suggestive evidence that a characteristic like marital status can indeed matter.

There are other demographic characteristics that could affect how quickly job-seekers revise their beliefs about their employment prospects on this platform, which we present in the appendix. For example, we might expect that job-seekers who are based in villages at baseline to have fewer social connections in the cities, and fewer outside options for finding employment opportunities. As such, they may react to the additional information from the portal by lowering their reservation wages and accepting jobs more quickly. Similarly, lower caste households are typically at a disadvantage relative to others when it comes to finding jobs. While they may be overly-optimistic about their prospects on the portal at first, they are likely to adjust relatively quickly because the cost of unemployment is high. Finally, we might expect education to matter as well: highly educated households are typically richer and may be able to afford to wait longer. Our evidence is largely consistent with these trends. Table C.6 suggests that the disemployment effect is less pronounced among treatment priority job-seekers living in villages at baseline (though the difference is not statistically significant). Similarly the coefficient on treatment priority for low caste job-seekers is more positive (by a factor of 10) than the effect for general caste (Table C.7). Finally, the impact of receiving more information from the portal is largely the same between high and and low educated job-seekers (Table C.8). This is perhaps not altogether surprising in our sample because the level of education remains low across the board.

3.4.4 Heterogeneity effects on Job Search

Overall, the employment results suggest that the heterogeneity we are seeing revolves around structural features of search rather than on just how well the job-seeker matches to the jobs on the portal. We investigate similar dynamics for job-search behavior. Specifically, we have characteristics that are typically associated with lower migration costs: namely age, marital-status, and caste. Young, single job-seekers typically do not have families that depend on them. Lower-caste households often have fewer job-prospects in any location, and may be more prepared to migrate in search of employment as the cost of unemployment is high. Our results demonstrate that the migration effects are stronger among these subgroups, but not significantly so. Younger job-seekers are 6 percentage point more likely to migrate than their older counterparts (Column 6, Table 3.14). Similarly, single job-seekers are 2 percentage points more likely to migrate than married job seekers (Column 6, Table 3.15). Finally, lower caste households are 1 percentage point more likely to migrate (Column 6, Table C.7). While these dimensions of heterogeneity are not as strong as they were for employment, they do suggest that job-seekers decision to migrate in response to the information from a job portal can depend on the their inherent characteristics. Any policy that tries to improve match rates will need to take this into consideration if it is to have it's intended impacts.

3.5 Conclusion

According to the International Labour Organization, global youth unemployment rates currently stand at 13%. Finding jobs for this segment of the labor market is of growing concern worldwide, but primarily in developing countries where the consequences of unemployment are often more severe.[2] Paradoxically in countries like India, employers also complain that they cannot fill positions. According to the Federation of Indian Chambers of Commerce and Industry (2011), employers frequently complained about the difficulty of filling vacant positions despite pervasive unemployment among semi-skilled laborers, and the glut of recent technical and vocational graduates.

A candidate explanation for this skills gap is search frictions: firms may not succeed in hiring qualified workers because of the difficulties generating a match between a worker and a job. Traditionally job seekers have had to rely on social connections and informal networks to find employment opportunities. These job search methods are costly and tend to favor the well-connected, thereby entrenching existing inequalities and preventing companies from finding the best worker for a vacancy (e.g. Beaman and Magruder, 2012; Loury, 2006; Ioannides and Loury, 2004; Magruder, 2010; Munshi 2003; Wang, 2011). Recently however, job portals - which connect prospective employees with potential employers - have emerged to provide a potentially new technological solution to this problem. Job portals in India work in a variety of ways: some allow applicants to post CVs online, or firms to post vacancies. Most create an algorithm to match workers with firms and then connect the two either on-line or through SMS. Compared to informal networks, which may exclude disadvantaged groups (Calvo-Armengol and Jackson, 2004) and which may fail to transmit information across geographical space, job portals potentially foster the connections for equitable access to growth. Given that India has one of the largest and fastest growing populations of internet users in the world (an estimated 37% of Indians use the internet, up from 7% in 2001) there is reason to believe that job portals may help smooth search frictions sharply, reducing vacancy rates and increasing access to jobs for the poorly connected.

To date, there has been little focus on understanding how job portals actually benefit job seekers in the developing world. This project provides a rigorous assessment of the value of job portals - examining their effect on labour market outcomes for job seekers, as well as on job search behaviors. To this end, we designed a randomized control trial that brings a randomly selected group of recent vocational training graduates onto a job portal. A second group of graduates are uploaded and given priority access on the portal, which means they received more text messages

[2]The region will continue to record the second highest percentage of youth in working poverty, after sub-Saharan Africa, at close to 50 per cent in 2016 (ILO, 2016)

about available job opportunities. A priori, it is difficult to predict the impact that the job-portal should have. On the one hand, portals should increase the number of job opportunities people see, thereby reducing search frictions in the labour market. In that case, we anticipate job-seekers will engage more actively in job-search activities and find employment more easily. On the other hand, the impacts of the platform will depend on it's ability to generate a match, on the types of people who are searching on these platforms, and their expectations about what the platform will do for them. These factors could potentially lead to negative or no impacts on employment and job-search. The literature is scarce but finds more evidence of the latter. Our results are consistent with these papers. We find that being uploaded to the job portal has a negative impact on the probability of being employed, but no significant impacts on job-search. The treatment priority group experiences a less strong disemployment effect, but is much more likely to respond to the treatment by migrating to urban centers. These results differ by job-seekers' observable characteristics such as age and martial status. We interpret our results as evidence that the impact of job-portals depends significantly on job-seekers beliefs about what the job-portal can do for them, rather than just on how well the platform can match the job-seeker with a particular job.

From a policy standpoint these results suggest that more needs to be done to set expectations for job-seekers joining these platforms, and to encourage job-seekers to be more active on these platforms so they can learn about the quality of jobs, and the distribution of offer rates, in the locations and trades they are interested in. Work by Kuhn and Mansour (2014) also suggests that as these portals improve their search algorithms, and job-seekers learn how to use them more effectively, their overall ability to generate positive matches will improve significantly.

Tables

Table 3.1: SMS receipt

	(1) Number SMS Received	(2) =1 SMS Received	(3) =1 Reported SMS Receipt
Treatment	1.437***	0.322***	0.040*
	(0.096)	(0.014)	(0.021)
Priority Treatment	16.405***	0.267***	0.044**
	(0.977)	(0.021)	(0.021)
Respondent Fixed Effects	Yes	Yes	No
Survey Round Fixed Effects	Yes	Yes	Yes
Number of Observations	7986	7986	6729

The dependent variables are an indicator for whether the respondent received an SMS from JS (column 1), the number of SMS sent by JS (column 2), the number of SMS they reported receiving (column 3). Standard errors are clustered at the respondent level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table 3.2: SMS receipt by Geo-Zone

	(1) Number SMS Received	(2) =1 SMS Received	(3) =1 Reported SMS Receipt
Treatment DelhiNCR	6.132^{***}	0.728^{***}	0.240^{***}
	(0.503)	(0.034)	(0.042)
Treatment North	1.043^{***}	0.319^{***}	0.063^{**}
	(0.101)	(0.024)	(0.028)
Treatment East	0.758^{***}	0.288^{***}	-0.020
	(0.087)	(0.024)	(0.029)
Treatment SouthWest	0.307^{***}	0.143^{***}	-0.008
	(0.059)	(0.021)	(0.032)
Priority Treatment	48.039^{***}	0.179^{***}	0.013
DelhiNCR	(4.537)	(0.041)	(0.061)
Priority Treatment	17.429^{***}	0.467^{***}	0.049
North	(1.130)	(0.032)	(0.036)
Priority Treatment	11.634^{***}	0.206^{***}	0.058
East	(1.279)	(0.038)	(0.035)
Priority Treatment	3.627^{***}	0.102^{***}	0.031
SouthWest	(0.863)	(0.037)	(0.043)
Respondent Fixed Effects	Yes	Yes	Yes
Survey Round Fixed Effects	Yes	Yes	Yes
Number of Observations	7986	7986	6729

The dependent variables are an indicator for whether the respondent received an SMS from JS (column 1), the number of SMS sent by JS (column 2), the number of SMS they reported receiving (column 3). Standard errors are clustered at the respondent level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table 3.3: Balance

	(1) Control	(2) Treatment	(3) Priority Treat-ment	(4) (1) vs. (2), p-value	(5) (1) vs. (3), p-value	(6) (2) vs. (3), p-value	(7) Joint test
=1 if male	0.86	0.89	0.88	0.04	0.20	0.52	0.13
Age	23.93	23.84	24.13	0.73	0.49	0.28	0.55
Education (Years)	14.17	14.22	14.29	0.66	0.30	0.50	0.58
Married Y/N	0.27	0.28	0.26	0.47	0.58	0.19	0.41
Religion=Hindu	0.92	0.94	0.94	0.07	0.19	0.70	0.18
Religion=Muslim	0.08	0.05	0.06	0.08	0.22	0.64	0.20
=1 if ST/SC caste	0.38	0.34	0.35	0.04	0.17	0.53	0.11
=1 if OBC caste	0.29	0.34	0.35	0.01	0.01	0.76	0.01
=1 if general caste	0.33	0.32	0.30	0.64	0.18	0.34	0.39
Father's education>0	0.80	0.83	0.81	0.19	0.66	0.39	0.40
Mother's education>0	0.55	0.58	0.52	0.36	0.30	0.04	0.13
=1 if live in village	0.49	0.48	0.48	0.68	0.98	0.70	0.90
Received formal skills training	1.32	1.34	1.34	0.29	0.54	0.68	0.56
=1 if currently employed	0.30	0.34	0.32	0.08	0.53	0.29	0.21
=1 if looking for job	0.65	0.66	0.65	0.69	0.95	0.74	0.91
Access to Internet Y/N (clean)	0.76	0.80	0.80	0.03	0.06	0.87	0.07
Reservation wage (winsorized)	12332.19	12305.14	13282.10	0.94	0.01	0.00	0.01
= 1 if Telecom	0.38	0.38	0.38	0.92	0.97	0.94	0.99
= 1 if Logistics	0.36	0.36	0.36	0.96	0.99	0.98	1.00
= 1 if SalesMarketing	0.18	0.18	0.18	0.99	0.99	0.98	1.00
= 1 if Security	0.08	0.09	0.08	0.90	0.99	0.91	0.99

Standard errors are clustered at the respondent level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table 3.4: Job Characteristics

	N = 88327
Wage	
mean	10025.8784719291
sd	2602.36182023563
median	9500
min	4000
max	90000
Education	
10th Pass	32,612 (36.92%)
12th Pass	12,530 (14.19%)
Graduate	2,070 (2.34%)
Missing	40,947 (46.36%)

Table 3.5: Employment

	(1) Employed	(2) Log(Wage)	(3) Wage	(4) Earnings
Treatment	-0.092***	-0.041	-1535.257*	-1203.818***
	(0.022)	(0.071)	(869.750)	(379.156)
Priority Treatment	0.048**	0.069	441.202	537.997
	(0.021)	(0.078)	(628.434)	(337.995)
Respondent Fixed Effects	Yes	Yes	Yes	Yes
Survey Round Fixed Effects	Yes	Yes	Yes	Yes
Number of Observations	6866	2311	2326	6475

Column 1, 4 include all respondents in the sample while column 2,3 only includes respondents who were employed at the time of survey. The dependent variables are an indicator for whether the respondent is employed (column 1), the log of wage (column 2), wage winzorised at the 1% (column 3), and earnings (imputing a zero for job-seekers who aren't working) (column 4). Standard errors are clustered at the respondent level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table 3.6: Beliefs

	(1) Reservation Wage	(2) Expected Wage	(3) Prob10	(4) Prob16	(5) Prob20
Treatment	187.539	108.962	-0.010	-0.165	0.099
	(278.021)	(503.221)	(0.216)	(0.192)	(0.175)
Priority Treatment	-582.858**	-2906.627	-0.399*	-0.239	-0.252
	(278.150)	(2250.205)	(0.207)	(0.187)	(0.178)
Respondent Fixed Effects	Yes	Yes	Yes	Yes	Yes
Survey Round Fixed Effects	Yes	Yes	Yes	Yes	Yes
Number of Observations	6504	6442	6337	6179	6097

Table 3.7: Migration

	(1)
	In-City
Treatment	-0.020
	(0.024)
Priority Treatment	0.060***
	(0.022)
Respondent Fixed Effects	Yes
Survey Round Fixed Effects	Yes
Number of Observations	6889

Column 1 includes all respondents in the sample. The dependent variables are an indicator for whether the respondent is currently living in a village (column 1). Standard errors are clustered at the respondent level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table 3.8: Job Search

	Unemployed			Employed		
	(1) Searching	(2) Hours	(3) Applications	(4) Searching	(5) Hours	(6) Applications
Treatment	-0.019	-1.357	-0.056	-0.023	0.022	-0.073
	(0.036)	(0.931)	(0.315)	(0.051)	(0.782)	(0.434)
Priority Treatment	-0.000	0.538	0.015	-0.024	-0.183	0.068
	(0.035)	(0.879)	(0.327)	(0.050)	(0.798)	(0.408)
Respondent Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Survey Round Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	4134	3848	3870	2694	2596	2505

Columns 1, 2 and 3 include all unemployed respondents in the sample, while columns 4, 5 and 6 include all employed respondents in the sample. The dependent variables are an indicator for whether the respondent is actively searching for employment (column 1/4), the number of hours spent searching in the past week - where people who aren't searching are assigned a value of 0 hours (column 2/5), the number of job applications submitted in the last 3 months (column 3/6). Standard errors are clustered at the respondent level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table 3.9: Employment by Geo-Zone

	(1) Employed	(2) Log(Wage)	(3) Wage	(4) Earnings
Treatment East	-0.051	0.135	-774.347	-424.456
	(0.037)	(0.258)	(2697.790)	(526.487)
Treatment DelhiNCR	-0.112*	-0.204	-966.679	-1764.012
	(0.059)	(0.143)	(2151.995)	(1560.625)
Treatment North	-0.087**	-0.057	-1645.476	-1604.703**
	(0.039)	(0.139)	(1759.242)	(713.562)
Treatment SouthWest	-0.144***	-0.023	-1997.511*	-1411.199**
	(0.046)	(0.090)	(1134.047)	(660.371)
Priority Treatment East	0.075**	0.147	1381.046	537.017
	(0.033)	(0.156)	(2179.476)	(492.131)
Priority Treatment DelhiNCR	-0.013	-0.007	-615.323	-793.814
	(0.055)	(0.069)	(1149.475)	(1224.249)
Priority Treatment North	0.016	0.189	1177.870	756.536
	(0.039)	(0.211)	(1458.986)	(644.556)
Priority Treatment SouthWest	0.089**	-0.021	4.454	956.908
	(0.044)	(0.091)	(696.587)	(634.144)
Respondent Fixed Effects	Yes	Yes	Yes	Yes
Survey Round Fixed Effects	Yes	Yes	Yes	Yes
Geo-Specifc Time Trend	Yes	Yes	Yes	Yes
Number of Observations	6866.0000	2311.0000	2326.0000	6475.0000

Column 1, 4 include all respondents in the sample while column 2, 3 only includes respondents who were employed at the time of survey. The dependent variables are an indicator for whether the respondent is employed (column 1), the log of wage (column 2), wage winzorised at the 1% (column 3), and earnings (imputing a zero for job-seekers who aren't working) (column 4). Standard errors are clustered at the respondent level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table 3.10: Migration by Geo-Zone

	(1)
	In-City
Treatment East	0.002
	(0.044)
Treatment DelhiNCR	0.046
	(0.032)
Treatment North	-0.038
	(0.043)
Treatment SouthWest	-0.061
	(0.048)
Priority Treatment East	0.029
	(0.042)
Priority Treatment DelhiNCR	0.014
	(0.033)
Priority Treatment North	0.120***
	(0.040)
Priority Treatment SouthWest	0.042
	(0.048)
Respondent Fixed Effects	Yes
Survey Round Fixed Effects	Yes
Geo-Specifc Time Trend	Yes
Number of Observations	6889

Column 1 includes all respondents in the sample. The dependent variables are an indicator for whether the respondent is currently living in a village (column 1). Standard errors are clustered at the respondent level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table 3.11: Employment controlling for TP0

	(1) Employed	(2) Log(Wage)	(3) Wage	(4) Earnings
Treatment East	-0.051	0.134	-777.520	-424.485
	(0.037)	(0.258)	(2700.079)	(526.614)
Treatment DelhiNCR	-0.112*	-0.204	-966.679	-1764.012
	(0.059)	(0.143)	(2153.395)	(1560.988)
Treatment North	-0.086**	-0.058	-1648.515	-1604.182**
	(0.039)	(0.140)	(1760.498)	(713.802)
Treatment SouthWest	-0.144***	-0.023	-1997.524*	-1411.344**
	(0.046)	(0.090)	(1134.794)	(660.406)
Priority Treatment	0.087**	0.278	1801.050	477.762
East	(0.039)	(0.207)	(2376.764)	(492.599)
Priority Treatment	-0.013	-0.007	-615.323	-793.814
DelhiNCR	(0.055)	(0.069)	(1150.223)	(1224.533)
Priority Treatment	0.011	0.058	757.796	632.780
North	(0.040)	(0.162)	(1450.042)	(656.438)
Priority Treatment	0.199***	0.050	138.606	2582.342**
SouthWest	(0.063)	(0.129)	(857.419)	(1290.245)
Priority Treatment	0.000	0.000	0.000	0.000
DelhiNCR, 0 SMS	(.)	(.)	(.)	(.)
Priority Treatment	-0.025	-0.242	-775.905	132.628
East, 0 SMS	(0.050)	(0.164)	(1320.196)	(561.692)
Priority Treatment	0.038	0.801*	2563.435	879.434
North, 0 SMS	(0.068)	(0.451)	(1759.033)	(886.533)
Priority Treatment	-0.152**	-0.104	-195.492	-2206.719*
SouthWest, 0 SMS	(0.063)	(0.137)	(810.754)	(1229.903)
Respondent Fixed Effects	Yes	Yes	Yes	Yes
Survey Round Fixed Effects	Yes	Yes	Yes	Yes
Geo-Specifc Time Trend	Yes	Yes	Yes	Yes
Number of Observations	6866	2311	2326	6475

Column 1, 4 include all respondents in the sample while column 2, 3 only includes respondents who were employed at the time of survey. The dependent variables are an indicator for whether the respondent is employed (column 1), the log of wage (column 2), wage winzorised at the 1% (column 3), and earnings (imputing a zero for job-seekers who aren't working) (column 4). Standard errors are clustered at the respondent level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table 3.12: Migration controlling for TP0

	(1)
	In-City
Treatment East	0.002
	(0.044)
Treatment DelhiNCR	0.046
	(0.032)
Treatment North	-0.038
	(0.043)
Treatment SouthWest	-0.061
	(0.049)
Priority Treatment East	0.032
	(0.050)
Priority Treatment DelhiNCR	0.014
	(0.033)
Priority Treatment North	0.113***
	(0.039)
Priority Treatment SouthWest	0.110*
	(0.065)
Priority Treatment DelhiNCR, 0 SMS	0.000
	(.)
Priority Treatment East, 0 SMS	-0.007
	(0.062)
Priority Treatment North, 0 SMS	0.049
	(0.089)
Priority Treatment SouthWest, 0 SMS	-0.093
	(0.071)
Respondent Fixed Effects	Yes
Survey Round Fixed Effects	Yes
Geo-Specifc Time Trend	Yes
Number of Observations	6889

Column 1 includes all respondents in the sample. The dependent variables are an indicator for whether the respondent is currently living in a village (column 1). Standard errors are clustered at the respondent level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table 3.13: Controlling for Predicted Geo-zone

	Main			Predict		
	(1) Emp	(2) City	(3) Hours	(4) Emp	(5) City	(6) Hours
Treatment East	-0.051	0.002	-1.470	-0.278	-1.438***	3.373
	(0.037)	(0.044)	(1.172)	(0.329)	(0.232)	(9.551)
Treatment DelhiNCR	-0.112*	0.046	2.534*	-0.325	-1.097***	5.998
	(0.059)	(0.032)	(1.513)	(0.321)	(0.219)	(9.120)
Treatment North	-0.087**	-0.038	-0.010	-0.314	-1.402***	4.006
	(0.039)	(0.043)	(1.082)	(0.333)	(0.230)	(9.516)
Treatment SouthWest	-0.144***	-0.061	-0.042	-0.354	-1.421***	4.227
	(0.046)	(0.048)	(1.239)	(0.332)	(0.227)	(9.625)
Priority Treatment	0.075**	0.029	-0.245	0.372	0.326	-12.411
East	(0.033)	(0.042)	(1.170)	(0.437)	(0.350)	(12.402)
Priority Treatment	-0.013	0.014	0.613	0.264	0.362	-9.742
DelhiNCR	(0.055)	(0.033)	(1.305)	(0.425)	(0.334)	(12.031)
Priority Treatment	0.016	0.120***	0.397	0.320	0.478	-11.002
North	(0.039)	(0.040)	(0.985)	(0.442)	(0.346)	(12.313)
Priority Treatment	0.089**	0.042	-0.995	0.335	0.429	-13.012
SouthWest	(0.044)	(0.048)	(1.230)	(0.440)	(0.344)	(12.583)
Respondent Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Survey Round Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	6866	6889	6444	5478	5479	5187

Columns 1, 2 and 3 represent the results from regressions of our main outcomes of interest (Employment, City, and Hours worked). Columns 4,5,6 represent the results from these same regressions when we control for the predicted geo-zone that job-seekers are in. Standard errors are clustered at the respondent level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table 3.14: Emp/City (Age)

	(1) Emp (AgeH)	(2) Emp (AgeL)	(3) Emp	(4) City (AgeH)	(5) City (AgeL)	(6) City
Treatment	-0.101***	-0.088***	-0.073***	0.005	-0.036	-0.031
	(0.037)	(0.026)	(0.024)	(0.038)	(0.030)	(0.027)
Priority Treatment	0.092***	0.021	0.021	0.019	0.087***	0.087***
	(0.034)	(0.026)	(0.026)	(0.034)	(0.029)	(0.029)
Treatment*AgeH			-0.050*			0.029
			(0.028)			(0.029)
Priority Treatment*AgeH			0.072*			-0.068
			(0.043)			(0.045)
Respondent Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Survey Round Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	2713	4144	6857	2723	4156	6879

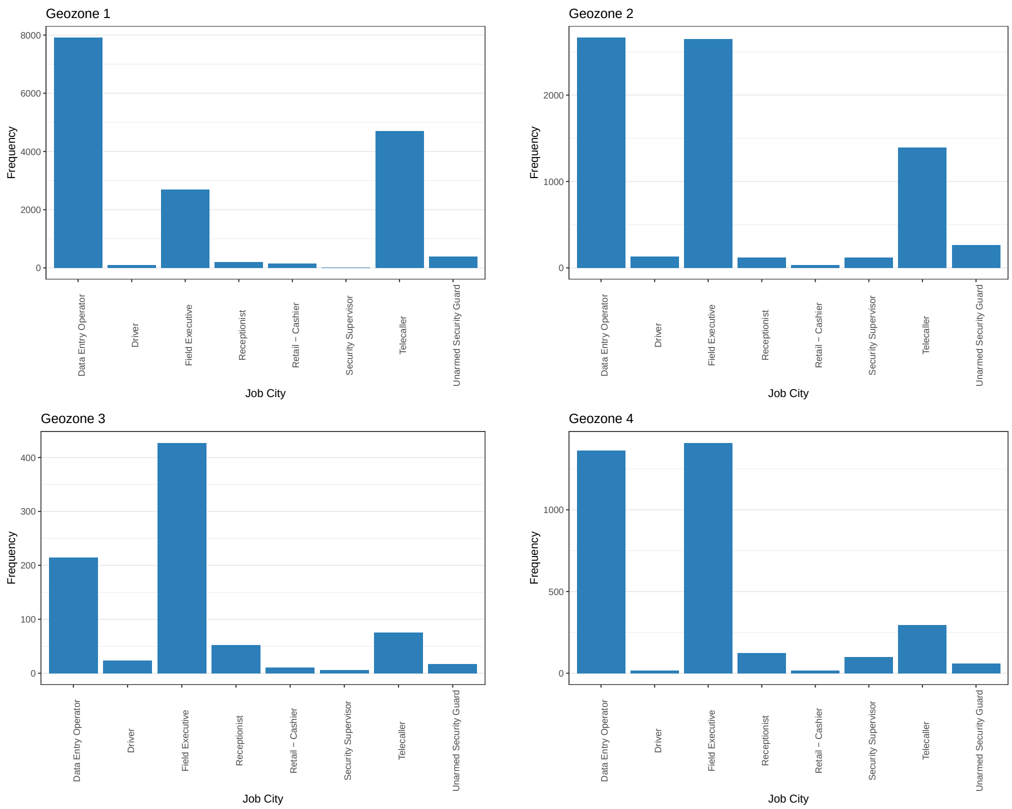
Table 3.15: Emp/City (Married)

	(1)	(2)	(3)	(4)	(5)	(6)
	Emp (Mar)	Emp (Single)	Emp	City (Mar)	City (Single)	City
Treatment	-0.134***	-0.075***	-0.073***	-0.012	-0.023	-0.039
	(0.041)	(0.026)	(0.024)	(0.044)	(0.028)	(0.025)
Priority Treatment	0.096**	0.030	0.030	0.041	0.069***	0.069***
	(0.040)	(0.024)	(0.024)	(0.043)	(0.026)	(0.026)
Treatment*Mar			-0.066**			0.073**
			(0.029)			(0.031)
Priority Treatment*Mar			0.065			-0.029
			(0.047)			(0.051)
Respondent Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Survey Round Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	1850	5006	6856	1857	5021	6878

Figures

Figure 3.1: Job types by Geozone

(a) DelhiNCR (b) North

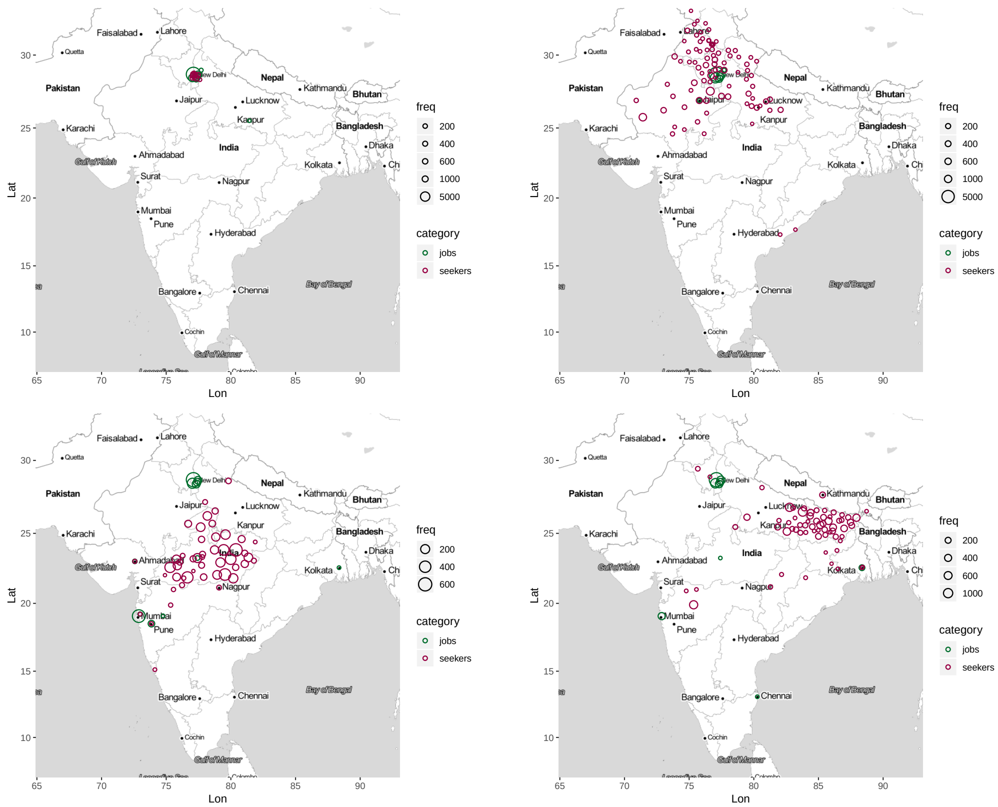


(c) South (d) East

Figure 3.2: Job seekers and job offers by Geozone - Pincode

(a) DelhiNCR

(b) North

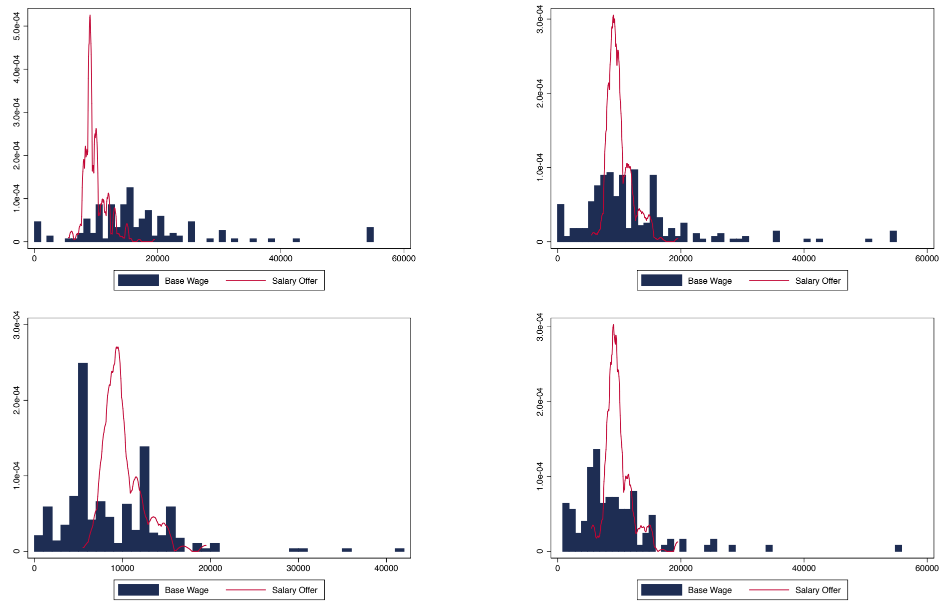


(c) South

(d) East

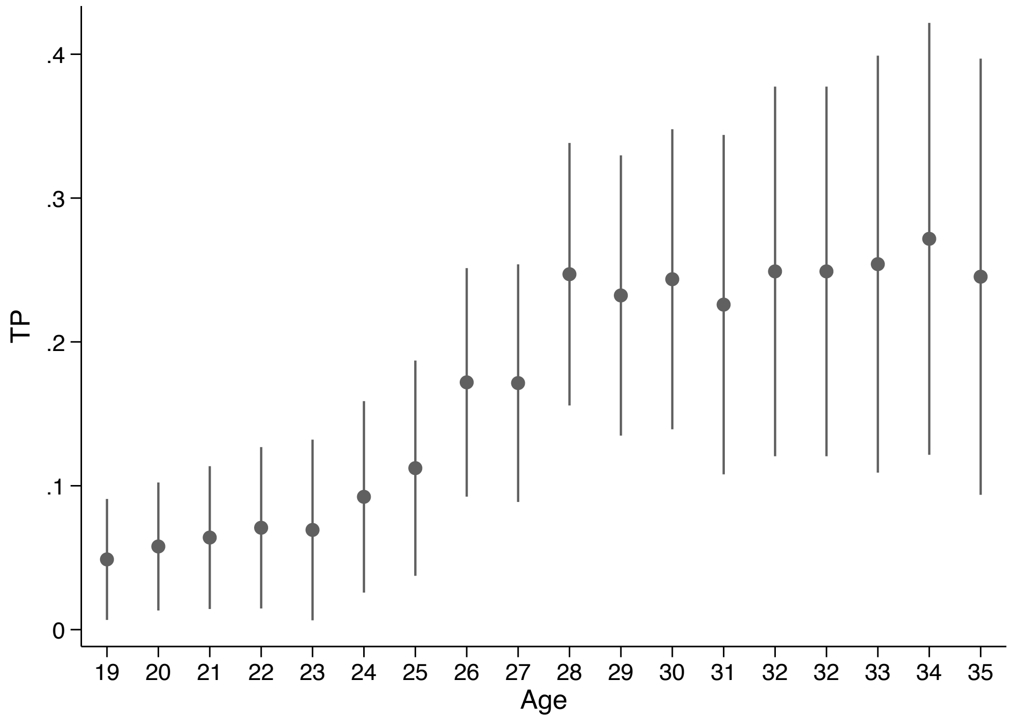
Figure 3.3: Base Wages and Salary Offers by Geozone

(a) DelhiNCR (b) North



(c) South (d) East

Figure 3.4: TP by age



Bibliography

Abebe, Girum, Stefano Caria, Marcel Fafchamps, Paolo Falco, Simon Franklin, and Simon Quinn. 2018a. "Anonymity or Distance? Job Search and Labour Market Exclusion in a Growing African City." *Working Paper* .

Abebe, Girum, Stefano Caria, Marcel Fafchamps, Paolo Falco, Simon Franklin, Simon Quinn, and Shilpi Forhad. 2018b. "Job Fairs: Matching Firms and Workers in a Field Experiment in Ethiopia." *Working Paper* .

Abel, Martin, Rulof Burger, Eliana Carranza, and Patrizio Piraino. 2019. "Bridging the Intention-Behavior Gap? The Effect of Plan-Making Prompts on Job Search and Employment." *American Economic Journal: Applied Economics* 11 (2):284–301.

Acemoglu, Daron, Kostas Bimpikis, and Asuman Ozdaglar. 2014. "Dynamics of information exchange in endogenous social networks." *Theoretical Economics* 9 (1):41–97.

Akbarpour, Mohammad, Suraj Malladi, and Amin Saberi. 2018. "Just a Few Seeds More: Value of Network Information for Diffusion." *Unpublished* .

Anderson, Jock R and Gershon Feder. 2007. "Agricultural extension." *Handbook of Agricultural Economics* 3:2343–2378.

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.

Bandiera, Oriana and Imran Rasul. 2006. "Social Networks and Technology Adoption in Northern Mozambique." *The Economic Journal* 116 (514):869–902.

Banerjee, Abhijit, Emily Breza, Arun Chandrasekhar, and Benjamin Golub. 2018a. "When Less is More: Experimental Evidence on Information Delivery During India's Demonitization." *Unpublished* .

Banerjee, Abhijit, Arun G Chandrasekhar, Esther Duflo, and Matthew O Jackson. 2013. "The diffusion of microfinance." *Science* 341 (6144):1236498.

———. 2018b. "Using Gossips to Spread Information: Theory and Evidence from two Randomized Controlled Trials." *Unpublished* .

Banerjee, Abhijit and Gaurav Chiplunkar. 2018. "How Important Are Matching Frictions in the Labor Market? Experimental & Non-Experimental Evidence from a Large Indian Firm." *Working paper* .

Beam, Emily A. 2016. "Do Job Fairs Matter? Experimental Evidence on the Impact of Job-Fair Attendance." *Journal of Development Economics* 120:32–40.

Beaman, Lori, Ariel BenYishay, Mushfiq Mobarak, and Jeremy Magruder. 2015. "Can Network Theory based Targeting Increase Technology Adoption?" *Unpublished* .

Beaman, Lori and Andrew Dillon. 2017. "Diffusion of Agricultural Information within Social Networks: Evidence on Gender Inequalities from Mali." *Unpublished* .

BenYishay, Ariel and A Mushfiq Mobarak. 2015. "Social Learning and Incentives for Experimentation and Communication." *Unpublished* .

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.

Birkhaeuser, Dean, Robert E Evenson, and Gershon Feder. 1991. "The economic impact of agricultural extension: A review." *Economic Development and Cultural Change* :607–650.

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.

Bryan, Gharad, Shyamal Chowdhury, and Ahmed Mushfiq Mobarak. 2014. "Underinvestment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh." *Econometrica* 82 (5):1671–1748.

Bursztyn, Leonardo, Florian Ederer, Bruno Ferman, and Noam Yuchtman. 2014. “Understanding mechanisms underlying peer effects: Evidence from a field experiment on financial decisions.” *Econometrica* 82 (4):1273–1301.

Cai, J, A de Janvry, and E Sadoulet. 2015. “Social Networks and the Decision to Insure.” *American Economic Journal: Applied Economics* 7 (2):81–108.

Calvó-Armengol, Antoni, Joan Martí, and Andrea Prat. 2015. “Communication and influence.” *Theoretical Economics* 10 (2):649–690.

Carranza, Eliana, Robert Garlick, Kate Orkin, and Neil Rankin. 2018. “Job Search and Hiring with Two-Sided Limited Information about Workseekers’ Skills.” *Working paper* .

Centola, Damon. 2010. “The spread of behavior in an online social network experiment.” *Science* 329 (5996):1194–1197.

Chandrasekhar, Arun and Randall Lewis. 2016. “Econometrics of sampled networks.” *Unpublished* .

Chandrasekhar, Arun G, Benjamin Golub, and He Yang. 2016. “Signaling, Stigma, and Silence in Social Learning.” *Unpublished* .

Chernozhukov, Victor, Mert Demirer, Esther Duflo, and Ivan Fernandez-Val. 2018. “Generic machine learning inference on heterogenous treatment effects in randomized experiments.” Tech. rep., National Bureau of Economic Research.

Cole, Shawn, Xavier Giné, and James Vickery. 2017. “How does risk management influence production decisions? Evidence from a field experiment.” *The Review of Financial Studies* 30 (6):1935–1970.

Conley, Timothy G and Christopher R Udry. 2010. “Learning about a new technology: Pineapple in Ghana.” *American Economic Review* :35–69.

Crepon, Bruno, Romain Aeberhardt, Vera Chiodi, Mathilde Gaini, Anett John, and Augustin Vicard. 2019. “You Get What You Pay For: Evidence from a Jobseeker Conditional Cash Transfer Program in France.” *Working Paper* .

Crépon, Bruno and Gerard J. van den Berg. 2016. “Active Labor Market Policies.” *Annual Review of Economics* 8 (1):521–546.

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.

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.

Duflo, Esther, Michael Kremer, and Jonathan Robinson. 2011. "Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya." *American Economic Review* 101:2350–2390.

Emerick, Kyle, Alain de Janvry, Elisabeth Sadoulet, and Manzoor H Dar. 2016. "Technological Innovations, Downside Risk, and the Modernization of Agriculture." *American Economic Review* 106 (6):1537–1561.

Foster, Andrew D and Mark R Rosenzweig. 1995. "Learning by doing and learning from others: Human capital and technical change in agriculture." *Journal of Political Economy* 103 (6):1176–1209.

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.

Golub, Benjamin and Matthew O Jackson. 2012. "How Homophily Affects the Speed of Learning and Best-Response Dynamics." *Quarterly Journal of Economics* 127 (3):1287–1338.

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

Guiteras, Raymond, James Levinsohn, and Ahmed Mushfiq Mobarak. 2015. "Encouraging sanitation investment in the developing world: a cluster-randomized trial." *Science* 348 (6237):903–906.

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.

Hanna, Rema, Sendhil Mullainathan, and Joshua Schwartzstein. 2014. "Learning through noticing: Theory and evidence from a field experiment." *The Quarterly Journal of Economics* 129 (3):1311–1353.

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.

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.

Jovanovic, Boyan. 1979. "Job Matching and the Theory of Turnover." *Journal of Political Economy* 87 (5):972–990.

Kantar IMRB, Ltd. 2018. "Digital Adoption & Usage Trends." Report 21st Edition, Kantar.

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.

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

Kim, David A, Alison R Hwong, Derek Stafford, D Alex Hughes, A James O'Malley, James H Fowler, and Nicholas A Christakis. 2015. "Social network targeting to maximise population behaviour change: a cluster randomised controlled trial." *The Lancet* 386 (9989):145–153.

Kondylis, Florence, Valerie Mueller, and Jessica Zhu. 2017. "Seeing is believing? Evidence from an extension network experiment." *Journal of Development Economics* 125:1–20.

Kroft, Kory and Devin G. Pope. 2014. "Does Online Search Crowd Out Traditional Search and Improve Matching Efficiency? Evidence from Craigslist." *Journal of Labor Economics* 32 (2):259–303.

Krueger, Alan B. and Andreas I. Mueller. 2016. "A Contribution to the Empirics of Reservation Wages." *American Economic Journal: Economic Policy* 8 (1):142–179.

Kuhn, Peter and Hani Mansour. 2014. "Is Internet Job Search Still Ineffective?" *The Economic Journal* 124 (581):1213–1233.

Kuhn, Peter and Mikal Skuterud. 2004. "Internet Job Search and Unemployment Durations." *American Economic Review* 94 (1):218–232.

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.

McCall, J. J. 1970. "Economics of Information and Job Search." *The Quarterly Journal of Economics* 84 (1):113–126.

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).

McKenzie, David. 2017. “How Effective Are Active Labor Market Policies in Developing Countries? A Critical Review of Recent Evidence.” *The World Bank Research Observer* 32 (2):127–154.

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.”

Miguel, Edward and Michael Kremer. 2004. “Worms: identifying impacts on education and health in the presence of treatment externalities.” *Econometrica* 72 (1):159–217.

Mobius, Markus, Tuan Phan, and Adam Szeidl. 2015. “Treasure hunt: Social learning in the field.” Tech. rep., National Bureau of Economic Research.

Munshi, Kaivan. 2004. “Social learning in a heterogeneous population: technology diffusion in the Indian Green Revolution.” *Journal of Development Economics* 73 (1):185–213.

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

National Knowledge Comission, Gvt. 2009. “Report to the Nation.” Tech. rep., Government of India, New Delhi.

National Skill Development Corporation, Gvt. 2019. “National Skill Development Corporation Schemes and Initatives.” https://nsdcindia.org/.

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.

Planning Comission, Gvt. 2012. “Eleventh Fifth Year Plan.” Tech. Rep. 11, Government of India.

Reed, Rachel C. 2018. “Transportation Turned Performance Art: Nairobi’s Matatu Crews.” *New York Times* .

Rigol, Natalia, Reshmaan Hussam, and Benjamin Roth. 2017. “Targeting High Ability Entrepreneurs Using Community Information: Mechanism Design In The Field.” *Unpublished* .

Sacerdote, Bruce. 2001. “Peer effects with random assignment: Results for Dartmouth roommates.” *The Quarterly Journal of Economics* 116 (2):681–704.

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

Slater, Joanna. 2019. “India’s Job Crisis Is Worse than People Thought — and Its Government Tried to Squelch the Data.” *Washington Post* .

Suri, Tavneet. 2011. “Selection and comparative advantage in technology adoption.” *Econometrica* 79 (1):159–209.

Valente, Thomas W. 2012. “Network interventions.” *Science* 337 (6090):49–53.

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.

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

World Bank Data, Statistics. 2017. “Individuals Using the Internet (% of Population) | Data.” https://data.worldbank.org/indicator/IT.NET.USER.ZS?locations=IN.

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

A.1 Tables

Table A.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 A.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 A.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 A.4: Deviceon

	(1) Hours device on	(2) Hours device on	(3) Hours 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 A.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 A.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 A.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 A.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 A.9: Under-report

	(1)	(2)	(3)
	shade	shade	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 A.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 A.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.

A.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

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

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

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

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

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

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

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

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:

$$\begin{aligned}
\frac{\partial^2 L}{\partial r^2} &= 0 - \frac{\partial}{\partial r} \left(\frac{q^* - e}{r} \right) \left(\frac{q^* - e}{r} \right) f_\varepsilon \left(\frac{q^* - e}{r} \right) - h''_{rr} - 2\beta \\
&= \left(\frac{q^* - e}{r^2} \right) \left(\frac{q^* - e}{r} \right) f_\varepsilon \left(\frac{q^* - e}{r} \right) - h''_{rr} - 2\beta \\
&= \underbrace{\frac{1}{r} \left(\frac{q^* - e}{r} \right)^2 f_\varepsilon(\cdot) - h''_{rr}}_{S.O.C \, < \, 0} \\
\frac{\partial^2 L}{\partial e^2} &= f_\varepsilon(\cdot) \left(\frac{1}{r} \right) - h''_{ee} \\
&= \underbrace{\frac{1}{r} f_\varepsilon(\cdot) - h''_{ee}}_{S.O.C \, < \, 0} \\
\frac{\partial L}{\partial e \partial r} &= 0 - \left[-\frac{1}{r} \left(\frac{q^* - e}{r} \right) f_\varepsilon \left(\frac{q^* - e}{r} \right) \right] - h''_{er} \\
&= \underbrace{\left(\frac{q^* - e}{r^2} \right) f_\varepsilon(\cdot) - h''_{er}}_{< \, 0} \\
\frac{\partial L}{\partial r \partial e} &= \underbrace{-f_\varepsilon(\cdot) \left(\frac{q^* - e}{r^2} \right) - h''_{er}}_{< \, 0} \\
\frac{\partial L}{\partial T \partial e} &= -\frac{1}{r} f_\varepsilon \left(\frac{q^* - e}{r} \right) \\
&= \underbrace{-\frac{1}{r} f_\varepsilon (\cdot)}_{< \, 0}
\end{aligned}$$

$$
\begin{aligned}
\frac{\partial L}{\partial T \partial r} &= 0 - \frac{1}{r}\left(\frac{q^* - e}{r}\right) f_\varepsilon\left(\frac{q^* - e}{r}\right) \\
&= \underbrace{-\left(\frac{q^* - e}{r^2}\right) f_\varepsilon(\cdot)}_{>\ 0}
\end{aligned}
$$

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:

$$
\begin{aligned}
\frac{\partial e}{\partial T} &= -\underbrace{\frac{1}{\text{Determinant}}}_{+}\left[\overbrace{\frac{\partial^2 L}{\partial r^2}}^{-} \cdot \overbrace{\frac{\partial L}{\partial T \partial e}}^{-} - \overbrace{\frac{\partial L}{\partial r \partial e}}^{-} \cdot \overbrace{\frac{\partial L}{\partial T \partial r}}^{+}\right] \\
&< 0 \\
\frac{\partial r}{\partial T} &= -\underbrace{\frac{1}{\text{Determinant}}}_{+}\left[- \overbrace{\frac{\partial L}{\partial e \partial r}}^{-} \cdot \overbrace{\frac{\partial L}{\partial T \partial e}}^{-} + \overbrace{\frac{\partial^2 L}{\partial e^2}}^{-} \cdot \overbrace{\frac{\partial L}{\partial T \partial r}}^{+}\right] \\
&> 0
\end{aligned}
$$

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

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

Computing the additional terms

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

Putting it altogether:

$$\frac{\partial e}{\partial \alpha} = -\frac{1}{\underbrace{\text{Determinant}}_{+}}\left[\overbrace{\frac{\partial^2 L}{\partial r^2}}^{-}\cdot\overbrace{\frac{\partial L}{\partial \alpha \partial e}}^{-}-\overbrace{\frac{\partial L}{\partial r \partial e}}^{-}\cdot\overbrace{\frac{\partial L}{\partial \alpha \partial r}}^{+}\right]$$
$$> 0$$

$$\frac{\partial r}{\partial \alpha} = -\frac{1}{\underbrace{\text{Determinant}}_{+}}\left[-\overbrace{\frac{\partial L}{\partial e \partial r}}^{-}\cdot\overbrace{\frac{\partial L}{\partial \alpha \partial e}}^{-}+\overbrace{\frac{\partial^2 L}{\partial e^2}}^{-}\cdot\overbrace{\frac{\partial L}{\partial \alpha \partial r}}^{+}\right]$$
$$< 0$$

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

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

Putting it altogether:

$$\frac{\partial e}{\partial \beta} = -\frac{1}{\underbrace{\text{Determinant}}_{+}}\left[\overbrace{\frac{\partial L}{\partial r \partial e}}^{-}\cdot\overbrace{1}^{+}\right]$$
$$> 0$$

$$\frac{\partial r}{\partial \beta} = -\frac{1}{\underbrace{\text{Determinant}}_{+}}\left[\overbrace{\frac{\partial^2 L}{\partial e^2}}^{-}\cdot\overbrace{(-1)}^{-}\right]$$
$$< 0$$

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

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

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

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

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

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

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

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

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

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

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

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:

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

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

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

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

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

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

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

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

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:

$$
\begin{aligned}
\frac{\partial L}{\partial \beta \partial \lambda} &= \int_{\frac{q^*-e^*}{r^*}}^{\infty} -r f_{\varepsilon}(\varepsilon) d\varepsilon - \cancel{\frac{\partial}{\beta}\left(\frac{q^* - e^*}{r^*}\right)(q^* - T - \beta r)}\, f_{\varepsilon}(\cdot) \\
&\quad + \int_{0}^{\frac{q^*-e^*}{r^*}} -r f_{\varepsilon}(\varepsilon) d\varepsilon + \cancel{\frac{\partial}{\beta}\left(\frac{q^* - e^*}{r^*}\right)\left(\frac{1}{\alpha} - \beta r\right)} f_{\varepsilon}(\cdot) \\
&\quad = -r
\end{aligned}
$$

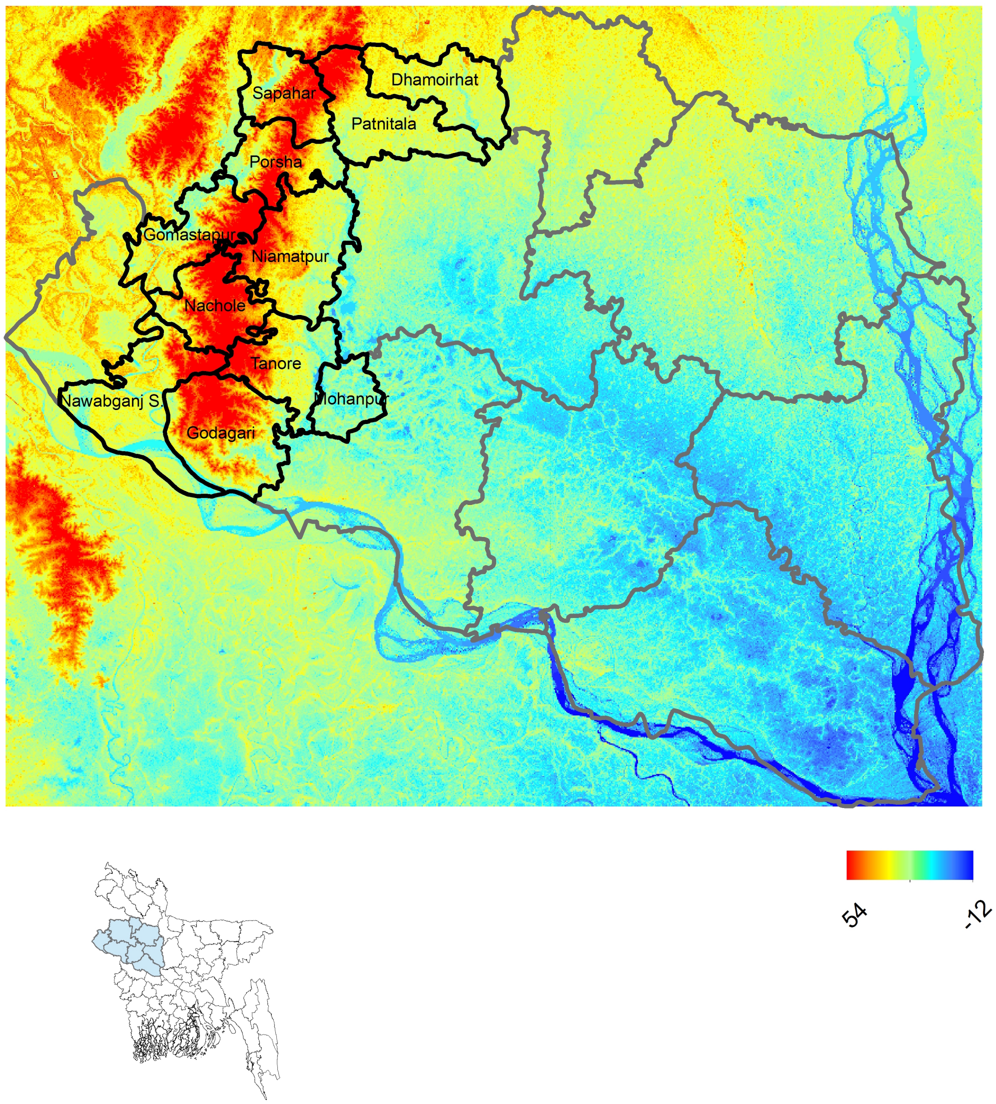
Putting it all together:

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

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

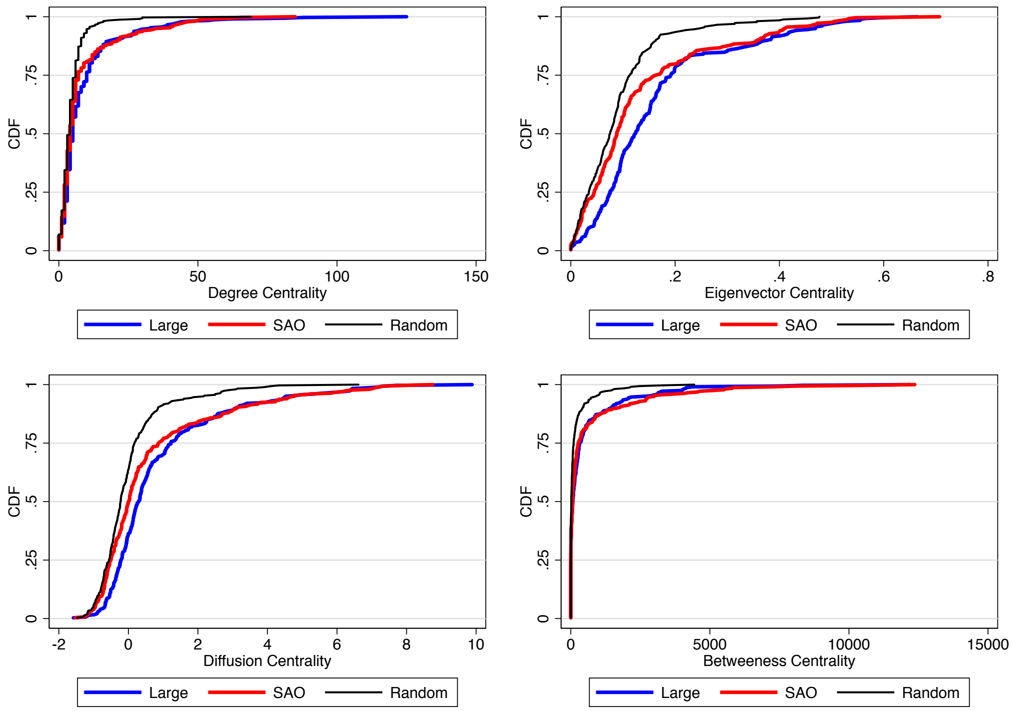
B.1 Figures

Figure B.1: Location of study area



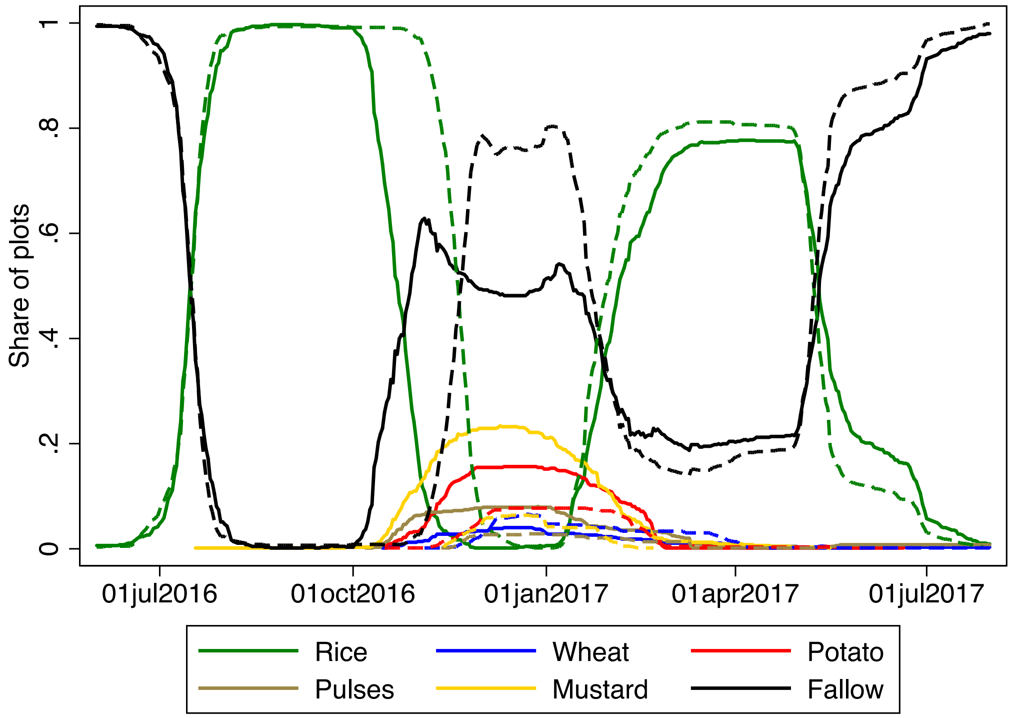
Notes: The figure shows the location of the 11 study Upazilas within Rajshahi district of Bangladesh. The shading corresponds to elevation, measured in meters.

Figure B.2: Cumulative distributions of network statistics for different types of entry points



Notes: Each graph shows the cumulative distribution function of the relevant network statistics, separately for the three different types of entry points. The network centrality measures are calculated for each entry point using the baseline social network survey.

Figure B.3: Annual land allocation for plots grown with either BD56 or BD51



Notes: The data are for the plots where either BD56 or BD51 was planted by the entry points. The vertical axis gives the share of plots that were allocated to the crop on the date corresponding to the horizontal axis. The sold lines are for the treatment (BD56) farmers and the dashed lines are for the control (BD51) farmers.

B.2 Tables

Table B.1: Balance of household characteristics across treatment arms

		Treatment Arm:						
	Control	Large	SAO	Random	Large + Demo	SAO + Demo	Random + Demo	Joint p-value
Education	4.235	3.951	4.663	4.604	4.216	4.368	4.768	0.381
	(4.314)	(4.202)	(4.521)	(4.293)	(4.125)	(4.466)	(4.235)	
Age	41.356	41.908	41.660	41.650	41.820	40.938	41.261	0.829
	(12.407)	(11.926)	(12.348)	(12.300)	(12.143)	(12.023)	(12.039)	
Owns Shallow Tubewell	0.103	0.149	0.160	0.085	0.072	0.086	0.115	0.356
	(0.304)	(0.356)	(0.367)	(0.278)	(0.259)	(0.280)	(0.319)	
Aman Rice Area (Bigah)	4.071	4.293	5.029	4.221	4.678	4.775	4.265	0.770
	(5.656)	(5.550)	(12.429)	(6.027)	(5.530)	(5.983)	(5.152)	
Aman Other Crop Area (Bigah)	0.348	0.375	0.319	0.462	0.300	0.236	0.395	0.682
	(1.567)	(1.893)	(1.478)	(9.638)	(1.676)	(0.794)	(1.412)	
Boro Rice Area (Bigah)	3.328	3.002	3.812	3.344	2.515	3.289	2.889	0.383
	(4.264)	(4.296)	(5.847)	(5.312)	(4.171)	(4.699)	(4.060)	
Boro Other Crop Area (Bigah)	1.125	1.252	1.332	1.140	1.478	1.100	1.346	0.840
	(2.454)	(2.513)	(3.325)	(2.362)	(3.043)	(2.219)	(2.358)	
Aman Urea Fertilizer (KG per Bigah)	21.427	21.953	21.260	22.161	20.644	21.313	20.588	0.843
	(15.109)	(15.459)	(21.673)	(21.759)	(17.417)	(25.641)	(14.648)	
Aman DAP Fertilizer (KG per Bigah)	15.834	16.110	15.519	16.435	14.889	15.813	15.157	0.326
	(11.660)	(15.169)	(20.073)	(13.662)	(6.739)	(15.346)	(10.332)	
Aman Rice Yield (KG per Bigah)	17.756	17.499	18.063	17.989	17.477	17.570	17.837	0.927
	(3.814)	(4.180)	(3.263)	(3.089)	(3.280)	(3.851)	(3.483)	
Grows Short-Duration Rice	0.011	0.035	0.007	0.004	0.007	0.018	0.008	0.831
	(0.102)	(0.185)	(0.085)	(0.066)	(0.081)	(0.135)	(0.088)	
Grows Wheat	0.236	0.260	0.243	0.190	0.372**	0.231	0.286	0.353
	(0.425)	(0.439)	(0.429)	(0.393)	(0.483)	(0.422)	(0.452)	
Grows Mango	0.086	0.063	0.071	0.093	0.076	0.076	0.063	0.958
	(0.280)	(0.242)	(0.256)	(0.290)	(0.265)	(0.266)	(0.244)	
Grows Potato	0.083	0.052	0.086	0.082	0.074	0.061	0.077	0.872
	(0.275)	(0.222)	(0.281)	(0.274)	(0.261)	(0.239)	(0.266)	
Grows Pulses	0.047	0.103	0.095	0.076	0.077	0.075	0.048	0.518
	(0.212)	(0.305)	(0.293)	(0.265)	(0.266)	(0.264)	(0.215)	
Grows Onion	0.049	0.037	0.050	0.021*	0.047	0.039	0.057	0.275
	(0.215)	(0.188)	(0.219)	(0.144)	(0.211)	(0.194)	(0.232)	
Grows Garlic	0.017	0.009	0.014	0.013	0.009	0.017	0.006	0.429
	(0.128)	(0.096)	(0.118)	(0.113)	(0.095)	(0.130)	(0.078)	

The summary statistics are calculated using the door-to-door census with 21,926 households. Each column shows mean values of each variable for either the control group or one of the six treatment groups. Standard deviations are reported in parentheses below each mean value. Asterisks indicate a statistically significant difference (1% ***, 5% **, and 10% *) between that arm and the control arm, where p-values are calculated by regressing each variable on a constant and indicators for each of the six treatment groups (standard errors adjusted for clustering at the village level). The final column shows the joint p-value of each of these regressions. Aman refers to the wet season prior to the door-to-door baseline (2015) and Boro refers similarly to the most recent dry season (2015-2016). 1 Bigah = 0.33 Acres.

Table B.2: Balance of household characteristics for entry points

	Control (BD51)	BD56 Treatment	p-value
Education	5.361	5.392	0.832
	(4.620)	(4.643)	
Age	43.326	43.690	0.577
	(12.585)	(12.225)	
Owns Shallow Tubewell	0.178	0.191	0.528
	(0.383)	(0.393)	
Aman Rice Area (Bigah)	8.553	8.977	0.523
	(11.097)	(10.494)	
Aman Other Crop Area (Bigah)	0.514	0.616	0.415
	(1.648)	(2.037)	
Boro Rice Area (Bigah)	6.483	6.247	0.795
	(7.781)	(8.606)	
Boro Other Crop Area (Bigah)	2.063	2.299	0.454
	(3.709)	(3.951)	
Aman Urea Fertilizer (KG per Bigah)	21.879	21.316	0.541
	(24.672)	(15.565)	
Aman DAP Fertilizer (KG per Bigah)	16.363	15.978	0.735
	(17.281)	(18.909)	
Aman Rice Yield (KG per Bigah)	17.888	17.566	0.207
	(3.797)	(3.881)	
Grows only rice	0.389	0.337	0.132
	(0.488)	(0.473)	
Grows Short-Duration Rice	0.017	0.024	0.488
	(0.129)	(0.153)	
Grows Wheat	0.304	0.312	0.946
	(0.460)	(0.464)	
Grows Mango	0.126	0.117	0.792
	(0.332)	(0.322)	
Grows Potato	0.121	0.098	0.359
	(0.327)	(0.297)	
Grows Pulses	0.078	0.111	0.127
	(0.268)	(0.314)	
Grows Onion	0.048	0.068	0.323
	(0.215)	(0.252)	
Grows Garlic	0.013	0.028	0.065
	(0.115)	(0.166)	

The analysis uses the door-to-door census conducted at the beginning of the experiment. Data are limited to the 1,747 entry points that consented to participate. Each column shows mean values and standard deviations are reported in parentheses below. The final column shows the p-value for the comparison of means, based on a regression of each characteristic on the treatment indicator and Upazila (strata) fixed effects. Standard errors are clustered at the village level.

Table B.3: Differences between SAO selected and random farmers, adjusting for farm size

	Degree		Eigenvector		Betweenness	
	(1)	(2)	(3)	(4)	(5)	(6)
SAO-based selection	3.582***	1.917**	0.042***	0.025**	394.084***	281.922***
	(1.044)	(0.838)	(0.012)	(0.011)	(103.649)	(89.698)
Farm Size		0.284***		0.003***		19.124***
		(0.053)		(0.000)		(4.634)
Mean in random group	4.56	4.56	0.09	0.09	164.19	164.19
Number of Observations	639	639	511	511	639	639
R squared	0.037	0.221	0.036	0.175	0.033	0.094

The data are limited to the 640 selected entry points in the random and SAO villages. The dependent variables are degree centrality (columns 1-2), eigenvector centrality (columns 3-4), and betweenness centrality (columns 5-6). Farm size is the total sum of cultivated area (across all three agricultural seasons). The omitted group in each regression is the villages where demonstrators were selected randomly. The standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table B.4: Analysis of take up by entry points

	(1)	(2)	(3)
Treatment village	-0.025		
	(0.035)		
Demo Village		0.060	
		(0.039)	
Random + Demo			0.085
			(0.062)
SAO			0.076
			(0.066)
SAO + Demo			0.067
			(0.067)
Large			0.053
			(0.067)
Large + Demo			0.157**
			(0.068)
Strata (Upazila) Fixed Effects	Yes	Yes	Yes
Mean in omitted group	0.71	0.65	0.69
Number of Observations	1795	953	953
R squared	0.046	0.059	0.064

The data are from the first midline with 1,795 entry points. Column 1 uses all observations and columns 2 and 3 use only observations from treatment (BD56) villages. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table B.5: Cultivation practices by treatment

	(1) Harvest Date	(2) 2nd Crop	(3) Boro Crop	(4) N Crops
Treated Village	-25.350***	0.278***	-0.035	0.243***
	(1.384)	(0.035)	(0.039)	(0.046)
Strata fixed effects	Yes	Yes	Yes	Yes
Mean in Control	8.64	0.24	0.82	2.06
Number of Observations	1242	1242	1242	1242
R squared	0.381	0.284	0.257	0.278

The data are limited to the plots where either BD56 or BD51 was planted by the entry points. The dependent variable in column 1 is the date of the harvest, measured in days after November 10, 2016. The dependent variables in columns 2 and 3 are indicators for whether the plot was sown with the Rabi (in-between) crop and the Boro (dry-season) crop. The dependent variable in column 4 is the total number of crops grown across all seasons. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table B.6: Cultivation practices by treatment and type of entry point

	(1) Harvest Date	(2) 2nd Crop	(3) Boro Crop	(4) N Crops
Treated Village	-25.315***	0.175***	-0.073	0.102
	(1.603)	(0.048)	(0.051)	(0.072)
Treatment Village * SAO	-0.219	0.110*	0.100*	0.210***
	(2.518)	(0.061)	(0.056)	(0.079)
Treatment Village * Large	0.533	0.173***	0.002	0.174**
	(2.302)	(0.065)	(0.062)	(0.085)
SAO	1.535*	-0.040	-0.047	-0.087**
	(0.796)	(0.033)	(0.029)	(0.035)
Large	0.784	-0.051	-0.014	-0.065*
	(0.801)	(0.031)	(0.029)	(0.034)
Mean in Control	8.64	0.24	0.82	2.06
Number of Observations	1242	1242	1242	1242
R squared	0.382	0.291	0.262	0.286

The data are limited to the plots where either BD56 or BD51 was planted by the entry points. The dependent variable in column 1 is the date of the harvest, measured in days after November 10, 2016. The dependent variables in columns 2 and 3 are indicators for whether the plot was sown with the Rabi (in-between) crop and the Boro (dry-season) crop. The dependent variable in column 4 is the total number of crops grown across all season. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table B.7: Profitability of BD56 and BD51 plots

	(1) Aman	(2) Rabi	(3) Boro	(4) Total
Treated Village	-4576.411***	1518.881***	-882.748	-4161.562***
	(248.751)	(451.867)	(619.309)	(767.897)
Strata fixed effects	Yes	Yes	Yes	Yes
Mean in Control	10309.28	2508.40	11900.65	24954.19
Number of Observations	1200	1205	1228	1156
R squared	0.396	0.438	0.314	0.380

The data are limited to the plots where either BD56 or BD51 was planted by the entry points. The dependent variables are profits per bigah, measured in Bangladeshi Taka (BDT). Approximately 80 BDT=1USD and 3 bigah = 1 acre. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table B.8: Peer effects on the number of conversations with entry points, separate for villages with and without demonstration plots

	(1)	(2)	(3)	(4)
Peer connections w/ entry points	0.050	0.179^{*}		
	(0.064)	(0.105)		
Peer connections w/ entry points * Demonstration Village		-0.261^{*}		
		(0.137)		
Connected to at least 1 entry point			0.093	0.284^{*}
			(0.098)	(0.147)
Connected to at least 1 entry point * Demonstration Village				-0.383^{*}
				(0.215)
Number of connections	0.001	-0.001	0.001	0.005
	(0.005)	(0.011)	(0.004)	(0.009)
Number of connections * Demonstration Village		0.003		-0.005
		(0.011)		(0.009)
Demonstration Village		0.176^{**}		0.220^{**}
		(0.087)		(0.087)
Strata fixed effects	Yes	Yes	Yes	Yes
Mean of Dep Variable	0.78	0.78	0.78	0.78
Number of Observations	636	636	636	636
R squared	0.301	0.310	0.301	0.312

The dependent variable in all regressions is the number of entry points that the respondent spoke to about BD56. The data are limited to the 64 villages where entry points where chosen randomly and peer effects can therefore be causally identified. The variable *Peer connections w/ entry points* is the number of entry points (from 0 to 5) that the farmer is connected with while *Connected to at least 1 entry point* is an indicator variable for being connected to at least one of the entry points. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table B.9: Effects of hitting a large-farmer entry point

	(1) All Villages	(2) Random Villages
At least 1 large entry point	0.079**	0.123*
	(0.032)	(0.071)
At least 1 large entry point * Demonstration Village	-0.101**	-0.131
	(0.048)	(0.104)
Demonstration Village	0.078**	0.119**
	(0.038)	(0.047)
Strata fixed effects	Yes	Yes
Mean of Dep Variable	0.67	0.64
Number of Observations	1919	639
R squared	0.171	0.189

The data are for the 10 random farmers per village that were selected for the information survey. Column 1 is for all 192 BD56 villages, column 2 is for the 64 villages where entry points were selected randomly. *At least 1 large entry point* is an indicator for villages where one of the five largest farmers was selected as an entry point. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

C | How do Online Job Portals affect Employment and Search outcomes? Evidence from India

C.1 Tables

Table C.1: Attrition

	(1)	(2)
	Midline	Endline
Treatment	0.017	-0.017
	(0.017)	(0.021)
Priority Treatment	0.024	0.018
	(0.017)	(0.021)
Number of Observations	2662	2662
R squared	0.0019	0.0004

Standard errors are clustered at the respondent level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table C.2: Balance (with controls)

	(1) Control	(2) Treatment	(3) Priority Treat-ment	(4) (1) vs. (2), p-value	(5) (1) vs. (3), p-value	(6) (2) vs. (3), p-value	(7) Joint test
=1 if male	1.26	0.02	0.01	0.19	0.52	0.53	0.42
Age	25.21	-0.20	0.14	0.37	0.57	0.13	0.31
Married Y/N	0.54	0.01	-0.01	0.66	0.60	0.31	0.60
Religion=Hindu	0.10	0.02	0.03	0.07	0.05	0.82	0.10
Religion=Muslim	0.90	-0.02	-0.03	0.07	0.06	0.85	0.11
=1 if ST/SC caste	0.81	-0.02	-0.02	0.23	0.26	1.00	0.42
=1 if OBC caste	0.01	0.05	0.06	0.02	0.01	0.58	0.01
=1 if general caste	0.19	-0.03	-0.04	0.17	0.07	0.56	0.17
Father's education>0	0.04	0.03	0.02	0.23	0.47	0.66	0.48
Mother's education>0	1.05	0.03	-0.02	0.34	0.39	0.06	0.18
=1 if live in village	-0.17	-0.01	0.01	0.61	0.74	0.38	0.67
Received formal skills training	1.94	0.03	0.02	0.17	0.40	0.64	0.39
=1 if currently employed	0.87	0.03	0.02	0.16	0.35	0.69	0.37
=1 if looking for job	0.33	0.02	0.01	0.40	0.80	0.57	0.69
Access to Internet Y/N (clean)	0.67	0.04	0.02	0.03	0.28	0.33	0.10
Reservation wage (winsorized)	19376.31	-398.80	753.96	0.21	0.03	0.00	0.00

Standard errors are clustered at the respondent level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table C.3: Summary Statistics by Geo-Zone

	(1) DelhiNCR	(2) North	(3) SouthWest	(4) East	(5) (1) vs. (2), p-value	(6) (1) vs. (3), p-value	(7) (1) vs. (4), p-value
=1 if male	0.81	0.89	0.87	0.90	0.00	0.01	0.00
Age	23.12	22.81	26.75	23.46	0.39	0.00	0.33
Education (Years)	14.78	14.28	14.54	13.70	0.00	0.10	0.00
Married Y/N	0.17	0.21	0.41	0.28	0.15	0.00	0.00
Religion=Hindu	0.86	0.91	0.95	0.97	0.01	0.00	0.00
Religion=Muslim	0.14	0.08	0.04	0.03	0.01	0.00	0.00
=1 if ST/SC caste	0.16	0.27	0.27	0.59	0.00	0.00	0.00
=1 if OBC caste	0.20	0.37	0.45	0.24	0.00	0.00	0.12
=1 if general caste	0.64	0.36	0.27	0.17	0.00	0.00	0.00
Father's education>0	0.89	0.80	0.82	0.79	0.00	0.03	0.00
Mother's education>0	0.77	0.53	0.61	0.46	0.00	0.00	0.00
=1 if live in village	0.06	0.46	0.46	0.69	0.00	0.00	0.00
Received formal skills training	1.30	1.28	1.42	1.35	0.43	0.00	0.10
=1 if currently employed	0.45	0.32	0.48	0.16	0.00	0.38	0.00
=1 if looking for job	0.54	0.64	0.65	0.72	0.00	0.00	0.00
Access to Internet Y/N (clean)	0.96	0.83	0.79	0.66	0.00	0.00	0.00
Reservation wage (winsorized)	17536.76	12375.24	12116.93	11063.11	0.00	0.00	0.00

Standard errors are clustered at the respondent level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table C.4: Job Search by Geo-Zone

	Unemployed			Employed		
	(1) Searching	(2) Hours	(3) Applications	(4) Searching	(5) Hours	(6) Applications
Treatment East	-0.077	-2.445*	0.682	-0.203	-3.054**	-1.826*
	(0.050)	(1.474)	(0.516)	(0.135)	(1.528)	(1.011)
Treatment DelhiNCR	0.021	0.629	-1.263	0.110	3.023	0.011
	(0.150)	(2.122)	(1.061)	(0.118)	(1.874)	(1.209)
Treatment North	-0.007	-0.146	-0.641	0.094	1.225	0.758
	(0.065)	(1.671)	(0.502)	(0.096)	(1.165)	(0.744)
Treatment SouthWest	0.072	-2.769	-0.239	-0.116	-1.170	0.023
	(0.089)	(2.280)	(0.745)	(0.079)	(1.484)	(0.690)
Priority Treatment East	-0.011	-0.274	-0.947*	0.102	1.335	-0.568
	(0.052)	(1.434)	(0.539)	(0.132)	(2.220)	(1.046)
Priority Treatment DelhiNCR	0.039	0.346	0.679	-0.055	1.386	1.061
	(0.119)	(1.968)	(1.032)	(0.124)	(1.680)	(1.340)
Priority Treatment North	0.030	2.647*	0.634	-0.121	-1.618	0.180
	(0.066)	(1.477)	(0.513)	(0.096)	(1.206)	(0.686)
Priority Treatment SouthWest	-0.066	-1.700	0.767	0.029	-0.280	-0.212
	(0.077)	(2.374)	(0.827)	(0.077)	(1.469)	(0.543)
Respondent Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Survey Round Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Geo-Specifc Time Trend	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	4134	3848	3870	2694	2596	2505

Columns 1, 2 and 3 include all unemployed respondents in the sample, while columns 4, 5 and 6 include all employed respondents in the sample. The dependent variables are an indicator for whether the respondent is actively searching for employment (column 1/4), the number of hours spent searching in the past week - where people who aren't searching are assigned a value of 0 hours (column 2/5), the number of job applications submitted in the last 3 months (column 3/6). Standard errors are clustered at the respondent level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table C.5: Job Search controlling for TP0

	Unemployed			Employed		
	(1) Searching	(2) Hours	(3) Applications	(4) Searching	(5) Hours	(6) Applications
Treatment East	-0.202	-3.045**	-1.828*	-0.077	-2.443*	0.682
	(0.135)	(1.530)	(1.012)	(0.050)	(1.474)	(0.516)
Treatment DelhiNCR	0.110	3.023	0.011	0.021	0.629	-1.263
	(0.118)	(1.875)	(1.210)	(0.150)	(2.123)	(1.062)
Treatment North	0.094	1.229	0.757	-0.007	-0.149	-0.642
	(0.096)	(1.166)	(0.745)	(0.065)	(1.672)	(0.503)
Treatment SouthWest	-0.116	-1.170	0.023	0.072	-2.769	-0.239
	(0.079)	(1.485)	(0.691)	(0.089)	(2.281)	(0.746)
Priority Treatment East	0.027	-0.099	-0.168	0.007	0.681	-1.085
	(0.158)	(2.769)	(0.945)	(0.061)	(1.673)	(0.671)
Priority Treatment DelhiNCR	-0.055	1.386	1.061	0.039	0.346	0.679
	(0.124)	(1.681)	(1.341)	(0.119)	(1.969)	(1.032)
Priority Treatment North	-0.126	-1.199	0.107	0.034	2.887**	0.686
	(0.100)	(1.204)	(0.708)	(0.067)	(1.463)	(0.537)
Priority Treatment SouthWest	-0.001	-0.241	0.053	-0.024	-1.409	1.114
	(0.113)	(1.432)	(0.408)	(0.124)	(4.216)	(0.778)
Priority Treatment DelhiNCR, 0 SMS	0.000	0.000	0.000	0.000	0.000	0.000
	(.)	(.)	(.)	(.)	(.)	(.)
Priority Treatment East, 0 SMS	0.154	2.865	-0.709	-0.042	-2.194	0.333
	(0.152)	(2.618)	(1.163)	(0.077)	(1.723)	(0.851)
Priority Treatment North, 0 SMS	0.028	-2.926	0.470	-0.032	-1.839	-0.433
	(0.149)	(2.171)	(0.695)	(0.133)	(4.230)	(0.450)
Priority Treatment SouthWest, 0 SMS	0.043	-0.059	-0.371	-0.053	-0.363	-0.461
	(0.117)	(1.357)	(0.473)	(0.124)	(4.432)	(0.838)
Respondent Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Survey Round Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Geo-Specifc Time Trend	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	2694	2596	2505	4134	3848	3870

Columns 1, 2 and 3 include all unemployed respondents in the sample, while columns 4, 5 and 6 include all employed respondents in the sample. The dependent variables are an indicator for whether the respondent is actively searching for employment (column 1/4), the number of hours spent searching in the past week - where people who aren't searching are assigned a value of 0 hours (column 2/5), the number of job applications submitted in the last 3 months (column 3/6). Standard errors are clustered at the respondent level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table C.6: Emp/City (Village)

	(1) Emp (Vil)	(2) Emp (City)	(3) Emp	(4) City (Vil)	(5) City (City)	(6) City
Treatment	-0.078***	-0.106***	-0.097***	-0.037	0.021	-0.293***
	(0.029)	(0.032)	(0.026)	(0.027)	(0.024)	(0.024)
Priority Treatment	0.068**	0.028	0.028	0.068**	0.030	0.030
	(0.028)	(0.030)	(0.030)	(0.026)	(0.021)	(0.021)
Treatment*Vil			0.011			0.565***
			(0.027)			(0.022)
Priority Treatment*Vil			0.040			0.038
			(0.041)			(0.034)
Respondent Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Survey Round Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	3371	3495	6866	3382	3507	6889

Table C.7: City (Caste)

	(1)	(2)	(3)	(4)	(5)	(6)
	Emp (cstH)	Emp (cstL)	Emp	City (cstH)	City (cstL)	City
Treatment	-0.126***	-0.075***	-0.092***	-0.028	-0.016	0.011
	(0.039)	(0.026)	(0.024)	(0.034)	(0.031)	(0.026)
Priority Treatment	0.006	0.058**	0.058**	0.054	0.063**	0.063**
	(0.040)	(0.024)	(0.024)	(0.034)	(0.028)	(0.028)
Treatment*Cst			0.004			-0.097***
			(0.029)			(0.029)
Priority Treatment*Cst			-0.052			-0.009
			(0.047)			(0.045)
Respondent Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Survey Round Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	2079	4667	6746	2088	4677	6765

Table C.8: Emp/City (Education)

	(1)	(2)	(3)	(4)	(5)	(6)
	Emp (EducH)	Emp (EducL)	Emp	City (EducH)	City (EducL)	City
Treatment	-0.085***	-0.108**	-0.081**	-0.005	-0.050	0.013
	(0.025)	(0.043)	(0.032)	(0.028)	(0.046)	(0.036)
Priority Treatment	0.049**	0.040	0.040	0.049*	0.091**	0.091**
	(0.024)	(0.043)	(0.043)	(0.026)	(0.044)	(0.044)
Treatment*EducH			-0.016			-0.044
			(0.032)			(0.035)
Priority Treatment*EducH			0.008			-0.041
			(0.049)			(0.051)
Respondent Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Survey Round Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	4959	1885	6844	4976	1891	6867