Monitoring and Intrinsic Motivation: Evidence from Liberia's Trucking Firms*

(Job Market Paper)

Golvine de Rochambeau†

November 19, 2017

Abstract

Severe information asymmetries are thought to make contracting particularly difficult within (and across) firms in developing countries. Standard principal-agent theory predicts that a new monitoring technology provided at zero cost should be widely adopted and unambiguously raise workers' effort. I test this classical prediction using a field experiment with trucking companies in Liberia. The treatment offered to install GPS tracking devices on randomly selected trucks at no cost. Treatment-on-the-treated estimates reveal that the tracking devices increased monitored drivers' average speeds by 58 percent, without leading to higher accident rates or maintenance costs. Despite this, managers declined to install the devices on 35 percent of the trucks selected for treatment. Using a model of intrinsic motivation, I show that it may be optimal not to monitor workers who are intrinsically motivated to work hard. While monitoring technologies increase agents' extrinsic incentives to provide effort, they also do not allow worker to show that he or she does not need these incentives to work hard, which can crowd out effort. I provide three pieces of evidence in support of this explanation. First, Liberian trucking company managers choose to install tracking devices only on the trucks of drivers who perform less well at baseline. Second, the treatment effect on speed for monitored drivers is greater the lower the performance of the driver at baseline. Finally, I show that for drivers who performed well at baseline, the treatment has a negative effect on the relation between the manager and the driver, and on the driver's propensity to follow the rules of the business. Overall, this paper demonstrates that, while new monitoring technologies can dramatically raise some workers' productivity in settings where employment contracts are difficult to enforce, their use may lower the productivity of some workers - those who are intrinsically motivated to work hard.

*I would like to thank Eric Verhoogen, Jonas Hjort and Andrea Prat for their guidance, support and valuable advice throughout this project. I would also like to thank seminar participants and fellow PhD students at Columbia for their helpful comments and suggestions, particularly Matthieu Teachout and Nandita Krishnaswamy. I am thankful to Paul de Rochambeau for his help. This project was supported by the International Growth Center (IGC), the Private Enterprise Development in Low-Income Countries Initiative (PEDL), the Center for Development Economics and Policy (CDEP), and the Development Colloquium at Columbia. USAID-Liberia financed the GPS trackers used in the experiment. I am particularly grateful to Jerrre Manarolla and Lee Rosner of USAID. I am also thankful for the excellent services provided by the team of enumerators from the Liberian firm "Q&A, Inc", led by Alpha Simpson. All errors are my own.

†Columbia University, Department of Economics, gd2382@columbia.edu.

1 Introduction

Severe information asymmetries are thought to make contracting particularly difficult within (and across) firms in developing countries. Increasingly available and affordable monitoring technologies have the potential to greatly reduce such assymetries. Standard principal-agent theory predicts that these technologies should raise worker effort. However, increased monitoring may also undermine informal arrangements that have developed exactly to substitute for the effort incentives provided by formal contracting institutions. If so, the impact of modern monitoring technologies on worker effort is theoretically ambiguous.

In this paper, I study both a firms' decision of whether to employ a new monitoring technology and the corresponding impact on worker effort. I present results from an experiment in the context of Liberia's trucking industry. The treatment offered managers the opportunity to install GPS tracking devices on randomly selected trucks at zero cost. The devices allow managers (principals) to better monitor drivers (agents) by reporting the position of trucks in real time. I first show how installed devices affect drivers' speeds without leading to higher accident rates or maintenance costs. I then present a simple model that illustrates how the level of effort provided by the worker with no monitoring can alter managers' optimal monitoring decisions across workers. Finally, I test the model's predictions by analyzing (i) which workers managers choose to monitor, (ii) the heterogeneity of the effect of monitoring across drivers, and (iii) whether monitoring can adversely affect the effort of some workers.

Liberia's trucking industry is an ideal setting in which to study firms' monitoring choices and worker effort for several reasons. First, the combination of dirt roads and heavy rains makes route completion times variable and unpredictable. This unpredictability means that drivers can easily shirk without raising suspicion. Second, drivers take a set of decisions that affect the firm's performance, and that the manager cannot observe perfectly. These include the number and length of breaks, and whether to transport additional goods or individuals without the manager's approval. In the absence of a monitoring technology, the only information a manager can use to try to infer such driver "input" choices are "outputs" such as the total travel time.[1]

I conducted the experiment on a sample of 150 trucks based in Liberia's main transport hubs. GPS tracking devices were offered free of charge.[2] The devices send the real-time positions of trucks to an online server through a mobile network.[3] Managers can then access the online server

[1]Managers were unable to effectively align incentives of drivers with the firm's objectives before the experiment. Only two trucks in the sample used GPS trackers before the experiment. The literature has shown that performance pay may increase incentives of employees (e.g. Jensen and Murphy, 1990; Foster and Rosenzweig, 1994; Lazear, 2000; Lavy, 2009; Muralidharan and Sundararaman, 2011). However, only 21 percent of the truck drivers in the sample receive performance pay, and for these drivers the variable part of their pay only represents less than 5 percent of their total wage. While this might seem surprising at first, a risk-averse worker may not agree to a performance pay in a case where many factors affect the output and are out of his control. In fact, when asked if they would like a performance-pay, the drivers answered that they would find it unfair.

[2]The mechanic's salary and proper guidance to use the device also were provided.

[3]The combination of mobile networks used by the GPS tracking devices covers most of the road network,

on a computer or smartphone.[4] The role of the GPS tracking devices and the managers' ability to track the truck was clearly explained to drivers at baseline.[5]

Evidence from the experiment shows that monitored drivers completed their tasks 58 percent faster. This occurs in part because monitored drivers take shorter breaks,[6] but do not have more accidents or higher truck maintenance costs. These results hold whether managers' or drivers' estimates of speed and break times are used.[7]

Despite the dramatic improvement in monitored drivers' performance, take-up of the monitoring devices was far from complete through the end of my data collection, a year after the start of the experiment. Managers chose not to install the devices on 35 percent of the trucks selected for treatment.

I show that incomplete take-up is consistent with a model of intrinsic motivation. Building on Bénabou and Tirole (2006), an intrinsically motivated agent provides effort irrespective of her extrinsic incentives to do so and values the principal's belief about her motivation for providing effort. The key insight of the model is that, while increased monitoring provides additional extrinsic incentives to provide effort, it does so at the cost of reducing a worker's ability to signal that she does not need such explicit incentives to work hard. Monitoring can therefore crowd out effort for workers who value the principal's opinion so much that the additional incentives from monitoring are outweighed by the workers' reduced ability to signal their "type".

The model predicts that the higher a worker's baseline effort, the lower the performance benefits from monitoring will be; the effect on performance might become negative for high levels of effort. I show three pieces of evidence that provide empirical support for this interpretation of my findings and the particular form of worker motivation that generates the model's predictions.

First, I show that managers decide to install GPS tracking devices on the trucks of drivers who perform less well at baseline. Drivers who receive a GPS tracking device ex ante complete their routes less fast, are more likely to have accidents, and are more likely to break the rules set out by their managers.[8] Additionally, managers report worse relationships[9] at baseline with the drivers they later decided to monitor. These drivers are typically from another county than the manager.[10]

Second, the treatment effect on speed for monitored drivers is greater the lower the perfor-

ensuring that the position of the truck is precisely reported at all times.

[4]Most of the managers in the study own or have access to a smartphone or computer.

[5]Not informing the drivers about the GPS tracking devices would have put drivers at risk of being reprimanded or disciplined by their managers.

[6]Estimated breaks are times when the driver stops for being stuck in the mud, for deliveries or for personal reasons.

[7]GPS tracking devices are only installed on trucks of the treatment group, so the data collected by the devices cannot be used for measuring treatment effects. The estimates presented in this paper are based on interviews with drivers and managers.

[8]At baseline, managers were asked which rules they clearly asked the drivers to follow. These rules include "do not carry unauthorized passengers or goods" and "do not use the truck for personal reasons".

[9]Managers are asked to rate their relationships with a particular driver on a scale from 0 to 10.

[10]County of origin plays a role similar to ethnic group in Liberia.

mance of the driver at baseline. Drivers who at baseline complete their routes fast, and are better at following the rules of the business, show little to no effect of the monitoring device on performance.

Third, the treatment effect on measures of performance that are not directly monitored by the GPS tracker is adverse for drivers who provided high levels of effort at baseline. Monitoring the drivers significantly decreased the propensity of these drivers to follow the rules of the business and to take good care of their trucks. Additionally, the drivers who provided high levels of effort at baseline reported a significant deterioration in their job satisfaction and in their relationship with the managers after monitoring was introduced.

These last two pieces of evidence support the model's prediction that the more the driver provides effort at baseline, the lower the benefits from monitoring. They also suggest that monitoring such drivers would be counter-productive for the firm which helps explain why managers on average choose not to install monitoring devices on such drivers' trucks. This evidence that Liberian managers on the whole accurately judge which workers to monitor – the combination of the first and second two pieces of evidence – is important: it suggests that simply making monitoring technologies available can increase firm productivity.[11] It also increases our confidence that the model of intrinsic motivation provides an accurate characterization of Liberian trucking firms' challenges.

Overall, this paper demonstrates that, while monitoring technologies can dramatically increase the productivity of certain drivers, their use may be counter-productive for other workers. Thus, blind application of monitoring devices to the entire worker-base may produce a suboptimal effect on overall productivity, contrary to the predictions of classical principal-agent theory. This paper contributes to literatures on the agency problem in the workplace, on the adverse effect of incentives, on contingent management and on the understanding of high transport costs in developing countries.

An existing literature on the agency problem has shown that information asymmetries in the workplace can in some contexts partly be overcome with informal arrangements, such as relational contracts (e.g. Greif, 1993; Brown et al., 2004; Macleod, 2007; Macchiavello and Morjaria, 2015[12]). In this paper, I study how monitoring can help solve a key agency problem when informal arrangements fail. This relates to the literature specifically on the effect of monitoring on an agent's effort. Monitoring has been shown to have a positive effect on the performance of workers in different contexts in developing countries, such as health care provision (Björkman

[11] Evidence suggests that management choices are not always optimal in developing countries (Bloom et al., 2013). My findings show that in choosing which drivers they wanted to monitor, managers on average made the optimal choice.

[12] Greif (1993) shows that Maghribi traders in the 11th century relied on a reputation mechanism to counter asymmetric information and limited contract liability. Brown et al. (2004) show that relational contracts emerge in the absence of third party enforcement contracts. Macleod (2007) presents a survey of formal and informal mechanisms. Macchiavello and Morjaria (2015) show that in the Kenyan rose export industry, buyers rely on the reputation to overcome poor contract enforceability.

and Svensson, 2009) or education (Duflo et al., 2012).[13]

The line of work that is the closest to this project is the work by George Baker and Thomas Hubbard. In several papers (Hubbard, 2000, 2003; Baker and Hubbard, 2003, 2004) they study the impact of the introduction of On-Board Computers (OBCs)[14] in trucking companies in the Unites States. The authors use historical data to show that the introduction of the technology helped solve the moral hazard issue within trucking companies, that it changed resource allocation inside the firm and truck-ownership incentives. My work differs from theirs in that in addition to directly measuring effort,[15] I study the introduction of monitoring in a setting where informal arrangements partially substitute for formal contract enforcing institutions. This is important because in such settings monitoring may undermine informal arrangements and may be counter-productive. In the context of this study I find counter-productive effects of monitoring that were not considered in previous literature.

In that respect, this paper adds to the literature on motivational crowding out and adverse effects of extrinsic incentives. Models explaining the possibility of adverse effect of incentives were developed in theoretical papers, such as Bénabou and Tirole (2003), Sliwka (2007) and Ellingsen and Johannesson (2008).[16] In this paper, I follow Bénabou and Tirole (2006). While existing studies[17] document motivational crowding out in blood donations (Mellström and Johannesson, 2008), day-care pick-ups (Gneezy and Rustichini, 2000a), and the lab (Falk and Kosfeld, 2006; Dickinson and Villeval, 2008; Békir et al., 2015; Belot and Schröder, 2016), there is to my knowledge no existing evidence on its implications for firms[18].

My findings suggest two things: (i) monitoring employees can crowd out their effort and (ii) managers are aware of and act on the phenomenon. In this regard, this paper contributes to the literature on contingent management practices. The employer-worker relation and how optimal

[13]This has also been shown in developed countries, for example (Jackson and Schneider, 2015) show that monitoring has a positive impact on the effort of workers in auto repair shops.

[14]GPS tracking devices are a particular case of OBCs. OBCs can be more sophisticated, and in some cases are able to provide additional information than the position of the truck, such as fuel and battery levels, or other measures of driver performance.

[15]They show that trucking companies are more likely to use GPS tracking devices when information assymetries are particularly strong (when perquisite-taking is attractive to drivers, driver effort is important, and verifying drivers' actions to insurers is valuable). Using direct measured of effort, I show that using GPS tracking devices increases the effort of drivers.

[16]In the case were the principal has private information about the task to be performed by the agent, Bénabou and Tirole (2003) show that incentives can adversely impact an agent's perception of the task or of his abilities, and act as negative reinforcers in the long run. In the model described by Sliwka (2007), some agents are "conformists" and will adhere to the social norm. By choosing to trust the agent, the principal can signal her conviction that most people are fair. On the other, if she chooses to control the agent, she reveals her pessimism about the social norm and this may lead conformists to become selfish. Ellingsen and Johannesson (2008) show that if there are different types of principals, those that are worth impressing, and those that are not, choosing to monitor or not (or to give incentives or not) is a signal about the type of the principal.

[17]See Frey and Jegen (2000) for a survey of empirical evidence in the fields of both economics and psychology.

[18]Gneezy and Rustichini (2000b) show that paying high-school students to do voluntary work can crowd out their effort. That I know of, this is the closest evidence in the literature on the adverse effect of incentives in an employer-employee relationship.

management practices are affected by this relation is studied by Blader et al. (2016). In my study, managers based their decision to monitor drivers or not on their relationship with drivers. This confirms the findings of Grund and Harbring (2013), who show that in an employment relationship, control is negatively correlated with trust[19].

Finally, this paper relates to the literature on transport costs. Authors have shown that transport costs are high in developing countries, particularly in Africa (Teravaninthorn and Raballand, 2009). Atkin and Donaldson (2015) and Bergquist (2016) use a new method to estimate trade costs from commodity prices and shed light on the role of intermediaries to explain high transport costs. Lall et al. (2009) show how bad roads and low competition in the transport sector contribute to high transport costs. This paper is the first to document the role of moral hazard in explaining high transport costs.

The paper is organized as follows. Section 2 presents the context of the experiment, section 3 details the experimental design and Section 4 describes the effect of the treatment on treated drivers. Section 5 describes the model and the predictions it entails. Section 6 presents empirical evidence that supports the predictions of the model. Section 7 concludes.

2 Background

Liberia's trucking industry is particularly prone to asymmetries of information between drivers and managers.

2.1 Liberia

In the early nineteenth century, thousands of freed African-Americans emigrated from the United States to the African continent. They arrived on the continent's West Coast and created Africa's first independent republic: Liberia. In 2003, after decades of civil war, Charles Taylor, a prominent warlord, was ousted from power, ending one of Africa's most violent wars. The 2005 elections were won by the continent's first woman president, Ellen Johnson Sirleaf.

Since the end of the war, Liberia's economy, which was entirely put to the ground during the war, has been slowly recovering. More than 200,000 Liberian refugees who had fled the war came back to their homeland. As a result of the war, Liberia's economy is very young.[20] Workers - who have either fought or lived in refugee camps during the war - are often not well educated.

Trust in the government's stability was slowly restored[21], but the country still suffers from poor institutions. Corruption is said to be common at any levels of the government and it is

[19] Grund and Harbring (2013) define trust as measured by the trust question: "Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?".

[20] According to a 2013 census, a minority of businesses operating today were already operating before the war. In 2013 businesses had been operating for an average of 5.7 years.

[21] In 2016, foreign investments had come back to their pre-war levels.

difficult to bring a disagreement to court. Often no formal contract is signed between employers and workers[22], and in most cases workers are hired under a verbal agreement.

In such a context, managers rely extensively on their personal networks to hire employees. Liberia is divided into fifteen different counties, and while close to one quarter of Liberians live in Monrovia, the capital, most of them identify themselves as originating from one of the fifteen counties. Liberians originating from the same county are more likely to have family or friends in common, and it is common for firms to hire employees in such networks.

2.2 The transport industry in Liberia

Liberia is connected to the international market through its main port of entry based in the capital, Monrovia, where most trucking companies are based. These companies find clients (mainly traders or producers), distribute imported goods to inland markets and collect local goods for international exports. A few other companies operate from other cities, mainly transporting goods that are locally produced and consumed. Since railroads and maritime routes do not exist for transporting containers[23], road transport is the only mode of transportation.

However, Liberia's road network is not well developed (see the lower panel of figure 1 for a map of Liberia's main roads).[24] The country's main axis connects the three main cities but still leaves a large share of the population unconnected to a paved road.[25] Liberia ranks 14th in terms of average rainfall per year.[26] As a result, vehicles get often stuck in the mud (figure 2 shows a common situation after a strong rain).

The combination of heavy rains and dirt roads makes travel times very variable and transport companies have a hard time predicting the time it will take for goods to be delivered.[27]

2.3 Information asymmetries in trucking firms

Information asymmetries arise when managers are not able to perfectly monitor workers. When trucks leave the parking lot, managers are not able to know precisely the location of the truck until it reaches its destination.[28] During that time, drivers choose the speed of the truck, the number and length of breaks without managers being able to see or control these decisions.

[22] In the rare cases where contracts are signed, a third person known to both parties signs as well to ensure enforcement.

[23] Only one railroad is still operating in Liberia, but the mining company who restored it after the war has a monopoly over its use. Liberia has four ports (Monrovia, Buchanan, Harper and Greenville), and Monrovia's Freeport is the only one to be fully equipped for container and cargo handling.

[24] On the map of figure 1, the width of roads is proportional to their state, where 0 is a very damaged road and 10 is an all-season paved road.

[25] According to the World Bank, only 11% of the population is connected to a an all-season paved road.

[26] World Bank average precipitation in depth (mm/year).

[27] In contracts between transport companies and their clients, very few stipulate a date of arrival. Most of the clients pay a flat rate for the goods to be delivered in a "timely manner".

[28] Once the truck reached the destination, the client can confirm he received the goods.

Discussions with the managers before the beginning of the experiment revealed different approaches to monitoring drivers prior to the experiment. Most of the managers I talked to mentioned that they couldn't rely on any information, and as such had no choice but to trust their drivers. To keep track of where the goods are, most managers require their drivers to call them regularly and update them on their position. There are three network providers in Liberia, which together cover almost entirely the road network, but separately each one only covers some areas. A truck driver, who often relies on only one network, lacks a signal in some sections of the journey. This provides him with a great reason to not continuously update his manager on the location of the truck (or to avoid answering the phone). Another obvious issue with this approach is that even in the case in which he can get in touch with the driver, the manager is not able to confirm the information that he is given.[29]

Very few companies used GPS tracking devices prior to the experiment. Information and cost barriers combined with the poor development of the mobile network contribute to the low take-up of the technology in the past. Local GPS trackers are expensive[30] and have not proven to be fully functioning, in part because the mobile network necessary to use the device has been very slow to develop and until recently did not cover the main roads. At baseline, only 25% of managers knew what a GPS tracker was before the interview.

In other contexts, information asymmetries have been shown to be at least partially solved with performance pay. In the context of trucking companies, drivers receiving a performance pay have a wage based on the time they take to complete a trip, the number of trips completed per month, or other observable measures of performance. Only 21 percent of the drivers in my sample received a wage based on performance, and for these drivers the variable part of their wage only represents less than 5 percent of their total wage. While this might seem surprising at first, a risk-averse worker may not agree to a performance pay in a case where many factors affect the output and are out of his control. In fact, when asked if they would like to be paid based on their performance, drivers answered that they would find it unfair. In the point of view of the manager, performance pay may not be optimal: providing incentives to drivers for one factor of effort, such as the total travel time, might distort drivers' incentives on other factors, such as taking good care of the truck.[31]

In the past, managers have not been able to perfectly align drivers' incentives with the optimal actions for the benefits of the firm. The experiment I describe in the next section measures to what extent the introduction of a monitoring device solves the agency problem.

[29] One manager said he had developed a system to monitor his driver. He had asked businesses on the road – managed by family, or friends – to watch the road and notify him when the truck went by. While this approach could in theory solve at least in part the agency problem, it requires an extensive network of friends and family along the road, and to trust the businesses giving the information.

[30] A quick survey of the local options for GPS trackers determined that the cheapest device available in Liberia is priced at US$ 1,500.

[31] This multi-tasking interpretation was developed by Holmstrom and Milgrom (1991).

3 Experimental Design

In this section I describe the recruitment of trucking companies, the randomization of the treatment and its implementation as well as the data collection. The empirical evidence on the effect of the treatment is presented in the following section.

3.1 Sample of Companies

Firms were recruited in the sample using a combination of different methods.

1. The *Liberian Business Registry (LBR)* provided the contact information of 57 firms registered under the freight transport sector. These included different types of firms (not only trucking) such as customs clearing.[32] The list only included businesses that had been registered at one point including businesses that had already closed. Most of the businesses on this list were either closed, not in the trucking business or not reachable[33].

2. *Building Markets* - an NGO that works with more than 3,000 Liberian businesses - has a publicly available directory of firms by sector, with their contact information.[34]

3. The majority of the firms were recruited directly at the *main transport hubs*. Enumerators were assigned to different areas in Monrovia, and sometimes travelled to other cities[35]. At transport hubs, they either directly talked to managers, or asked drivers for the contact information of their manager.[36]

4. Monrovia has a *Port Trucker Union* that brings together small trucking companies that have access to the port's container unloading area. Contact was made with the transport union and I was invited to a meeting that gathered around 50 drivers and managers. Enumerators had the opportunity to explain the project and took the contact information of interested managers. While this method seemed to be working, some of the managers were contacted later, were skeptical and later refused to answer interviews. According to local informal discussions, it appeared that some firms in that group thought I was working with the government.

5. Lastly, enumerators asked the firms they had recruited for contact information of other firms in their sector. Some firms were able to point us to other firms they knew, but at this

[32]When they register, businesses are asked what sector they are in and LBR never verifies that sector. This leads to businesses being registered under wrong categories.

[33]That I know of, LBR's contact information of businesses is never updated which results in many contact information being obsolete.

[34]Building Market's contact information of firms is regularly updated.

[35]Enumerators were sent to four other cities to recruit firms: Ganta, Saclepea, Saniquellie and Karnplay.

[36]At this point, the drivers were not informed about the GPS trackers. They were told that I was conducting a study on the trucking industry. Enumerators did not report any driver who refused to give out the contact information of the owner of the truck, or other superior.

point we already had the contact information of most of these firms.

All the firms that had at least one truck were given the opportunity to participate in the study. Out of the 76 firms for which we had contact information of, 62 agreed to participate in the experiment (82%). These 62 companies represent 152 trucks and 160 drivers[37]. Since there is no census of trucking companies in Liberia, it is hard to know what share of the total population the sample represents, but I am confident that it represents a significant share of the overall universe of trucking companies in Liberia.[38]

Table 1 presents summary statistics. 98% of drivers are always assigned to the same truck, which means that in most cases treatment is at the driver level. A "trip" is defined as the journey from origin to destination (the return journey is not considered as being part of the trip). On average, trips in the sample last 26 hours, a little more than a day, but the time spent on a trip varies a lot among trips. The average speed of drivers at baseline (which took place during the rainy season) is 18 kilometers per hour (around 11 mile per hour). However, this only reflects the average speed of trucks during the rainy season, and the average speed of the control group during the dry season is significantly higher, at 28 kilometers per hour.

3.2 Treatment

Trucks in the treatment group were randomly assigned a GPS tracker.

3.2.1 Randomization

The treatment was rolled out in two phases.

In August 2016, firms were recruited and the owner[39] of the truck was asked to sign a Memorandum of Understanding giving the authorization to install a GPS tracking device on all the trucks in his company.[40] Managers and drivers were then interviewed, and upon completion of the interviews of all manager and drivers of a company, the company's trucks were randomly assigned to treatment and control groups, according to the following procedure:

1. Each firm was assigned a random share of trucks to be treated, with an average across firms of 2/3 of treated trucks.[41]

[37] Here, only drivers that are not managers are counted. These are the drivers of interest for the experiment.

[38] According to Dablanc (2010), Mali had 1,864 trucks in 2006, or 0.13 trucks per thousand people or 0.35 trucks per $ million GDP. Scaled to Liberia in 2016, this is equivalent to 600-735 trucks.

[39] For legal reasons, the owner's approval is necessary to start the installation. Most of the time the owner of the truck is the manager.

[40] The MoU detailed all the interviews the drivers and the managers would have to go through, as well as the procedure of GPS installation. For the firm to participate in the study, the owner had to give his approval for the interviews, and for the potential installation of GPS trackers. Drivers were individually asked for consent to be interviewed.

[41] Firms are assigned a random number - the share of treated trucks - according to a normal distribution centered in 2/3 and a standard deviation of 0.1. All firms assigned a number above 1 are assigned the number 1 and all

2. Within each firm, treated trucks were randomly selected according to the assigned share.

3. If a truck was on the threshold then it was randomly assigned to treatment or control such that the expected value of treated trucks was equal to the randomly assigned share.

At the end of this randomization, 104 trucks (68% of all trucks) where assigned to the treatment group and 48 trucks where in the control group.

Despite the treatment being entirely free for the firms and extensive efforts from the mechanics to install the devices, by March 2017 (7 months after the beginning of treatment) only 55 of the 104 trackers had been installed (53% of the treatment group).

Given the slow take-up of treatment a second round of randomization was done in April 2017. The second randomization followed the procedure of the first randomization, and assigned an additional one third of the control group to the treatment group. At the end of the second randomization, 19 more trucks were assigned to the treatment, bringing the total number of treatment trucks to 123 (81% of all trucks).

By the end of April, the number of trackers installed reached 80 out of the 123 assigned to treatment (65%). Despite strong efforts to install GPS trackers, no trackers were installed between May and September 2017.

Figure 3 shows the timeline of installation of the GPS trackers. The blue dashed line shows the number of trackers assigned to treatment, while the plain red line shows the number of trackers installed. While there appears to be some progression in the first months after the assignment of trucks to treatment, after a year, the number of installations seems to have reached a limit.

In the rest of the paper I will call treatment and control the groups after the second randomization as follows: the "control group" is the group of trucks that were not selected for treatment on either of the randomizations, and the "treatment group" is the group of trucks that where selected on either the first or the second randomization. Table 1 shows that treatment and control groups are balanced. The balance table for the first randomization can be found in the Appendix.

3.2.2 GPS trackers and installation

All the trucks in the treatment group were assigned a GPS tracker with a label stating the license plate of the truck they were assigned to (or other form of truck identification) as well as the name of the driver. To make sure that a truck in the control group did not wrongly receive a GPS tracker, the mechanic was asked to verify the license plate of the truck and the name of the driver before starting the installation.

GPS trackers used for the purpose of this study are small black boxes of around 3 inches width by 2 inches length (figure 4 shows a picture of a tracker). To be properly functional, GPS trackers

firms assigned a number below 1/3 are assigned 1/3. This random assignment of shares of treated trucks was initially done to investigate the extensive vs. intensive margin of treatment. However, results don't show any significant variation along the intensive margin of treatment, and won't be presented in this paper.

must be installed inside the truck's dashboard by a professional mechanic. The GPS tracker is directly connected to the battery of the truck, which allows it to stay on even when the engine is turned off. The tracker turns off soon after the battery of the truck is disconnected from the truck (which is pretty common in Liberia when the truck is parked, to avoid battery theft).[42]The installation takes about 30 minutes to 2 hours, depending on the truck. The mechanics were hired by the research team, and were sent to the truck's parking space, so that the tracker and the installation were completely free of charge for the company. Figure 4 shows a mechanic working on the dashboard to install the tracker.

Once the GPS tracker is properly installed on the truck, it sends the position of the truck through a mobile network to an online server. The combination of mobile networks used covers most of the road network, ensuring that the position of the truck is precisely reported over time. The server stores the history of recorded positions. Codes of access for the online platform to access the data were given to the owner of the truck.[43] A printed manual for using the online platform was also provided to the owner, as well as the telephone numbers that they could reach if they needed help. In the cases in which the owner was not the direct manager of the driver, the owner shared access codes with the manager. The bottom picture of figure 4 shows an example of trip.

It quickly became apparent that the drivers were unhappy about the GPS tracking devices. To ensure that the devices were installed on the trucks according manager's willingness, and not the driver's, the mechanics were asked to follow this procedure:

1. Mechanics contacted the driver assigned to the truck, and scheduled a meeting for the installation.

2. If the first step failed and the tracker could not be installed, the mechanics contacted the manager and asked him to schedule the meeting with the driver.

3. If the two first steps failed, the mechanics repeated the two first steps the following week.

The mechanics received a wage partly based on the number of devices installed. In addition, an independent enumerator verified that the mechanic contacted the drivers and showed up at meetings.[44] When a tracker failed[45], it was immediately replaced by a new tracker.

[42]Other more complex trackers are able to give other type of information such as the oil level. The GPS tracking devices used in this study do not; they only give information on the location of the truck.

[43]Driver were not given the codes to access the the online platform.

[44]While the mechanics sometimes failed to show up at meetings, significant effort was done to correct these errors. By the end of the treatment period, all failures of installation where due to a failure from the company's side.

[45]Fewer than five trackers were reported to fail during the experiment.

3.3 Data

Data collection was done both before and after the installation of the GPS trackers. Data collection was completed electronically by a team of twelve enumerators (one supervisor, ten interviewers, and one back-checker[46]).

Two types of interviews were completed: manager and driver interviews.

3.3.1 Manager interviews

Managers were first interviewed at baseline, upon recruitment of the firms. These rounds of interviews lasted from the end of August to the end of October 2016. After the baseline, follow-up interviews were completed regularly. Three rounds of follow-ups were completed: January 2017, February 2017 and March 2017.

The baseline and follow-up interviews of managers were very similar. They included questions on:

- The manager (the interviewee), including age, nationality and county of origin.
- The business, including registration status, number of employees, number of trucks and sector.
- The transport sector, including the state of roads, estimated travel times, the competition and the difficulty of finding loads.
- The truck, including type, age and price when bought.
- Each driver. Managers were asked to rank on a scale from 0 to 10 a set of statements such as their relationship to the driver or the care the driver takes of the trucks.
- The drivers' trips. For each trip completed in the past month that they could remember[47], the managers were asked questions about the origin and destination, estimated time of completion, the types and length of breaks, the commodity transported, technical issues and whether there was another truck from the same company on the trip.

The baseline interviews were all conducted in person. Follow-up interviews were done in person or on the phone at the convenience of the manager. The baseline interviews lasted up to one hour and thirty minutes, and follow-up interviews lasted on average half an hour.[48]

[46]Interviewers complete interviews on the phone or in-person. The supervisor makes sure that the interviewers are accurately completing interviews. The back-checker contacts a randomly chosen sample of 10 percent of interviewees and corroborates answers collected by interviewers.

[47]To limit the length of the interview, we limited the number of trips to three per driver. If the driver completed more than three trips in the last month, the manager was asked about the three main trips completed.

[48]Since interviews were relatively long and frequent, we prevented manager's exhaustion by interviewing them at their own convenience. Also, the enumerators were assigned to companies so that the managers knew the enumerators calling them.

3.3.2 Driver interviews

Drivers were interviewed at baseline, in the same period that the manager baseline was completed. There was one follow-up interview for drivers, in April 2017.

Driver interviews were shorter than managers interviews and included questions on:

- The drivers (the interviewee), including their wage, their nationality and their county of origin. Drivers were asked to rank on a scale from 0 to 10 a set of statements which mimicked the managers' interviews. Statements included their relationship with their manager and the care they take of the trucks.
- Their trips. For each trip completed in the past month[49], the drivers were asked the same questions as the managers: origin and destination, estimated time of completion, the types and length of breaks, the commodity transported, technical issues and whether the was another truck from the same company on the trip.[50]

In addition to the questionnaire, enumerators where trained to explain the role of the GPS tracking device and the manager's ability to track the truck at baseline.[51] Driver interviews were all carried out in person during the baseline and follow-up, except when an in-person was not feasible. Interviews usually lasted between thirty and forty minutes.

Figure 5 shows the timeline of interviews.

3.3.3 Distance and speed

Given that the GPS trackers are only installed on the trucks of the treatment group, their data cannot be used for measuring treatment effects.

For this reason, I use data collected from interviews to estimate speed of drivers, and the length of breaks. For each trip, interviewees are asked about origin and destination, the time it took them to complete, and an estimate of the time they spent on breaks. To calculate speed, I additionally used a distance measure that is based on the distance between origin and destination locations as calculated by Google Maps.[52]

Providing a precise estimation of the length of a trip or a break is not an easy exercise, which led to a lot of measurement error. Observations with unrealistic speed estimates were dropped.[53]

[49] To limit the length of the interview, we limited the number of trips to three per driver. If the driver completed more than three trips in the last month, the manager was asked about the three main trips completed.

[50] Since driver interviews and the manager interviews are not overlapping in time (except for the baseline), it is not possible to directly compare the data on trips from drivers and managers interviews.

[51] Not informing the drivers about the GPS tracking devices would have put drivers at risk of being reprimanded or disciplined by their managers.

[52] Given Liberia's road network, there is often only one route to go from one city to another. In case Google Maps offered two itineraries, the most common route was confirmed by managers of trucking companies. In those cases, the distance difference was not high and the results are unaltered using one route or the other.

[53] See Appendix for more details on the cleaning procedure.

The reliability of the data from interviews and how it affects the results are discussed in a greater length in the next section, which explains the effect of the treatment on treated drivers.

4 Results of the Experiment

4.1 Econometric Specification

The key variation that I exploit is within drivers over time. I collected data before and after the treatment event, which allows me to use a difference-in-difference regression framework with driver fixed effects. To capture seasonal variation and infrastructure differences among roads I also use month and road fixed effects. I estimate both the intent-to-treat estimate and the local average treatment effect.

4.1.1 Specification at the trip level

The first specification I use is at the trip level. The main regression follows:

$$y_{itr} = \alpha_i + \beta_r + \gamma_t + X_{itr} + \delta T_{it} + \epsilon$$

Where y_{itr} is the output for driver i at time t on road r,
α_i, β_r, γ_t are driver, road, and time fixed effects,
X_{itr} are trip controls (the type of goods transported, technical issues during the trip, and whether there were several trucks from the same company on the trip),
T_{it} is the treatment variable, which takes the value one when the truck of driver i was assigned to treatment at t,
and δ is the coefficient of interest.

The results of these regressions yield the intent-to-treat estimates. To compute the local average treatment effect on the treated, I instrument the treatment take-up with the assignment to treatment.[54]

4.1.2 Specification at the driver level

One concern is that treatment affects the number of trips the driver completed, and biases the number of observations in each group. To avoid that, in a second specification I reduce the number of observations at the driver-period level (rather than at the trip level). I apply a fixed effects method in a first stage, and use the fixed-effects of the first stage to run the second stage.

First stage:

$$y_{itr} = \lambda_{it} + \beta_r + X_{itr} + \epsilon$$

[54] In all tables, two stage regressions are done using 2SLS.

Where y_{itr} is the output for driver i at time t on road r,
β_r is a road fixed effect,
X_{itr} are trip controls (the type of goods transported, technical issues during the trip, and whether there were several trucks from the same company on the trip),
and λ_{it} is a driver-period[55] fixed effect

Second stage:

I recover the driver-period fixed effects from the first stage, and use it as the output for the second stage. Now the observations are at the driver level.

$$\hat{\lambda_{it}} = \alpha_i + \gamma_t + \delta T_{it} + \epsilon$$

Where T_{it} is the treatment variable, which takes the value one when the truck of driver i was assigned to treatment at t,
and α_i and γ_t are driver and period fixed effects.

4.2 Effect of Monitoring on Driver Performance

A monitoring device should increase workers' efficiency along measures of effort monitored by the device. In this section, I study the effect of GPS tracking devices on speed and length of breaks.

4.2.1 Effect on Speed

The first output I explore is the average speed of drivers. By definition of the treatment only the treated group has a GPS tracker, so that the average speed is estimated with information from interviews (completion time, destination and origin). I compute separate estimates based on drivers and managers interviews.

In this case, the output (y_{itr} in the previous section) is the average speed (of driver i, on road r, which happened at time t).

For clarity, all tables in this section present the specifications in the same order. Columns (1), (2) and (3) show "Instrumental Variable" estimates, which is the local average treatment effect on the treated. Columns (4), (5) and (6) show "Reduced Form" estimates, also called intent-to-treat estimates, which is the effect of assignment to treatment. Columns (1) and (4) have no fixed effects, columns (2) and (5) have period, driver and road fixed effects, and columns (3) and (6) have the same fixed effects as (2) and (5), and additional trip controls (fixed effect on types of goods transported, technical issues during the trip, and whether another truck from the same company was on the trip).

Tables 3 and 4 show the effect of the treatment on speed of drivers, at the trip level. Table 3 was computed using data from the manager interviews. It shows that the installation of the GPS

[55] Two periods are defined: before and after the randomization.

tracker significantly increase the average speed of the drivers, as estimated by their managers. The local average treatment estimate indicate that, according to managers, monitoring devices increased the speed of treated drivers by about 21 kilometers per hour. Due to the low take-up of the treatment - which is explored in the next section - the intent-to-treat estimates are smaller. Being assigned to treatment increased the average speed of the driver by 9.6 kilometers per hour.

To ensure that these results do not come from bias in the managers' interviews, I compute estimates based on drivers interviews. Managers might be biased if in order to keep good relationships with the research team[56], they over-estimate the effect of the tracker[57]. If this is the case, estimates from table 3 may be an over-estimation of the effect. Table 4 shows the same specification, but in this case speed was estimated from the drivers interviews. While drivers can also be biased in answering the interviews, they on the contrary would want to under-estimate the effect of the tracker.[58] In this case, estimates from table 4 would be an under-estimation. Estimates from table 4 show that the effect of the treatment on speed is very similar than the ones from table 3. The local average treatment estimate indicated that, according to drivers, monitoring devices increased the speed of drivers by about 18 kilometers per hour, while the intent-to-treat estimate is 10 kilometers per hour.

The specification at the driver level (described in section 4.1.2) gives very similar estimates, presented in tables 5 and 6. The preferred specification is the estimation of the effect of treatment at the driver level, using driver interviews. This estimation shows that the treatment increased the speed of treated drivers by 14 kilometers per hour. This is the specification that will be used in the remainder of the paper. Other specifications are available in the Appendix.

4.2.2 Effect on Breaks

The installation of monitoring seems to significantly increase the speed of treated drivers, however if drivers go faster on the road, this may have unexpected consequences (such as more accidents on the road).

In this section, I explore the effect of the treatment on the length of the breaks taken by the driver during a trip. I measure the length of breaks from interviews both in hours and as a percentage of the total trip. Here, breaks include all sorts of reasons to stop, such as deliveries, stops because of mud on the road, or personal breaks.

y_{its} is the length of breaks of the total time of the trip, for driver i, on trip t, which happened at time s.

Tables 7 and 8 show the effect of the treatment on the length of breaks. Table 7 shows the

[56] Given the positive and significant effect of the trackers on the drivers' efficiency, it would not be surprising for the companies to want trackers on their trucks in the control group as well.

[57] Note that nor the managers nor the drivers were told what the purpose of the research, but the managers could have guessed that the speed was one of the outputs of interest.

[58] Given that, as seen in section 6, drivers on average prefer not to be monitored, they would want to convince the research team that monitoring drivers have no effect on their performance.

estimate of the effect of monitoring on breaks in hours, and table 8 shows estimates on breaks as a percentage of the total trip. Both tables show that GPS trackers significantly reduced the length of breaks during a trip. Estimates show that drivers with GPS trackers spend around 25% of their trip less on breaks.

These results suggest that the GPS trackers increase speed by decreasing the number of breaks the driver takes.

4.2.3 Effect on accidents, maintenance costs and technical problems

To further show that the increase in speed did not have a detrimental effect, table 9 presents estimates of the treatment effect on the number of accidents, maintenance costs and technical problems. The coefficients are very small and not significant.

Estimates reveal that the tracking devices significantly increases monitored drivers' average speeds, without leading to higher accident rates or maintenance costs. In the following sections, I explore why managers refused to install monitoring devices on selected trucks.

5 Model and Predictions

Despite the fact that GPS trackers were offered at zero cost, and clear evidence of benefit on monitored drivers, the take-up rate of monitoring devices remains low even a year after the start of the experiment. Managers deliberately declined to install the devices on 35 percent of the trucks selected for treatment. In this section I describe a mechanism based on Bénabou and Tirole (2006) that can account for the incomplete take-up rate. The model has two predictions which will be tested empirically in the following section.

5.1 Model

The agent takes action a, which sends a signal $s = a + \epsilon$ (with $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon)$) to the principal. The key element of the model is that when taking action a, the agent has an intrinsic motivation for taking that action, and he values the principal's opinion of that intrinsic motivation.

The principal rewards the agent for his action and also chooses whether he want to monitor the agent or not. Monitoring allows the principal to reduce the variance of the signal s and better monitoring allows him to better provide incentives to the agent. This means that a tighter monitoring, m, gives more incentives to the agent. [59]. The agent values these extrinsic incentives based on the signal s with v_m ($v_m > 0$). The agent also has an intrinsic motivation for contributing to the task, v_a ($v_a > 0$).

[59] If the principal can better measure the output of the driver, he (i) can provide "better" bonuses – bonuses that better reflect the agent's effort – and (ii) has a higher probability of catching shirking behavior. In both cases, the agent has higher incentives to provide effort.

v_a and v_m are private information of the agent, and over the population:

$$\begin{pmatrix} v_a \\ v_m \end{pmatrix} \sim \mathcal{N}\left(\begin{matrix} \bar{v_a} \\ \bar{v_m} \end{matrix} , \begin{bmatrix} \sigma_a^2 & 0 \\ 0 & \sigma_m^2 \end{bmatrix} \right)$$

The action a has a cost for the agent, $C(a) = \frac{1}{2}ka^2$. The agent's direct benefit can be written:

$$B = w + v_a a + v_m m \cdot s - C(a)$$

Additionally, the agent values the principal's opinion of his intrinsic motivation. We suppose that that valuation depends linearly on the principal's posterior observation:

$$\mu \mathbb{E}[v_a|s, m]$$

Where μ is known to both the agent and the principal. This parameter is key in this model. It represents how much the agent cares about the principal's opinion. For example, the agent will value the principal's opinion if she cares about her reputation (for future employment or recommendations), or if the principal can retaliate through their common network (in the case where they do have a common network). Note that μ is a value of the principal-agent pair. How the agent cares about the principal's opinion depends on the identity of the principal. This parameter represents an informal contract between the principal and agent, and we suppose that the value of μ is common knowledge.

An agent with preferences $\boldsymbol{v} \equiv (v_a, v_m)$ and "image concerns" μ has the utility:

$$U = w + v_a a + v_m m \cdot s - C(a) + \mu \mathbb{E}[v_a|s, m]$$

There is uncertainty about the signal, s, and we suppose for simplicity that the agent is risk-neutral so that he maximizes his expected utility:

$$\max_a \mathbb{E}_s[U|a] = w + v_a a + v_m m \cdot a - C(a) + \mu \mathbb{E}_s \left[\mathbb{E}[v_a|s, m] | a \right]$$

Which yields:

$$k \cdot a - \frac{\partial R(a, m)}{\partial a} = v_a + v_m \cdot m \tag{1}$$

Where $R(a, m) = \mu \mathbb{E}_s \left[\mathbb{E}[v_a|s, m] | a \right]$

Bayesian updating (with no prior on the action a) yields:

$$\mathbb{E}[a|s, m] = s$$

Given that $\begin{pmatrix} v_a \\ v_m \end{pmatrix} \sim \mathcal{N}\left(\begin{array}{c} \bar{v_a} \\ \bar{v_m} \end{array}, \begin{bmatrix} \sigma_a^2 & 0 \\ 0 & \sigma_m^2 \end{bmatrix} \right)$, and $\mathbb{E}_s\left[\mathbb{E}[v_a|s,m]|a=s\right] = \mathbb{E}[v_a|s,m]$:

$$\mathbb{E}[v_a|s,m] = \frac{\bar{v_a}\sigma_m^2 m^2 + \left(k \cdot s - \frac{\partial R(s,m)}{\partial a} - \bar{v_m} m\right)\sigma_a^2}{\sigma_a^2 + \sigma_m^2 m^2}$$

Now from the perspective of the agent, his expected reputation is:

$$\mathbb{E}_s\left[\mathbb{E}[v_a|s,m]|a\right] = \frac{\bar{v_a}\sigma_m^2 m^2 + \left(k \cdot a - \frac{\partial R(a,m)}{\partial a} - \bar{v_m} m\right)\sigma_a^2}{\sigma_a^2 + \sigma_m^2 m^2} \tag{2}$$

Note that, by definition

$$\frac{1}{\mu}R(a,m) = \mathbb{E}_s\left[\mathbb{E}[v_a|s,m]|a\right]$$

So, equation 2 is a linear differential equation in $R(a,m)$, which can be solved:

$$\frac{\partial R(a,m)}{\partial a} = \mu k \frac{1}{1 + m^2 \frac{\sigma_m^2}{\sigma_a^2}}$$

And from equation 1:

$$a = \frac{v_a + v_m \cdot m}{k} + \mu \frac{1}{1 + m^2 \frac{\sigma_m^2}{\sigma_a^2}}$$

So on average over the population, the optimal level of effort chosen by an agent with image concerns μ is:

$$\bar{a} = \frac{\bar{v_a} + \bar{v_m} \cdot m}{k} + \mu \frac{1}{1 + m^2 \frac{\sigma_m^2}{\sigma_a^2}} \tag{3}$$

Equation 3 has two terms. The first term, $\frac{\bar{v_a} + \bar{v_m} \cdot m}{k}$ is the result of explicit incentives: intrinsic and extrinsic incentives. This term is increasing in monitoring, m. The second term, $\mu \frac{1}{1 + m^2 \frac{\sigma_m^2}{\sigma_a^2}}$, represents the adverse effect of monitoring and is decreasing in monitoring. The intuition for this last term is that tighter the monitoring undermines the agent's ability to signal his intrinsic motivation. If the second term's negative effect over-compensates the first term's additional incentives, the effect of monitoring might be detrimental to the level of effort provided by the agent.

The model presented here is similar to the model of Bénabou and Tirole (2006), with two differences. First, while the agent does want to appear "prosocial" ($\mu > 0$), she does not care about looking "greedy"(she does not care about the expected value of v_m). The reason this assumption was dropped is that I don't need it to explain the phenomena observed in this study. However, adding it to the model would not change the conclusions or the interpretations. Second, the principal here does not observe directly the effort of the agent. This assumption had to be added since the principal here is deciding to monitor or not, which would not be a option if he observed the agent's effort.

When determining if he should monitor or not the agent, the principal has to determine if the agent's optimal effort is increasing or decreasing in monitoring. The parameter μ (known to both the agent and the principal) is key in this setting. The interpretation and predictions of the model are presented in the next section.

5.2 Predictions

Figure 6 shows how the population-average optimal level of effort ($\bar{a}$) varies with the level of monitoring, for agents with different levels of μ. Given that μrepresents how the agent values the principal's opinion of her intrinsic motivation, μ will be referred to as "image concerns". The figure shows that while for low levels of image concerns (low μ), effort is an increasing function of monitoring, for some high levels of image concerns, the optimal level of effort is decreasing and then increasing in monitoring m. The figure illustrates two main predictions.

First, monitoring increases effort for agents with low image concerns and is counter-productive for agents with high image concerns. In figure 6, two levels of monitoring are presented, $m = 0.5$ and $m = 1.5$. For agents with low image concerns ($\mu = 0$ or $\mu = 1$), transitioning from $m = 0.5$ to $m = 1.5$ increases the optimal level of effort. Conversely, for agents with high image concerns ($\mu = 3$ or $\mu = 4$), transitioning from $m = 0.5$ to $m = 1.5$ decreases the optimal level of effort. If, in choosing whether to monitor or not, the manager observes μ, he will monitor agents with low image concerns.

Second, the effect of monitoring is greater the lower are image concerns (the lower μ). In figure 6, transitioning from $m = 0.5$ to $m = 1.5$ increases the optimal level of effort for both agents with $\mu = 0$ and $\mu = 1$. However the impact of monitoring (the difference between the optimal level effort at $m = 0.5$ and the new optimal level at $m = 1.5$) is greater for the agent with $\mu = 0$.

These two predictions depend on image concerns. In order to test them, I need a measure of image concerns. In figure 6, it is clear that for any given level of monitoring, higher image concerns predict a higher level of effort. Therefore, a proxy for image concerns is the baseline level of effort. Other characteristics of the manager-driver pair may also predict higher image concerns, such as a matching county of origin: if the drivers come from the manager's county, he is more likely to care about his reputation and will have higher image concerns.

The two predictions of the model become:

1. The effect of monitoring is positive for agents who show low effort at baseline and monitoring is counter-productive for agents that showed high effort at baseline.

2. The effect of monitoring is decreasing in baseline effort and can become negative for drivers who show high effort at baseline.

The next section tests these predictions.

6 Empirical Evidence

6.1 Baseline characteristics on treatment take-up

Managers decide which drivers will and will not receive a GPS tracker. According to the model, they base their decision on their priors about the driver. In this section, I show that the managers decided to install a GPS tracker on the trucks of drivers that showed lower levels of effort at baseline, indicating lower image concerns.

Figure 7 shows how drivers in the treatment group compare at baseline according to treatment take-up. On every panel, the top line presents the average and the 95% confidence interval for drivers for whom the manager refused to install a GPS tracking device. The bottom line presents the same information but for drivers who received a tracker. The figure is based on interviews from managers.

The figure clearly shows that managers chose to install GPS tracking devices on drivers who showed lower performance at baseline. At baseline, according to managers, drivers who received the tracker were slower, were less able to follow the rules of the firm (such as not transporting unauthorized passengers or goods, or not using the truck for personal reasons), and were more likely to have accidents. During interviews, managers are asked to rate their relationship with their drivers on a scale from 0 to 10. Managers decided to install monitoring devices on drivers who received a lower "relationship index". Managers are also more likely to install a monitoring device on drivers from a different county of origin.

This evidence confirms the hypothesis that managers are more likely to install monitoring devices on drivers with whom they don't have "informal arrangements". In particular, the last piece of evidence (the fact that manager are less likely to monitor someone from their county of origin) is consistent with the interpretation of image concerns: drivers from the same county of origin than their managers are more likely to have high reputation concerns.

Figure 7, together with the results from section 4, show that the manager selected the driver who performed less well at baseline, and that on those drivers, the effect of monitoring was positive. I now explore the heterogeneity of the treatment effect.

6.2 Treatment heterogeneity on speed

In this section I examine the heterogeneity of the treatment on speed.

I use the specification at the trip level presented in section 4, with an interaction-term:

$$\hat{\lambda_{it}} = \alpha_i + \gamma_t + \pi T_{it} + \delta T_{it} \times base_i + \epsilon$$

where all the variables are defined as in section 4 and $base_i$ is a baseline proxy of effort. The coefficient of interest is δ, it indicated whether the baseline proxy of effort is positively or negatively correlated with the effect of treatment.

The baseline proxies for effort include the average time spent on breaks (as a percentage of total trip), the propensity of the driver to follow the rules of the business, the relation between the driver and the manager (as rated from 0 to 10), and whether they are from the same ethnicity.

Table 10 presents the results. Columns (1) to (5) explore the heterogeneity of the effect of treatment with respect to baseline proxies of effort. Drivers who completed trips faster at baseline, who respected the rules of the business, and who took good care of their truck, showed a higher effect of the treatment. This effect also appears for driver who were given a high "relationship index" by their manager, and who have the same county of origin.

I combine the different baseline characteristics of drivers into a "propensity to be treated" index. This index is the fitted value of a logit regression of the variable "received treatment" on baseline characteristics, such as baseline proxies for effort, "relationship index", or matching origins.[60] The index takes values from 0 to 1 and increases with the probability of being treated. Figure 8 shows the histogram of treated and not treated drivers with respected to the "propensity to be treated" index. As expected, treated drivers are especially numerous in high levels of the index, while drivers who were not treated tend to be in lower levels of the index. Following the predictions of the model, a high propensity to be treated reflects a driver with low image concerns.

Column (5) of table 10 shows the heterogeneity effect of treatment with respect to a "propensity to be treated" index. The effect is consistent with the previous columns: a higher propensity to be treated at baseline increases the effect of treatment.

6.3 Treatment heterogeneity on other outputs

Using the same specification than in the previous section, I explore the heterogeneity of the effect with respect to the "propensity to be treated" index.

Table 11 presents the results. I explore two different measures of effort. First, I explore the propensity of the driver to follow the rules of the business. Managers were asked to rank on a scale from 0 to 10, whether they believed the driver followed the rules of the business (Fully Disagree - 0 to Fully Agree -10). The results from this measure are presented in column (1). While treatment seems to have a positive effect on the propensity of the driver to follow the rules

[60] This regression is available in the appendix.

of the business for drivers with a high propensity to be treated, this effect disappears for drivers with a low propensity to be treated.

Second, I explore the propensity of the driver to take care of the trucks. Managers were also asked to rank on the same scale whether they believed the driver took good care of the truck. The results from this measure of effort are very similar, and are presented in column (2).

To explore whether the treatment was counter-productive for some drivers, I divide drivers in three groups according to their tercile in the "propensity to be treated" index. Figure 9 shows the treatment coefficients for each group. Clearly, the coefficient for the first tercile (the group with the lowest propensity to be treated) is negative: according to managers, treatment had a negative impact on the propensity of this group to follow the rules of the business and to take care of the trucks. This confirms the hypothesis that monitoring is counter-productive for drivers who provided a high level of effort at baseline. This also confirms the findings of Belot and Schröder (2016) who show that monitoring increases cheating on dimensions that are not monitored. Here I find that it lowers the level of effort on these other dimensions, but only for one group of drivers (drivers who have a low propensity to be treated).

6.4 Robustness Checks

6.4.1 Effect of monitoring on relationships

According to the mechanism explored in this paper, if an employer decides to monitor the worker with whom he had a prior informal arrangement, their relationship deteriorates.

During interviews, managers are asked to rank between 0 and 10 their relationship with their drivers. While on average the treatment did not seem to have any effect on this "relationship index", the impact is very heterogenous. Table 12 shows the heterogeneity of the effect with respect to the "propensity to be treated" index. For drivers with a high propensity to be treated index have a positive impact of monitoring on their relationship with their manager, the impact is negative for drivers with a low index. Separating the sample into three terciles of this index, shows that effect is in fact negative for drivers with a low score. Figure 10 shows the graph of the coefficients for these different groups.

Managers are also asked to rank between 0 and 10 their trust in each of their drivers. Results for this "trust index" are similar than for the relationship index, confirming that the decision to monitor is a signal of distrust to the drivers.

6.4.2 Treatment heterogeneity according to drivers

All the results presented until now (expect results on speed and breaks) were based on manager interviews. In this section I will present results based on driver interviews.

During each interview, drivers are asked to rank some statements on a scale from 0 (Disagree) to 10 (Agree). I explore the effect of treatment on three different statements: "I like working for this firm", "I like being a driver", "I am a good driver". Drivers are also asked to rank their relationship with their manager on a scale from 0 to 10.

The heterogeneity of each of this measures with respected to the propensity to be treated index is presented in table 13 and figure 11. Results are consistent with manager interviews: drivers with a low propensity to be treated are more negatively affect by the treatment. They are less likely to like their firm, less likely to enjoy their work, and less likely to be good drivers. The impact on their relationship with their manager goes in the same directing, though less significantly. [61]

6.4.3 An estimate of the benefits from monitoring

If the benefits from installing a GPS trackers were small, a small but positive cost of installation would explain why managers chose not to install GPS tracking devices on some trucks.[62] In this section, I present an back-of-the-envelope estimate of the benefits from installing the monitoring device on monitored drivers.

After being explained what a GPS tracking device is, managers pointed out three main reasons why it could be beneficial to their firm. First, the devices would allow managers to better monitor drivers. Second, the devices would allow managers to estimate the time of arrival at the destination, as well as the time of return to the base, which would allow them to better allocate goods to trucks. Finally, in case the vehicle is stolen, the company would be able to locate the truck and possibly recover it.[63] In this project, when calculating the benefits of the GPS tracking devices, I focus only on the first effect, the monitoring of the driver. Benefits from this effect are directly measurable, as opposed to the two other effects. These two other effects also contribute to the overall effect of the monitoring device on transport costs, so in that regard, my estimates of the benefits of monitoring devices under-estimate the overall effect.

Table 14 shows estimates of trucks' marginal cost per kilometer. For reference, the same estimates are shown for the United States. The cost per kilometer is more than 2.6 times higher in Liberia than it is in the United States.[64]

As shown in section 4, monitoring devices significantly increase the speed of drivers, by 58 percent. Most of the costs per kilometer will not be affected by the number of trips completed by the truck. However, the price of the truck as well as the permits and licenses are fixed costs, and their impact will lower if the number of kilometers completed by the firm changes.

[61] Note that the number of observations from drivers' interviews is much smaller, since they only had one follow-up, while managers had three.

[62] Given that in this experiment, GPS trackers are given at zero cost (including the mechanic's pay), the "cost of installation" entirely comes from the opportunity cost of the driver and the truck not transporting goods during the installation.

[63] While this seemed to be a concern for managers at baseline, none of the trucks in my sample were stolen during the experiment.

[64] This coefficient is comparable to what Atkin and Donaldson (2015) find for Nigeria (2.2 times higher than the United States) and Ethiopia (3.3 times higher).

The last line of table 14 shows that the number of kilometers completed by a truck in the United States is almost 20 times higher than the number of kilometers completed by trucks in Liberia. If the speed of the truck significantly increases, the trucks in Liberia will be able to complete more trips and increase the number of kilometers they travel in a year. If the number of trips directly adapts to the increased speed of driver [65], the overall number of kilometers will increase by a factor 1.58, which will bring the overall number of km completed to 9,070. The price of the truck and the price of permits and licenses per kilometer will decrease, by the same factor. Assuming that the prices of fuel, repairs, bribes, wages and benefits remain constant per km, the overall marginal cost will come down to 2.16 $ per km. Given that on average a truck covers 5,742 kilometers per year, in one year a monitoring device would save USD 976.

If the benefits of monitoring drivers was constant across drivers, the cost of installation for the drivers that did not receive a tracking device despite being in the treatment group, should be higher than one thousand dollars, or 162 driver days[66]. This seems unlikely, and confirms the hypothesis that the benefits of monitoring are not constant across drivers.

7 Conclusion

The experiment demonstrates that the introduction of GPS monitoring to Liberian trucking firms results in significantly increased route completion speeds for monitored drivers. However, the effect of monitoring varies from driver to driver. These effects can be explained by segmenting the drivers into certain heterogeneous groups. Specifically, drivers who do not value the manager's opinion about them have high productivity returns on monitoring treatment. On the other hand, drivers who value the manager's opinion have low returns on treatment, with some cases exhibiting counter-productive effects on individual productivity. I show that these effects are consistent with the theory of motivation crowding.

This study indicates that productivity gains from technology adoption can be challenged by the presence of informal arrangements between principals and agents. This effect is especially pronounced in developing countries, where parties extensively rely on informal arrangements. However, this study should not be interpreted as evidence that the introduction of technology is counter-productive across-the-board. The conclusion to be drawn here is that blind application of monitoring technology may produce a sub-optimal effect on overall productivity. To maximize productivity gain, the choice of which drivers to monitor should take into consideration the motivations of each individual driver. In the long term, this experiment indicates that GPS tracking devices could allow firms to hire drivers outside of their network without a corresponding loss of productivity.

In addition, this study contributes to the explanation of high transport costs in developing

[65] This is assuming that the trucking companies are facing unlimited demand.
[66] Drivers are paid on average USD 6 per day.

countries by demonstrating that there is a substantial cost due to information asymmetry between managers and drivers. Resolving this information asymmetry (via GPS monitoring) can significantly lower the marginal costs of these firms, reducing overall transport costs.

References

Atkin, D. and Donaldson, D. (2015). Who's Getting Globalized? The Size and Nature of Intranational Trade Costs.

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

Baker, G. P. and Hubbard, T. N. (2004). Contractability and Asset Ownership: On-Board Computers and Governance in U.S. Trucking. *The Quarterly Journal of Economics*, 119(4):1443–1479.

Békir, I., Harbi, S. E., Grolleau, G., Mzoughi, N., and Sutan, A. (2015). The Impact of Monitoring and Sanctions on Cheating: Experimental Evidence from Tunisia. *Managerial and Decision Economics*, 37:461–473.

Belot, M. and Schröder, M. (2016). The Spillover Effects of Monitoring: A Field Experiment. *Management Science*, 62(1):37–45.

Bénabou, R. and Tirole, J. (2003). Intrinsic and Extrinsic Motivation. *The Review of Economic Studies*, 70(3):489–520.

Bénabou, R. and Tirole, J. (2006). Incentives and Prosocial Behavior. *The American Economic Review*, 96(5):1652–1678.

Bergquist, L. F. (2016). Pass-through, Competition, and Entry in Agricultural Markets: Experimental Evidence from Kenya. Job Market Paper.

Björkman, M. and Svensson, J. (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.

Blader, S., Gartenberg, C., and Prat, A. (2016). The Contingent Effect of Management Practices. Discussion Paper Series, Industrial Organization, Discussion Paper No. 11057.

Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., and Roberts, J. (2013). Does Management Matter? Evidence from India. *The Quarterly Journal of Economics*, 128(1):1–51.

Brown, M., Falk, A., and Fehr, E. (2004). Relational Contracts and the Nature of Market Interactions. *Econometrica*, 72(3):747–780.

Dablanc, L. (2010). Freight Transport for Development Toolkit : Urban Freight.

Dickinson, D. and Villeval, M.-C. (2008). Does Monitoring Decrease Work Effort? The Complementarity Between Agency and Crowding Out Theories. *Games and Economic Behavior*, 63:56–76.

Duflo, E., Hanna, R., and Ryan, S. P. (2012). Incentives Work: Getting Teachers to Come to School. *American Economic Review*, 102(4):1241–1278.

Ellingsen, T. and Johannesson, M. (2008). Pride and Prejudice: The Human Side of Incentive Theory. *The American Economic Review*, 98(3):990–1008.

Falk, A. and Kosfeld, M. (2006). The Hidden Cost of Control. *The American Economic Review*, 96(5):1611–1630.

Foster, A. D. and Rosenzweig, M. R. (1994). A Test for Moral Hazard in the Labor Market: Contractual Arrangements, Effort, and Health. *The Review of Economics and Statistics*, 76(2):213.

Frey, B. S. and Jegen, R. (2000). Motivation Crowding Theory. *Journal of Economics Surveys*, 15(5):589–611.

Gneezy, U. and Rustichini, A. (2000a). A Fine is a Price. *Journal of Legal Studies*, 29(1):1–17.

Gneezy, U. and Rustichini, A. (2000b). Pay Enough or Don't Pay at All. *The Quarterly Journal of Economics*, 115(3):791–810.

Greif, A. (1993). Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition. *The American Economic Review*, 83(3):525–548.

Grund, C. and Harbring, C. (2013). Trust and Control at the Workplace: Evidence from Representative Samples of Employees in Europe. *Journal of Economics and Statistics*, 233(5/6):619–637.

Holmstrom, B. and Milgrom, P. (1991). Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. *Journal of Law, Economics, and Organization*, 7:24–52.

Hubbard, T. N. (2000). The Demand for Monitoring Technologies: The Case of Trucking. *The Quarterly Journal of Economics*, 115(2):533–560.

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

Jackson, C. K. and Schneider, H. S. (2015). Checklists and Worker Behavior: A Field Experiment. *American Economic Journal: Applied Economics*, 7(4):136–168.

Jensen, M. C. and Murphy, K. J. (1990). Performance Pay and Top-Management Incentives. *Journal of Political Economy*, 98(2):225–264.

Lall, S. V., Wang, H., and Munthali, T. (2009). Explaining High Transport Costs within Malawi: Bad Roads or Lack of Trucking Competition. Policy Research Working Paper No. 5133.

Lavy, V. (2009). Performance Pay and Teachers' Effort, Productivity, and Grading Ethics. *The American Economic Review*, 99(5):1979–2011.

Lazear, E. P. (2000). Performance Pay and Productivity. *The American Economic Review*, 90(5):1346–1361.

Macchiavello, R. and Morjaria, A. (2015). The Value of Relationships: Evidence from a Supply Shock to Kenyan Rose Exports. *The American Economic Review*, 105(9):2911–2945.

Macleod, B. (2007). Reputations, Relationships, and Contract Enforcement. *Journal of Economic Literature*, 45(3):595–628.

Mellström, C. and Johannesson, M. (2008). Crowding Out in Blood Donation: Was Titmuss Right? *Journal of the European Economic Association*, 6(4):845–863.

Muralidharan, K. and Sundararaman, V. (2011). Teacher Performance Pay: Experimental Evidence from India. *Journal of Political Economy*, 119(1):39–77.

Sliwka, D. (2007). Trust as a Signal of Social Norm and the Hidden Costs of Incentive Schemes. *The American Economic Review*, 97(3):999–1012.

Teravaninthorn, S. and Raballand, G. (2009). *Transport Prices and Costs in Africa: A Review of the Main International Corridors*. Washington, DC: World Bank.

Figure 1: Liberia

Figure 2: Liberia's Roads

Figure 3: Timeline of GPS tracker installation

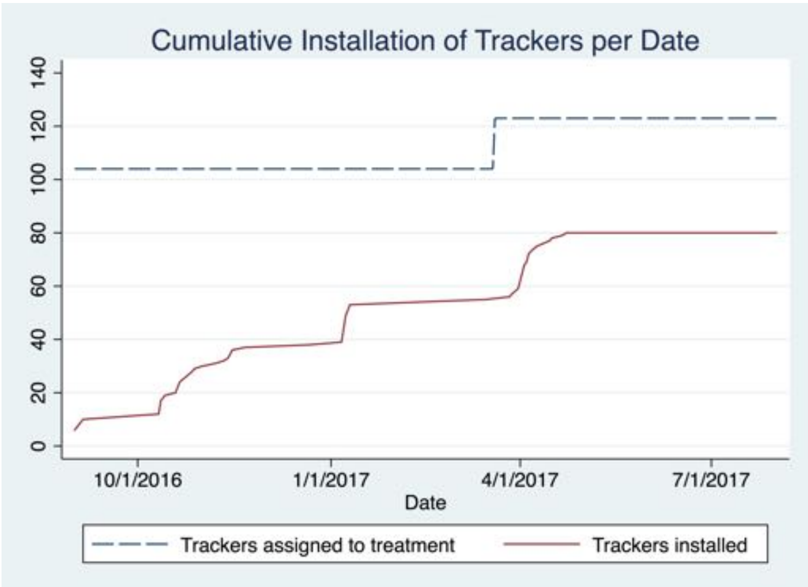

Figure 4: GPS Tracker and Truck Dashboard

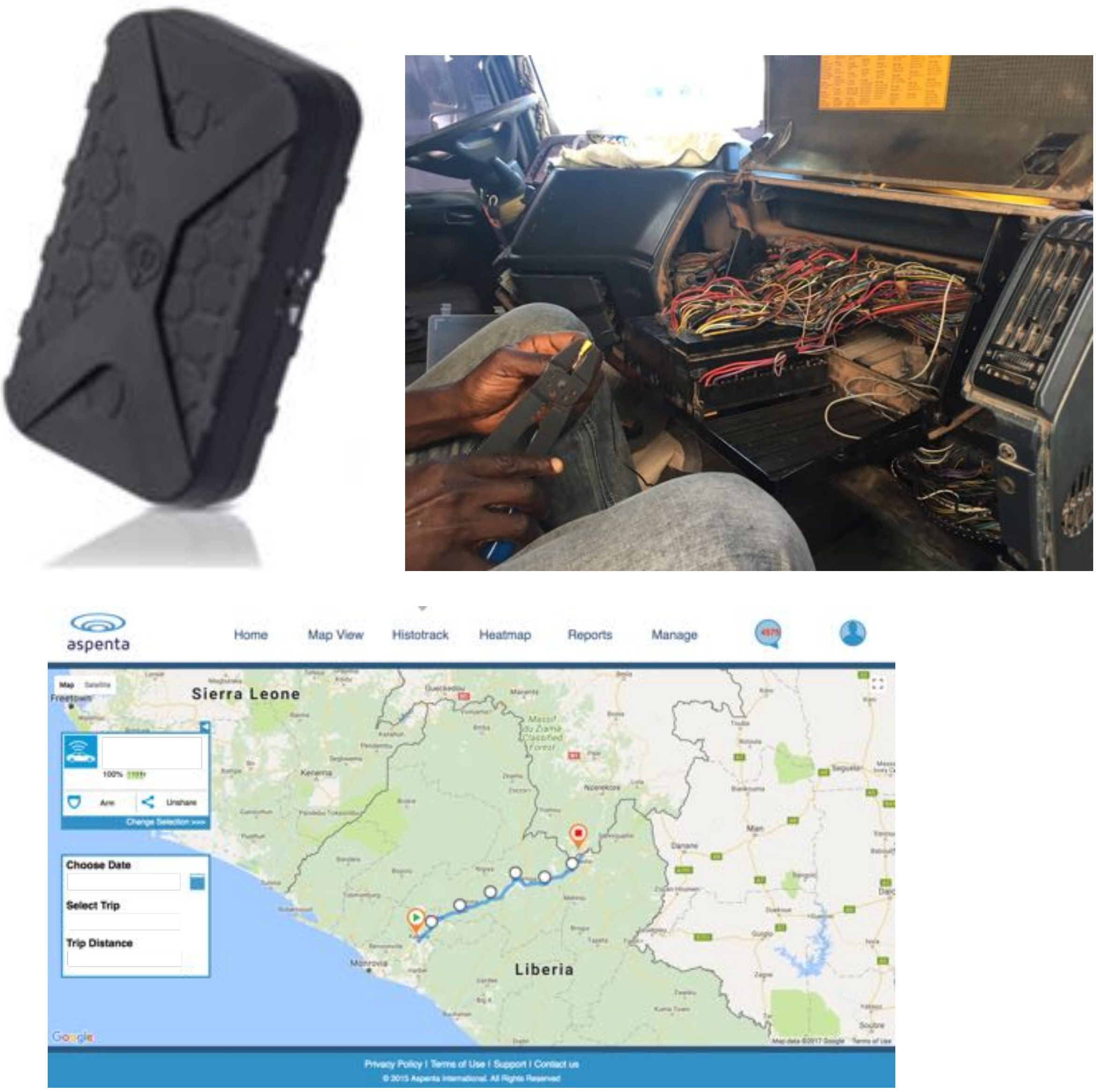

Figure 5: Timeline of Interviews

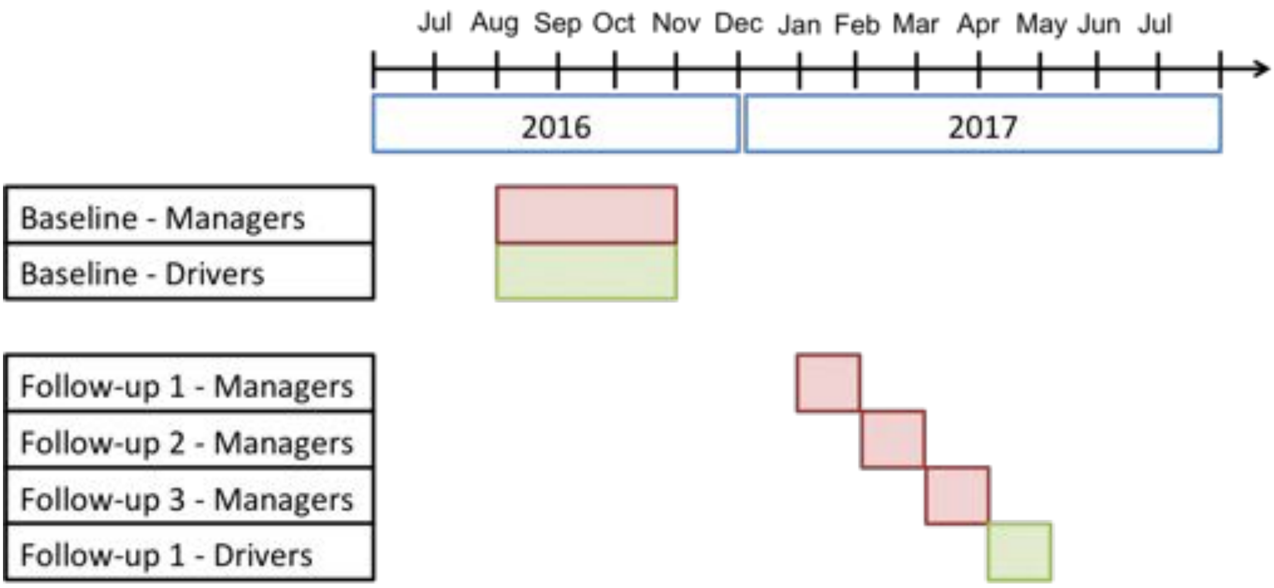

Figure 6: Optimal Effort Level as a Function of monitoring

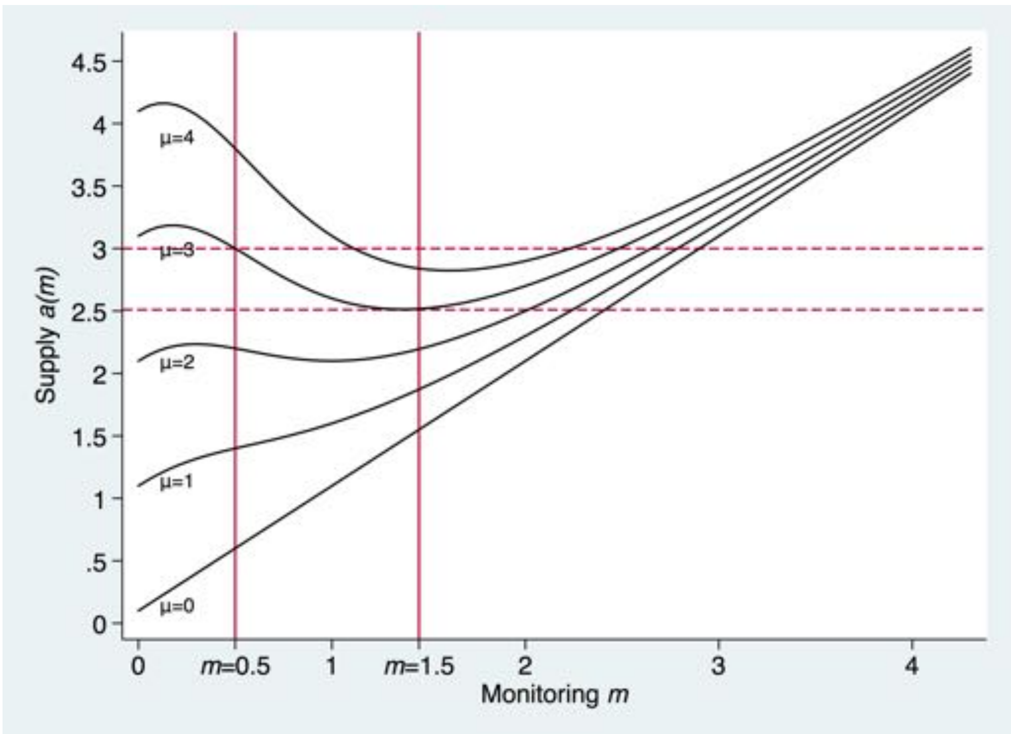

Figure 7: Baseline Comparison of Speed by Treatment Take-Up

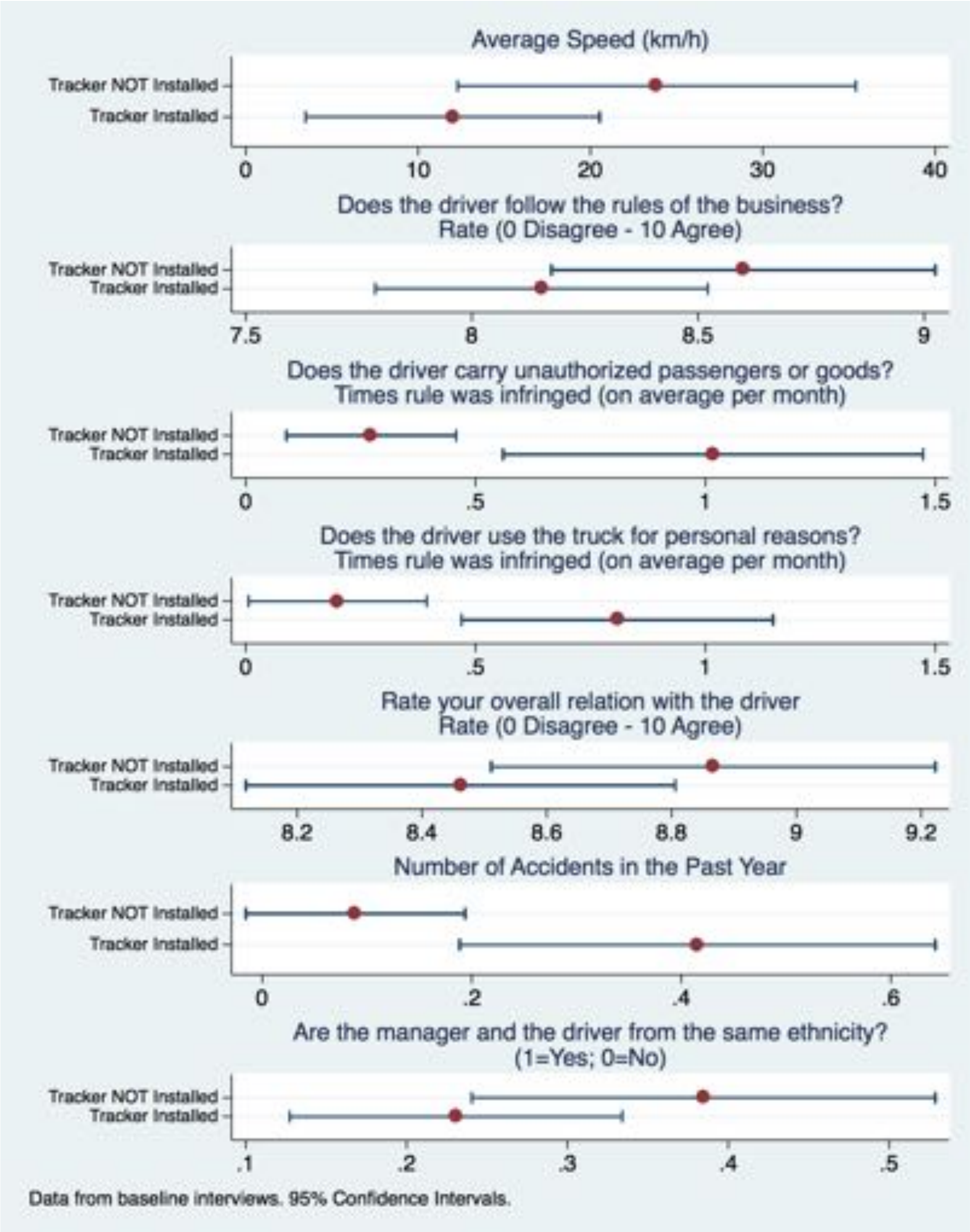

Figure 8: Propensity to be Treated

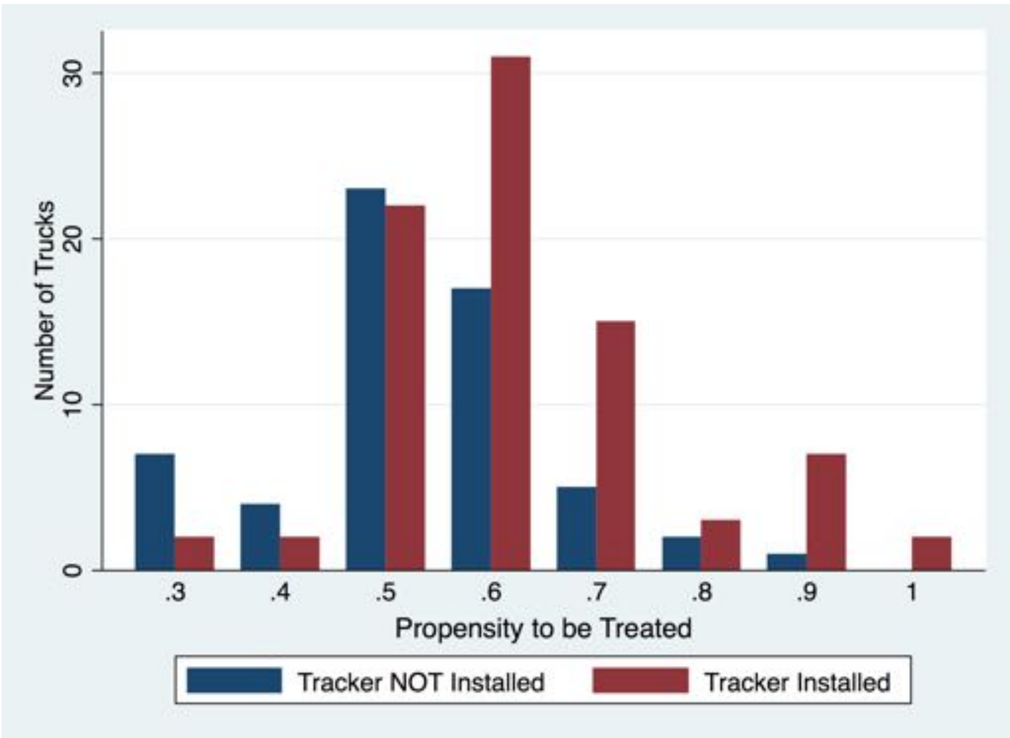

Figure 9: Effect of Treatment by Quartiles of "Propensity to be Treated" Index

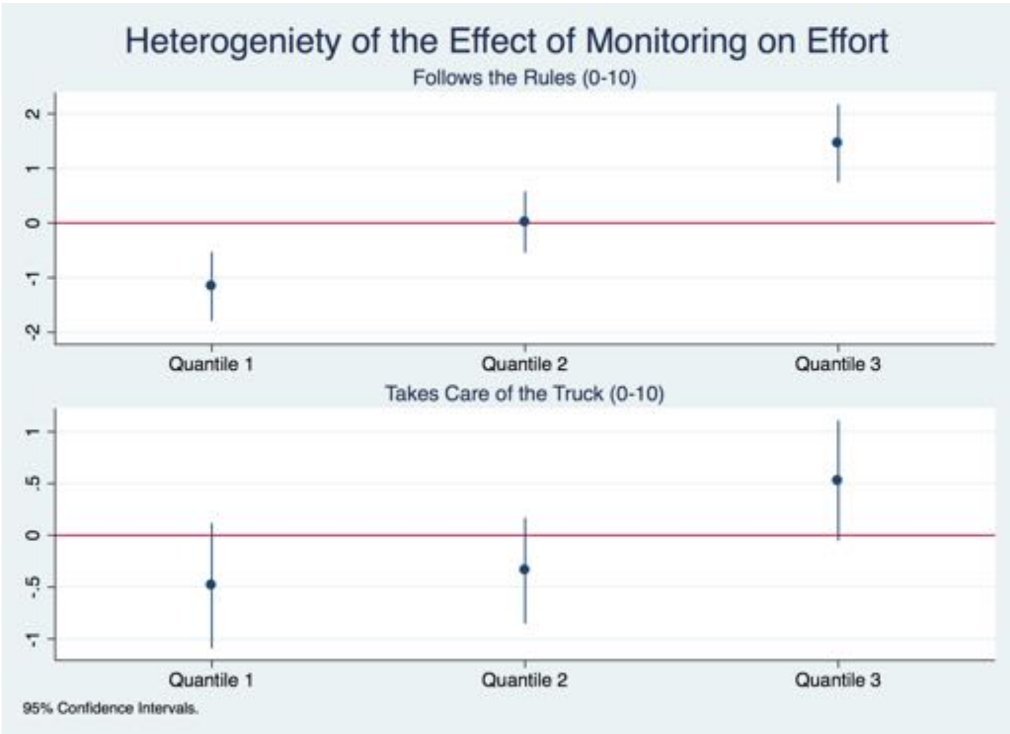

Figure 10: Effect of Treatment by Quartiles of “Propensity to be Treated” Index

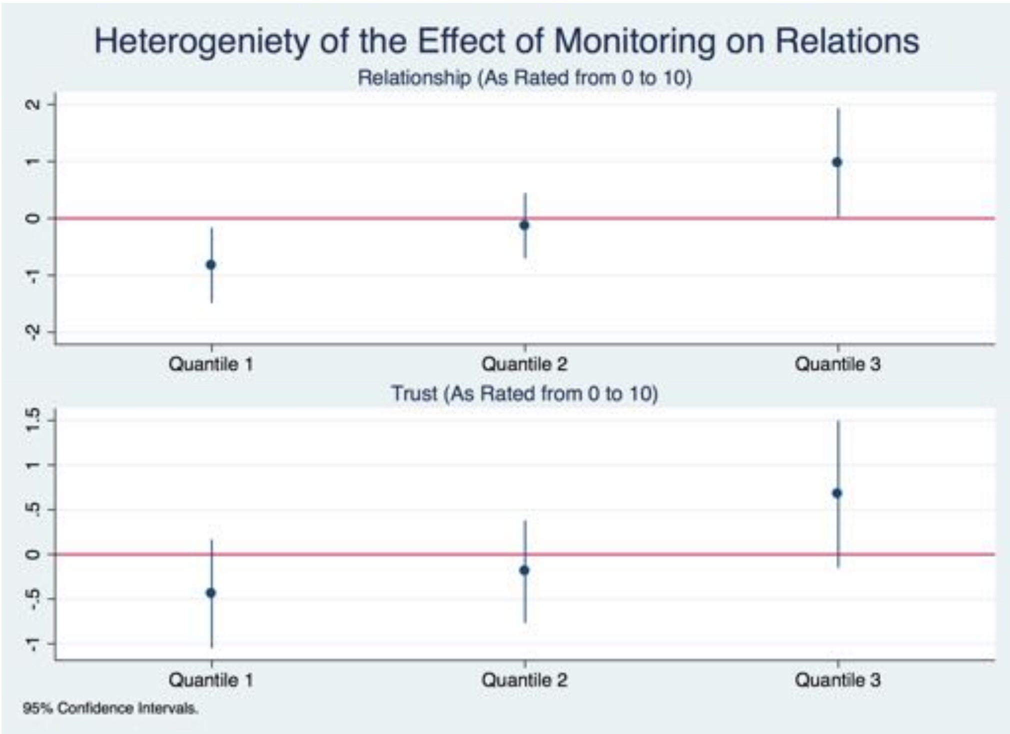

Figure 11: Effect of Treatment by Quartiles of "Propensity to be Treated" Index

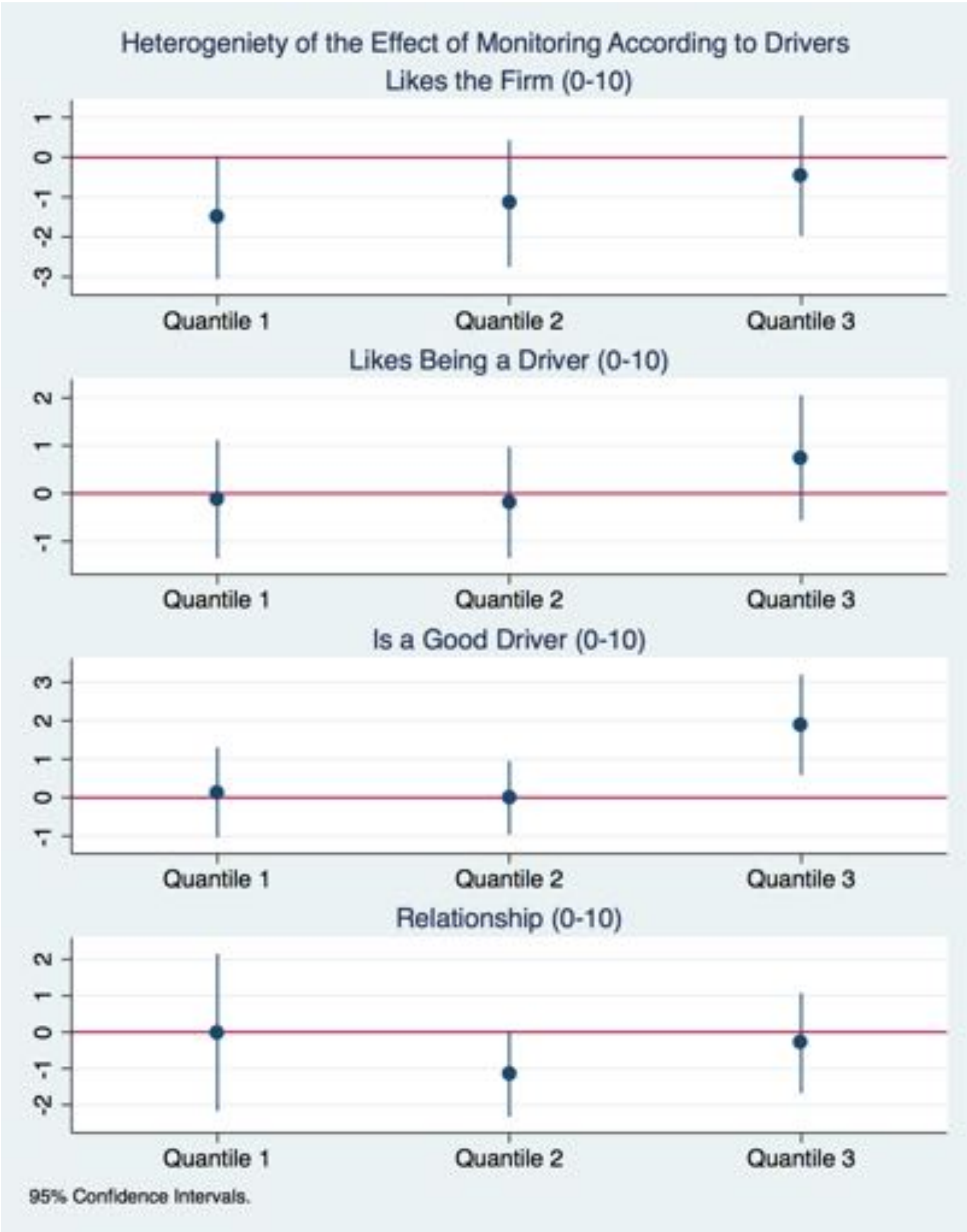

Table 1: Summary Statistics at Baseline

	Mean	Std. Dev.	Min	Max	Count
Managers					
How many trucks does the firm own?	3.98	5.92	1.00	33.00	62
Is transport the main activity?	0.76	0.43	0.00	1.00	62
Is the firm officially registered?	0.71	0.46	0.00	1.00	62
What is the total number of employees?	16.18	31.87	1.00	188.00	62
Does the firm deliver goods outside the country?	0.21	0.41	0.00	1.00	62
Gender of the interviewee (Male=0; Female=1)	0.02	0.13	0.00	1.00	62
Is the manager Liberian?	0.92	0.27	0.00	1.00	62
Is the manager a driver?	0.44	0.50	0.00	1.00	62
Does your business hire carboys?	0.56	0.50	0.00	1.00	62
How many tires did your business purchased last month?	4.15	4.30	0.00	20.00	62
Did the firm ever pay for overloading?	0.27	0.45	0.00	1.00	62
Has this firm ever been a victim of theft?	0.35	0.48	0.00	1.00	62
Is the firm based in Monrovia?	0.69	0.46	0.00	1.00	62
Drivers always assigned to same truck	0.98	0.23	0.00	2.00	55
Trucks					
Did the firm import this truck? (No=0; Yes=1)	0.65	0.48	0.00	1.00	152
Truck was second-hand when bought	0.86	0.35	0.00	1.00	152
Maintenance in the past month (USD):	609.09	1112.07	0.00	11000.00	152
Price of truck when bought (1k USD)	23.81	21.31	2.30	120.00	91
Drivers					
Did you ever had an accident?	0.42	0.49	0.00	1.00	142
Driver wage (in USD per day):	6.18	2.20	1.05	13.33	138
Trips					
Distance of trip (km)	275.04	209.96	4.00	753.00	177
Time spent on trips (hours)	26.40	28.80	0.50	144.00	177
Average speed (km/h) - Rainy Season	18.13	12.01	5.05	44.36	177
Time of trip spent on breaks (hours)	9.68	22.86	0.00	144.75	156
Percentage of trip spent on breaks (%)	12.99	15.23	0.00	60.00	143

This table was computed with data collected during baseline interviews. Statistcs on managers and trucks were computed using data from manager interviews, and summary statistics on drivers and trips were computed using data from driver interviews. Baseline interviews were collected between August and October 2016, which corresponds to the rainy season. In the case where several managers were interviewed in the same firm, the answers from the oldest manager were retained.

Table 2: Balance Table

	Diff(T-C)	Std. Err.	T-Stat	P-Val
Managers				
How many trucks does the firm own?	-3.42	2.27	-1.50	0.13
Is transport the main activity?	0.09	0.08	1.16	0.25
Is the firm officily registered?	0.02	0.08	0.33	0.75
Does the firm deliver goods outside the country?	-0.12	0.10	-1.19	0.24
Gender of the interviewee (Male=0; Female=1)	-0.01	0.02	-0.48	0.63
Is the manager Liberian?	0.04	0.04	1.10	0.27
Is the manager a driver?	0.05	0.08	0.66	0.51
Does your business hire carboys?	0.12	0.10	1.13	0.26
How many tires did your business purchased last month?	0.20	1.09	0.18	0.86
Did the firm ever pay for overloading?	0.03	0.09	0.27	0.78
Has this firm ever been a victim of theft?	-0.02	0.10	-0.17	0.87
Is the firm based in Monrovia?	0.08	0.08	0.98	0.33
Drivers always assigned to same truck	-0.04	0.04	-0.90	0.37
Trucks				
Truck was second-hand when bought	0.02	0.07	0.22	0.82
Maintenance in the past month (USD):	76.59	230.22	0.33	0.74
Price of truck when bought (1k USD)	-4.44	5.40	-0.82	0.41
Drivers				
Did you ever had an accident?	-0.10	0.11	-0.87	0.39
Driver wage (in USD per day):	-0.53	0.48	-1.10	0.28
Trips				
Distance of trip (km)	-13.37	26.82	-0.50	0.62
Time spent on trips (hours)	-6.93	5.45	-1.27	0.21
Average speed (km/h)	2.04	2.30	0.89	0.38
Time of trip spent on breaks (hours)	3.95	4.53	0.87	0.38
Percentage of trip spent on breaks (%)	2.53	3.32	0.76	0.45

This table was computed with data collected during baseline interviews. Statistcs on managers and trucks were computed using data from manager interviews, and summary statistics on drivers and trips were computed using data from driver interviews. Baseline interviews were collected between August and October 2016, which corresponds to the rainy season. In the case where several managers were interviewed in the same firm, the answers from the oldest manager were retained.

Table 3: Effect of Monitoring on Speed using Managers Interviews

	Instrumental Variable			Reduced Form		
	(1) Speed (km/h)	(2) Speed (km/h)	(3) Speed (km/h)	(4) Speed (km/h)	(5) Speed (km/h)	(6) Speed (km/h)
Treatment	15.71***	20.24***	21.04***			
	(5.765)	(5.202)	(5.357)			
Assignment to Treatment				6.167**	8.923**	9.646**
				(2.359)	(4.340)	(4.631)
Round FE		✓	✓		✓	✓
Driver FE		✓	✓		✓	✓
Road FE		✓	✓		✓	✓
Trip Controls			✓			✓
Observations	243	243	243	243	243	243

Standard errors are in parentheses and are clustered at the company-round level.

* p<.1, ** p<0.05, *** p<0.01

Table 4: Effect of Monitoring on Speed using Drivers Interviews

	Instrumental Variable			Reduced Form		
	(1) Speed (km/h)	(2) Speed (km/h)	(3) Speed (km/h)	(4) Speed (km/h)	(5) Speed (km/h)	(6) Speed (km/h)
Treatment	21.58***	18.31***	17.94***			
	(3.853)	(6.315)	(6.300)			
Assignment to Treatment				12.30***	10.41**	10.17**
				(2.531)	(4.011)	(3.995)
Round FE		✓	✓		✓	✓
Driver FE		✓	✓		✓	✓
Road FE		✓	✓		✓	✓
Trip Controls			✓			✓
Observations	349	349	349	349	349	349

Standard errors are in parentheses and are clustered at the company-round level.

* p<.1, ** p<0.05, *** p<0.01

Table 5: Effect of Monitoring on Speed Fixed Effects using Managers Interviews

	Instrumental Variable			Reduced Form		
	(1) Speed FE (km/h)	(2) Speed FE (km/h)	(3) Speed FE (km/h)	(4) Speed FE (km/h)	(5) Speed FE (km/h)	(6) Speed FE (km/h)
Treatment	18.25***	24.65***	24.65***			
	(7.051)	(9.012)	(9.012)			
Assignment to Treatment				6.263***	7.418*	7.418*
				(2.341)	(3.840)	(3.840)
Round FE		✓	✓		✓	✓
Driver FE		✓	✓		✓	✓
Road FE in First Stage		✓	✓		✓	✓
Trip Controls in First Stage			✓			✓
Observations	148	148	148	148	148	148

Standard errors in parentheses.

* p<.1, ** p<0.05, *** p<0.01

Table 6: Effect of Monitoring on Speed Fixed Effects using Drivers Interviews

	Instrumental Variable			Reduced Form		
	(1) Speed FE (km/h)	(2) Speed FE (km/h)	(3) Speed FE (km/h)	(4) Speed FE (km/h)	(5) Speed FE (km/h)	(6) Speed FE (km/h)
Treatment	17.76***	15.87***	14.34***			
	(3.508)	(4.112)	(4.261)			
Assignment to Treatment				11.17***	9.423*	8.513
				(2.314)	(5.278)	(5.487)
Round FE		✓	✓		✓	✓
Driver FE		✓	✓		✓	✓
Road FE in First Stage		✓	✓		✓	✓
Trip Controls in First Stage			✓			✓
Observations	176	176	176	176	176	176

Standard errors in parentheses.

* p<.1, ** p<0.05, *** p<0.01

Table 7: Effect of Monitoring on Breaks Fixed Effects using Driver Interviews

	Instrumental Variable			Reduced Form		
	(1) Breaks (hours)	(2) Breaks (hours)	(3) Breaks (hours)	(4) Breaks (hours)	(5) Breaks (hours)	(6) Breaks (hours)
Treatment	-7.343***	-12.43**	-12.31**			
	(2.550)	(5.640)	(5.597)			
Assignment to Treatment				-4.801***	-6.800	-6.734
				(1.650)	(6.364)	(6.366)
Season FE		✓	✓		✓	✓
Driver FE		✓	✓		✓	✓
Road FE in First Stage		✓	✓		✓	✓
Trip Controls in First Stage			✓			✓
Observations	143	143	143	143	143	143

Standard errors in parentheses.

* p<.1, ** p<0.05, *** p<0.01

Table 8: Effect of Monitoring on Breaks Fixed Effects using Driver Interviews

	Instrumental Variable			Reduced Form		
	(1) Breaks (%)	(2) Breaks (%)	(3) Breaks (%)	(4) Breaks (%)	(5) Breaks (%)	(6) Breaks (%)
Treatment	-7.363**	-24.83***	-25.25***			
	(3.292)	(6.156)	(6.000)			
Assignment to Treatment				-4.814**	-13.58*	-13.81*
				(2.106)	(7.318)	(7.463)
Season FE		✓	✓		✓	✓
Driver FE		✓	✓		✓	✓
Road FE in First Stage		✓	✓		✓	✓
Trip Controls in First Stage			✓			✓
Observations	143	143	143	143	143	143

Standard errors in parentheses.

* p<.1, ** p<0.05, *** p<0.01

Table 9: Effect of Monitoring on Accidents, Maintenance Costs and Technical Problems

	Instrumental Variable			Reduced Form		
	(1) Number of Accidents	(2) Maintenance Costs (USD)	(3) Technical Issues	(4) Number of Accidents	(5) Maintenance Costs (USD)	(6) Technical Issues
treated	0.0762	-33.42	-0.0288			
	(0.0971)	(255.3)	(0.124)			
intenttt				0.0288	-15.68	-0.0137
				(0.0367)	(119.6)	(0.0588)
Round FE	✓	✓	✓	✓	✓	✓
Driver FE	✓		✓	✓		✓
Truck FE		✓			✓	
Observations	755	687	730	755	687	730

Standard errors in parentheses. Standard errors are robust

* p<.1, ** p<0.05, *** p<0.01

Table 10: Heterogeneity of the Effect of Monitoring

	Reduced Form					
	(1) Speed FE (km/h)	(2) Speed FE (km/h)	(3) Speed FE (km/h)	(4) Speed FE (km/h)	(5) Speed FE (km/h)	(6) Speed FE (km/h)
Treatment	10.50	36.11**	22.34*	52.08**	10.10**	-19.57**
	(3.626)	(1.507)	(3.370)	(1.409)	(0.475)	(0.607)
Treatment * Average Speed	-0.423					
	(0.121)					
Treatment * Respects Rules		-3.466**				
		(0.190)				
Treatment * Takes Care of the Truck			-1.604			
			(0.524)			
Treatment * Relation Score				-5.273**		
				(0.109)		
Treatment * Same Origin					-4.710	
					(1.352)	
Treatment * Propensity to be Treated						50.79**
						(0.926)
Round and Driver FE	✓	✓	✓	✓	✓	✓
Road FE and Trip Cont. in FS	✓	✓	✓	✓	✓	✓
Observations	146	162	162	162	162	173

Standard errors in parentheses.

* p<.1, ** p<0.05, *** p<0.01

Table 11: Heterogeneous Effect of Monitoring on Other Measures of Effort

	Manager Interviews	
	(1) Takes Care of the Truck Rate (0-10)	(2) Follows the Rules Rate (0-10)
Assignment to Treatment	-3.890***	-1.929***
	(0.850)	(0.697)
Assignment to * Propensity to be Treated	6.670***	3.043***
	(1.431)	(1.063)
Round FE	✓	✓
Driver FE	✓	✓
Observations	622	622

Standard errors are in parentheses and are clustered at the company-round level.
* p<.1, ** p<0.05, *** p<0.01

Table 12: Heterogeneous Effect of Monitoring on Manager-Driver Relations

	Manager Interviews	
	(1) Takes Care of the Truck Rate (0-10)	(2) Follows the Rules Rate (0-10)
Assignment to Treatment	-2.898***	-2.065***
	(0.738)	(0.695)
Assignment to * Propensity to be Treated	4.853***	3.468***
	(1.326)	(1.153)
Round FE	✓	✓
Driver FE	✓	✓
Observations	622	622

Standard errors are in parentheses and are clustered at the company-round level.
* p<.1, ** p<0.05, *** p<0.01

Table 13: Heterogeneous Effect of Monitoring on Manager-Driver Relations

	Manager Interviews			
	(1) Likes the Firm Rate (0-10)	(2) Likes Being a driver Rate (0-10)	(3) Is a Good Driver Rate (0-10)	(4) Relationship Rate (0-10)
Treatment	-3.445**	-1.202	-1.973	-0.0232
	(1.403)	(1.335)	(1.349)	(2.502)
Treatment * Propensity to be Treated	4.055**	2.226	4.370**	-0.907
	(2.018)	(2.066)	(2.089)	(3.692)
Round FE	✓	✓	✓	✓
Driver FE	✓	✓	✓	✓
Observations	249	259	259	245

Standard errors are in parentheses and are clustered at the company-round level.
* p<.1, ** p<0.05, *** p<0.01

Table 14: Comparison of Marginal Costs of Freight Transport Companies

Marginal Cost	US		Liberia	
	$ per km	%	$ per km	%
Fuel Costs	0.25	25	0.84	32
Truck purchase or lease	0.143	14	0.449	17
Truck Repair/Maintenance	0.097	10	0.345	13
Truck Insurance Premiums	0.057	6	N/A	0
Permits and Licenses	0.012	1	0.074	3
Tires	0.027	3	0.072	3
Tolls	0.012	1	N/A	0
Bribes	N/A	0	0.043	2
Driver Wages	0.31	31	0.68	26
Driver Benefits	0.081	8	0.1	4
TOTAL	**0.989**	100	**2.603**	100

km completed by a truck in a year	98,543	5,742

Data for the United States comes from the American Transportation Research Institute Survey of 2016. Data for Liberia comes from interviews of managers of trucking companies.

Appendix

A Balance Table

Table B.1 is the balance table for treatment and control groups after the first randomization.

The balance table for the first randomization shows that treatment and control groups are balanced. Average speed seems to be slightly different between treatment and control, though at a low level of significance (10%). The difference in average speed is of 3 kilometers per hour, which is low compared to the estimated effects of treatment.

B Additional Specifications

B.0.4 Effect by type of breaks

Comparing the data from the driver interviews and the manager interviews shows some information asymmetry. While tables 3 and 4 show similar estimates, tables B.2 and B.3 show that the managers under estimate the effect of the GPS trackers on the length of breaks. Without a GPS tracker, the total time of a trip is easy to estimate for a manager (because he knows when the driver leaves and when the driver arrives). But the time spent on breaks is a pure guess. In this section, we explore this further, by looking at the reason for breaks.

Evidence confirms a dis-alignment of incentives between the driver and the manager.

Table B.4 shows the effect of the GPS tracker on breaks by reasons for stopping. As before, columns (1), (2) and (3) show the IV estimate, and columns (4), (5) and (6) show the Intent to Treat estimates. Column (1) and (4) show the effect of treatment on "mud breaks" - when the truck has to stop because he is stuck in mud (or the truck in front of him is stuck in the mud). Column (2) and (5) show the effect of treatment "delivery breaks", and (3) and (6) on personal breaks.

Estimates of table B.4 show that most of the effect comes from the drivers taking less personal breaks. It also shows that part of the effect comes from being less stuck in the mud. This could seem surprising since there is no reason why the treatment would affect the time spent in the mud. However, drivers are probably adjusting their answers when they have a GPS tracker.

B.0.5 Propensity to be treated score

Table B.5 presents the logit regression of treatment take-up on baseline characteristics of drivers or of manager-driver pairs (such as matching county of origin). The fitted values of this regression generate the "propensity to be treated" score, used in the heterogeneity analysis.

B.0.6 Effect of treatment on productivity and relations

Columns (1) and (2) of table B.6 show the overall effect of the assignment to treatment on the propensity of the driver to follow the rules of the business and to take care of the truck (as measured by managers). Columns (3) and (4) of table B.6 show the overall effect of the assignment to treatment on the relationship between the driver and the manager, and on trust (as measured by the manager). All these coefficients are negative but small and not significant. This table shows that by not measuring the heterogeneity among workers (which is explored in the main text), the treatment appears to have no effect on these variables.

Table B.7 shows the effect of treatment on the driver's perception of his work in the firm. Except for the first column, results are small and not significant. The first column, however, shows that treatment has an overall negative impact (and significant at the 10% level) on the driver's perception of the firm.

Table B.1: Balance Table

	Diff(T-C)	Std. Err.	T-Stat	P-Val
Managers				
How many trucks does the firm own?	-2.10	1.93	-1.09	0.28
Is transport the main activity?	-0.04	0.07	-0.67	0.50
Is the firm officily registered?	0.08	0.06	1.23	0.22
Does the firm deliver goods outside the country?	-0.12	0.08	-1.52	0.13
Gender of the interviewee (Male=0; Female=1)	0.02	0.01	1.48	0.14
Is the manager Liberian?	0.05	0.03	1.55	0.12
Is the manager a driver?	-0.04	0.07	-0.64	0.52
Does your business hire carboys?	0.08	0.09	0.95	0.34
How many tires did your business purchased last month?	1.09	0.92	1.19	0.24
Did the firm ever pay for overloading?	0.03	0.08	0.42	0.67
Has this firm ever been a victim of theft?	0.12	0.09	1.35	0.18
Is the firm based in Monrovia?	0.05	0.07	0.77	0.44
Drivers always assigned to same truck	-0.03	0.04	-0.95	0.34
Trucks				
Truck was second-hand when bought	0.01	0.06	0.13	0.90
Maintenance in the past month (USD):	-18.59	194.67	-0.10	0.92
Price of truck when bought (1k USD)	-2.73	4.73	-0.58	0.57
Drivers				
Did you ever had an accident?	-0.04	0.09	-0.46	0.65
Driver wage (in USD per day):	-0.19	0.40	-0.46	0.64
Trips				
Distance of trip (km)	-6.29	23.03	-0.27	0.79
Time spent on trips (hours)	-6.28	4.78	-1.31	0.19
Average speed (km/h)	3.62	2.00	1.81	0.07
Time of trip spent on breaks (hours)	3.93	3.99	0.98	0.33
Percentage of trip spent on breaks (%)	2.40	2.90	0.83	0.41

This table is based on data collected during interviews of managers at baseline. Baseline interviews were collected between August and October 2016, which corresponds to the rainy season. In the case where several managers were interviewed in the same firm, the answers from the oldest manager were retained.

Table B.2: Effect of Treatment on Breaks (as a % of trip time) - MANAGER Interviews

	Instrumental Variable			Reduced Form		
	(1) Breaks (%)	(2) Breaks (%)	(3) Breaks (%)	(4) Breaks (%)	(5) Breaks (%)	(6) Breaks (%)
Treatment	-13.43*	-16.01**	-13.14*			
	(7.013)	(7.911)	(7.017)			
Assignemnt to Treatment				-5.140*	-8.642	-7.469
				(2.697)	(6.445)	(6.436)
Season FE		✓	✓		✓	✓
Driver FE		✓	✓		✓	✓
Road FE		✓	✓		✓	✓
Trip Controls			✓			✓
Observations	188	188	188	188	188	188

Standard errors are in parentheses and are robust.
* p<.1, ** p<0.05, *** p<0.01

Table B.3: Effect of Treatment on Breaks (as a % of trip time) - DRIVER Interviews

	Instrumental Variable			Reduced Form		
	(1) Breaks (%)	(2) Breaks (%)	(3) Breaks (%)	(4) Breaks (%)	(5) Breaks (%)	(6) Breaks (%)
Treatment	-6.049**	-19.88***	-21.36***			
	(2.621)	(5.347)	(5.653)			
Assignemnt to Treatment				-3.447**	-9.913*	-10.14*
				(1.447)	(5.284)	(5.441)
Season FE		✓	✓		✓	✓
Driver FE		✓	✓		✓	✓
Road FE		✓	✓		✓	✓
Trip Controls			✓			✓
Observations	291	291	291	291	291	291

Standard are errors in parentheses and robust.
* p<.1, ** p<0.05, *** p<0.01

Table B.4: Effect of Treatment on Breaks Reason (as a % of trip time) - DRIVER Interviews

	Instrumental Variable			Reduced Form		
	(1) Mud (%)	(2) Delivery (%)	(3) Personal (%)	(4) Mud (%)	(5) Delivery (%)	(6) Personal (%)
Treatment	-8.761*	2.651	-15.25***			
	(4.527)	(2.713)	(4.890)			
Assignemnt to Treatment				-4.158	1.258	-7.237
				(4.051)	(2.465)	(4.817)
Season FE	✓	✓	✓	✓	✓	✓
Driver FE	✓	✓	✓	✓	✓	✓
Road FE	✓	✓	✓	✓	✓	✓
Trip Controls	✓	✓	✓	✓	✓	✓
Observations	291	291	291	291	291	291

Standard errors are in parentheses and are robust.

* p<.1, ** p<0.05, *** p<0.01

Table B.5: Baseline Characteristics on treatment take-up

	(1) Logistic Regression of Treatment Take-up
treated	
Manager and driver have the same origin	-0.549
	(0.419)
Number of accidents past year	0.846**
	(0.422)
Relation with driver (0-10)	-0.0712
	(0.171)
Driver follows rules (0-10)	-0.176
	(0.148)
mean_avg_speed	0.00157
	(0.0234)
Other Baseline Characteristics	✓
Observations	143

Standard errors in parentheses.

* p<.1, ** p<0.05, *** p<0.01

Table B.6: Effect of Monitoring on Manager-Driver Relations

	Manager Interviews			
	(1) Takes Care of the Truck Rate (0-10)	(2) Follows the Rules Rate (0-10)	(3) Relationship Rate (0-10)	(4) Trust Rate (0-10)
Assignment to Treatment	-0.167	-0.0272	-0.0879	-0.0563
	(0.198)	(0.291)	(0.230)	(0.217)
Round FE	✓	✓	✓	✓
Driver FE	✓	✓	✓	✓
Observations	622	622	622	622

Standard errors are in parentheses and are clustered at the company-round level.
* p<.1, ** p<0.05, *** p<0.01

Table B.7: Effect of Monitoring on Manager-Driver Relations

	Manager Interviews			
	(1) Likes the Firm Rate (0-10)	(2) Likes Being a driver Rate (0-10)	(3) Is a Good Driver Rate (0-10)	(4) Relation Rate (0-10)
Treatment	-1.029*	0.123	0.629	-0.564
	(0.520)	(0.526)	(0.389)	(0.516)
Round FE	✓	✓	✓	✓
Driver FE	✓	✓	✓	✓
Observations	249	259	259	245

Standard errors are in parentheses and are clustered at the company-round level.
* p<.1, ** p<0.05, *** p<0.01