Asymmetric Information and Digital Technology Adoption: Evidence from Senegal*

Deivy Houeix†

[Click here for the most recent version]

July 14, 2025

Abstract

Digital technologies have the potential to increase firm productivity. However, they often increase data observability, which can be a double-edged sword. Observability reduces information frictions and can increase efficiency, but some agents may lose their informational rent and thus resist adoption. I explore this trade-off between observability and adoption through two field experiments, conducted over nearly two years, where I introduce digital payments to the Senegalese taxi industry in partnership with the country's largest payment company. In the first experiment, I randomize access to digital payments for drivers (employees) and transaction observability to taxi owners (employers). I find that digital payments reduce drivers' cash-related costs by about half but also serve as effective monitoring tools for taxi owners. Transaction observability increases driver effort, contract efficiency, and the duration of owner-driver relationships. However, 50% of drivers—the worst-performing and poorest—decline to adopt digital payments when transactions are observable. The second experiment shows that the adoption rate doubles when drivers are assured that owners will not be able to observe their transactions. A relational contract framework helps interpret these results, which I combine with the experimental variation to estimate the welfare effects of policy counterfactuals. I show that removing transaction observability would maintain moral hazard problems but broaden adoption and thus increase overall welfare—an approach ultimately implemented by the payment company. These findings highlight the crucial role of information embedded in digital technologies, as it magnifies gains for adopting firms but can deter initial adoption.

*Former Job Market Paper. I am deeply grateful to my PhD advisors, David Atkin, Esther Duflo, Ben Olken, and Tavneet Suri, for their unwavering guidance and support throughout this project. For very helpful discussions and suggestions, I would also like to thank Charles Angelucci, Abhijit Banerjee, Bob Gibbons, Frank Schilbach, Robert Townsend, and participants at the MIT behavioral, development, digital, labor, organizational, theory, and trade economics lunches, WEFIDEV 2023, and various external seminars for valuable suggestions and comments. I am also indebted to my amazing research team, composed of over 130 people throughout the two years of this project. To name a few: Koffi Adjassou, William Amedanou, Elisée Amewouame, Justin Chery, Martin Duflo, Cheikh Gaye, Ismael Keita, Loic Laurent, Sara Niang, Yacine Ouahioune, Francois Soleiman for their outstanding field assistance, as well as Gabriel Davis-Hollander, Alexis Yi, Lawrence Yong at MIT for their excellent research assistance. I am especially thankful to Drew Durbin, Joyce Keeley, Lincoln Quirk, and the numerous people involved at Wave Mobile Money, the implementing partner company, for trusting me to carry out this study independently. I gratefully acknowledge financial support from JPAL-DigiFI Africa, PEDL, The Weiss Fund, and Shultz and data collection support from the SDI. I received IRB approval from MIT, protocol #2012000286. This RCT was registered as AEARCTR-0009155.

†Harvard University. deivyhoueix@fas.harvard.edu

"Digital technologies are ushering in a new era in development. ... They have dramatically changed how we conduct business and our interaction with the environment. The international community has an unprecedented opportunity to help developing countries reap the benefits of digitalization while mitigating the risks and ensuring that ... we can close the digital divide."

– Axel van Trotsenburg, *World Bank Senior Managing Director*, 2024

1 Introduction

The rapid spread of digital technologies in lower-income countries has the potential to increase productivity by fundamentally reshaping how firms operate. Yet, these technologies often come bundled with data observability and this can be a double-edged sword. Observability reduces information frictions (Holmström, 1979), essential in settings with limited liability, weak contract enforcement, and low monitoring capacities. However, agents may fear losing informational rents, deterring adoption.

Digital payments are one such setting where these dynamics may operate. They enable employers to monitor employees' sales more easily than cash transactions, reducing moral hazard and improving efficiency (Kelley et al., 2024). But for employees, digital payments have mixed effects: while they can reduce cash-related costs, some employees may fear losing the ability to avoid work or misreport revenue, and thus resist adoption.

This trade-off between observability and adoption could contribute to the "digital divide" across businesses—a growing concern for development institutions, who view digital technologies as crucial drivers of firm growth (World Bank, 2023). Understanding how digital technologies reshape organizations—and how these changes may hinder adoption—is crucial to unlocking their potential for addressing core development challenges. Yet, evidence is scarce. In this paper, I address two central questions: How do digital technologies impact information frictions, contracts, and firm performance? To what extent might these impacts hinder technology adoption by agents in the first place?

I combine two field experiments in Senegal's informal taxi industry and structural estimates to show that digital technologies can increase productivity but inherently bundle observability, which may deter adoption by some workers. To conduct the experiments, I partnered with the largest mobile money and payment company in Senegal, Wave, and together we developed a novel digital payment solution for the taxi industry at the onset of digital payments expansion. Like most digital technologies, payment systems are not explicitly designed as monitoring tools. However, they generate data that can be leveraged for monitoring. The main goal of the payment technology is for taxi drivers to securely and easily receive digital payments from customers into their business mobile money wallets. Prior to this study, cash transactions were ubiquitous in Senegal. Thus, the shift to digital payments can directly lower cash-related transaction costs. By increasing observability and reducing moral hazard, these gains will be magnified among firms that use the technology, but at the cost of potentially discouraging initial adoption.

The taxi industry provides an ideal setting to examine the impact of digital technologies on within-firm relationships, as it exemplifies common principal-agent challenges faced by small firms in lower-income countries. In the Senegalese taxi industry, the typical arrangement involves a car owner (employer) and a single driver (employee) linked by a relational contract. The driver keeps any revenue exceeding a fixed rental fee, which is paid at the end of the week, and in certain arrangements, also receives a basic upfront payment from the owner. However, due to limited liability, the driver can default on the rent by claiming low earnings. Importantly, the owner has no way to observe whether this is due to bad luck, lack of effort from the driver, or if the driver misreports revenue. This creates scope for moral hazard in both driver's effort and output reporting and allows drivers to capture informational rents, contributing to inefficiencies. Default may lead to (costly) termination of the relationship as a way to mitigate moral hazard.

The distinctive feature of my experimental design is the ability to randomly vary and separate the "observability" effect of drivers' digital transactions to taxi owners, which allows me to quantify its impact on both contractual relationships *and* employees' adoption decisions. Guided by contract theory, I implemented two experiments: the "impact experiment" to identify the technology's effect on firm performance and contracts, and the "adoption experiment" to estimate the role of observability on technology adoption.

In the "impact experiment," I identified drivers willing to adopt and varied two dimensions: (i) access to digital payment technologies among 1,891 drivers—including drivers who are owner-operators to increase precision—and (ii) transaction observability for taxi owners (employers) among 613 owner-driver pairs. Taxi owners were randomly assigned to one of three observability levels regarding their drivers: *Granular Observability*, typically offered to all business owners by default, allowing the taxi owner to observe all digital transactions; *Coarse Observability*, where the owner could observe the driver's digital collection up to a cutoff, allowing the driver to signal low-output periods, but not the full transaction history and effort; and *No Observability*—to estimate the benefits of digital payments unrelated to observability. The goal of the observability treatments is to test a key prediction in contract theory: the observability—embedded in payments—should alleviate moral hazard frictions, increase agent's effort, and improve contract efficiency.

In the "adoption experiment," I measure how observability may also impede technology adoption by drivers. To adopt the technology, the payment company required taxi drivers to provide their employers' contact information. I followed up with 433 drivers who refused to share their employers' information, preventing adoption. My goal was to isolate the role of observability from other factors, such as concerns about owners' privacy, in explaining their reluctance.[1] I offered the technology to drivers again and randomly varied whether the taxi owner could observe their digital transactions. As opposed to the first experiment, drivers received this information *before* making the adoption decision. I then assessed the drivers' willingness to share employers' information, allowing them to adopt the technology.

[1]The company first approached drivers with the technology, as owners were hard to reach due to the lack of a formal registry. Drivers who refused to share employers' information were not allowed to adopt the technology.

I use three key sources of data to measure outcomes. First, I conducted five rounds of surveys with owners and drivers separately over nearly two years, achieving a 95% follow-up rate. The key innovation of these surveys is to track employer-employee informal contract data over time, enabling an analysis of how technology influences contract dynamics. This unique two-year survey panel dataset offers a rare opportunity to study contracts in developing economies, where the relational, verbal, and informal nature of contracts has historically hindered progress in the literature. Second, I measured drivers' effort by conducting mystery passenger audits across the city of Dakar. In this exercise, about twenty surveyors hailed 7,897 taxis and secretly recorded their license numbers, which I then matched to the experimental sample. This provides information on how often taxis are observed working on the road. Third, I obtained comprehensive payment and transfer data from the payment company, covering nearly all adults in Senegal, including all study participants. This data offers insights into daily transaction patterns at the driver level.

I have four main findings: (i) digital payments cut drivers' cash-related costs by half; (ii) transaction observability reshapes employer-employee contracts and increases drivers' effort and firm performance; (iii) observability deters low-ability workers from adopting the technology; and (iv) while the technology improves overall welfare by providing employers better information, it may also widen welfare inequality across and within businesses.

First, digital payment technologies significantly reduce cash-related costs, such as the time spent searching for small change or losing customers who prefer to pay digitally. These costs are significant for small and medium-sized enterprises like taxis, representing about 9% of profits at baseline, as self-reported by drivers.[2] Mystery passenger audits cross-validate these survey reports and show a 43% reduction in price distortions from small-change shortages, as drivers no longer round fares down when they use digital payments. Overall, this decrease in the cost of using cash is particularly striking as digitalization covers only a limited portion of drivers' revenue (about 2 customers a day, i.e., 13%). After 7–9 months, drivers express a high willingness to pay for the technology—equivalent to a week's worth of their profits—showing that the technology has substantial benefits unrelated to its observability feature.

Second, transaction observability reduces information frictions and reshapes contracts by mitigating moral hazard. Drivers under *Granular Observability* are seen on the road 34% more often in mystery passenger audits (p=0.003) and process 35% more digital transactions (p=0.015). Owners who observe digital transactions experience 31% fewer rent defaults (p=0.078). To compensate for drivers' increased effort, the owners change the contract structure: they are 18% more likely to provide an upfront monthly payment to their drivers (p=0.003)—that they refer to as a "salary"—in addition to the rent they collect. Overall, observability increases taxi owners' profits by 8%, though this estimate is not statistically significant.

Furthermore, transaction observability improves worker retention. Throughout the experi-

[2]The costs involved in cash payments do not entirely disappear because customer adoption of mobile money wallets remains incomplete—see Higgins (2024) for insights on the demand side of payment system adoption, and Alvarez and Argente (2022); Crouzet et al. (2023) for the dynamics between cash and digital payments. I explore the diffusion of digital payments across the Senegalese economy in a companion paper (Houeix, 2024).

ment, 34% of owner-driver pairs overall separated after nine months, and 62% after two years—a high turnover that constitutes a critical challenge for many industries in lower-income countries (McKenzie and Paffhausen, 2019). Both drivers and owners report losing an average of 33–34 days finding replacements. Transaction observability reduces turnover by 30% after nine months (p=0.034), especially among non-family pairs. The effect is concentrated in the *Granular Observability* treatment, with some suggestive impacts observed under *Coarse Observability* after two years. These findings indicate that moral hazard in effort and output reporting are significant constraints for owners, with effort-related issues particularly mitigated under *Granular Observability*. This effect shows how digital technologies can reduce information frictions and support business growth, something particularly relevant for industries dominated by family businesses precisely due to trust challenges between employers and employees (Bertrand and Schoar, 2006).

Third, observability is an important barrier to technology adoption for the worst-performing and poorest workers. Initially, 50% of drivers refused to adopt the technology, citing various privacy concerns for not sharing owners' contact information, a prerequisite for adoption. To characterize selection, I follow Karlan and Zinman (2009) and compare drivers in the "impact experiment"—who opt into potential observability but whose owners are randomly assigned not to receive it—to reluctant drivers in the "adoption experiment." Reluctant drivers are 83% more stressed at work, perform significantly worse, e.g., have fewer passengers (-11%) and work fewer hours (-4%), but are also significantly poorer and 28% less likely to have attended primary school. This shows that increased monitoring deters adoption by a potentially policy-relevant group, the poorest and least educated. In the "adoption experiment", I find that randomly assuring drivers that their digital transactions will not be observable by their employer nearly doubles adoption rates—with effects more than twice as large for the worst-performing and poorest—revealing the importance of observability in driver's adoption decision.

I develop a framework to formalize the mechanisms underlying the findings and guide the structural estimation. I begin by rationalizing the contract observed empirically between a taxi owner (the principal) and a driver (the agent), in the context of a simple framework incorporating agent's limited liability and both unobservable effort and output, which generate informational rent for the agent. I analyze both the impact and adoption of the technology in a relational contract where the principal cannot commit to contract terms once the agent adopts the technology.

The framework formalizes three key predictions regarding the effects of digital payments. First, the technology increases observability of the agent's effort through his digital transaction history, including timestamps, and thereby reduces moral hazard in effort. This observability leads to increased agent effort, fewer rent defaults, and contract changes—an upfront payment to compensate for the higher driver effort. The upfront nature of this payment helps address concerns about the principal reneging in this relational contract. Second, the technology provides the principal with information on the agent's total digital payments received. These two mechanisms—imperfect observability of effort and output—can both enhance trust and reduce termination, as the principal can now differentiate between bad luck, low effort, or misreporting of output, relaxing the incentive and truth-telling constraints. Third, when deciding whether to adopt digital payments, the agent anticipates this "observability effect" and weighs the technology's benefits (reduced cash-related costs) against potential unfavorable contract changes (increased effort and lower informational rent) given that the principal cannot commit to initial terms in a relational contract setting. Transaction observability will particularly deter technology adoption among low-ability drivers, for whom higher effort is more costly and unprofitable for the principal to compensate.

The framework highlights the key ingredients that generate these predictions: *limited liability* and *information asymmetries*, which provide informational rent to the agent and create incentive problems, and *weak contract enforcement*, which explains both the change in contractual form—an increased payment upfront rather than an ex-post rent reduction—and the principal's inability to commit to contract terms, hindering adoption for some agents. These three frictions are common in lower-income contexts.

The results show that digital payments have nuanced welfare effects: employee adopters remain with their employers longer but also exert more effort. To quantify drivers' disutilities of effort, I leverage the theoretical framework and apply the generalized method of moments (GMM), using the randomized observability treatment and the characteristics of non-adopters. The technology generates efficiency gains through reduced moral hazard, contract adjustments, and lower cash-related costs. Notably, 68% of these gains accrue to high-ability drivers, who benefit from reduced costs. Assigning equal weight to both owners and drivers,[3] I find that a counterfactual mandating adoption increases overall welfare, though less so than the status quo where only high-ability drivers adopt. This stems from a 12% decrease in the welfare of low-ability drivers, who are induced to exert higher effort by owners. In contrast, redesigning the technology to remove transaction observability for employers, though preserving the underlying information frictions, results in Pareto improvements since all workers can then adopt the technology.

Taken together, these findings highlight that digital technologies can support business growth and mitigate information frictions, but their impact on contracts may hinder adoption. These results thus emphasize the importance of careful technology design. Given the finding that observability is a barrier to adoption for many workers, a real-world policy impact of this study was the partner company's decision to make non-observable transactions the default to increase technology adoption in the taxi industry, though this decision may maintain contract inefficiencies.

To my knowledge, this is the first paper to randomize features of a digital technology across firms to capture both the resulting changes in firm practices and performance and the barriers to adoption. This approach delivers three key contributions to the fields of development and organizational economics.

First, I isolate a novel mechanism in the emerging literature investigating the impact of digital technologies on firms. While studies have focused on their role in reducing transaction costs

[3] Welfare calculations consider only the employer and employee, excluding the consumer's welfare from paying digitally. Also, I verify a "no-deviation" condition on δ under which the owner has no incentive to terminate a low-type agent and incur replacement costs in the hope of finding a high-type willing to adopt the technology.

for households (Aker et al., 2016; Jack and Suri, 2014, 2016), I quantify how digital payments can reduce these costs for firms, showing the importance of cash-related transaction costs (Beaman et al., 2014). Importantly, unlike the recent literature that emphasizes channels such as customer acquisition or credit access (Agarwal et al., 2019; Dalton et al., 2023; Higgins, 2024; Riley, 2024), I reveal the "observability effect" embedded in digital payments and its impact on contracts. A separate literature examines "observability" more directly, by testing the canonical prediction that monitoring technologies can mitigate moral hazard, as shown in seminal work on the U.S. trucking industry (Baker and Hubbard, 2003, 2004; Hubbard, 2000, 2001, 2003) or more recent experimental work in developing contexts on GPS tracking (de Rochambeau, 2021; Kelley et al., 2024). My paper combines these two strands to make a distinct point: despite not being designed for monitoring, many digital technologies like payment systems embed observability as a byproduct, which improves firm performance but can *also* hinder technology adoption.

Second, I contribute to the literature on technology adoption by uncovering a new mechanism behind firms' uneven uptake of digital technologies (World Bank, 2023). Academic research on the role of transparency in digital technology adoption has remained largely descriptive, highlighting issues with patient data sharing in healthcare (Goldfarb and Tucker, 2012; Derksen et al., 2024) or firms' reluctance to digitize payments due to tax compliance (Brockmeyer and Somarriba, 2024). Outside of digital technologies, Atkin et al. (2017b) studies how misaligned incentives hinder adoption of a soccer-ball technology. Here, I am able to analyze the trade-off between observability and technology adoption by running two experiments within the same industry and technology, thus directly linking technology impact to adoption barriers. Furthermore, I contribute to the expanding literature on technology adoption among small urban firms, which represent an increasing share of employment in Africa and exhibit different behaviors from rural farmers.[4]

Third, I contribute to the growing empirical literature on the distinct features that shape organizations in lower-income contexts (Macchiavello and Morjaria, 2015, 2021). By bringing theory to data and simplifying equilibrium assumptions, I estimate relationship values, thus expanding the limited research on contract valuation. In doing so, I follow Macchiavello and Morjaria (2023) that advocates for empirically testing theory to deepen our understanding of within-firm contracts. I achieve this by running field experiments with firms (Bandiera et al., 2005, 2007, 2009; de Mel et al., 2008; Atkin et al., 2017a; Burchardi et al., 2018; Boudreau, 2024) to estimate key theoretical parameters and assess the welfare and distributional impacts of observability, contributing to the recent studies that apply structural methods in development economics (e.g., Bergquist and Dinerstein (2020); Bai (2024); Kelley et al. (2024)). The experimental variation and the theory offer insights into important policy counterfactuals, such as mandating adoption or redesigning digital technologies. This research diverges from U.S. studies (Lazear, 2000; Blader et al., 2019) by showing how weak enforcement and limited liability, in settings with information frictions, shape employer-employee relationships and can hinder the adoption of profitable technologies.

The rest of this paper proceeds as follows. Section 2 reviews the setting and key stylized facts.

[4]See Suri and Udry (2024) for an extensive review of research on technology adoption among rural farmers.

Section 3 outlines the experimental design. Section 4 discusses the data collection process. Section 5 details the experimental results and mechanisms. Section 6 presents the theoretical framework. Section 7 explores counterfactuals based on structural estimations. Section 8 concludes.

2 Study Background

2.1 The Digital Payment Technology

This research leverages a new digital payment technology that I developed in partnership with Wave Mobile Money, the largest mobile money company in Senegal and Francophone Africa's first unicorn. Wave operates in six countries and is one of only seven unicorns across Africa. It serves about 80% of the adult population (7.2 million active users) who use it for peer-to-peer (P2P) transfers at the time of the study (2022).

Working with them, I helped design a new peer-to-business (P2B) technology to digitize payments in the Senegalese taxi industry.[5] I contributed to each stage of the technology's development, from conducting market research to adapting Wave's existing core digital business payment app for the specific needs of the taxi industry. Working closely with software engineers, we refined the mobile application and tested the QR code system for taxis—a printed QR code hung from the rear-view mirror—that enables passengers to securely pay via mobile money, with drivers paying a 1% transaction fee (see Figure 1).[6] This system introduces two key differences from peer-to-peer (P2P) transfers: (i) *irrevocability*—passengers cannot reverse the transactions after the ride, which is possible with a simple phone call for personal mobile money transfers—and (ii) *convenience*—payment is made by scanning a QR code instead of entering the driver's phone number into a mobile money account.

Collaborating with a company that leads the payment market in the country was crucial, as the widespread adoption of mobile money among customers created strong incentives for businesses to adopt digital payments, and therefore an opportunity to focus on the business-side frictions discussed in this paper. At the start of the study, the mobile money company had just begun digitizing payments for about 10,000 non-taxi businesses and reached nearly 200,000 businesses in the first year of the study. The technology progressively gained popularity, with over two million unique Senegalese users making business payments.[7] Despite this growth, digitalization is only partial, as cash still remains the predominant mode of transaction.

The technology was offered to both taxi owners and taxi drivers. When provided with the technology, both groups received training on how to use the product, and the corporate mobile application was installed on their smartphones. Taxi owners who did not drive but hired a driver were briefed on the technology at the end of a baseline survey.

[5] Electronic merchant payment solutions have been implemented in other regions of Africa before; a notable example is *Lipa na M-PESA*, launched by Safaricom in Kenya in 2013, which also aims to digitalize business payments.

[6] For the first 50,000 FCFA (USD 85), we waived the fee for all drivers in this experiment.

[7] For interested readers, I leverage the experimental variation in digital payment access to analyze its diffusion across businesses in the entire Senegalese economy in a companion paper (Houeix, 2024).

The base technology, used by businesses outside the taxi industry, embeds an observability feature by default, which provides employers with detailed information about employees' digital transactions, including timestamps and amount collected—see Figure 1(c). This observability feature is inherent in most digital technologies, which generate and store data available to stakeholders. To isolate its impact and the barriers to adoption, I use this private sector collaboration to randomize both access and the observability feature, as described later in Section 3.2.

2.2 Stylized Facts

2.2.1 Taxi Industry in Senegal

This project focuses on private taxi businesses in Dakar, Senegal. Taxis are ubiquitous in Dakar, employing about 4-6% of the adult urban male population, with about 21,000 active taxis in 2019.[8] The sector is largely informal and includes two main stakeholders: the taxi business owner (principal) and the taxi driver (agent). Both are typically middle-aged (30-50 years old) urban male workers.

The taxi industry provides a compelling setting to study the implications of digital payments. First, it has potential for productivity gains, especially given the substantial costs of using cash. Second, the sector exhibits significant information asymmetries, with firm owners often unable to observe both the effort exerted and the revenue collected by their drivers. Digital payment technologies, if adopted, promise to alleviate both of these challenges.

Stakeholders in the Taxi Industry: Owners and Drivers. In the baseline survey, described in more detail in Section 4.1, I identify three main types of taxi owners, distributed almost equally: sole proprietors who drive their own taxis without employees; owners who do not drive their taxis but employ drivers; and owners who alternate between driving and renting out their taxis. Most owners possess one taxi (92%) and employ one driver, and the industry often operates as a family business, with about 51% of taxi owners and drivers belonging to the same family. Most taxi owners and drivers have not completed primary education (68%). About half of the drivers were unable to save any money over the past three months. Most are the primary earners in their households and can be classified as urban poor (see Tables B1 and B2).[9]

Prevalence of Cash Payments at Baseline. Passengers typically negotiate fares at the start of a ride and pay primarily in cash. While personal mobile money payments as individuals (peer-to-peer P2P) are an option, they are infrequently used due to revocability concerns. Passengers can easily reverse P2P transactions after the ride, posing a significant risk to drivers, which is an issue

[8] This figure is from the CETUD and is especially high for a city of 3 million people, with 53% under 19 years old. Since the CETUD data include only formally licensed taxis, this employment share is likely a lower bound. I also excluded other transportation-related industries like minibuses. For comparison, New York City, with 8 million residents, has 13,000 licensed taxis.

[9] I use the Poverty Probability Index (PPI) specific to Senegal to measure wealth. The average PPI score of 63 indicates a 50% likelihood of living below 200% of the 2011 National Poverty Line.

the technology solves. Consequently, drivers received an average of only six P2P transactions resembling taxi payments, based on their value, in the three months prior to the study, representing an insignificant share of their total transactions.

Large Costs of Using Cash at Baseline. Table B3 outlines the substantial costs of using cash across four categories I identified through background research and interviews with taxi drivers. (i) *Time Lost*: 86% of drivers report losing time searching for small change, spending at least 10 minutes about 1.52 times weekly, with 6% of drivers experience this daily. (ii) *Refused Customers*: Drivers often refuse passengers with no small change because P2P digital transfers can be easily reversed, leading to missed transactions. About 60% of drivers reported this issue, turning down 3.60 customers weekly, which is 3.79% of their passengers.[10] (iii) *Reduced Price*: Due to small change shortages, drivers often price rides at the nearest lower bill (CFA 500). Digital payments eliminate this issue, allowing precise pricing (e.g., CFA 700), smoothing the demand curve and increasing efficiency. 92% of drivers reported having to reduce prices due to that cash problem at least once weekly. (vi) *Mistake Giving Change*: Miscalculations often lead drivers to give incorrect change. About 41% of drivers reported making such mistakes weekly. In addition, drivers report anxiety about theft related to carrying cash, and electronic theft—when customers pay digitally through personal transfers and then leave the taxi—is a major concern. This is one of the main reasons why most drivers do not use digital personal transfers at baseline but would prefer irreversible business payments.

The weekly monetary loss is about USD 6.66, which represents about 9% of their effective profit. This imputed loss is calculated by assigning a monetary value to each reported cost, based on fieldwork with a subset of drivers prior to the experiment. For the past 7 days, each time lost and mistake giving change is estimated at CFA 500 (USD 0.8), refused customers at CFA 1,500 (USD 2.5), and reduced price at CFA 800 (USD 1.3). Note that this loss does not even include some additional costs of using cash, such as theft, anxiety about theft, and difficult record-keeping, that are more challenging to quantify in monetary terms.

Absence of Monitoring at Baseline. *GPS tracking* is virtually nonexistent in the taxi sector, and owner awareness of such tools is minimal. Important barriers to adoption are the high acquisition and maintenance costs—GPS trackers cost owners about a month of profit along with recurring fees—as well as strong resistance from drivers, who find them too disruptive. Moreover, most taxis are old (about 20 years on average, some up to 30-40 years), poorly maintained, and malfunctioning *odometers* in the majority of taxis hinder the monitoring of miles driven. At the start of this study in 2022, *ride-hailing* technology was largely absent with very few drivers using such technologies ($< 4\%$): digital payments are typically the first digital technology firms adopt and often serve as a stepping stone to the expansion of other digital technologies (World Bank, 2024).

[10] Quantifying lost customers requires knowing the time it takes for drivers to get another request after refusing one. Estimates show drivers spend about half their working time without a customer, indicating that refusing a customer is a substantial cost.

2.2.2 Taxi Owner-Driver Contractual Relationships

In this section, I present four facts about the taxi industry that motivate both the experimental design and the theoretical framework that follow.

Fact 1 - Owner-Driver Contract Structure: Rental Contracts with Possible Upfront Payment. Taxi owners use rental contracts and in some cases provide upfront payments to drivers. The rental contract is constituted of an agreed-upon target transfer to be paid to the owner at the end of the week, on average about CFA 60,000 (see Table B4). This informal contract allows for partial default on the agreed transfer on low-revenue weeks, with possible contract termination if the driver defaults. This structure is akin to other contracts in lower-income contexts, such as those in the Kenyan minibus industry described in Kelley et al. (2024). However, it differs substantially from U.S. taxi contracts (Angrist et al., 2021), as each of the four stylized facts underscores key differences in the economic environments that distinguish the two contexts.

In addition to the fares kept by drivers, around 53% of owners pay their drivers an upfront payment—that taxi owners and drivers referred to as a "salary"—, ranging from CFA 40,000 to 50,000 ($65 to $85) each month. In 90% of cases, this payment is made upfront and does not depend on whether the driver pays the full rent. Conditional on getting one, it represents about 18% of a driver's total compensation at baseline, besides driver surplus, i.e., revenue collected minus the rent (about CFA 60,000 per week) and the costs (e.g., fuel, food consumption, police bribes, minimal maintenance costs).[11] The primary motivation for this payment, cited by 84% of owners, is that it commits the owner to a minimum payment to their driver, even when no fares are collected. Other reasons include preventing poaching and reducing risk-taking behavior, though the latter is not the focus of this paper. Finally, owners cover major taxi maintenance costs.

Fact 2 - Limited Liability and High Prevalence of Default on Rental Transfers to Taxi Owners. Each period, the driver reports the revenue collected to their employer and pays the full rent if the revenue collected is enough and partially defaults otherwise. The main reasons cited by drivers for defaulting are demand-related factors beyond their control, such as accidents, traffic congestion, or a lack of customers. Limited liability, a typical constraint in lower-income contexts, prevents drivers from paying the rent in advance and is consistent with the lack of credit access in this setting. Table B4 shows that only 8% of drivers had access to any form of loan—whether formal or informal—in the three months prior to the baseline survey, despite a high demand for credit. Less than half were able to save any money in the past three months. In line with this, 70% of drivers experienced at least one instance of default in the preceding three months and 48% defaulted at least once a month. However, owners cannot easily verify their driver's effort nor reports of low-output. For these two reasons—moral hazard in effort and in output reporting—defaults have been identified as a frequent source of disputes between owners and drivers for 65% of owners and are an important source of stress for 48% of drivers.

[11] The precise method of calculating profit is detailed further in Section 5.2.

Fact 3 - Large Information Asymmetries between Owners and Drivers in the Taxi Industry. Taxi owners often have incomplete and inaccurate information about their drivers' work habits. In Table B5, I compare driver and owner's separate survey reports within the same pairs, and show that owners are more likely to *underestimate* rather than overestimate key performance indicators of their driver, such as hours worked, days worked, revenue collected, and number of passengers. Only 39% of owners accurately know their drivers' working hours (+/- 2 hours), and around 26% accurately know daily earnings or passenger count. Furthermore, only 46% of owners correctly estimate the number of days their drivers work each week, while 33% underestimate it. Additionally, 68% of drivers park the taxi away from their taxi owner's home, preventing owners from verifying work days.

One key explanation for these information asymmetries is that drivers' daily output is not only unobservable but also is highly variable. Table B6 estimates a model to predict daily output using the specific days of the week and month, calendar date fixed effects (FEs), and the driver's hours worked. The analysis shows that these data poorly predict daily output, with an adjusted R^2 of 22%, and including driver FEs only doubles the R^2 to 45%, all of which indicate the significant role of demand variation on revenue.

The biased beliefs and limited information are crucial to understand the potential impact of transaction observability embedded in digital payments. Given owners' limited and often inaccurate baseline information, digital transaction observability could greatly improve owners' knowledge of drivers' effort, including days and hours worked, and output levels.

Fact 4 - High Worker Turnover. The taxi industry is characterized by high turnover rates.[12] At the start of the study, the median duration of owner-driver relationships was approximately 1.5 years. On average, taxi drivers had worked for three owners at the start of the study. Drivers' limited liability and private information significantly contribute to conflicts between owners and drivers, sometimes leading to the termination of the relationship. Specifically, 65% of owners cite issues like drivers defaulting on rental payments as among primary causes for separations.

Over the 9-month study period, 34% of pairs separated overall, and 62% after two years. I manually coded open-ended responses to reveal the reasons for these separations. Even though both owners and drivers were often reluctant to talk about the actual reasons for separations, Table B7 shows that the predominant reasons stated for separations include: "drivers decided to leave" (29%), "drivers were explicitly fired" (21%), and "owners decided to sell the taxi" (20%). I recorded the exact wording used by owners and drivers about their separations. For example, one owner stated, *"My driver did not respect his promises: he told me that 2 days after the Tabaski (Eid-Al-Adha) he will come back to work, but I suspect he did not, so I fired him."* The high turnover rate negatively affects both drivers and owners, who report spending about 33–34 days finding replacements.

[12]The high turnover rate in this industry is comparable to exit rate and employee turnover documented in other industries in lower-income countries: Fajnzylber et al. (2006); Nagler and Naudé (2017) discuss factors explaining small firm death in lower-income economies. McKenzie and Paffhausen (2019) reviewed sixteen panel surveys from twelve developing countries, estimating an average small firm death rate of 8.2% per year. However, few studies document employee turnover within medium and small firms, likely due to the lack of within-firm data in informal contexts.

3 Experimental Design

3.1 Experiments Guided by Contract Theory

I designed two experiments within the taxi industry to test the paper's core idea: digital technologies can increase productivity but inherently bundle observability, which may prevent their adoption. Specifically, I test two key components within the context of a principal-agent framework: (1) Increased observability—embedded in digital payments—can reduce *moral hazard in effort* and *in output reporting*, increase agent's effort, and improve contract efficiency, in line with the canonical prediction in contract theory and (2) due to the nature of *relational contracts*, the principal cannot credibly commit to the initial contract terms, which may hinder technology adoption by some agents.

I conducted the following two experiments: (1) the "impact experiment" that causally tests how digital payments affect firm performance, focusing on (i) drivers' cash-related costs and (ii) the observability effect on employer-employee relationships, including contract structure, trust, and relationship duration, and (2) the "adoption experiment" that tests how observability influences technology adoption by some workers in this relational contract setting. The first experiment targets drivers willing to adopt the technology by sharing their employers' information, while the second focuses on those who refused, thereby preventing adoption. The complete experimental design is summarized in Figure 2 and detailed below. The theoretical framework in Section 6 formalizes the mechanisms behind the treatments and guides the structural estimation.

3.2 Impact Experiment: Digital Technology Access and Observability

Drivers were listed and invited to participate in the study to adopt a newly developed digital payment technology. After drivers provided their owner's contact information, I randomized access to the technology at the *taxi business owner* level, with the sample primarily consisting of owners with only one taxi. In this setting, while technology adoption decisions are often initiated by drivers, they require approval from the taxi owners. Typically, customers decide whether to pay for the ride digitally or in cash.

Technology Access. I invited interested drivers to come to three different locations in Dakar along with their taxi owners to be surveyed separately (see Section 4.1). At the end of the baseline surveys, drivers were randomly assigned to receive the digital payment technology and were given specific training from the payment company on how to use the mobile app and the QR hanging card. Note that sole proprietors who drive their own taxis without employees were also randomized access to the technology primarily to improve the precision of estimates related to the technology's impact on cash-related costs.[13]

[13] As discussed later in Section 5, these owner-operators were also randomized into one of the observability treatments, which would only apply if they hired employees in the future. They are excluded from the analysis of observability effects, except when investigating how the technology's observability affects their hiring decisions.

The Observability Treatments. The base digital payment technology, available to a range of businesses outside the taxi industry, includes an observability feature that allows employers to monitor their employees' digital transactions. By partnering with the payment company, I am able to tweak the design of the digital payment technology to isolate the role of observability from all the other effects the technology might have on firm performance.

To conduct the research, I designed three observability options within the owner-driver pairs who got access to the technology. I vary the amount of information available to taxi owners regarding their driver's digital transactions. Each option targets a specific type of information asymmetry—moral hazard in effort and output reporting—which I formally model in Section 6. In particular, the three following observability levels are randomized:

1. *No Observability (N-O)*—20% of taxi businesses. Owners do not observe their drivers' transactions while drivers are provided with the technology.

2. *Granular Observability (G-O)*—20%. The owners get access to a mobile application detailing the agent's complete transaction history, including timestamps (see an illustration in Figure 1(c))—the default option typically offered to businesses. These metrics include customer count, value and time of each transaction, and the total digital revenue of their driver. Empirically, Figures A3 and A4 illustrate the data available to taxi owners with this treatment. Owners also receive daily SMS updates at midnight showing the driver's total digital collections. This treatment provides owners with a far more comprehensive view of the driver's work, in contrast to the previous complete lack of observability. This data can mitigate information frictions and allows monitoring by providing the principal with a signal for driver's *effort*—through the hours and the days worked indicated in timestamps—and *output*—through the digital revenue collected. Since digital transactions represent only a portion of total transactions,[14] this information may be particularly useful when drivers exert high effort, allowing owners to observe substantial transactions at various points throughout the workday, enabling drivers to credibly signal effort to owners.

3. *Coarse Observability (C-O)*—20%. This treatment arm aims to provide owners with a signal of low-output periods. To achieve this, owners receive the daily SMS updates on digital collections up to a pre-set threshold of CFA 5,000 (USD 8), the average daily *digital* collection from pilot data. Surplus amounts above CFA 5,000 are not disclosed to the owners.[15] It primarily serves as a signal that output was low and holds policy relevance by reducing misreporting concerns, benefiting both owners and drivers. Additionally, it is designed to test for potential ratchet effect, where drivers could strategically minimize their digital transactions to avoid revealing high output. Fewer transactions under *Granular Observability* compared to *Coarse Observability* could indicate an attempt to hide high-output periods from owners.

[14] Owners were explicitly reminded that digital transactions would represent a limited share of their driver's total revenue, given the use of cash.

[15] The CFA 5,000 threshold is less than the average daily collection of CFA 30,000, representing about 17% of the total, reflecting the limited share of digital transactions in the driver's total earnings.

4. *Pure Control*—40%. No technology was provided to drivers in the control group in the first nine months of the experiment. They were placed on a waitlist.[16]

The three distinct levels of digital observability measure the "observability effect" on contracting in distinct ways. The difference between *Granular Observability* and *Coarse Observability* aims to isolate the effect of increased information on effort (reducing moral hazard in effort). In contrast, the comparison between *Granular Observability* and *No Observability* aims to capture the combined effect of the effort and low-output signals. *Coarse Observability* mainly reflects days with low total digital payments, but provides limited information into effort, as it does not reveal customer count, transaction values, or work indicators like hours worked.

3.3 Adoption Experiment: Digital Technology Adoption

The majority of drivers approached during the listing expressed high interest in adopting the technology (about 89%), but about half of them refused to provide their taxi owner's contact information, which was a necessary step for adoption. I followed up with drivers up to three times and included them in the impact experiment if they eventually agreed to provide their employers' contact information.

For the reluctant drivers, I conducted the "adoption experiment" to disentangle the factors behind their refusal and quantify the role of digital observability. In particular, (i) I re-offered the technology to a sub-sample of drivers who continued to refuse to share their employers' information and (ii) I randomized owner's observability *before* drivers decide to adopt (see the bottom half of Figure 2). The adoption experiment examines how observability affects drivers' willingness to share their employers' information. The goal is to determine if drivers change their decisions when they no longer fear their employers seeing their digital transactions. These drivers, along with others, were followed through mid- and long-term surveys to measure their performance and contract changes, and they could access the technology after the adoption experiment.

3.4 Randomization Procedure

Randomization for the impact experiment was conducted at the owner level using a computer program, across twelve batches. The STATA command `randtreat` was used to stratify the randomization to ensure balance. Batches were designed to minimize the time between listing and baseline surveys, reducing respondent attrition.

The randomization was stratified on three key dimensions of heterogeneity: proxy for *baseline digital usage*—whether drivers made more than the median number of six personal transactions in the taxi price range over the past three months; proxy for *baseline relationship*—relationship length

[16]The control group received the technology at the end of the mid-term survey, with a randomly assigned option of *Granular Observability* or *No Observability*. The observability group was further split into two: drivers *or* owners were *nudged* about potential contractual changes due to observability. This follow-up experiment aimed to measure what triggers contractual changes between owners and drivers. The nudge content was based on this paper's results and will be detailed in a companion paper. Therefore, the control group is excluded from the longer-term analysis.

(whether the relationship exceeded two years, the median), business type (whether the owner also drives the taxi), and the number of taxis (whether the owner has one or multiple taxis); proxy for *baseline risk aversion*—whether the taxi drivers operate throughout the city or wait for passengers at a fixed location.

The stratification serves several purposes. First, it ensures that treatment assignment is balanced across key variables. Tables B1 and B2 test the randomization balance across the two samples: the set of owner-driver pairs and the set of taxi drivers, including owners driving alone used for the effect of cash-related costs. Second, it facilitates the heterogeneity analysis of observability effects on relevant subgroups, such as recent relationships.

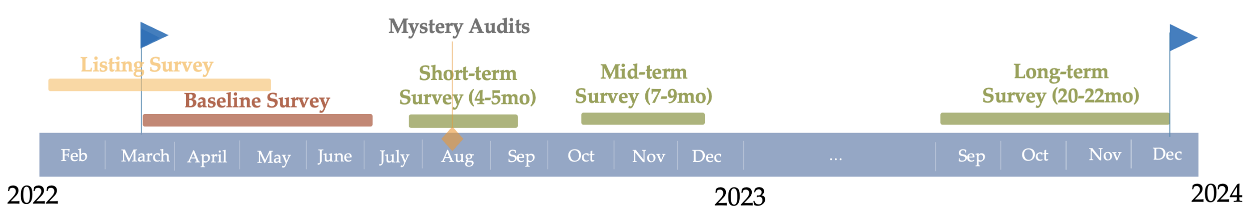
For the adoption experiment, randomization of the observability treatment arm was conducted among drivers still interested in the technology at follow-up. This was not stratified and I include multiple specifications and controls in the following analysis to ensure robustness.

4 Data: Measuring Informal Contracts and Worker Behavior

This section details the three primary sources of data used in this study: survey responses, mystery audits, and payment data.

4.1 Survey Data Collections

I collected five rounds of survey data: listing, baseline, short-term (after 4–5 months), mid-term (7–9 months), and long-term (20–22 months)—see the timeline below.



Listing Survey. The experimental sample includes all drivers and owners recruited through a listing survey from March to May 2022 (see the map in Figure A1). Drivers were recruited at garages, car wash stations, meeting points, and on the streets of Dakar. Owners not driving their taxi were primarily recruited by obtaining their contact information from their drivers during the listing survey, as required for adopting the technology. In total, about 3,600 eligible owners and drivers were listed, including the employed drivers who refused to give owners' contacts.[17]

[17]Non-eligible taxis included drivers who did not have a smartphone during the listing (approximately 15%) since the technology requires drivers to have an Android smartphone to utilize the business app. Other ineligible categories include drivers who were unreachable after several attempts to contact them (approximately 10%), and owners with more than four taxis (approximately 2%).

Baseline Survey (Willing to Adopt). I contacted 3,026 owners and drivers, including 881 pairs in twelve batches to be surveyed in three locations in Dakar. The baseline survey took place from the end of March to June 2022. Both owners and drivers were asked to come to the survey locations to separately answer a 45-minute survey. The baseline survey included both taxi owners driving alone with no employee and owners and drivers in pairs. The former category is exclusively used for analyzing the impact of digital payments on drivers' hassle costs of using cash, while the pairs are used to study the impact of digital observability on owner-driver relationships. The attrition rate from invitation to survey participation in the experiment was about 25% among all taxi stakeholders (owners and drivers), and 30% among pairs (since both owners and drivers had to be surveyed to be included in the study).[18] Respondent treatment arms were revealed at the end of the baseline survey to rule out concerns of differential attrition across groups at the baseline stage. The top rows of Tables B1 and B2 show that attrition rates are not statistically different across the treatment arms. Table 1 displays the baseline sample size composed of 613 taxi pairs and 2,269 taxi owners and drivers, including 1,891 drivers (the rest being owners not driving their taxis).

Adoption Survey (Reluctant Drivers). Drivers who refused to provide owners' contact information pre-randomization were unable to adopt the technology, as providing this information was a condition for adoption. To study this differential adoption, I followed up with 433 of these drivers over the phone by the end of the baseline survey, from June to early July 2022, and re-offered them the technology to quantify the role of observability in their adoption decision.

Short-Term, Mid-Term, Long-Term Surveys. All owners and drivers in the study were surveyed three times post-intervention, approximately five, nine, and twenty months later. The first follow-up survey occurred from July to September 2022, the second from October to December 2022, and the third from September to December 2023. Attrition rates were low (see Table 1). The survey team was able to follow up, after almost two years, with 82.2% of all drivers, 95.1% of pairs, and 84.8% of non-adopters. Since many outcomes focus on owner-driver contracts, I can still leverage survey responses from just one party to uncover key information about the business. Table B8 shows no differential attrition rates across observability treatments. The three rounds of follow-up surveys were conducted to constitute an exhaustive panel data of firms, including changes in business activity, worker behavior, retention, and contracts over time.

To ensure survey data accuracy and reliability, interviews were carried out by trained enumerators experienced in surveying businesses. After each survey, a team of back-checkers reviewed approximately half of the collected data by asking respondents key questions and cross-checked responses. I implemented several quality checks, such as evaluating the completion time for each module, and conducting targeted reviews of entire surveys. Follow-up phone surveys adhered to a detailed callback procedure, including night and weekend surveyor shifts to accommodate

[18] Some drivers came to be surveyed—about one-quarter of this attrition—but were not eligible to receive the product because they did not have an Android smartphone, a driver's license, or an ID card, which were required for adoption. Also, 130 owners, mostly not driving their taxi, did not want to come to the designated locations, so a short phone survey was administered to them, and their treatment was revealed at the end.

drivers, use of WhatsApp message reminders, and reaching out to respondent's friends to recover updated phone numbers. While discrepancies between owner and driver reports were limited, any inconsistencies prompted a third review by a senior field coordinator to determine its cause.

4.2 Mystery Passenger Audits

In August 2022, I conducted mystery audits to measure (i) drivers' behavior related to digital payments and pricing, and (ii) drivers' effort based on their presence on the road. I trained twenty mystery passengers to hail taxis throughout Dakar, Senegal, following a strict procedure to mimic typical price bargaining. Over two weeks, they systematically rotated across six high-traffic locations each day, capturing a broad sample of taxis and driver behavior over a meaningful time-frame. Surveyors asked questions and secretly recorded taxi license numbers. They pretended they had to leave after a pre-set bargaining process—primarily to increase the sample size and reduce field costs (see the map in Figure A2). The activity was repeated a sufficient number of times to match taxi drivers with their license numbers in the experimental sample. Specifically, mystery passengers adhered to the following steps: (1) Memorize the randomized destination and pre-specified price on their data collection application, (2) Stop a taxi, (3) Ask the driver's initial price, (4) Suggest the pre-specified low price, (5) Listen to the driver's counteroffer and ask their last price, (6) Suggest a non-rounded price, (7) Ask to use digital payments. Detailed data was recorded on a tablet once the taxi left about each step of the process. The bargaining process averaged just 1-2 minutes to minimize any negative impact on the driver's activities.

Over the two weeks of audits, mystery passengers audited 7,897 taxis, which allowed recovery of information from 503 taxis from the study (including approximately 41% of the taxis in pairs).[19]

4.3 Mobile Money and Payment Data

Administrative mobile money data was made available by the mobile money partner and contains the universe of transaction-level data for most of the adult population in Senegal, including all study participants. This data is key to measure driver-level digital payment usage across treatment arms. The data includes pre-, during, and post-study periods, with pre- being only personal transactions and during/post including both business payments and personal transactions.

5 Experimental Results

5.1 Empirical Specifications

Main Specifications This section describes the estimation strategy, following the main specifications outlined in the pre-analysis plan (PAP) under AEA registry ID #0009155. The "impact

[19] About 30% of the taxis recovered were audited more than once. I kept all the times they were encountered in the following analysis, cluster standard errors at the business level, and used the number of times seen on the road for the following section to measure driver's effort.

experiment" employs two regression specifications: (1) business-level linear regressions to measure the impact of digital observability on owner-driver relationships, and (2) driver-level linear regressions to assess the impact of digital payments on drivers' cash-related costs, unrelated to observability.

Business-level regressions

$$y_j = \beta_0 + \beta_1 T_j^{GranularObs} + \beta_2 T_j^{CoarseObs} + \beta_3 T_j^{NoObs} + \alpha_s + \epsilon_j \tag{1}$$

Taxi driver-level regressions

$$y_{ij} = \beta_0 + \beta T_{ij}^{Access} + \alpha_s + \epsilon_{ij} \tag{2}$$

where i indexes drivers, and j indexes businesses (taxi owners). y_j and y_{ij} denote business-level and driver-level outcomes, respectively. No covariates are included in the main regressions, unless specified. α_s are strata fixed effects. $T^{GranularObs}$, $T^{CoarseObs}$, and T^{NoObs} represent the three different observability treatment arms (granular observability, coarse observability, and no observability) described earlier. T_{ij}^{Access} denotes the pooled treatment of receiving access to digital payments for the driver level outcomes. ϵ is the error term.

These regressions include all eligible firms with near-complete treatment compliance: almost no control group received the technology and all eligible treated firms received it.[20] Firms that did not participate or were ineligible are simply excluded, as they were not made aware of their treatment status in advance (ensuring no differential attrition), hence ITT=TOT.

Each coefficient estimates the average causal effect of the treatment arms on the outcome of interest. Equation (1) is used to capture the observability effects on business outcomes, and I add the F-test of the differences between β_1 and β_3 to isolate the impact of observability from the impact of technology itself on owner-driver relationships.

The taxi-level regression (2) measures the effect on driver's cost associated with using cash, including both employed drivers and owners driving their taxis to increase sample size and statistical power. Standard errors are clustered at the business level in taxi-level regressions, as recommended in Abadie et al. (2022). The results section also includes specifications for heterogeneous treatment effects and Poisson regressions for count data outcomes.

Outcomes from short-, mid-, and long-term surveys are examined. A key empirical challenge is the varying separation rates among owner-driver pairs across treatment arms, with separation as an outcome variable. To address this, I analyze data at the business level, including both existing and newly formed pairs where drivers were replaced. Owners with granular observability accessed digital transaction histories for both current and new drivers. This approach reduces bias from separation effects. Results for pairs that remained together are shown in Appendix B.6. The focus is on mid-term results, allowing enough time for contracts to evolve while ensuring that a significant number of owner-driver pairs remain together. After two years, the main outcome of

[20]Over the nine-month period when the control group was on a waitlist (and prevented from accessing the technology), only 9 control drivers (less than 1%) managed to access the technology by changing their phone numbers.

interest is the separation rate, given that around 35% of owners exiting the industry and a smaller share becoming sole proprietors. This turnover complicates the assessment of long-term effects on contracts, but the main results remain robust, as shown in Appendix B.7.

The specification for the "adoption experiment" is more straightforward. As explained below, I regress the impact of observability on the likelihood of drivers providing the owner's contact information and compare the characteristics of those willing to adopt versus those who are not.

Econometrics Robustness Checks I show results remain unchanged while accounting for contamination bias correction in the case of multiple treatments to address non-convex averages of the effects of other treatments, as recently recommended in Goldsmith-Pinkham et al. (2024).

5.2 Impact Experiment: Digital Payments Reduce Costs of Using Cash

In this section, I examine the impact of digital payments on various costs of using cash. I include all taxi drivers, including taxi owners driving without employees. To underscore the relevance of the trade-off between observability and adoption, I first establish that the technology delivers tangible benefits to drivers. In particular, I quantify the cost reduction and highlight that drivers have a high willingness to pay for the technology by the end of the study.

Costs of Using Cash. In Table 2, I pool across observability treatments and document that the digital payment technology reduces several costs of using cash by approximately half. Column (1) shows that 44% of drivers in the control group spend more than 10 minutes looking for small-change at least once in the past 7 days in the short-term. For the treated drivers with digital payments, this share is reduced by 21 percentage points (pp)—nearly half. Similar reductions are seen in other cost categories, such as the share of drivers refusing customers or making change mistakes. Specifically, the share of drivers reducing prices decreases by 12 pp in the short-term. As a result, the imputed loss from these small-change issues is nearly halved, decreasing from USD 5 (approximately 6.2% of profit, i.e., about a 3% increase). This reduction represents a modest but meaningful share of drivers' profits, as handling cash is an important issue for drivers (see Section 2.2.1). Finally, digital payments decrease digital theft by approximately 71%. I cross-validated some of these results by conducting mystery audits on taxi drivers to check for social desirability bias, as described below.

This reduction in costs of using cash is striking given that the digitalization covered a limited share of drivers' revenue: 2.2 customers per day on average (up to 6 for the 95pp). To compute this, I asked treated drivers to self-report their revenue sources over the three days prior to the survey, both in the mid-term and long-term. On average, digital payments account for 13% of the total revenue among users, with variation in digital usage across drivers (up to 40% in the 95pp).

The impact of digital payments evolves over time: Table 2 shows that the treatment effect is greater after 5 months compared to 9 months. Interestingly, the control group also experiences improvements, as costs like time lost decrease significantly, resulting in a substantial reduction in their imputed loss. Yet, the treatment effect remains significant. Analysis of transfer data provides additional insights into this trend: as digital payments become more widespread in Dakar, the control group drivers increasingly accept P2P transfers from passengers, despite the revocability concerns. This shift highlights how the broader adoption of digital payment systems indirectly benefits even those outside the treatment group.

Prices. I use the mystery passenger audits on taxi drivers to cross-validate the self-reports. In Table 3, Panel A, I find that treated taxis are 31 pp more likely to accept digital payments, above the 44% control mean, with control drivers not having the technology but risking transaction revocation by using their personal mobile mobile money wallet as previously discussed. The treatment group does not reach 100% acceptance for at least two reasons: (1) drivers may still prefer cash-on-hand and (2) the license number matching may not perfectly identify the experimental driver if the car was shared with an occasional driver on the audited day (measurement error).

I examine the impact of digital payments on pricing strategy in two ways. First, I run the OLS specification in Panel B and find that digital payments increase the likelihood of accepting non-rounded prices by 28% compared to control. Second, given imperfect compliance, I use an IV to measure the local average treatment effect (LATE) on drivers accepting digital payments. The exclusion restriction states that treatment status affects the outcome "accepting a non-rounded price" only through accepting digital payments. Panel C shows a higher impact on non-rounded prices than OLS: drivers accepting digital payments are almost twice as likely to accept a non-rounded price, above the 32% control mean or a reduction in the price distortions induced by small change by 43%, even when accounting for surveyor and origin-destination fixed effects. In line with this, about 30% of prices in the treated group's administrative payment data are also non-rounded. The impact on prices provides casual evidence that cash payments distort pricing, pushing businesses to use stepwise pricing at the lowest available bill. Digital payments allow for more flexible pricing, reducing this dead-weight loss.

Driver Profit. I connect the previous results on costs of using cash with the overall profitability of the taxi drivers. Measuring profits in informal settings is challenging and often results in noisy data. To address this, I use two strategies. The first measure, derived from extensive fieldwork, details the taxi production function, itemizes costs, and measures total revenue over the last 3 working days of each survey round. The drivers' profit function consists of the revenue collected, coming from a unique source (passengers), minus costs such as fuel, transfers to the owner, small repairs, police bribes, washing, tolls, association contributions, and food/drinks consumed during work. The second measure is a self-report estimated average daily profit over the past 30 days, following De Mel et al. (2009). As evidence that the profit measure is noisy, the last 6-day profit (measured twice, in the short- and mid-term) tends to be higher than the last 30-day daily average, although there is a strong positive correlation between the two profit measures—see Figure A5.

I am unable to detect a significant increase in drivers' daily profit from the technology. In Table B9, Columns (1) and (2) show that profit does not significantly increase. While there might be a reduction in the number of customers in the short-term with an increase in productivity, measured by customers or revenue per hour, this effect is not significant and disappears in the mid-term. Although the technology reduces the costs of using cash, this change may not be large enough to significantly impact reported profit (or to be detected). To investigate whether reported profit is too volatile to detect effects, I run power calculations and compute the minimum detectable effect (MDE) for profit-related outcomes, with a power of 80%. Table B10 shows that I am not sufficiently powered to detect small changes in profits, with MDEs on profit of $\approx$ 6-12%, above the estimated impact of the technology. The relatively high MDEs indicate that while the technology significantly reduced the hassle costs of cash, the sample size and noise in profit measures, commonly noted in the literature, may limit the ability to detect corresponding profit changes.

Additional Outcomes: Theft Anxiety, Record-Keeping, and Savings. Table B11 explores other outcomes, outlined in the PAP, including theft anxiety, record-keeping, and savings. Digital payments significantly improve drivers' record-keeping by about 3-5pp (over a control mean of 2%) and reduce theft anxiety by 6-8%. Additionally, digital payments lower "luxury" purchases by 13% in the short term, although this is not sufficient to increase drivers' savings. Anecdotally, digital payments motivated some drivers to formalize their status by acquiring the necessary paperwork to access the technology, highlighting the formalization benefit of digitalization.

Willingness to Pay (WTP). I elicited WTP for both treated and control groups at baseline and in the mid-term. Baseline WTP was elicited for a lump sum payment using the Becker–DeGroot–Marschak (BDM) procedure (Becker et al., 1964) for a technology covering the QR hanging card, mobile application, and training on how to use it. Respondents were informed that their valuation could be used to determine their access to the technology.[21] Mid-term WTP elicitation was not incentivized, as removing the technology from drivers was not desirable for the payment company. Treated drivers were asked their WTP to keep the technology, and control drivers about their valuation to access it.[22]

Among drivers willing to adopt the technology, I show WTP over time across the two main treatment arms, which reveals three key findings (see Figure 3). First, the endline driver's WTP is substantial, approximately one week's profit, for a technology provided for free. Second, WTP increases substantially over time, from about USD 17 to USD 57 (an increase of 235%). While this was not an incentivized measure and should be interpreted with caution, it may in fact understate the true mid-term valuation. This change is possibly due to network effects from the citywide adoption of the technology, which is discussed in a companion paper (Houeix, 2024). Third, both

[21] To preserve randomization, a lottery was conducted on 5% of the treated sample with a predetermined random number between 0 and CFA 3000, typically below the baseline WTP, resulting in the exclusion of only one treated driver from receiving the product.

[22] If anything, the mid-term WTP may be biased downward, as drivers might understate their true WTP, fearing the partner company could use this information to set future prices.

treated and control drivers value the technology equally, with no significant difference between the groups (p=0.431), suggesting that "learning by using" does not determine willingness to pay in this setting.

Finally, I demonstrate that WTP is meaningful by showing that baseline cash transaction costs significantly predict WTP. Table B12 shows a strong correlation between higher costs and WTP across all cost types. For instance, experiencing delays due to small change shortages increases the WTP by 17%. In these regressions, I control for WTP for a benchmark good—a bottle of water—to reduce noise, as recommended in Dizon-Ross and Jayachandran (2022), though removing this control does not affect the results. This indicates that drivers genuinely value the technology's benefits in reducing the hassle costs of cash and are therefore willing to pay to acquire it.

5.3 Impact Experiment: Digital Observability Improves Efficiency

This section tests the key hypothesis that making digital transactions observable reduces information asymmetries among adopters, improves contract efficiency, and affects the relationship between taxi owners and drivers. The underlying theoretical mechanisms are formalized in the framework presented in Section 6.

Because owners with observability gained access to the digital transaction history of both current and newly hired drivers, the analysis is conducted at the owner level, including existing and newly formed owner-driver pairs, if any. For robustness, I also present an analysis restricted to baseline pairs remaining together in Appendix B.6, showing similar effects.

Addressing Information Asymmetries. I demonstrate that digital payment technology provides valuable information to owners about their drivers' effort and digital revenue. Figures A3 and A4 illustrate the data available to the taxi owners with full digital observability, such as transaction timestamps, which can indicate driver's days worked and possibly working hours. To test this "first-stage" more formally, I employ two strategies. First, I directly ask owners whether they use the technology to monitor their drivers and measure the treatment effect of transaction observability on their knowledge. Table B13 shows that owners with *Granular Observability* are significantly more likely to claim they know their driver's digital revenue by 15 pp (a 72% increase, as shown in Column 3) and 8 pp (29%) more likely to correctly guess the number of days worked by their driver in the past week, though the latter is not significant. Owners with *Coarse Observability* also gain some knowledge about the driver's work, but the impact is lower due to the less informative nature of their treatment. The positive but limited increase in knowledge suggests that digital transactions provide only *imperfect* information on drivers' work.[23] Overall, Columns 4 and 5 show that 34% of owners with *Granular Observability* use the technology to observe how much their driver works, and 43% look at the driver's transactions daily or weekly.[24]

[23] A cultural and religious norm in Senegal—particularly the Islamic principle that one should not speak without certainty—led many owners to respond "don't know" when asked about their perceptions of the driver's work.

[24] Anecdotally, one owner stated in an open-ended response "I'm thrilled because the digital payment app gives me a clearer view of my driver's work by tracking the exact number of transactions he completes."

Second, I analyze the information contained in the digital transaction data. In Table B14, I compare drivers' self-reported effort and output to their digital transactions using administrative data. There is a strong positive correlation between drivers' effort, output, and use of digital payments. Among treated drivers, the number of digital transactions and revenue collected are highly correlated with the number of days worked. For example, one additional digital transaction predicts an increase of 0.8 in the total number of passengers that day.

Taken together, these results show that owners gain valuable information about their drivers from transaction observability, especially when transactions are fully observable.

Increase in Worker Effort. Measuring effort in an informal sector is challenging due to the inherent difficulty for the principal to observe it. Moreover, self-reported effort is often subject to social desirability bias and recall bias—drivers often struggle to recall performance metrics, with 41% reporting finding it complicated to remember their performance in the past 3 days—and working hours are generally a noisy measure.[25]

I leverage three strategies to measure efforts: (a) mystery audits, (b) drivers' default, and (c) digital usage. Altogether, these metrics consistently indicate higher effort from drivers.

First, I use the mystery passenger audits to measure driver's effort directly, as described in Section 4.2. The underlying assumption is simple: the more frequently a taxi driver is spotted on the road, the more likely they are exerting higher effort. Table 4 uses Poisson regressions and shows that taxis under *Granular Observability* are seen on the road 34% ($\beta = 0.29$) more often compared to the control group (p=0.003). This number is even higher (47%) when I restrict to pairs where owners are not driving (where it is more certain the driver is indeed the one working), although the coefficient for *No Observability* is also positive there. Most effects are observed on the extensive margin (number of unique days observed on the road). I do not find significant effects under the *Coarse Observability* treatment. This positive effect of *Granular Observability* indicates that drivers' effort increases, as they seem to work more intensively over the two-week audit survey period. Mystery audits can also be used to verify the absence of manipulation in digital payment usage across treatments. I find that drivers consistently respond to mystery passengers requesting digital payments, with no evidence of manipulation, as shown and discussed in Figure A6.

Second, I measure drivers' monthly default rate. To mitigate drivers' self-report biases, I asked both owners and drivers about the frequency of default on the rental fee over the past three months, and created a dummy variable for whether drivers defaulted at least once a month (the case for 31% of drivers in the mid-term). In Table 5, I find that the partial default likelihood decreases by 10 pp, a 31% reduction under *Granular Observability* (p=0.078). This reduction, although suggestive, highlights a clear gain for owners in monitoring their employees: they can encourage increased effort, thereby raising the frequency of high-revenue and reducing default. As a

[25] Despite these challenges, I still collected self-reported performance data over the three days prior to the short-term and mid-term surveys. Fixed effect models were used to reduce measurement noise. As shown in Table B19, there is a small but significant effect on hours worked in the mid-term compared to the control group, though not statistically different from the *No Observability* group. No significant differences were observed in the number of customers, total revenue, or days or hours worked across groups.

result, owner's profit under *Granular Observability* increases by 8% (Column 3), though the effect on the profit measure is not statistically significant (p=0.240). Note that this effect accounts for the increase in the upfront payment to drivers described below, but does not directly include the positive worker retention effect and reduced cost of finding a new match discussed below.

Third, I compare the frequency of digital transactions received by treated drivers in their business wallets across observability treatments. Table B18 shows Poisson regressions indicating that drivers whose owners can observe their transactions use the digital payment technology significantly more. On average, drivers under *Granular Observability* receive 35% (p=0.015) more transactions than under *No Observability*, on 16% more weeks, and 24% more days throughout the first 9 months of the experiment. Controlling for pre-trends using pre-experiment peer-to-peer transactions does not substantially alter these results (see Columns 1 and 2). Figure A9 illustrates that digital usage of drivers under *Coarse Observability* also increases, though this effect diminishes after 6-7 months. This suggests drivers might have initially attempted to signal effort, but stopped when it did not lead to contract changes. There is no evidence of ratchet effect: drivers do not use digital transactions more under *Coarse Observability* compared to *Granular Observability*.[26]

Taken together, these findings indicate that drivers exert more effort under *Granular Observability*, consistent with the fact that this treatment reveals a signal of worker effort to owners, contrary to the *Coarse Observability*.

Contract Changes. I show the impact of digitalization on the contractual form. First, I examine the impact of observability on two key contract parameters: the upfront fixed payment from owners and the rental transfer from drivers. Table 6 and Figure A7 both show that owners with *Granular Observability* are significantly more likely to offer an upfront payment to their drivers.[27] In the mid-term, 75% of owners tend to offer an upfront payment to their drivers, but that share increases by 13 pp or 18% when they can observe digital transactions (p=0.003).[28] This is supported by the F-test for the difference between *Granular Observability* and *No-Observability* (F-stat = 7.2). This change primarily occurs at the extensive margin, resulting in a 20% increase in payment values over the control group, and less affecting the drivers already receiving the upfront payment

[26] As an additional robustness check, I analyze drivers' strategic behavior by examining P2P transactions within the typical range of taxi transactions, termed "taxi-like" transactions. This is to detect if drivers make personal transactions appear as business ones to please their employers, which would negatively affect P2P "taxi-like" transactions. A null or positive effect would indicate increased actual effort, as these transactions should (if anything) positively correlate with business transactions. Table B21 argues against strategic misreporting. Panel A shows that the number of P2P "taxi-like" transactions is about three times smaller than the business transactions on average (22.35 P2P vs. 59.18 business transactions), but they still increase by 14%. This positive effect, though not significant, is observed across different outcomes, with all coefficients below the effect on business transactions. Panel B shows that all non "taxi-like" transactions also increase, though not significantly, which is more challenging to interpret since drivers under observability sometimes receive their upfront payment via mobile money.

[27] Through this section, I refer to this contract component as a "upfront payment" for clarity. The exact term used in the taxi industry is a "salary".

[28] Note that the higher control mean after 9 months might be attributed to an overall increase in drivers' outside options, possibly influenced by the ongoing construction work related to the bus rapid transit system across the city of Dakar during the experiment.

(see Figure A8).[29] Table B15 shows that the effect is slightly less pronounced among owner-driver pairs that stayed together. This suggests that observability also impacts contracts for newly hired drivers, as treated owners can now monitor transactions of both new and existing drivers (since remaining together is also an outcome, I interpret this effect with caution).

The upfront payment may compensate for the higher driver effort when adopting the technology. I do not find any change in the rental fee (see Table 6, Column 3); owners keep the agreed rent parameter constant as they further provide this upfront payment—only rent default decreases, as previously shown.[30] The absence of change in the rental fee can be explained by the relational nature of the contract: providing the payment upfront prevents the principal from reneging on his commitment, as formalized in Section 6. Ultimately, this upfront payment means drivers' profits increase on average (see Table B22)—though this measure excludes the non-monetary cost of effort. I estimate the welfare effects, accounting for effort costs, in Section 7.

As a robustness check, I show that the treatment effect on the salary is more pronounced when drivers use the digital platform more intensively. In Table B20, I regress the contract outcomes on the observability treatments, interacting with whether the driver is in the top 50% of digital usage in terms of daily digital transactions received. The evidence is suggestive because the regressor is endogenous, but reassuring that the contract effect is more concentrated in pairs where drivers process digital transactions more intensively (by 13 pp further than the other drivers, although the interaction is not significant).

Reduction in Owner-Driver Separations. I first document a high separation rate in the taxi industry. Throughout the first nine months of the experiment, 34% of pairs overall split for various reasons, as summarized in Table B7, ranging from those directly related to information asymmetries, such as drivers being fired for defaulting on the rental fee, to indirectly related reasons, such as owners selling the taxi or the taxi undergoing long maintenance (see Fact 4).[31]

Table 6 and Figure A10 show that separation rates are lower across pairs with *Granular Observability*, especially after nine months and nearly two years. I find that pairs with a driver using the technology are less likely to split, primarily driven by the *Granular Observability* arm. Pairs under *Granular Observability* are 11 pp (p=0.034) less likely to break than the control group, a reduction of 30%. Although large, this difference is not statistically significant compared to *No Observability* (F-test p=0.16).[32] The effect of observability on separation, primarily concentrated in the *Granular Observability* treatment arm can be explained by the fact that the technology enabled better

[29] These contract changes remain significant when accounting for possible contamination bias correction for multiple treatments (see Table B17).

[30] Note that the taxi industry is characterized by a rigid norm around the agreed rental transfer, fixed at FCFA 60,000 (USD 100) per week, which is the case for 75% of pairs in the mid-term.

[31] Since separations occur for various reasons, I analyze overall separation since analyzing treatment effects on specific reasons did not reveal significant differences across treatment arms.

[32] As a robustness check, with the same endogeneity caveat as before, I show that the treatment effect on separation is more pronounced when driver's digital usage is high in Table B20. Drivers using the digital platform more intensively with an owner observing their transactions tend to remain in their pairs longer than those who use it less. The effect is 3 pp larger, although not significant.

monitoring of effort, reduced moral hazard, and consequently led to fewer separations. I find some evidence suggesting that drivers under *Coarse Observability* are 4 pp more likely to stay with their owners after two years (see Table B16). Although the estimate is highly suggestive, this may indicate that low output observability may become beneficial as the app gets used more.

Digital Observability Increases Hiring by Owners. I investigate whether transaction observability impacts hiring decisions to causally test the hypothesis that effort contractibility influences firm size (Baker and Hubbard, 2004). To do so, I combine the sample of owner-driver pairs with owners driving their taxis alone. The latter were similarly randomly assigned to one of the three observability treatments when given technology access, with the observability applying to any potential future hired drivers.[33] By doing so, I increase the sample size and test how observability might influence hiring decisions, particularly for owners typically reluctant to hire drivers.

In Table B23, I find that owners under *Granular Observability* are 7 pp and 4 pp more likely to hire a driver after nine months and nearly two years, respectively (corresponding to 43% and 14% increase) compared to those under *No Observability*. This table also includes long-term results because the impact on hiring decisions is likely to take time to manifest. For comparability, the pure control group is excluded from the analysis as they received treatment after 9 months. The effect is mostly driven by owners who already have a driver at baseline. While significant only in the mid-term, the positive effect suggests that observability could also facilitate business expansion.

Increase in Trust and Value of Relationships. I show that observability increases trust in the relationship between owners and drivers. Given the multifaceted nature of trust, I employ two strategies. First, I measure treatment effects on owners and drivers' survey responses, especially using trust questions from the World Values Survey, Wave 7 (2017). Second, I estimate heterogeneous treatment effects using variation in baseline business characteristics.

First, I asked owners about their perceptions of their drivers' moral hazard. Owners are significantly 44.1% less likely to attribute low earnings to low driver's effort vs. bad luck (Column 1 of Table 7). Second, I collected a revealed preference measure of trust regarding the driver's effort: whether the driver is allowed to park the taxi at their own house. In this context, owners who lack trust in their drivers often require them to park the car at a specific location they can monitor each day. Column 2 shows that drivers under *Granular Observability* are 27.6% significantly more likely to park the taxi at their own place. Additionally, owners' self-reported trust in their drivers

[33] For instance, taxi owners with no employees randomized into *Granular Observability* were informed: "Your option with this technology is *Granular Observability*. As taxi owner, you will have access to the digital transaction history of any driver you hire in the future and receive an SMS indicating their total daily transactions." Conversely, owners with *No Observability* were explicitly told that they would not have access to any new driver's transaction history in the future.

increases under *Granular Observability* among treated pairs.[34] Overall, these measures suggest that owners trust their drivers more when they can better observe their transactions and effort.

Second, I measure the heterogeneous treatment effect of observability on separation rates based on baseline business characteristics related to trust. In Table B24, I investigate three heterogeneity dimensions: long relationships, family businesses, and risk-averse agents.[35] Two findings directly relate to trust. On the one hand, the top row of Table B24 shows that separation rates are generally higher in non-family businesses and in recent owner-driver relationships, characteristics negatively correlated with trust in the data. On the other hand, the treatment effect of observability appears stronger in these same pairs, with interaction terms of -3 pp and -3 pp for recent relationships, and -3 pp and -14 pp for non-family pairs at mid- and long-term, respectively.[36] These findings, although only suggestive, suggest that owners are more likely to retain non-family and recent employees compared to the *No Observability* group.[37]

Taken together, the impact of observability on trust, especially in low-baseline trust pairs, underscores the critical role of monitoring technologies in lower-income countries. Information frictions and limited access to these technologies may partly explain the persistence of family businesses. These findings have development implications, as family businesses are frequently associated with efficiency losses (Bertrand and Schoar, 2006; Chandrasekhar et al., 2020).

Taking Stock: Digital Payments as Effective Monitoring Technologies. Switching to digital payments provides owners with additional information about their drivers, leading to several outcomes: increased effort, reduced default, contract change, and increased worker retention. These effects suggest that moral hazard in effort is a key binding constraint in this sector. Most of the effects are in the *Granular Observability* treatment. I only observe changes under *Coarse Observability* after two years, likely because digital payments need to be used more extensively to provide sufficient information about revenue to also reduce moral hazard in output reporting. Since many drivers still use digital payments for only a limited portion of their total revenue (13%

[34] Table 7 shows that trust levels, while not significantly different from the control group, are significantly higher than in the *No Observability* treatment arm, even when controlling for general trust in drivers (Columns 3 and 4). This finding suggests that not revealing digital transactions to owners could have negatively affected their trust. However, I interpret this with caution, as the stated trust on the 0-10 scale is high on average and the effect size is small. As expected, there is no significant effect on driver's trust in their employers (Column 5).

[35] These heterogeneity dimensions were pre-specified in the pre-analysis plan under AEA registry ID #0009155, and two of them are used to stratify the randomization. Additional heterogeneity analyses were also discussed but did not yield significant differences, hence they are not reported in this paper.

[36] I exclude the control group here, as they were treated after 9 months, to perform a clean comparison between *Granular Observability* and *No Observability* at mid- and long-term. Analyzing longer-term effects is particularly useful for this heterogeneity analysis because differential effects on sub-groups tend to increase over time. However, the limited sample size and high p-values on interacted terms lead to results that should be interpreted with caution.

[37] The last two columns on risk aversion are discussed in this footnote, as they are not directly related to trust. I elicited the risk aversion coefficient for all owners and drivers at baseline using an incentivized game with simple choice tasks à la Holt and Laury (2002), tailored to the taxi industry. Drivers were then assigned a relative risk-aversion score, with those above 1 (CRRA utility function) defined as non-risk-averse agents. The impact of observability is greater for non-risk-averse agents: -3 pp and -14 pp at mid- and long-term. This result is consistent with a similar logic as previously discussed: pairs with risk-averse agents are more likely to be in established low-risk contracts with low separation rates already.

on average), the technology may be more effective in informing owners about drivers' effort (e.g., how much they work on a given day and week) and thus mitigating moral hazard in effort in the short-term, rather than fully revealing total revenue collected. The digital payment technology thus expands the production possibility frontier by reducing cash transaction costs and moral hazard, thereby increasing efficiency for businesses whose agents adopt the technology.

5.4 Adoption Experiment: Digital Observability is a Barrier to Technology Adoption

This section analyzes the technology's observability feature on adoption, as detailed in Section 3.3. The hypothesis tested is that observability could lower the technology's total benefits by discouraging initial adoption by agents.

Barriers to Adoption of Digital Payments. During the listing survey, most approached drivers showed interest in the new technology. However, 50.2% of these declined to provide the contact information of their taxi owners after three follow-ups, thereby preventing them from adopting the technology. Drivers often mentioned that they needed to talk to the taxi owners before sharing contact information. After follow-ups, only 5.0% of drivers changed their minds and were thus included in the impact experiment. Table B25 summarizes the reasons drivers cited for withholding the owner's contact information: 48% of the drivers raised concerns over owner's privacy, with the need to consult with the vehicle owners before sharing their details, while 15% claimed that the owner was unavailable or uninterested. About 20% of drivers explicitly mentioned apprehension about owners accessing their digital transaction history.

Adoption Experiment: The Role of Observability in Technology Adoption. To isolate and quantify the impact of digital observability on the driver's adoption decision, I re-offered the technology to reluctant drivers a month after their initial refusal, as described in Section 3.3 and in the bottom half of Figure 2.

Table 8 shows a large positive effect of removing observability on driver's adoption. In the control group, where drivers were informed that their employer would have access to their transaction history, 14% of drivers changed their decision over time and agreed to provide their employers' contact, allowing them access to the technology. However, removing observability nearly doubles this share, with an increase of 80% (p=0.003). This result remains robust across various specifications and controls, including surveyor fixed effect, privacy controls (e.g., driver's willingness to share alternative contacts like that of their association president or closest friend), and the length of their relationship with their employer. The effect is more than twice as large among the worst-performing and poorest drivers—which I examine in greater detail in the next paragraph.

This finding provides empirical evidence that, although observability can increase profits and improve contractual efficiency, it may also discourage adoption among certain drivers. This is particularly relevant in sectors where employees have some degree of autonomy in deciding whether to adopt the technology. For instance, employees of small businesses may play a crucial role in informing their employers about new technologies available in the market. However, assuring employees that digital technology won't bundle data sharing significantly increases uptake.

The effect of removing observability might even be underestimated, as drivers may not have fully trusted the research team to keep their transactions private after adoption, which could partly explain why most drivers still refused to provide their owner's information. To support this, I run regressions for two groups: (1) those who cited observability as the reason for withholding their owner's contact information, and (2) those who cited other reasons during the listing survey (see Table B26). The results show that (a) the treatment effect is significant for both groups, including those who did not raise observability as an issue, and (b) adoption remains incomplete for those concerned about observability, even when it is removed. This suggests the treatment effect may be underestimated if some drivers are still uncertain whether observability will actually be removed. Additionally, recent administrative data show that 88% of these previously reluctant drivers ultimately adopted the technology by 2024 after the payment company made non-observability the default option following the experiment, as further explored in Section 7.4.3.

Profile of Reluctant Drivers: High-Disutility of Work, Low-Performing, and Poorest Drivers θ^l. To characterize selection, I rely on insights from Karlan and Zinman (2009)'s methodology by comparing drivers in the "impact experiment"—who agree to potential observability but whose owners are randomly assigned not to receive it, i.e., *No Observability* or *Control* referred to as "willing to adopt"—to "reluctant drivers". In the latter, I exclude the 14% of drivers who provided the required information at the one-month follow-up under *Granular Observability*, as described above. This approach is useful since only mid-term performance data could be collected. Many reluctant drivers refused to participate in a long one-month follow-up survey, as they were unwilling to share their owner's information and knew it would hinder adoption. The results remain robust when comparing only drivers randomized to the *No Observability* groups (see Table B27).

Table 9 shows that reluctant drivers perform worse along several dimensions: in the three days prior to the survey, they have fewer passengers (-11%), collect less money (-4%), and work fewer hours (-4%), resulting in a lower overall z-score performance index—see Panel A. They are also more likely to feel stressed at work about making rental payments (+83%) and to be risk-averse. They tend to already receive an upfront payment (likely due to limited liability and their lower income status—see below), with a similar separation rate in the mid-term (Panel B). This comparison suggests that digital observability embedded in the technology could reduce their informational rent (e.g., their ability to avoid work), thus discouraging their initial adoption.

Furthermore, reluctant drivers are substantially poorer across several dimensions. Table 9, Panel D, shows that the wealth index, developed by Innovations for Poverty Action (IPA) and tailored to Senegal, is significantly lower for reluctant drivers (-9%). They also tend to rely on other income sources further, although not significantly. Reluctant drivers are also 28% less likely to have attended primary school and 9% less likely to be able to read and write. This finding is particularly relevant for welfare, as the poorest drivers might derive the greatest marginal utility from reducing the hassle costs of using cash.

High-Performing Drivers Prefer Observability at Baseline. To provide additional evidence explaining the differential adoption across drivers, I show that high-performing drivers prefer digital observability. First, I document heterogeneity in observability preferences among drivers willing to adopt. Both owners and drivers ranked the three observability options (granular, coarse, and no observability) without knowing if their preferences would be considered to minimize any reporting bias (these rankings did not affect the random assignment at the survey's end). While most drivers preferred *No Observability* (59%), 23% preferred *Granular Observability*. Owners were split, with about half favoring *Granular Observability* and the other half preferring none, often due to concerns about drivers' reactions to being monitored.

Second, I use this heterogeneity and find that high-performing characteristics significantly predict drivers' preferences for observability. Table B28, Panel A, shows that higher performance indicators, such as average daily revenue, days worked per week, and fewer defaults on rental fees, are positively correlated with this preference. Drivers with further high-performance days are 10 pp more likely to prefer observability while low performers are 8 pp less likely to do so. In line with this, drivers preferring observability have longer relationships with their employers compared to those who prefer *No Observability* or initially refused to provide their employers' contact information. Finally, Panel B shows that owners with drivers preferring *Granular Observability* tend to underestimate their drivers' work more than owners of drivers preferring *No Observability*. This suggests that drivers may want to signal their effort to potentially biased owners, anticipating that increased observability would help build trust and increase retention.

Long-Term Worker Retention Across Groups. Figure A11 shows a higher separation rate after nearly two years among drivers who were reluctant to adopt the technology. Overall, 61% of owner-driver pairs separated during the experiment. The leftmost bar shows the separation rate of the reluctant drivers—those who refused to provide their employers' information explicitly due to observability concerns—have the highest separation rate (69%).[38] As previously discussed, pairs randomly assigned to *Granular Observability* have the lowest turnover rate (56%), followed by those in the *Coarse Observability* group, although these differences are not statistically significant given the limited sample size. The differential retention rates across experimental groups could have important welfare implications, as explored in the structural estimation section 7.

6 Theoretical Framework: Impact and Adoption of Digital Payments

The goal of the theory is to formalize the mechanisms underlying the experimental results and guide the structural estimation. Using insights from the relational contracting literature (Baker et

[38] The definition of reluctant drivers is based on their explicit refusal to share contact information because of concerns about transaction observability. Data further supports this classification, showing that drivers who refused for reasons unrelated to observability have lower separation rates.

al., 2002; Levin, 2003) and the literature on sharecropping (Banerjee et al., 2002), I describe a simple framework of the owner-driver relationship (principal-agent) in the informal taxi industry where I incorporate some key features of the setting, i.e., imperfect information on effort and output, limited liability, and the relational nature of the contracts. The framework generates two sets of predictions: (a) the impact of digital observability on the principal-agent contract when effort and output are imperfectly made observable, and (b) the differential adoption across drivers' types that may arise following the introduction of the technology. It provides the basis for the structural estimation of driver and owner welfare, which then allows to run policy counterfactuals.

The framework formalizes three key mechanisms underlying the impact of observability introduced by the digital payment technology. First, *Reduction of Moral Hazard in Effort*: the technology generates a signal of high effort when a sufficient volume of digital transactions is processed throughout the work period. Second, *Reduction of Moral Hazard in Output Reporting*: the technology provides information on the agent's digital output, representing a share of total revenue. Third, *Reluctance to Adopt for Low-type Agents*: by reducing moral hazard in effort, the technology would deter low-type agents (those with a high disutility of effort) from adopting. The key ingredients that generate these mechanisms are common frictions in lower-income contexts: *limited liability* and *information asymmetries*, which provide informational rent to the agent, creating incentive problems, and *weak contract enforcement*, which explains both the change in contractual form—an increased payment upfront rather than an ex-post rent reduction—and the principal's inability to commit to contract terms, hindering adoption for some agents.

6.1 Setup

Consider an environment with an infinite discrete time horizon, with periods indexed by $0, 1, \ldots, \infty$. Both the principal and the agent share a common discount factor, $\delta < 1$. Throughout this paper, "he" is used for the agent and "she" for the principal, by convention. The principal aims to incentivize the agent to exert effort on the job. The agent's effort takes discrete values, $e \in \{0, 1, 2\}$, and is unobservable to the principal at baseline. I assume that the output $y(e)$ is binary and follows:

$$y(e) = \begin{cases} Y & \text{with probability } q_e \\ X & \text{with probability } 1 - q_e \end{cases} \tag{3}$$

where $X, Y \in \mathbb{R}^+$, $X < Y$, and $q_2 > q_1 > q_0$. The agent's production function reflects the substantial output uncertainty inherent in the taxi industry, where an agent exerting low effort can achieve the same output as one exerting high effort, due to stochastic factors, as described in Section 2.2.2. For instance, a taxi driver might work all day yet struggle to find customers, while another may work fewer hours but finds high-paying fares.

There are two types of agents, low and high $\theta \in \{l, h\}$, which differ in their disutility cost of effort $\phi^\theta(e)$. The type of each agent is publicly known, where $\phi^l(e) > \phi^h(e)$, with $\phi^\theta(0) = 0 \; \forall \; \theta$. The public type assumption is made for simplicity and to highlight that adverse selection is not required to generate the results—but the framework could be extended to accommodate it. The agent is referred to as low- and high-ability type, respectively.[39] The agent's utility, denoted by U^θ, is defined as a function of the agent's total revenue collected y, minus the transfer t and the disutility cost of effort $\phi^\theta(e)$.

The baseline contract between the principal and the agent revolves around two key endogenous variables: the transfer $t(\tilde{y})$ from the agent to the principal at the end of each period and the continuation probability of their relationship $p(\tilde{y})$. Both these variables are functions of the self-reported output by the agent, $\tilde{y}$. Upon termination of their relationship, both the principal and the agent incur one-time replacement costs in the next period, K_p and K_a, respectively. Upon termination, I assume that both the principal and the agent get a new random draw of an agent and a principal, respectively, from an infinite pool of players. The pool of unmatched agents consists of a fraction μ of high-type agents, $1-\mu$ of low-type agents, with μ known to the principal. To keep the exposition simple, I take μ as constant over time, reflecting a setting where the stock of new agents is significantly larger than the number currently participating in the game. The agent always has the option to exit the taxi industry and take an outside option valued at $\bar{u} > 0$.[40] The principal always has the outside option to sell the car and stop working in the taxi industry.

6.2 Assumptions

Based on survey insights and empirical observations documented in Section 2.2.2, I make the following assumptions:

Assumption 1. ***Unobservability.*** *Agent's effort e and output y are unobservable to the principal.*

While unobservable effort is a standard assumption in contract theory, I also model for unobservable output because it is a prevalent feature in many informal settings. Supported by Fact 3, the principal lacks direct information about the agent's actions, relying solely on reported output $\tilde{y}$ (see the timeline of one period in Figure 4). This output reporting issue is discussed in static contract models in Townsend (1979); Lacker and Weinberg (1989); de Janvry and Sadoulet (2007).

Assumption 2. ***Limited Liability.*** *The agent faces a constraint $t(\tilde{y}) \leq \tilde{y} \leq y(e)$, ensuring that the transfer t does not exceed reported output and reported output does not exceed actual output.*

This constraint reflects Fact 2, which highlights that many drivers defaulted recently, have no savings, and have limited access to credit. By assumption, they cannot report more than what they collected. This framework uses insights from Innes (1990).

Assumption 3. ***Restrict to Stationary Equilibria.*** *When deciding the contract for the next period, the principal relies exclusively on current period reported output, with contract parameters constant over time.*

[39] For simplicity, this section does not delve into what determines these ability types. However, it is empirically observed that low-ability drivers are typically poorer, a population of high policy relevance due to the significant welfare implications, as discussed in the results section.

[40] In this Section, I assume that both types of agents have the same outside option for simplicity. In the structural estimation section, I relax this assumption and allow for $\bar{u}^h > \bar{u}^l$.

The contract retains the same contingent compensation and termination scheme each period without considering past actions. Given the complexities of implementing optimal equilibria under assumptions suited to lower-income settings—such as limited liability and unobservable output—I restrict attention to stationary equilibria.[41] I thus omit the time subscripts for simplicity. Empirically, the rental payment paid by drivers remains fixed over time for most pairs.

Assumption 4. ***Risk-neutrality***. *Both principal and agent are risk-neutral.*

This assumption maintains tractability of the analysis, as the theory literature has not extensively explored risk aversion within relational contracts.

The framework is a two-stage game solved using backward induction. In Stage 1 ("adoption"), *the agent* decides whether to adopt the new technology based on expected utility in Stage 2 ("impact"). In Stage 2, *the principal* offers a new contract, defined by a transfer schedule, $t(\tilde{y})$, and a continuation probability, $p(\tilde{y})$, subject to the agent's constraints. In line with practices in this industry, I assume that contracts are "take-it-or-leave-it" offers by owners, and both the owner's and driver's participation constraints must be satisfied for the relationship to form.

6.3 Baseline Contract Without Digital Payments

Figure 4 depicts the timing of events within a period of this dynamic game.[42] The principal seeks to maximize expected transfers and the future discounted value of the relationship. The objective functions of the principal V^θ when matched with agent of type θ can thus be written:

$$V^\theta = \max_{t,p,e} \mathbb{E}[t(\tilde{y}) + \delta[p(\tilde{y})V^\theta + (1-p(\tilde{y}))(-K_p + \mu V^h + (1-\mu)V^l)]|e] \tag{4}$$

This optimization is subject to the following constraints:

$$\begin{cases} \mathbb{E}[(y(e) - t(\tilde{y})) - \phi^\theta(e) + \delta U^\theta - \delta K_a(1-p(\tilde{y}))|e] \geq \max\{-\delta K_a + \delta U^\theta; \bar{u}\} & \text{Participation Constraint } (IR) \\ e \in \arg\max_{\tilde{e}\in\{0,1,2\}} \mathbb{E}[y(e) - t(\tilde{y}) + \delta U^\theta - \delta K_a(1-p(\tilde{y}))|\tilde{e}] - \phi^\theta(\tilde{e}) & \text{Incentive Compatibility } (IC) \\ t(\tilde{y}) \leq \tilde{y} \leq y(e) & \text{Limited Liability } (LL) \\ Y - t(Y) + \delta(U^\theta - K_a(1-p(Y))) \geq Y - t(X) + \delta(U^\theta - K_a(1-p(X))) & \text{Truth-Telling } (TT) \\ y(e) - t(\tilde{y}) + \delta(U^\theta - (1-p(\tilde{y}))K_a) \geq y(e) + \delta(U^\theta - K_a) & \text{Dynamic Enforceability } (DE) \end{cases}$$

The principal designs a contract that includes the driver's transfer t and the continuation probability p, while taking into account the eventual choice of effort e by the agent. Constraints include the participation (IR) and incentive compatibility (IC) that ensure both the agent's participation and appropriate effort level. The limited liability constraint prevents the agent from transferring

[41] Recent studies have examined optimal non-stationary relational contracts, e.g., Andrews and Barron (2016); Fong and Li (2017), but this is beyond the scope of this paper.

[42] This sub-section draws partly from Kelley et al. (2024). The key departures are: (1) I study a two-stage game and introduce agents' types θ to study selection and (2) I relax the role of risk-taking to focus on dynamics around effort.

and reporting more than the collected output. An immediate consequence is that the agent transfers the rental fee t at the end of the period, not upfront. The dynamic enforceability (DE) or non-reneging constraint ensures that the agent has dynamic incentives to come back to the owner at the end of the period, rather than terminating the relationship and leaving with the collected output. A parallel constraint applies to the owner. Finally, the truth-telling constraint ensures truthful reporting of collected output.[43] This latter constraint departs from standard principal-agent models, where output is typically observed by the principal.

Lemma 1 (in the Appendix) shows that, with full information, the owner would require the agent to exert their optimal level of effort, offering upfront fixed compensation that makes the agent indifferent between working and his outside option, with no termination occurring in equilibrium. I now derive the best stationary when effort and output are unobservable.

Result 1. (Baseline Contract Without Digital Payments) *Under Assumptions 1–4, $\exists\ \underline{K}_p < \bar{K}_p$, $\underline{\delta} < \bar{\delta}$, s.t. when $K_p \in (\underline{K}_p, \bar{K}_p)$ and $\delta \in (\underline{\delta}, \bar{\delta})$, the principal's best type-dependent stationary contract is:*

$$\bar{t}^\theta = \begin{pmatrix} t(Y) = R^\theta \\ t(X) = X \end{pmatrix} \quad \textit{and} \quad \bar{p}^\theta = \begin{pmatrix} p(Y) = 1 \\ p(X) = \bar{p}^\theta \end{pmatrix}$$

where $\bar{p}^\theta$ is the continuation probability for a low-output outcome and R^θ is the rental transfer for a high-output outcome for agent θ, with $\bar{p}^h > \bar{p}^l$, $R^h > R^l$. The agent induced effort is $e^l = e^h = 1 < 2$.

The contract parameters $(R^\theta, \bar{p}^\theta)$ ultimately depend on which of the incentive compatibility constraint (IC) or the truth-telling constraint on output (TT) binds for agent θ. See analytic derivations of $(R^\theta, \bar{p}^\theta)$ and proof in Appendix C.1. Intuitively, to incentivize the driver to exert effort *and* truthfully report output, the owner finds it best to retain the driver with probability 1, $p(Y) = 1$, when the reported output is high, as punishment is costly. Conversely, when the agent reports low output, the owner must terminate on path, with some probability $\bar{p}^\theta < 1$. This punishment is necessary due to the limited liability constraint *(LL)* (Assumption 2): following low output, the principal cannot extract money so the only way to discipline incentives requires inefficient punishment (Fuchs, 2007).[44] Limited liability also implies that the participation constraint is slack and the agent has an informational rent at baseline (see Appendix C.1).

The derived transfer schedule is in line with empirical Fact 1 to Fact 4, rationalizing the transfer from driver to owner, the possibility to default, and the high turnover rate in the taxi industry. Variation in outside options and limited liability among agents may explain the presence of upfront salaries at baseline, as some agents might require initial payments to start working.

[43] The limited liability constraints implies that the agent will never report more output than what is realized on path, meaning truth-telling for output X will always be verified on the equilibrium path.

[44] The principal's optimization implies that the agent's payoff should be minimized during low-output periods, resulting in termination with probability $\bar{p}^\theta$. To prevent the principal from reneging on the contract, I follow the literature—e.g., see Mailath and Samuelson (2006), Chapter 7—and assume that both players observe a public randomization device for p at the end of each period. The deviation, in which the principal does not follow through with the randomization device, is assumed to lead to an equilibrium where the agent misreports output and exerts no effort.

In summary, the baseline contract under limited liability is inefficient for two main reasons: (a) moral hazard in effort e and (b) in output reporting $\tilde{y}$. By using the transaction observability provided by digital payments, the principal can receive a signal of effort and output, thus reducing information frictions and potentially increasing the total surplus. I now examine this possibility.

6.4 Stage 2: Impact of Digital Observability

This section derives the comparative statics regarding the impact of digital payment technology on adopters by considering various information benchmarks, relaxing Assumption 1. For simplicity, I assume throughout the framework that digital payments offer no additional benefits to drivers (such as lower transaction costs from handling cash). However, these gains coming from the technology itself are incorporated in the structural estimation presented in Section 7. In addition, this framework is set in partial equilibrium, by assuming that the share of high-type agents, μ, remains constant before and after the introduction of the technology.[45]

I assume that digital payments provide the principal with *imperfect* information about their agent for two empirical reasons: (1) cash is still used since digitalization of payments is not complete, and (2) while digital transactions include timestamps and transaction values, these only partially reflect the agent's effort and output. The following result shows that such imperfect information can lead to contract changes, relying on the informativeness principle (Holmström, 1979). Let's define s the high-effort signal and $\kappa = P(s|e=2)$ and assume $P(s|e=1) = P(s|e=0) = 0$.

Result 2. (Imperfect Information on Effort) *Under Assumptions 1–4, when (IC) binds, for $K_p \in (\underline{K}_p, \bar{K}_p)$ and $\delta \in (\underline{\delta}, \bar{\delta})$, $\kappa > \bar{\kappa}$ for $\bar{\kappa} < 1$, and $\phi^\theta(2) < \tilde{\phi}\ \forall\ \theta$, the principal's best type-dependent stationary contract is:*

$$\bar{t}^\theta = \begin{pmatrix} t(Y) = R^\theta - W^\theta_{\tilde{e}=2} \\ t(X) = X - W^\theta_{\tilde{e}=2} \end{pmatrix} \quad \text{and}$$

$$p(\tilde{y}, s) = \begin{cases} 1 & \text{if } \tilde{y} = Y, \\ \bar{p}_{TT} & \text{if } \tilde{y} = X \text{ and } s \text{ is observed} \\ \bar{p}^{\theta\prime} < \bar{p}^\theta & \text{if } \tilde{y} = X \text{ and } s \text{ is not observed} \end{cases}$$

The agent θ induced effort is $e^\theta = 2$.

The principal can incentivize the agent to adopt the technology by offering an upfront payment $W^\theta_{\tilde{e}=2}$ each period and then induce the agent to exert high effort ($e = 2$). To do so, the principal increases the continuation probability during low-output periods, contingent on receiving the high-effort signal s; otherwise, the principal would terminate the agent with a higher probability than at baseline. In this relational contract framework, to mitigate concerns about reneging, the

[45] General equilibrium effects of the technology, such as shifts in market entry or exit among agents, are beyond the scope of this paper due to the experiment's time frame and empirical setting. Notably, the payment company removed observability as the default based on the research findings, which led to limited owner demand for observability.

principal incentivizes the agent to adopt the technology by offering the payment $W^{\theta}_{\tilde{e}=2}$ *upfront*, rather than reducing the target rental payment *ex-post*. See the formal proof in Appendix C.4.

Appendix Sections C.3, C.4, and C.5 derive various information benchmarks with comparative statics on the owner-driver relationship. Specifically, I formulate three additional lemmas. Lemma 2 and 3 show how (imperfect) information on output can benefit both the principal and the agent by relaxing the truth-telling constraint without revealing the agent's effort—following a similar logic as Result 2. I especially look at the impact of signaling low output to map to the experimental treatment *Coarse Observability*. Lemma 4 demonstrates that the agent has limited incentives to manipulate either the output or the effort information, as such manipulation would result in no contract changes or may lead to termination.

6.5 Stage 1: Differential Technology Adoption

Stage 1 examines the adoption decision by different types of agents. In line with the experiment's setting, I assume that the agent (driver) ultimately holds the *decision-making power* to adopt the technology and can keep the technology with them upon termination. Alternatively, this framework may be extended to the other perspective: the principal could adopt the technology and screen for drivers willing to use it, screening out others and achieving a similar separation.

I incorporate the results derived from *Stage 2* to formulate the following result:

Result 3. (Differential Adoption) *Under Assumptions 1–4, $\exists\, \underline{K}_p < \bar{K_p}, \underline{\delta} < \bar{\delta}^{tech}$ and $\bar{\phi^h} < \bar{\phi^l}$ s.t. if $K_p \in (\underline{K}_p, \bar{K_p})$ and $\delta \in (\underline{\delta}, \bar{\delta}^{tech})$, $\phi^l(2) > \bar{\phi^l}$, and $\phi^h(2) < \bar{\phi^h}$, then only high-ability agents adopt the technology, while low-ability agents opt not to adopt it.*

The intuition is as follows: (i) High-type agents benefit from adopting the technology because the new contract allows them to earn more to compensate higher effort, face a lower probability of termination, and keep them indifferent if $\phi^h(2) < \bar{\phi^h}$, using Result 2. (ii) Low-type agents choose not to adopt the technology because their disutility from exerting higher effort, $\phi^l(2)$, is too high. It would be unprofitable for the principal to fully compensate them to exert $e = 2$. Without formal commitment, the principal cannot credibly promise not to demand high effort once they access the information, and would fire the agent if the high-effort signal is not observed. Thus, a low-type agent does not adopt and his welfare remains unchanged. The formal proof is in Appendix C.6.

Both types of agents co-exist in equilibrium for a sufficiently high share of low-types or low discount rate. I specify the "no-deviation" condition C19 on $\delta < \bar{\delta}^{tech}$, which states that the owner should have no incentive to deviate by terminating the low-type agents, incur the replacement cost K_p, and recruit a new agent, with probability μ of being matched with a high-type who accepts the technology.

6.6 Comparative Statics: Impact and Adoption of Digital Payments

The framework's predictions are consistent with the experimental outcomes, justifying the modeling assumptions. It formalizes the mechanisms behind the findings and suggests the model is suitable for the structural estimation exercise that follows.

1. Observability Effect on Contract for Adopters

1a *Effort Effect* $e \uparrow$: Digital payments allow the principal to observe the driver's effort more effectively and incentivize higher effort. This, in turn, reduces the default rate and enables the principal to extract a higher average transfer from the agent.

1b *Contract Effect* $W_{\tilde{e}=2}$: The principal compensates the agent for increased effort using an upfront payment $W_{\tilde{e}=2}$.

1c *Worker Retention Effect* $\bar{p}^{\theta} \uparrow$: Imperfect information on output and effort from digital payments reduces moral hazard in effort and in output reporting, increasing the continuation probability in low-output periods.

2. Observability Effect on Adoption and Differential Adoption

2a *Characterization of Low-Types* θ^{l}: Low-type agents can be identified by comparing agents willing to adopt to the ones unwilling to adopt, focusing on characteristics related to higher disutility of work.

2b *Technology Adoption*: Observability and subsequent contractual changes requiring higher effort create a barrier to technology adoption for low-type agents.

Beyond the taxi owner-driver relationship, this framework elucidates the critical role of information asymmetries in shaping many employer-employee relationships in lower-income contexts, where limited liability and weak contract enforcement, as formalized within the framework, play a central role. I use the framework combined with the reduced-form results to quantify the distributional consequences of digital payments and estimate policy counterfactuals in Section 7.

7 Structural Estimation and Welfare Impacts of Digital Payments

I combine the framework with the reduced-form estimates to examine the welfare impacts of digital technologies in this context, focusing on two key objectives.

First, I quantify the welfare effects of digital payments, both the cost savings from reducing cash payments and the "observability effect" introduced by digital transactions. To achieve this, I estimate the disutility of effort for each type of driver and calculate the values of the relationship for the owner (V) and each driver type (U^{θ}).

Second, I estimate a series of counterfactuals to explore different policy and design choices: (1) mandating digital payment adoption by drivers, (2) redesigning the technology to remove the "observability" feature—now implemented as the default option by the payment company, and (3) the full-information benchmark. These counterfactuals allow me to quantify the magnitude of various frictions affecting owners and drivers, providing insights into how information asymmetries shape relationships and outcomes in informal labor markets in lower-income countries. This structural analysis offers a framework to help managers, policymakers, and innovators to navigate trade-offs in information disclosure and guide decision-making. I conclude by discussing potential policy implications.

7.1 Inputs Calibration

I apply the framework by calibrating some parameters directly from the survey data and estimating the remaining ones using the generalized method of moments (GMM). I use key moments from both the experimental variation and the survey. I summarize the framework inputs below, with further details in Appendix D.1 (see in particular Table B29).

From the Survey Data. The survey data collection was carefully designed to provide the necessary components for estimating key parameters and conducting welfare analysis. First, I use the baseline survey data relating working hours to output to estimate the simple binary production function with the probabilities of high and low outputs. To follow the simple two-output, three-effort framework, I define high and low output, Y and X, as the average output above or below median output when working more than the median hours, respectively, calibrated from the groups where the information frictions remain unchanged—*Control* and *No-Observability*—see Figure A12. The proportion of high-type drivers μ and the owner's replacement cost K_p are calibrated using owners' survey responses. I allow the outside options for the two types of drivers to vary, and I calibrate their values relying on a representative survey of small vendors I conducted in September 2022—a common outside option for drivers. I then use the difference in the wealth index to calibrate the two disutilities of work, $\bar{u}^h$ and $\bar{u}^l$. Robustness checks are conducted by (1) varying model inputs and computing bootstrapped standard errors, and (2) assessing the sensitivity of parameter estimates to each moment, following Andrews et al. (2017) as detailed in Appendix D.4.

From the Reduced-Form Estimates. I use reduced-form estimates to calibrate the following parameters: the driver's gains from reducing cash-related costs, G, for drivers with technology access (pooling across observability treatments); the production function for drivers under *Granular Observability*; and the contract characteristics under *Granular Observability*, in particular the observed upfront payment $W_{\tilde{e}=2}$ and the retention probability $p_{\tilde{e}=2}$.

From the Literature. I calibrate the discount rate δ using the closest relevant setting I could find: Yesuf and Bluffstone (2019) in Ethiopia, which reports a weekly discount rate of 0.99.

7.2 Estimation Procedure

In the theoretical framework in Section 6, I derive the best stationary contracts under different information benchmarks, from no observability to full-information. Using the framework derivations, I estimate key parameters to quantify the owner's and driver's contract valuations under various information benchmarks in order to quantify the technology's impact on welfare and study counterfactuals.

Identification. I estimate three key unknown parameters. Using the baseline data, I estimate the disutility of work of both types, $\phi^l(1)$ and $\phi^h(1)$; and using the mid-term data, I estimate the disutility of work of high-types with high-effort $e = 2$, $\phi^h(2)$, under the *Granular Observability* treatment arm. These three parameters are identified using the following eight target moments: the rehiring rate under three different experimental arms (reluctant drivers, control group, and granular observability)—intuitively, if the job becomes tougher, then the continuation probability in a low-output period would need to be reduced to incentivize effort without observability—, the perceived replacement cost, the driver's valuation of the contract, the target transfers for both types, and the upfront payment offered under *Granular Observability*. The model is thus over-identified. These empirical moments are based on the collected data regarding the production environment and contract characteristics. For each moment, I provide the theoretical formula, describe the corresponding empirical moment, and provide the intuition in Appendix D.2.

Parameter Estimation. I estimate the parameters of interest using a generalized method of moments (GMM) approach, which minimizes the distance between the structural and reduced-form components. My data $\mathbf{X}_i$ comprises eight empirical moments. I minimize the difference between the vector of empirical moments ($\bar{p}^l, \bar{p}^h, \tilde{p}_{\tilde{e}=2}, U^h, K, R^l, R^h, W_{\tilde{e}=2}$) and the vector of structural moments. The weighting matrix $\mathbf{W}$ consists of the inverse variance of the estimation moments. Detailed derivations are provided in Appendix D. To test the sensitivity of the estimation to the framework inputs, I estimate the standard errors for each parameter using a bootstrap procedure, resampling with 1,000 bootstrap replications of the survey data.

7.3 Parameter Estimates and Owner-Driver Contract Valuations

Table 10 summarize the estimation results and the matched moments. In Panel A, I find that the structural moments match the moments in the data reasonably well, e.g., the low- and high-type baseline continuation probability are both matched within a 1-2 percentage point difference, and the driver's replacement cost K_a, the target transfers R^h and R^l, and the upfront payment of adopters under observability $W_{\tilde{e}=2}$ match almost perfectly. Note that the driver's replacement cost is estimated to be \$445, which corresponds to approximately 33 days of lost profit. Although I did not collect contract valuations specifically for low-type drivers, I compare the structural moment with valuations from drivers willing to adopt (in the impact experiment), but who would have preferred not to have *Granular Observability*, and this comparison yields a close match.

I report the GMM estimates of the disutilities of work for both types of drivers. In Panel B, I estimate the parameters of interest using GMM and find that the weekly disutility of effort for low-type drivers is estimated to be \$21, while it is lower for high-type drivers (\$15). Lastly, the disutility of work for high-type drivers at $e = 2$ is estimated to be $\phi^h(2) =$ \$28, which is higher than their baseline disutility at $e = 1$. In Panel C, I compute the counterfactual lower bound for the driver's disutility of effort for $e = 2$, $\phi^l(2)$, had they adopted the technology—such that compensating for the utility loss upfront would not be profitable for the owner. This value is approximately \$39, as expected higher than $\phi^h(2)$.

I use these estimates to compute the owner's contract valuation and the total welfare from the contract at baseline and with the digital payment technology, assuming that the social planner maximizes a social welfare function simply equals to sum of the owner's and the driver's welfare (equal weight). At baseline, without the technology, the average owner's present-discounted contract value is approximately \$6,719,[46] and the average driver's present-discounted contract value is \$3,924, representing about 37% of the total welfare. Specifically, the high-type driver's contract value is \$4,274, while the low-type driver's contract value is lower (\$3,674) due to their higher disutility of effort. These contract valuation estimates broadly align with the emerging literature estimating the value of the contractual relationship between principal and agent in lower-income contexts. For instance, Kelley et al. (2024) estimate the value of the relationship to be between \$1,794 and \$2,753 on average for a minibus owner in Kenya, while the driver's valuation is \$507. In a different setting, the rose market in Kenya, Macchiavello and Morjaria (2015) estimate higher valuations, \$13,872 and \$22,127 for the seller and the buyer, respectively.[47]

7.4 Welfare and Distributional Consequences Under Different Scenarios

7.4.1 Without Policy Intervention

I begin by analyzing the status quo in the absence of social planner intervention, where only high-type drivers adopt the technology. Figure 5(a) plots contract valuations for two types of pairs: owners matched with high- and low-type drivers. It illustrates the outcomes with and without the digital payment technology. Since low-type drivers have no incentive to adopt the technology, their welfare remains unchanged. This analysis focuses on production-side welfare—the welfare of both owners and drivers—assigning equal weight to each and excluding consumer welfare, likely leading to an underestimate of the full technological welfare gains.

Throughout this structural estimation, I assume that the principal does not capture the direct reduced cash-related costs from accessing digital payments. This assumption is consistent with empirical evidence: as shown in Section 5.3, while the technology without the observability component provided gains to drivers, these gains did not translate into contract changes between owners and drivers (comparing the *No Observability* group to the control). The technology thus

[46] The owner's valuations for high-type and low-type driver contracts are nearly identical in this estimation, though this doesn't have to be the case. This similarity arises because the outside option for low-type drivers is lower such that the owner optimally offers a comparable baseline contract to both types.

[47] Kelley et al. (2024) assume a lower discount rate (0.99 daily, where a dollar today is worth 2 cents in a year, compared to 0.59 using 0.99 weekly). I use a 0.99 weekly discount rate, based on a recent estimate from Ethiopia (Yesuf and Bluffstone, 2019). Once adjusted for the difference in δ, the relationship values between owners and drivers remain comparable across both studies. In Macchiavello and Morjaria (2015), the relationship value is given by the maximum temptation to deviate between the Kenyan rose seller and the Dutch buyers, computed using the replication data from its Table 1.

increases the welfare of high-type drivers by reducing the costs associated with cash transactions, as quantified in Section 5.2. Additionally, it generates efficiency gains through reductions in moral hazard that are captured by owners, given the structure of the model. Overall, 68% of the technological gains (reduced cash-related cost + moral hazard) accrue to high-type drivers.

Without social planner intervention, the introduction of the technology exacerbates welfare inequality between high- and low-type businesses, as only high-type pairs benefit from its adoption. Figure 6, Panel 2, displays the total welfare for owners matched with high- and low-type drivers, respectively. The top of each bar represents the overall welfare increase in the economy, accounting for the distribution of high- and low-type drivers as reported by the owners. The total welfare in the economy increases by 0.5% following the adoption of the technology by high-type pairs, with equal weights assigned to owners, high-type drivers, and low-type drivers.

7.4.2 Counterfactual 1: Mandating Digital Payment Adoption

I consider a counterfactual in which the social planner observes that only high-type drivers adopt the technology and thus decides to mandate the adoption of digital payment technologies for all drivers. With governments worldwide contemplating the enforcement of the adoption of some digital technologies, this counterfactual examines the potential impact of such policies.[48]

I find that this mandate has significant welfare and distributional consequences—see Figure 5(b). For high-type pairs, the mandate redistributes gains from drivers to owners. By requiring drivers to adopt the technology, the mandate eliminates the need for owners to offer compensation for adoption: the high-type drivers can only retain the lower cash-related costs.[49]

The mandate substantially reduces the welfare of low-type drivers by 12%. These drivers are induced to exert higher effort upon adoption given the particular high-effort signal this technology provides, as owners may otherwise terminate the relationship. The gains from reduced cash-related costs are insufficient to offset the increased disutility from higher effort. In contrast, the owner's welfare in low-type matches increases by 4%, benefiting from reduced moral hazard in effort.

At the aggregate level, the mandate further exacerbates welfare inequality between and within high- and low-type businesses (see Figure 6, Panel 3). The overall welfare gain under the mandate is lower than in the no-intervention scenario (0.2% compared to 0.5%). This result suggests that, although the mandate might seem beneficial to increase adoption, it may have adverse welfare and distributional consequences, by inducing low-ability drivers to work more than optimal. The high-effort signal would thus benefit owners but not the drivers.

[48]Several African governments, such as those in Ghana and Nigeria, have explored mandates and restrictions on cash payments to promote digital economies.

[49]If I relax the assumption that high-type drivers retain the technological gains, the result—redistribution of gains from drivers to owners—is even stronger.

7.4.3 Counterfactual 2: Redesigning the Technology to Remove Observability

This counterfactual explores the impact of redesigning the technology to remove its observability feature. This scenario is particularly relevant for at least two reasons. First, it investigates the effect of the company's decision to implement a version of the technology without observability as the default option in the taxi industry, based on the study's findings.[50] Second, many technologies, beyond payment systems, default to observability but can often be redesigned to exclude it.

I find that removing observability fundamentally changes the welfare effects of the technology—see Figure 5(c). With this redesign, both low- and high-type drivers can now adopt the technology and benefit from reduced cash-related costs, thereby shifting the production possibility frontier upward. In this scenario, all welfare gains accrue to the drivers, who retain the full informational rent, while owners see no direct benefit. For example, the welfare share for a low-type driver and their employer shifts from 35% to 37%, thus reducing inequality within firms. However, this technology design choice introduces an efficiency trade-off for high-type businesses: total welfare gains for high-type pairs are smaller compared to the baseline scenario without policy intervention discussed in Section 7.4.1, with the possibility frontier lower than in the status quo with observability.

At the aggregate level, this policy significantly reduces welfare inequality between high- and low-type businesses, as all drivers can now adopt the technology (see Figure 6, Panel 4). The company chose to implement this version of the technology primarily to increase driver access, prioritizing adoption over efficiency coming from reduced moral hazard. Moreover, the overall welfare gain is greater than in the no-intervention scenario (0.7% compared to 0.5%), providing a possible justification for implementation.

7.4.4 Counterfactual 3: Full-Information Benchmark

I examine the welfare implications of the full-information benchmark. This scenario sheds light on the extent to which moral hazard impacts welfare in this economy and the distributional effects of removing it by leveraging digital technologies. Specifically, I consider a hypothetical scenario where the technology provides the same cost-saving benefits (i.e., reduced cash-related costs) and is universally adopted, but now fully reveals the agent's effort level, not just high effort as in the previous cases.[51] This counterfactual can also be interpreted as a mandate requiring digital technology that provides employers full information on effort. Under full information, the best stationary equilibrium is wage employment, as outlined in Lemma 1, where the owner compensates the agent just enough to cover their disutility of effort and their outside option.

Transitioning to the full-information benchmark—essentially wage employment—leads to an increase in owner welfare by 21%, and is the most efficient counterfactual compared to all other

[50] I assume that removing observability leads all drivers to adopt the technology in this counterfactual. Empirically, 88% of previously reluctant drivers adopted the technology after the payment company made non-observability the default post-experiment.

[51] A full-information benchmark with universal adoption but no cost-saving benefits would yield qualitatively similar results.

ones. However, drivers experience a significant welfare reduction, as they lose their informational rents entirely. Figure 5(d) illustrates this trade-off: driver welfare decreases by 18% for high-type drivers and 20% for low-type drivers, while overall welfare rises by 2.6% (Figure 6, Panel 5). This outcome highlights that although the full-information scenario pushes the production possibility frontier upward, it is not Pareto improving, thus possibly justifying policy interventions as these technologies become more widely adopted. Particularly in settings with credit constraints, employers may not be able to compensate some employees for the loss of informational rent needed to incentivize the adoption of welfare-enhancing technologies.[52]

7.5 Discussion: Trade-off Between Observability and Adoption

The structural estimation, together with the experimental results, reveals important policy implications. While both principals and agents can benefit from the technology—through reduced cash-related cost and lower information frictions—low-type agents are not adopting it. This exacerbates welfare inequality between high- and low-type workers and, within the framework, reduces aggregate welfare compared to a scenario without observability. To reduce this gap and increase both technology access and overall welfare, a social planner or technology designer with reasonable welfare weights on low-type businesses may thus be better off limiting the observability embedded in digital technologies. This would prevent information from being used against low-type agents, encouraging adoption and increasing overall welfare.

As a direct implication of these findings, the partner mobile money company chose to default to *No Observability* of digital transactions to owners (Counterfactual 2) to broaden technology access in the taxi industry and enable the low-type drivers to adopt, precisely due to the considerations highlighted in this structural estimation. By early 2024, the technology is widely used by over 16,000 taxi drivers in Dakar, representing about 75% of the industry, according to my best estimates, and the company has expanded it to a new market: Côte d'Ivoire. The company maintained observability in sectors where within-firm relationships made the adoption less problematic, such as supermarkets—typically using formal contracts—where managers now use digital transaction observability to monitor cashiers.

8 Conclusion

This study investigates the relationship between technology adoption and within-firm contracts in a lower-income setting. The global proliferation of digital technologies has drawn considerable interest from policymakers and the private sector due to their potential to enhance firm productivity and firm growth. Academic research examining their influence on private sector development and intra-firm organization is key to inform this discussion.

[52]This analysis assumes risk neutrality, as the theory literature has not extensively explored risk aversion within relational contracts. Conclusions might differ quantitatively since risk-averse agents may particularly value salaried employment due to reduced income volatility. I leave this theoretical and empirical consideration for future research.

To gain insights into the effects of digital technologies like payments on businesses, I conducted two randomized experiments in Senegal's taxi industry spanning nearly two years in partnership with the country's largest payment company. Relying on contract theory, the experiments aim to assess the influence of digital payments on employee selection and behavior, by exploring how the enhanced observability embedded in digital payments reduces information frictions and affects contracts within firms.

The study has four key findings. First, digital payment technologies benefit businesses by significantly reducing the costs associated with using cash, enhancing security, and improving earnings tracking. Second, business owners leverage digital payments as a monitoring tool, enabling contract changes and increased employee effort. As a result, these owner-driver relationships last significantly longer, and owners' trust in their drivers increases. Third, digital observability acts as a barrier to adoption for low-ability drivers, with differential adoption among workers. Fourth, the technology increases overall welfare by providing employers with better information on employee actions, although it exacerbates welfare inequality between adopters and non-adopters.

Taken together, these findings show how digital technologies can expand the production possibility frontier and increase total surplus. However, they may not be adopted in the first place, suggesting the need for policy interventions. These insights may extend to informal sectors often characterized by weak contract enforcement and limited liability, where monitoring technologies are often not adopted, or to contexts where digital technologies raise data privacy concerns.

Three features of this paper—the dual randomization of technology access and observability features, the comprehensive two-year panel data on employers and employees in informal firms, and the analysis of the interplay between technology adoption and within-firm interactions grounded in economic theory—aim to contribute to the literature that studies organizations in developing contexts. This study suggests the importance of further investigating how technology design shapes organizational structures within and across firms at various stages of economic development. Understanding these dynamics is an exciting avenue for future research.

References

Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey M Wooldridge, "When Should You Adjust Standard Errors for Clustering?*," *The Quarterly Journal of Economics*, 10 2022, *138* (1), 1–35.

Agarwal, Sumit, Wenlan Qian, Bernard Y. Yeung, and Xin Zou, "Mobile Wallet and Entrepreneurial Growth," *AEA Papers and Proceedings*, May 2019, *109*, 48–53.

Aker, Jenny, McClelland Amanda Boumnijel Rachid, and Niall Tierney, "Payment Mechanisms and Anti-Poverty Programs: Evidence from a Mobile Money Cash Transfer Experiment in Niger," *Economic Development and Cultural Change*, 2016, *65* (1), 1–37.

Alvarez, Fernando and David Argente, "On the Effects of the Availability of Means of Payments: The Case of Uber.," *The Quarterly Journal of Economics*, 2022, *137* (3), 1737–1789.

Andrews, Isaiah and Daniel Barron, "The Allocation of Future Business: Dynamic Relational Contracts with Multiple Agents," *American Economic Review*, September 2016, *106* (9), 2742–59.

_ , Matthew Gentzkow, and Jesse M. Shapiro, "Measuring the Sensitivity of Parameter Estimates to Estimation Moments*," *The Quarterly Journal of Economics*, 06 2017, *132* (4), 1553–1592.

Angrist, Joshua D., Sydnee Caldwell, and Jonathan V. Hall, "Uber versus Taxi: A Driver's Eye View," *American Economic Journal: Applied Economics*, July 2021, *13* (3), 272–308.

Atkin, David, Amit K. Khandelwal, and Adam Osman, "Exporting and Firm Performance: Evidence from a Randomized Experiment*," *The Quarterly Journal of Economics*, 02 2017, *132* (2), 551–615.

_ , Azam Chaudhry, Shamyla Chaudry, Amit Khandelwal, and Eric Verhoogen, "Organizational Barriers to Technology Adoption: Evidence from Soccer-Ball Producers in Pakistan," *The Quarterly Journal of Economics*, 2017, *132* (3).

Bai, Jie, "Melons as Lemons: Asymmetric Information, Consumer Learning, and Seller Reputation," *Review of Economic Studies*, 2024.

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

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

_ , Robert Gibbons, and Kevin J. Murphy, "Relational Contracts and the Theory of the Firm," *The Quarterly Journal of Economics*, 2002, *117* (1), 39–84.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul, "Social Preferences and the Response to Incentives: Evidence from Personnel Data," *Quarterly Journal of Economics*, 2005, *120* (3), 917–962.

— , — , and — , "Incentives for Managers and Inequality among Workers: Evidence from a Firm-Level Experiment," *Quarterly Journal of Economics*, 2007, *122* (2), 729–773.

— , — , and — , "Social Connections and Incentives in the Workplace: Evidence from Personnel Data," *Econometrica*, 2009, *77* (4), 1047–1094.

Banerjee, Abhijit V., Paul J. Gertler, and Maitreesh Ghatak, "Empowerment and Efficiency: Tenancy Reform in West Bengal," *Journal of Political Economy*, 2002, *110* (2), 239–280.

Beaman, Lori, Jeremy Magruder, and Jonathan Robinson, "Minding Small Change among Small Firms in Kenya.," *Journal of Development Economics*, 2014, *108*, 69–86.

Becker, GM, MH DeGroot, and J Marschak, "Measuring utility by a single-response sequential method," *Behav Sci*, 1964, *9* (3), 226–32.

Bergquist, Lauren Falcao and Michael Dinerstein, "Competition and Entry in Agricultural Markets: Experimental Evidence from Kenya," *American Economic Review*, December 2020, *110* (12), 3705–47.

Bertrand, Marianne and Antoinette Schoar, "The Role of Family in Family Firms.," *Journal of Economic Perspectives*, 2006, *20* (2), 73–96.

Blader, Steven, Claudine Gartenberg, and Andrea Prat, "The Contingent Effect of Management Practices," *The Review of Economic Studies*, 07 2019, *87* (2), 721–749.

Boudreau, Laura, "Multinational enforcement of labor law: Experimental evidence on strengthening occupational safety and health (OSH) committees," *Econometrica*, 2024, *92* (4), 1269–1308.

Brockmeyer, Anne and Magaly Sáenz Somarriba, "Electronic Payment Technology and Tax Compliance: Evidence from Uruguay's Financial Inclusion Reform," *American Economic Journal: Economic Policy*, 2024, (9947).

Burchardi, Konrad B, Selim Gulesci, Benedetta Lerva, and Munshi Sulaiman, "Moral Hazard: Experimental Evidence from Tenancy Contracts*," *The Quarterly Journal of Economics*, 09 2018, *134* (1), 281–347.

Chandrasekhar, Arun G, Melanie Morten, and Alessandra Peter, "Network-Based Hiring: Local Benefits; Global Costs," February 2020.

Crouzet, Nicolas, Apoorv Gupta, and Filippo Mezzanotti, "Shocks and Technology Adoption: Evidence from Electronic Payment Systems," *Journal of Political Economy*, 2023.

Dalton, Patricio, Haki Pamuk, Ravindra Ramrattan, Burak Uras, and Daan van Soest, "Electronic Payment Technology and Business Finance: A Randomized, Controlled Trial with Mobile Money," *Management Science*, 2023.

de Janvry, Alain and Elisabeth Sadoulet, "Optimal share contracts with moral hazard on effort and in output reporting: managing the double Laffer curve effect," *Oxford Economic Papers*, 04 2007, *59* (2), 253–274.

de Mel, Suresh, David J. McKenzie, and Christopher Woodruff, "Returns to Capital in Microenterprises: Evidence from a Field Experiment," *Quarterly Journal of Economics*, 2008, *123* (4), 1329–1372.

de Rochambeau, "Monitoring and Intrinsic Motivation: Evidence from Liberia's Trucking Firms," *Working Paper*, 2021.

Derksen, Laura, Anita McGahan, , and Leandro Pongeluppe, "Privacy at What Cost? Saving the Lives of HIV Patients With Electronic Medical Records," *Working Paper*, 2024.

Dizon-Ross, Rebecca and Seema Jayachandran, "Improving Willingness-to-Pay Elicitation by Including a Benchmark Good.," *AEA Papers and Proceedings*, 2022, *112* (551-55).

Fajnzylber, Pablo, William Maloney, and Gabriel Montes Rojas, "Microenterprise Dynamics in Developing Countries: How Similar are They to Those in the Industrialized World? Evidence from Mexico," *World Bank Economic Review*, 2006, *20* (3), 389–419.

Fong, Yuk-Fai and Jin Li, "Relational Contracts, Limited Liability, and Employment Dynamics," *Journal of Economic Theory*, 2017, *169*, 270–293.

Fuchs, William, "Contracting with Repeated Moral Hazard and Private Evaluations," *American Economic Review*, September 2007, *97* (4), 1432–1448.

Ghatak, Maitreesh and Priyanka Pandey, "Contract choice in agriculture with joint moral hazard in effort and risk.," *Journal of Development Economics*, 2000, *63* (2), 303–326.

Goldfarb, Avi and Catherine Tucker, "Privacy and Innovation," *Innovation Policy and the Economy*, 2012, *12* (12).

Goldsmith-Pinkham, Paul, Peter Hull, and Michal Kolesár, "Contamination Bias in Linear Regressions," *NBER Working Paper*, 2024, (30108).

Higgins, Sean, "Financial Technology Adoption: Network Externalities of Cashless Payments in Mexico," *American Economic Review*, November 2024, *114* (11), 3469–3512.

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

Holt, Charles A and Susan K Laury, "Risk aversion and incentive effects," *American Economic Review*, 2002, *92* (5), 1644–1655.

Houeix, Deivy, "Nationwide Diffusion of Technology Within Firms' Social Networks.," *Working Paper*, 2024.

Hubbard, Thomas, "The Demand for Monitoring Technologies: The Case of Trucking*," *The Quarterly Journal of Economics*, 05 2000, *115* (2), 533–560.

— , "Contractual Form and Market Thickness in Trucking," *RAND Journal of Economics*, 2001, *32* (2), 369–386.

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

Innes, Robert, "Limited liability and incentive contracting with ex-ante action choices," *Journal of Economic Theory*, 1990.

Jack, William and Tavneet Suri, "Risk Sharing and Transactions Costs: Evidence from Kenya's Mobile Money Revolution," *American Economic Review*, January 2014, *104* (1), 183–223.

— **and** — , "The Long-run Poverty and Gender Impacts of Mobile Money," *Science*, 2016, *354* (6317), 1288–1292.

Karlan, Dean and Jonathan Zinman, "Observing Unobservables: Identifying Information Asymmetries with a Consumer Credit Field Experiment," *Econometrica*, 2009, *77* (6), 1993–2008.

Kelley, Erin, Gregory Lane, and David Schönholzer, "Monitoring in Small Firms: Experimental Evidence from Kenyan Public Transit," *American Economic Review (Forthcoming)*, 2024.

Lacker, Jeffrey M. and John A. Weinberg, "Optimal Contracts under Costly State Falsification," *Journal of Political Economy*, 1989, *97* (6), 1345–1363.

Lazear, Edward P., "Performance Pay and Productivity," *American Economic Review*, December 2000, *90* (5), 1346–1361.

Levin, Jonathan., "Relational Incentive Contracts," *American Economic Review*, 2003, *93* (3), 835–857.

Macchiavello, Rocco and Ameet Morjaria, "The Value of Relationships: Evidence from a Supply Shock to Kenyan Rose Exports.," *American Economic Review*, 2015, *105* (9).

— **and** — , "Competition and Relational Contracts in the Rwanda Coffee Chain.," *The Quarterly Journal of Economics.*, 2021, *136* (2).

— **and** — , "Relational Contracts: Recent Empirical Advancements and Open Questions.," *Working Paper*, 2023.

Mailath, George J. and Larry Samuelson, *Repeated Games and Reputations: Long-Run Relationships*, Oxford University Press, 09 2006.

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

Mel, Suresh De, David McKenzie, and Christopher Woodruff, "Measuring Microenterprise Profits: Must We Ask How the Sausage Is Made?," *Journal of Development Economics*, 2009, *88:19-31.*

Nagler, Paula and Wim Naudé, "Non-farm entrepreneurship in rural sub-Saharan Africa: New empirical evidences," *Food Policy*, 2017, *67*, 175–191.

Riley, Emma, "Resisting Social Pressure in the Household Using Mobile Money: Experimental Evidence on Microenterprise Investment in Uganda," *American Economic Review*, May 2024, *114* (5), 1415–47.

Shavell, Steven, "Risk Sharing and Incentives in the Principal and Agent Relationship.," *The Bell Journal of Economics*, 1979, *10* (1), 55–73.

Shetty, Sudhir, "Limited Liability, wealth differences and tenancy contracts in agragian economies," *Journal of Development Economics*, 1988, *29*, 1–22.

Suri, Tavneet and Chris Udry, "Agricultural Technology in Africa," *VoxDev Vol 5*, *Issue 2*, March 2024.

Townsend, Robert M, "Optimal contracts and competitive markets with costly state verification," *Journal of Economic Theory*, 1979, *21*, 265 – 293.

World Bank, *Digital Africa: Technological Transformation for Jobs*, Washington, DC: World Bank, 2023.

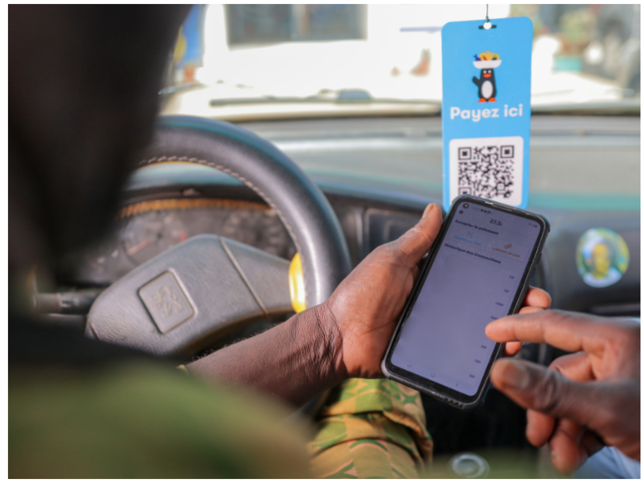
— , *Digital Opportunities in African Businesses*, Washington, DC: World Bank, IFC, 2024.

World Values Survey, Wave 7, *World Values Survey: Round Seven – Country-Pooled Datafile Version 6.0* 2017.

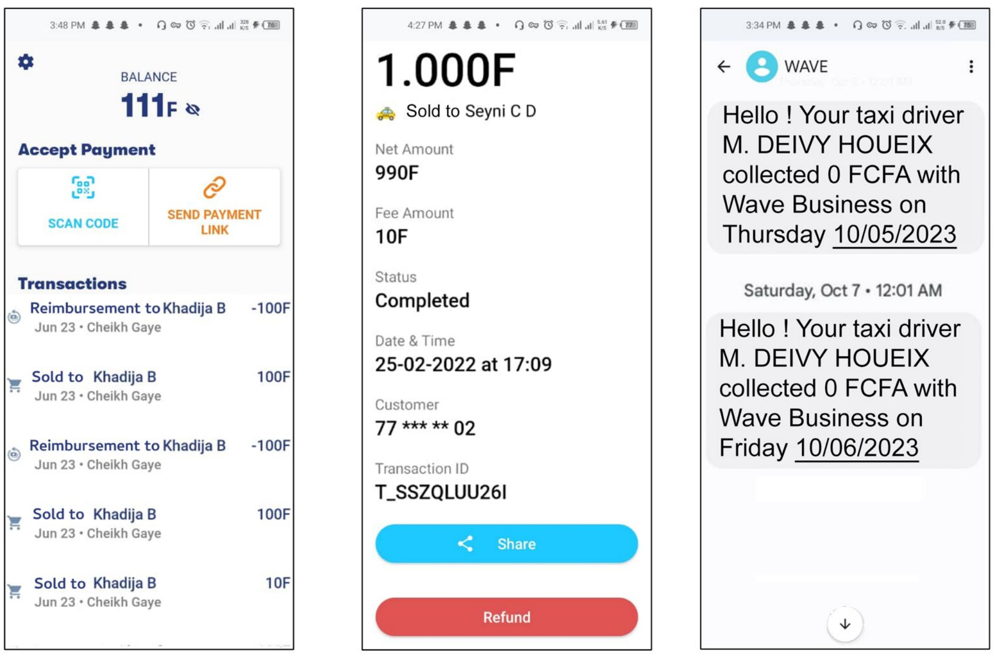
Yesuf, Mahmud and Randall Bluffstone, "Consumption Discount Rates, Risk Aversion and Wealth in Low-Income Countries: Evidence from a Field Experiment in Rural Ethiopia," *Journal of African Economies*, 06 2019, *28* (1), 18–38.

Figures and Tables

(a) Taxi Driver with Technology Access



(b) Digital Payment Interface for Drivers



(c) Owner's Observability of Driver Transactions and Daily SMS Updates

Figure 1: Digital Payments and Transaction Observability for the Taxi Industry in Sénégal

Notes: The first two pictures 1(a) and 1(b) illustrate the implementation of the new peer-to-business (P2B) technology that I designed with Wave Mobile Money for the taxi industry in Senegal. The technology utilizes a QR code visible from outside the vehicle—hung from the rear-view mirror—allowing potential customers to easily identify it. Drivers are provided with a sticker to place on their windshield, enhancing the visibility of the product. Additionally, drivers were equipped with the corporate mobile application and trained on how to use it by a dedicated mobile money agent. Owners under the randomized observability treatment were provided with a similar digital app with their driver's transaction history. Passengers can pay using their smartphones, or, if they don't have smartphones, drivers can scan their mobile money cards.
Figure 1(c) represents screenshots of a typical digital payment app, from the owner app (test account). In the first picture, one owner can view the driver's remaining balance, the transaction history including the day of the transaction, the name of the customer, and the amount of money collected. The second picture shows what appears when one clicks on the transaction in picture 1. In addition to the previous information, the owner can see the exact time of the transaction. Finally, the third picture shows the typical SMS sent to the owner at midnight each day. My name (Deivy Houeix) appears in these screenshots, as I coordinated the fieldwork and received daily SMS updates on the account initially used to pilot and fine-tune the technology. These screenshots were translated from French to English.

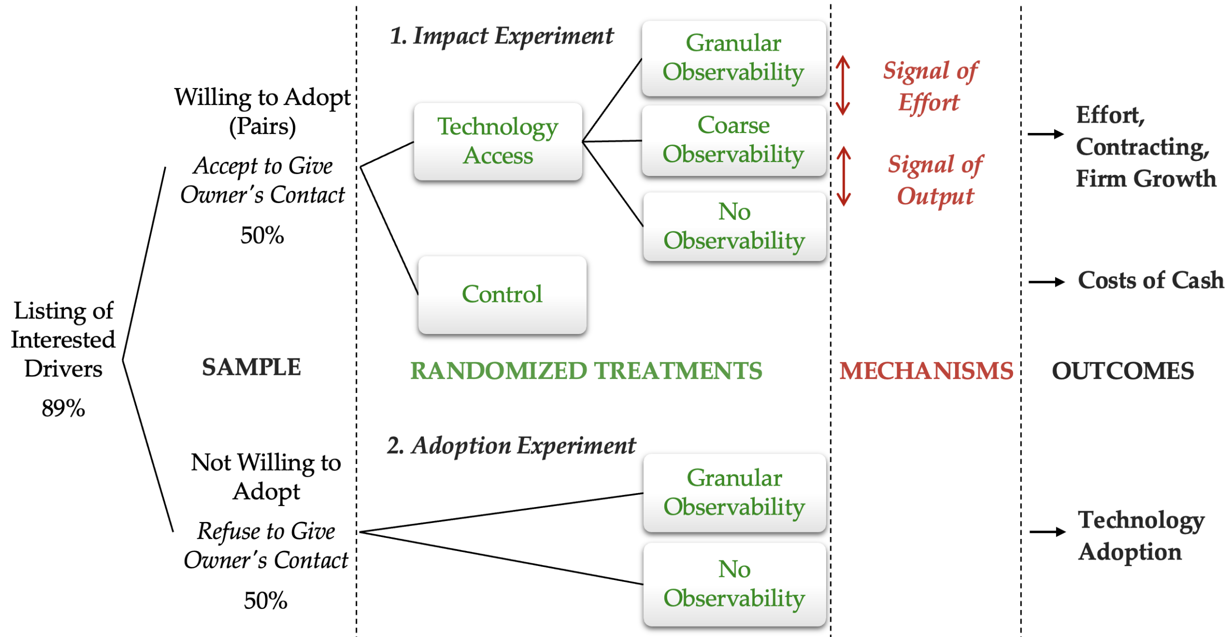


Figure 2: Experimental Design: Impact and Adoption Experiments

Notes: From the listing activity, where a majority of drivers (89%) expressed interest in adopting the technology, two distinct groups emerged. The experimental design includes two approaches for each of these driver groups:
(i) The **impact experiment** (top half of the figure) that randomized access to a digital payment technology among taxi drivers and owners, primarily focusing on owners with only one taxi. Drivers willing to adopt the technology were first randomized access to measure the impact of the technology on cash-related costs and then their owners was further assigned to one of three observability treatments: *No Observability*, where owners received no information on transactions; *Granular Observability*, where owners had full access to the driver's transaction history, including timestamps, customer count, digital revenue, and working hours—signal of effort and output; and *Coarse Observability*, where owners received limited information via daily SMS updates capped at a threshold—signal of low-output periods. This experiment aimed to test how varying levels of observability influenced drivers' effort and owner-driver relational contracts, particularly by mitigating moral hazard in effort and/or output reporting.
(ii) The **adoption experiment** (bottom half of the figure) targeted drivers who initially refused to adopt the technology due to their reluctance to share their owner's contact information—a prerequisite for adoption. About a month after the baseline, these drivers were re-offered the technology with a randomized observability condition but, as opposed to the impact experiment, *before* making their adoption decision. The goal was to assess whether concerns about employer monitoring deterred adoption and to identify the role of transaction observability as a barrier. I also compare characteristics between drivers in the two experiments to explore selection and explain the differential adoption.
Both experiments were followed by mid- and long-term surveys (nine months and nearly two years) to track changes in drivers' performance, contracts, and owner-driver relationships more generally. This two-part design captures the trade-off between the benefits of observability in digital technologies and the adoption barriers it may create.

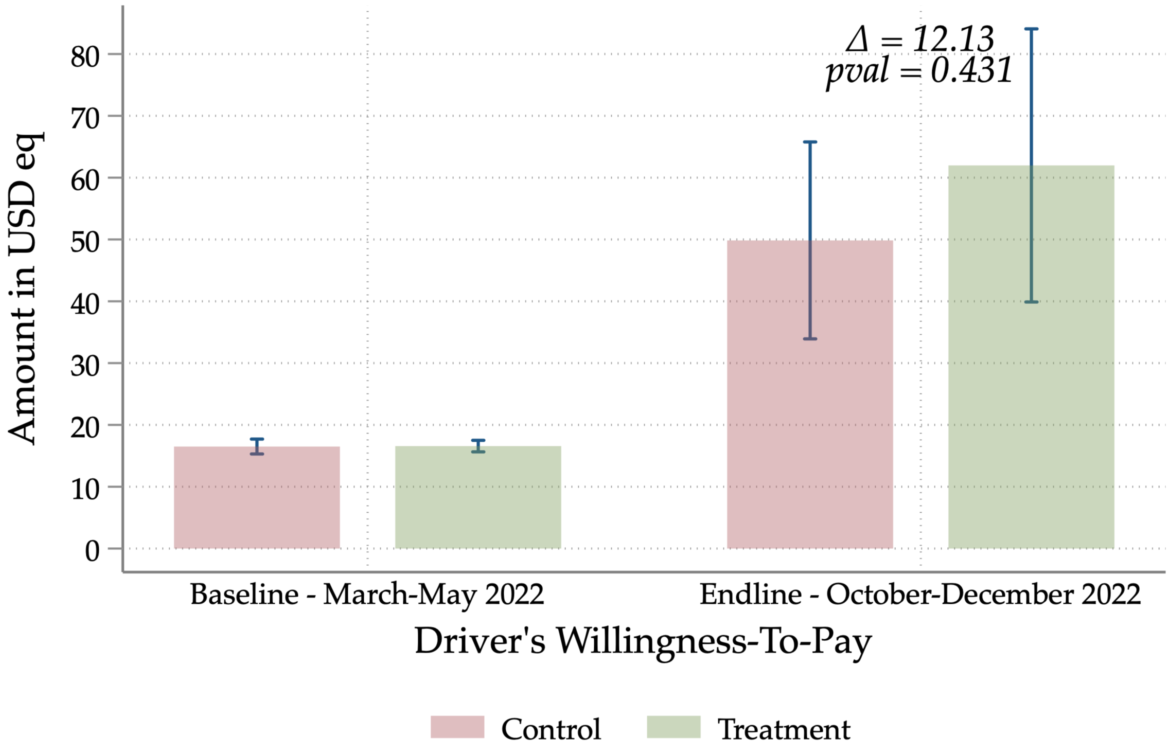


Figure 3: Driver's Willingness-To-Pay for Digital Payments

Notes: Driver's willingness to pay (WTP) was elicited at baseline following the Becker–DeGroot–Marschak (BDM) procedure in an incentivized way (Becker et al., 1964). To preserve the randomization treatment arms, the lottery was run on 5% of the treated sample only drawing a random number from an uniform distribution such that most of the 5% of treated drivers were actually given the product below their WTP. I also measured WTP for a benchmark good at baseline, a bottle of water, to reduce noise, as recommended in Dizon-Ross and Jayachandran (2022). At mid-term, given that incentivizing drivers was not feasible—the technology was available for free for drivers outside the experimental sample—the treated drivers were asked the following: "We are trying to understand your valuation of the digital payment technology. To *keep access* to the technology in your taxi, how much would you be willing to pay at most?" and similarly the control drivers were asked: "We are trying to understand your valuation of the digital payment technology. To *get* the technology in your taxi, how much would you be willing to pay at most?" The amount of money reported was converted from CFA to USD equivalent, by dividing by 600 ($1 $\approx$ CFA 600).

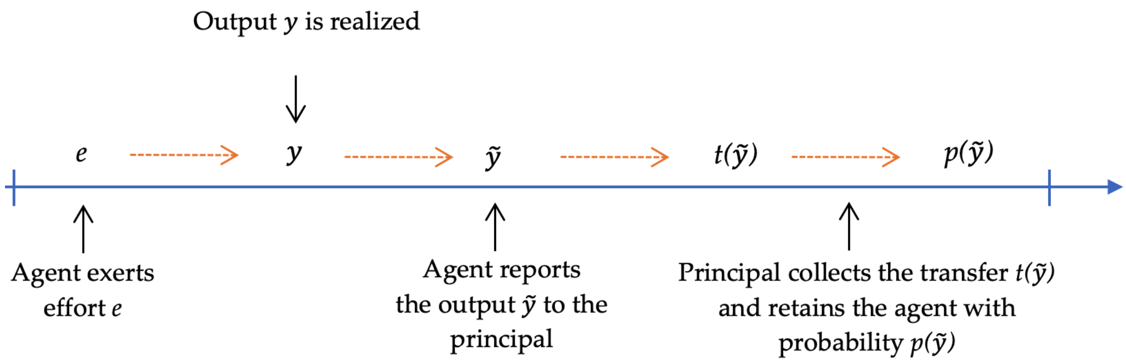
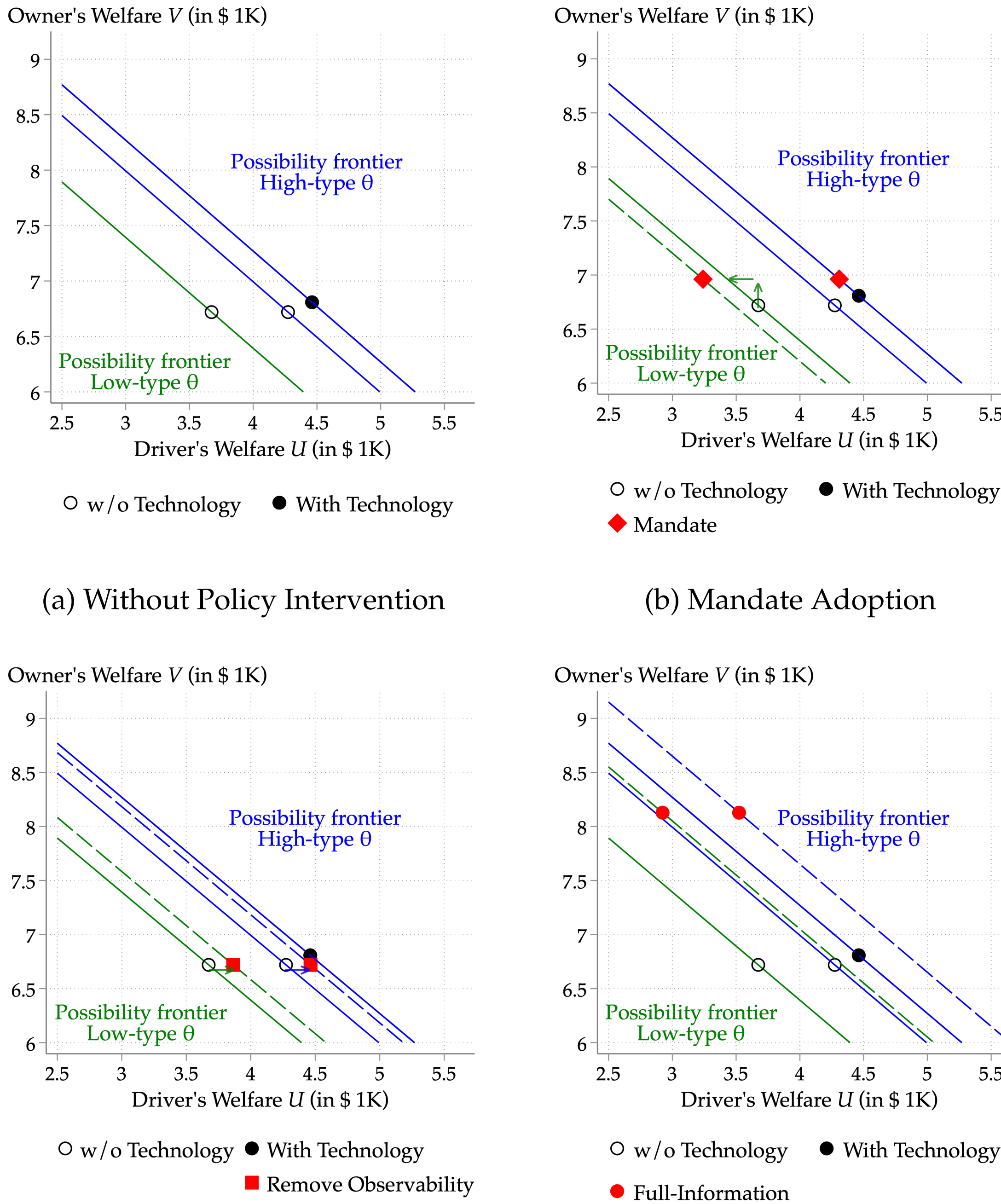


Figure 4: Timing of a Principal-Agent Relational Contract Period

Notes: This figure depicts the timing of events within a period of the dynamic game. Consider an environment with an infinite discrete time horizon, with periods indexed by $0, 1, \ldots, \infty$. The principal aims to incentivize the agent to exert effort on the job to maximize expected transfers and the future discounted value of the relationship. The agent's effort, $e \in \{0, 1, 2\}$ is unobservable to the principal at baseline. The agent's production function reflects the substantial output uncertainty inherent in the taxi industry, where an agent exerting low effort can achieve the same output as one exerting high effort, due to stochastic factors. Because the output is unobservable to the principal, the agent reports $\tilde{y}$, with the possibility to misreport the actual output collected $\tilde{y} \neq y$. The reported output will lead to a transfer $t(\tilde{y})$ and a continuation probability $p(\tilde{y})$ to pursue the relationship.



(a) Without Policy Intervention

(b) Mandate Adoption

(c) Redesign Without Observability

(d) Full-Information Benchmark

Figure 5: Owner's and Driver's Welfare Under Different Counterfactuals

Notes: These figures present counterfactual analyses by plotting the contract valuations (welfare) for both the owner and the driver on the same graph. The solid lines depict the utility possibility frontiers for low- and high-type drivers, before and after the introduction of the technology. Dashed lines represent the utility frontiers under various counterfactual scenarios. The graph is re-scaled to zoom into the area of interest. In constructing these graphs, I assume a social planner who maximizes total welfare, defined as the sum of the owner's and driver's welfare (equal weight).

Panel (a) contrasts the baseline contract (without digital payments) with the *Granular Observability* group (with digital payments). In Panel (b), I analyze the effect of mandating digital payment adoption, which requires both low- and high-type drivers to adopt the technology and exert high-effort. Panel (c) examines a counterfactual where the technology is redesigned to remove the observability feature, allowing both types of drivers to adopt. Panel (d) explores a full-information benchmark, where the technology remains unchanged, is universally adopted, but now fully reveals the driver's effort level (and not only a high-effort signal) and output.

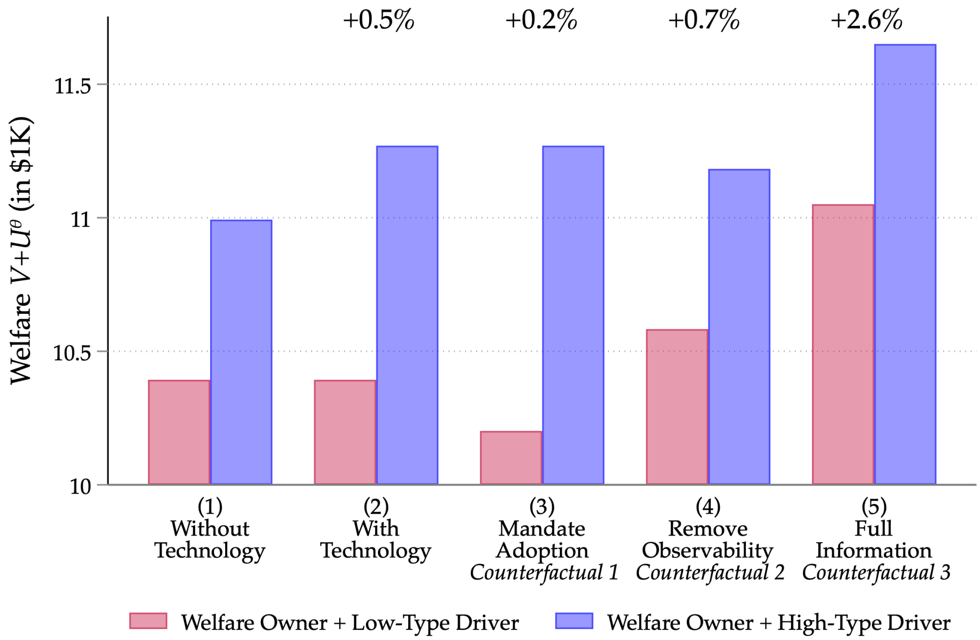


Figure 6: Total Welfare $U + V$ in the Economy Under Different Counterfactuals

Notes: This figure plots the total contract valuation (welfare) for both owners and drivers under different counterfactuals, combining both the owner and the high-type θ driver (in blue) and the low-type θ driver (in red). Each bar represents a different scenario, from no policy intervention—where the welfare of low-type pairs does not increase since they do not adopt the technology—to the counterfactual scenario of a complete shift to wage employment under a full-information benchmark. In this latter benchmark, the principal collects all the output (now observable) and provides a fixed wage to the employee.
The leftmost bar shows the total welfare with information frictions, absent the technology. The second bar represents the counterfactual scenario with no policy intervention where only high-type drivers adopt. The third bar plots welfare with a mandate on digital payments, requiring all drivers to adopt and the owners can receive the high-effort signal. The fourth bar plots welfare when the technology is redesigned to remove observability, thus enabling all drivers to adopt. The rightmost bar represents welfare in the full-information benchmark (wage employment) with effort and output being fully observable.
The welfare increase at the top of each bar reflects the total welfare change in the economy for both the high-type (with a share of μ) and low-type pairs (with a share of $1 - \mu$).

Table 1: Survey Rounds and Follow-up Rates

	(1)		(2)		(3)	
	Short-term (5mo)		Mid-term (9mo)		Long-term (20mo)	
	#	Rate (%)	#	Rate (%)	#	Rate (%)
Panel A. Taxi Businesses Overall						
Taxi Businesses (Any Owner or Driver) surveyed	2024	89.2	1945	85.7	1820	80.2
Drivers of a taxi cab surveyed	1714	90.6	1674	88.5	1555	82.2
# of baseline taxi businesses	2269		2269		2269	
# of baseline drivers of taxis	1891		1891		1891	
Panel B. Taxi Businesses with an Employee						
At least one surveyed	612	99.8	577	94.1	583	95.1
Owner surveyed	538	87.8	497	81.1	457	74.6
Driver surveyed	582	94.9	551	89.9	500	81.6
Both Owner and Driver surveyed	508	82.9	471	76.8	374	61.0
Relationship Outcomes Available	551	89.9	479	78.1	366	59.7
# of baseline taxi pairs	613		613		613	
Panel C. Taxi Drivers Non-Adopters						
Drivers surveyed			366	84.5	367	84.8
# of baseline taxi drivers non-adopters	433		433		433	

Notes: The survey data collection process is detailed in Section 4.1. Short-term survey data were collected from July to September 2022, approximately 5 months after the initial data collection. Mid-term data were collected from October to December 2022, about 7-9 months after the initial data collection. Long-term data were collected from September to December 2023, approximately 20-22 months after the baseline. Respondents independently consented to participate in the survey at each round.

For taxi businesses with an employee (pairs), I report whether the owner, the driver, both, or at least one of them was surveyed. They were always surveyed separately, enabling me to recover some of the relational contracts by surveying only one party. 'Relationship Outcomes Available' describes the taxi businesses for which relationship/contract information could be recovered. This indicates that either (1) the pair is still together and at least one responded or (2) the owner responded about the current or their new driver(s).

Baseline taxi drivers non-adopters were administered an adoption survey, and then followed up at mid-term and long-term.

Table 2: Impact of Digital Payments on Costs Associated with Cash Payments

	Any Time Lost	Had to Refuse Customers	Had to Reduce Price	Mistakes Giving Change	Imputed Loss	Electronic Theft
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Short-Term 5-Month Survey						
Technology Access	-0.214***	-0.174***	-0.121***	-0.044***	-1.887***	
	(0.023)	(0.021)	(0.025)	(0.015)	(0.241)	
Observations	1714	1714	1714	1714	1714	
Control Mean at Short-Term	0.44	0.31	0.49	0.11	4.67	
% Change T at Short-Term	-48.43	-56.25	-24.62	-38.82	-40.45	
Panel B. Mid-Term 9-Month Survey						
Technology Access	-0.109***	-0.143***	-0.034	-0.029**	-1.137***	-0.044***
	(0.020)	(0.021)	(0.022)	(0.012)	(0.196)	(0.011)
Observations	1674	1674	1674	1674	1674	1674
Control Mean at Mid-Term	0.25	0.31	0.32	0.08	3.05	0.06
% Change T at Mid-Term	-44.17	-46.33	-10.68	-35.99	-37.27	-70.98

Notes: Baseline survey data collected from March to June 2022. Short-term survey data collected from July-September 2022. Mid-term survey data collected from October-November 2022 (after 9 months). Driver-level regressions of the outcomes on the treatment (access to the technology), with the omitted category the pure control group. The model is the following: $y_{ij} = \beta_0 + \beta T_{ij}^{Access} + \epsilon_{ij}$, where y_{ij} is the outcome variable displayed at the top of the column, with i the driver in taxi business j. Cluster-robust standard errors at the business level j, and the significance of each coefficient is denoted by asterisks: $p < 0.1$ *, $p< 0.05$ **, $p< 0.01$ ***. The percent change between dividing the coefficient by the control mean is also reported. The sample includes all drivers that were surveyed at least once at short and/or mid-term — the case for about 90% of the sample — and replaced by missing and dummied out otherwise. Each regression includes controls for strata fixed effects and baseline controls when available.
The outcomes are dummies (0-1) coming from the survey data collected about the past 7 days prior to the survey date. They are constructed from the following questions:
- *Any Time Lost*: How many times have you wasted time (more than 10 minutes) looking for small-change during your work?
- *Refused Customers*: How many customers have you turned down because they only wanted to pay with electronic money, and not cash?
- *Reduced Price*: How many times have you reduced the price of the ride because of the small change problem?
- *Mistakes Giving Change*: How many times have you lost part of your collection with customers by giving change (e.g., miscalculation)?
- *Imputed Loss*: Money lost due to all these cash-related problems, imputed based on background interviews with taxi drivers.
- *Electronic Theft*: During the last 3 months, have you had part of your electronic collection (e.g., Wave) robbed?
All values are converted in USD when relevant (USD 1 = CFA 600).

Table 3: Impact of Digital Payments on Non-Rounded Prices

	Accept Digital Payments (1)	Non-Rounded Prices (2)	Non-Rounded Prices (3)
Panel A. First-Stage			
Treatment (Technology Access)	0.307***		
	(0.043)		
Panel B. OLS			
Treatment (Technology Access)		0.089**	0.080**
		(0.041)	(0.037)
Panel C. IV			
Accept Digital Payments		0.292**	0.258**
		(0.125)	(0.109)
Observations	710	710	710
Control Mean	0.440	0.323	0.323
Surveyor & OD FE	NO	NO	YES

Notes: Data on mystery passengers were collected over a two-week period in August 2022. Trained surveyors, acting as mystery passengers, audited taxis at various locations across Dakar. They engaged in a bargaining process with taxi drivers and discreetly recorded the license plates at the end of each interaction. These license plates were then matched to the corresponding taxi driver data and treatment group.

Each regression is performed at the passenger-taxi driver interaction level. Robust-heteroskedastic standard errors are reported, clustered at the business level. In the first column, I show the treatment effect on accepting digital payments, while in the two subsequent columns of the table, the OLS specification is run to measure the effect of treatment on accepting a non-rounded price. Controls always include strata fixed effects and surveyor, $Surveyor_i$, and origin-destination (OD), α_s, fixed effects when specified below the table. The surveyor and OD fixed effects are used to mitigate any noise related to systematic price differences across surveyors (e.g., taxi drivers tend to charge higher prices to male mystery passengers). The exact specification in Panel B is $NonRounded_i = \beta_0 + \beta T_i^{Access} + Surveyor_i + \alpha_s + \epsilon_i$

The last panel includes the IV specification, with the first-stage in Panel A. At baseline, 18 drivers refused to provide their license plate, and 3 had duplicate plates: these cases were excluded from the regressions. The outcome variables are coming from the following survey procedure:

- *Non-Rounded Prices*: After following the bargaining procedure, the mystery passenger offered the 'last price' to be driver's proposed price minus CFA 200, that is 33 cents. The goal was for the mystery passengers to suggest a non-rounded price (not a price rounded to 500, the lowest bill in this setting), and record whether the taxi driver accepted this request.
- *Accept Digital Payments*: The mystery passengers recorded whether the driver suggested or accepted to receive digital payments using the technology (business wallet QR code) during the bargaining process.

About 30% of the taxis recovered were found more than once, and I kept all the times they were encountered and use that number in Section 5.3 to measure driver's effort.

Table 4: Impact of Observability on Effort (Mystery Passengers Audit Survey)

	Count (1)	Count (2)	Unique Days (3)	Count Per Day (4)
Granular Observability	0.294***	0.385***	0.242***	0.063**
	(0.097)	(0.113)	(0.090)	(0.032)
Coarse Observability	-0.024	0.087	-0.086	0.055*
	(0.083)	(0.082)	(0.071)	(0.030)
No Observability	0.067	0.191**	0.051	0.017
	(0.086)	(0.095)	(0.078)	(0.031)
Observations	592	388	592	592
Only Owner Not Driving	NO	YES	NO	NO
Control Mean	0.368	0.375	1.333	1.010
Enumerator FEs	YES	YES	YES	YES
Chi-squared test Granular O = No O (p-value)	0.02	0.10	0.03	0.24

Notes: Business-level OLS regressions. The model is defined as $y_j = \beta_0 + \beta_1 T_j^{GranularObs} + \beta_2 T_j^{CoarseObs} + \beta_3 T_j^{NoObs} + \alpha_s + \epsilon_j$, where y_j is the outcome variable displayed at the top of the column, and α_s are the strata fixed effects. I also include enumerator fixed effects to remove any differences coming from enumerator's characteristics. Data on mystery passengers were collected over a two-week period in August 2022. Trained surveyors, acting as mystery passengers, audited taxis at various locations across Dakar, as described in Section 4.2. They engaged in a bargaining process with taxi drivers and discreetly recorded the license plates at the end of each interaction. These license plates were then matched to the corresponding taxi driver data and treatment group. In the first column of the table, the sample size consists of taxi businesses that agreed to provide their license plate number during the survey. In particular, at baseline, 18 experimental drivers refused to provide their license plate during the baseline survey, and 3 had duplicate plates: these cases were excluded from the regressions.
Business-level Poisson regressions were subsequently performed only on taxis driven by employees, to clearly distinguish the effects of owner and driver efforts. Controls included enumerator fixed effects for the drivers audited, and strata fixed effects. Heteroskedasticity-robust standard errors are reported and the significance of each coefficient is denoted by asterisks: $p < 0.1$ *, $p < 0.05$ **, $p < 0.01$ ***. The Chi-squared test for differences between the 'No Observability' and 'Granular Observability' coefficients is displayed at the bottom of the table. The following outcome variables are considered:
- *Count*: Represents the number of times a taxi, identifiable by its license plate, was audited, meaning it was observed on the road. I imputed 0 if the taxi was not audited.
- *Unique Days*: Indicates the number of different days a taxi was audited.
- *Count Per Day*: Reflects the average number of audits per unique day for each taxi.

Table 5: Impact of Observability on Default and Owner's Profit

	Monthly Default Rate	Transfer Value	Owner's Profit
	(1)	(2)	(3)
Granular Observability	-0.097*	27.458	19.573
	(0.055)	(17.328)	(16.642)
Coarse Observability	0.039	-12.406	-13.953
	(0.063)	(19.584)	(17.801)
No Observability	-0.004	0.005	0.506
	(0.060)	(18.957)	(17.879)
Observations	479	479	524
Control Mean	0.31	312.55	254.93
% Change Granular Observability	-31.46	8.79	7.68
F-test Granular O = No O (p-value)	0.16	0.19	0.34

Notes: Outcomes are regressed on the three treatment arms, with the omitted control group. Heteroskedasticity-robust standard errors (SE) are used, and the significance of each coefficient is denoted by asterisks: $p < 0.1$ *, $p < 0.05$ **, $p < 0.01$ ***. Controlling for strata fixed effects. The outcome is recorded at mid-term, after 7–9 months. Respondents who refused to respond or said they did not know were dummied out. Values are converted to USD (USD 1 = CFA 600). The default rate indicates the proportion of drivers who default at least once a month (in the past three months), according to either the owner's or their driver's reports. The Transfer Value was imputed as follows: total default on the rental payment is set to half of the expected payment. The default rate is defined as a dummy variable equal to 1 if the driver defaulted at least once in the month, as reported by either the owner or the driver. The owner's profit is computed as the monthly transfer value minus the upfront payment, where applicable. Even accounting for the upfront payment increase as a result of observability (see Table 6), the owner's profit still increases overall. This excludes costs not impacted by the intervention (e.g., maintenance costs) to minimize noise in the profit outcome. For the same reason, this regression controls for baseline profit.

Table 6: Impact of Observability on Contracts and Relationships

	Upfront Payment 'Salary' Dummy (1)	Upfront Payment 'Salary' Value (USD) (2)	Weekly Rent Target Value (3)	Separation (4)
Granular Observability	0.132^{***}	10.891^{***}	0.261	-0.107^{**}
	(0.044)	(3.394)	(1.087)	(0.050)
Coarse Observability	-0.012	0.503	1.579	-0.024
	(0.052)	(4.046)	(1.305)	(0.053)
No Observability	-0.008	0.223	0.216	-0.023
	(0.050)	(3.966)	(1.273)	(0.053)
Observations	479	479	479	577
Control Mean	0.75	55.34	100.60	0.35
% Change Granular Observability	17.57	19.68	0.26	-30.37
F-test Granular O = No O (p-value)	0.01	0.01	0.97	0.16

Notes: Business-level OLS regressions of the contract outcomes on the three treatment arms, with the pure control group as the omitted category. The model is defined as $y_j = \beta_0 + \beta_1 T_j^{GranularObs} + \beta_2 T_j^{CoarseObs} + \beta_3 T_j^{NoObs} + \alpha_s + \epsilon_j$, where y_j is the outcome variable displayed at the top of the column and $alpha_s$ are the strata fixed effects. These regressions are conducted for the mid-term period (approximately 9 months after the baseline survey). Heteroskedasticity-robust standard errors are used and the significance of each coefficient is denoted by asterisks: p < 0.1 *, p< 0.05 **, p< 0.01 ***.
- *Upfront Payment 'Salary' Dummy*: Whether the owner provides a monthly upfront fixed payment to their driver (also referred to as a 'salary' in the industry).
- *Upfront Payment 'Salary' Value (USD)*: Value of the monthly upfront fixed payment to the driver, including 0 if no payment is provided. Values converted in USD (USD 1 = CFA 600).
- *Weekly Rent Target Value*: Weekly rental target amount from driver to owner, as agreed between owner and driver. Values converted in USD (USD 1 = CFA 600).
- *Separation*: Owner and driver are not working together at the time of the survey.

Table 7: Impact of Observability on Taxi Owner's Trust

	Perceived Moral Hazard	Taxi Parked at Driver's House	Owner's Trust	Owner's Trust	Driver's Trust
	(1)	(2)	(3)	(4)	(5)
Granular Observability	-0.127**	0.140**	0.020	0.045	-0.086
	(0.052)	(0.057)	(0.183)	(0.181)	(0.139)
Coarse Observability	-0.074	-0.073	-0.342*	-0.298	0.043
	(0.058)	(0.059)	(0.185)	(0.181)	(0.153)
No Observability	0.079	0.044	-0.410*	-0.369	-0.159
	(0.066)	(0.062)	(0.236)	(0.235)	(0.165)
Observations	429	429	429	429	459
Control Mean	0.287	0.506	8.924	8.901	9.140
Relative Scale Control			NO	YES	YES
F-test Granular O = No O (p-value)	0.00	0.18	0.12	0.13	0.69
% Change Granular Observability	-44.1	27.6	0.2	0.5	-0.9

Notes: Mid-term survey data collected from October-December 2022, approximately 9 months later. The model is defined as $y_j = \beta_0 + \beta_1 T_j^{GranularObs} + \beta_2 T_j^{CoarseObs} + \beta_3 T_j^{NoObs} + \alpha_s + \epsilon_j$, where y_j is the outcome variable displayed at the top of the column and $alpha_s$ are the strata fixed effects. Heteroskedasticity-robust standard errors are used. The F-stat testing for the difference between the estimate of *No Observability* and *Granular Observability* is shown at the bottom of the Table. Respondents who refused to respond or said they did not know were dummied out.

The following questions were asked for each of the outcomes displayed:

- *Perceived Moral Hazard*: If your driver collects less than CFA 25,000 (USD 40) during a full day of work, do you believe this is due to bad luck, God's will, unpredictable aspects of the job, or the driver's low effort?
- *Taxi Parked at Driver's House*: Where is the taxi generally parked outside working hours? Whether the driver is allowed to park the taxi at their own house. In this context, owners who lack trust in their drivers often require them to park the car at a specific location they can monitor each day.
- *Owner's Trust*: How would you rate your trust in your current driver from 0 to 10?
- *Owner's Trust - Relative Scale Control*: Generally, if you had to rate your trust in *taxi drivers* (not your own, but *taxi drivers* more broadly), from 0 to 10, what score would you give them?
- *Driver's Trust*: How would you rate your trust in your current owner from 0 to 10?

Table 8: Impact of Observability On Technology Adoption

	Technology Adoption (Willing to Share Owner's Information)					
	(1)	(2)	(3)	(4)	(5)	(6)
Removing Observability	0.114***	0.111***	0.102***	0.100***	0.196***	0.158***
	(0.038)	(0.038)	(0.037)	(0.037)	(0.049)	(0.060)
Observations	433	433	433	433	204	159
Mean Under Observability	0.143	0.143	0.143	0.143	0.069	0.095
Enumerator FE	NO	YES	NO	YES	NO	NO
Privacy Concern Controls	NO	NO	YES	YES	NO	NO
Relationship Length Control	NO	NO	NO	YES	NO	NO
Sample	All	All	All	All	Poorest	Worst-Performing
% Change Removing Observability	80	77	71	70	284	166

Notes: Survey data were collected from June 15 to July 7, 2022, on drivers who refused to provide their owner's contact information during the listing. Driver-level regressions are performed. The outcome *Adoption* is whether the driver provided the owner's contact information to the surveyor, thus enabling them to adopt the digital payment technology. Controls include enumerator fixed effects, the length of the driver-owner relationship (years they spent working with their owner), and privacy concern controls, which are specified at the bottom of the table. Heteroskedasticity-robust standard errors are reported, and the significance of each coefficient is denoted by asterisks: $p < 0.1$ *, $p < 0.05$ **, $p < 0.01$ ***. The random assignment of removing the owner's observability of drivers' digital transactions was not stratified.

Privacy concern controls include the following three dummies: (i) whether the driver accepted to provide a friend's phone number for easier future contact, (ii) whether the driver accepted to provide their garage's or taxi association's name, when relevant, and (iii) whether the driver accepted to provide their license number. These controls aim to observe privacy concerns for sharing information more generally, beyond the impact of observability on the relationship. The 'poorest' drivers are defined as those below the median score of the wealth PPI index. The 'worst-performing' drivers are defined as those with an average z-score productivity index below 0.

Table 9: Reluctant Drivers Tend to Be Lower-Performing and Poorer

	Willing To Adopt (1)	Reluctant Drivers (2)	Difference $\beta_{ReluctantDrivers}$ (3)
Panel A. Driver's Performance			
Performance Index (Z-Score)	0.085	-0.122	(-0.207***)
	(0.709)	(0.857)	
Number of Passengers (3 Days)	43.819	39.128	(-4.692***)
	(13.021)	(13.587)	
Total Collection (3 Days, USD)	153.010	147.571	(-5.439*)
	(30.312)	(38.349)	
Effective Hours Worked (Avg 3 Days)	10.121	9.764	(-0.356**)
	(2.110)	(2.325)	
Total Work Time (End to Start - Avg 3 Days)	11.690	11.182	(-0.508**)
	(2.469)	(2.749)	
Panel B. Relationship and Contracts			
Monthly Upfront Payment 'Salary' (Dummy) W	0.746	0.883	(0.137***)
	(0.436)	(0.322)	
Weekly Rent Value R (USD)	101.030	101.806	(0.776)
	(10.559)	(11.280)	
Separation Rate p	0.354	0.356	(0.002)
	(0.479)	(0.479)	
Owner-Driver Relationship $>$ Two Years	0.425	0.415	(-0.010)
	(0.495)	(0.493)	
Panel C. Driver's Characteristics			
Risk-Averse Agents (CRRA$>$ 1)	0.460	0.590	(0.130***)
	(0.499)	(0.493)	
Often Stressed About Rent Transfers	0.081	0.148	(0.067***)
	(0.273)	(0.356)	
Panel D. Demographics and Poverty			
Education (At Least Primary)	0.302	0.216	(-0.086**)
	(0.460)	(0.412)	
Literacy (Reading and Writing)	0.645	0.587	(-0.059)
	(0.479)	(0.493)	
Wealth Index (PPI-IPA)	63.558	58.141	(-5.417***)
	(16.877)	(16.484)	
Additional Revenue Source	0.163	0.220	(0.057)
	(0.369)	(0.415)	
Observations	367	340	

Notes: The table summarizes the characteristics of drivers who were willing to adopt the digital payment technology compared to those who did not adopt because they refused to share their owner's contact information. Data, except for demographics and risk-aversion coefficients, were collected during the mid-term survey with drivers from October to December 2022. The first two columns present the mean values of each variable for drivers willing to adopt and non-adopters, with standard deviations in brackets below. To characterize selection, I compare drivers in the impact experiment—who agree to potential observability but whose owners are randomly assigned not to receive it—to reluctant drivers. This approach is used since only mid-term performance data could be collected as many reluctant drivers refused to participate in a long one-month follow-up survey due to their reluctance to provide their owner's contact, knowing it would prevent adoption. I compare performance between drivers 'willing to adopt'—assigned to *No Observability* or *Control* to increase sample size and power—and 'non-adopters'—the reluctant drivers, excluding the few drivers who provided the owner information at the one-month follow-up under *Granular Observability*. Robustness checks are conducted using only drivers randomized to the *No Observability* groups, and the results remain qualitatively similar, see Table B27. The third column shows the estimate from the regression of the variable on being a non-adopter, that is $Y_i = \beta ReluctantDrivers + \epsilon_i$. Heteroskedasticity-robust standard errors and the significance of each coefficient is denoted by asterisks: p $<$ 0.1 *, p $<$ 0.05 **, p$<$ 0.01 ***.
Values are converted to USD (USD 1 = CFA 600). In particular: The Z-Score Productivity Index is a combination of mean of the z-scores of the number of passengers, the total collected, and the hours worked. Risk-averse agents are defined as driver with a coefficient of relative risk aversion above 1 (CRRA utility function), as determined in the field in an incentivized game. The Wealth Index is defined using the methodology developed by IPA in Senegal to measure household wealth, referred to as the Poverty Probability Index (PPI) based on the poverty survey (ESPS-II) developed in 2011 in Senegal.

Table 10: Structural Estimation: Matched Moments and Parameter Estimates

Panel A: Reduced Form, Structural, and Matched Moments

Control group outcome	Reduced form	Structural	Difference
Targeted moments:			
Probability $\bar{p}^l$ from $min\{(IC_l),(TT_l)\}$	0.965	0.985	-0.020
	(0.004)	(0.010)	(0.011)
Probability $\bar{p}^h$ from $min\{(IC_h),(TT_h)\}$	0.968	0.985	-0.017
	(0.005)	(0.010)	(0.011)
Probability $\tilde{p}_{\tilde{e}=2}$ under Granular Observability	0.972	1.000	-0.028
	(0.009)	(0.000)	(0.009)
Driver's replacement cost K_a	436.11	444.54	-8.431
	(61.00)	(0.00)	(61.002)
High-θ Driver's contract value U^h	4103.80	4273.50	-169.705
	(124.37)	(37.81)	(127.664)
Transfer R_l from $min\{(IC_l),(TT_l)\}$	100.293	100.301	-0.007
	(0.274)	(0.273)	(-0.007)
Transfer R_h from $min\{(IC_h),(TT_h)\}$	100.293	100.298	-0.005
	(0.274)	(0.273)	(-0.005)
Salary of Adopters $W_{\tilde{e}=2}$	10.94	10.94	0.000
	(0.46)	(0.46)	(0.000)
Untargeted moment:			
Low-θ Driver's contract value U^l	3993.78	3673.62	320.158
	(115.66)	(37.81)	(120.650)
Baseline welfare estimates:			
Owner's contract value V_h	—	6719.42	—
		(316.72)	

Panel B: GMM Parameter Estimates

Input	Value	Interpretation
Low-θ Baseline driver disutility $\hat{\phi}^l(1)$	21.20	Driver disutility in USD
	(5.75)	
High-θ Baseline driver disutility $\hat{\phi}^h(1)$	15.20	Driver disutility in USD
	(5.82)	
High-θ Endline driver disutility $\hat{\phi}^h(2)$	28.49	Driver disutility in USD
	(7.27)	

Panel C: Computed Parameter Estimate

Input	Value	Interpretation
Low-θ Baseline driver disutility $\hat{\phi}^l(2)$	39.17	Driver disutility in USD 39.17
	(9.35)	

Notes: In Panel A, I use the generalized method of moments (GMM) with eight targeted moments: weekly rehiring probabilities for non-adopters, the control group, and the *Granular Observability* group; driver contract valuation; perceived driver replacement cost; agent transfers for both types; the upfront payment/salary under *Granular Observability*. The reduced form consists of observed empirical data, while structural represents the corresponding model predictions. The difference measures the gap between reduced form and structural moments. Reduced-form weekly continuation probabilities are computed by deriving observed mid-term probabilities (28 weeks after baseline) while accounting for effort differences q_1 and q_2. It is considered that the agent has a probability $(1-q_1)(1-\bar{p})$ of leaving each period when exerting effort $e=1$. The untargeted moment uses empirical contract valuation for adopting drivers who initially would have preferred not to have *Granular Observability* at baseline, as captured in the survey responses. This approach is used because baseline valuations for low-type drivers were not collected.
In Panel B, I use GMM to estimate driver disutilities ϕ for each type of driver with $e=1$, and $e=2$ for a high-type driver (upon adoption of the technology).
In Panel C, I estimate the counterfactual lower bound for the driver's disutility of effort for $e=2$, $\phi^l(2)$. Since this parameter is empirically unobserved, as low-type drivers did not adopt the technology, I obtain a lower bound using the following model intuition: a low-type θ driver would need a high enough salary $W^l_{\tilde{e}=2} > W$, (for a given $\tilde{p}$) to exert high effort $e=2$. I compute the minimum value for $\phi^l(2)$ such that at $W^l_{\tilde{e}=2}$, the owner would be better off in the status quo (not profitable to compensate adoption).
Standard errors for each parameter are shown in brackets, estimated using a bootstrap procedure with 1000 replications based on the empirical distributions of the framework inputs.

Appendix

Table of Contents

A Additional Figures

A.1 Listing Survey and Mystery Passenger Audit Maps

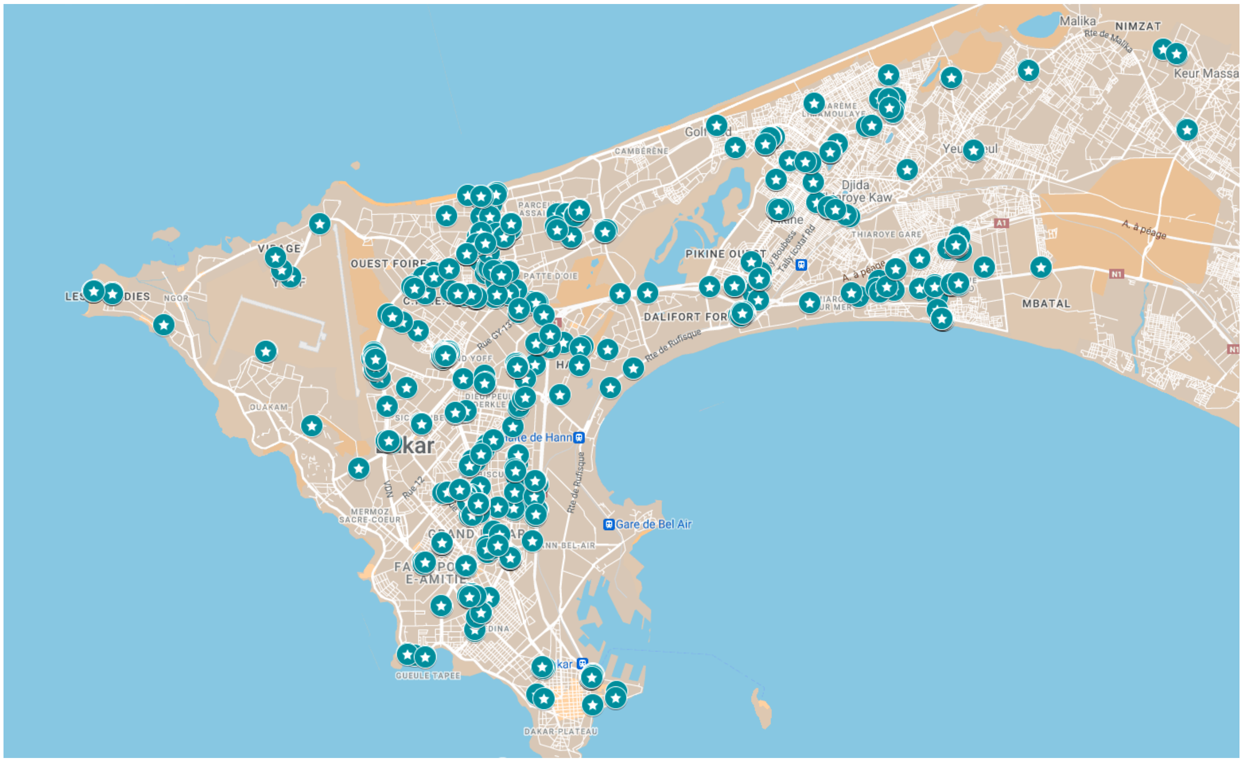


Figure A1: Listing Survey Activity in Dakar - March to May 2022

Notes: The figure displays a map of Dakar, the capital city of Senegal, where the experiment took place. Each blue dot represents the GPS location of the listing survey conducted to constitute the experimental sample. GPS coordinates from a random subset are displayed. Drivers were recruited in garages, car wash stations, meeting points, and on the streets of Dakar. The extent of the listing was broad enough to cover most parts of the city as described in Section 4.1. Owners not driving were primarily recruited by asking drivers about owners' contact information during the listing survey. Basic characteristics were collected during the listing survey to stratify the randomization.

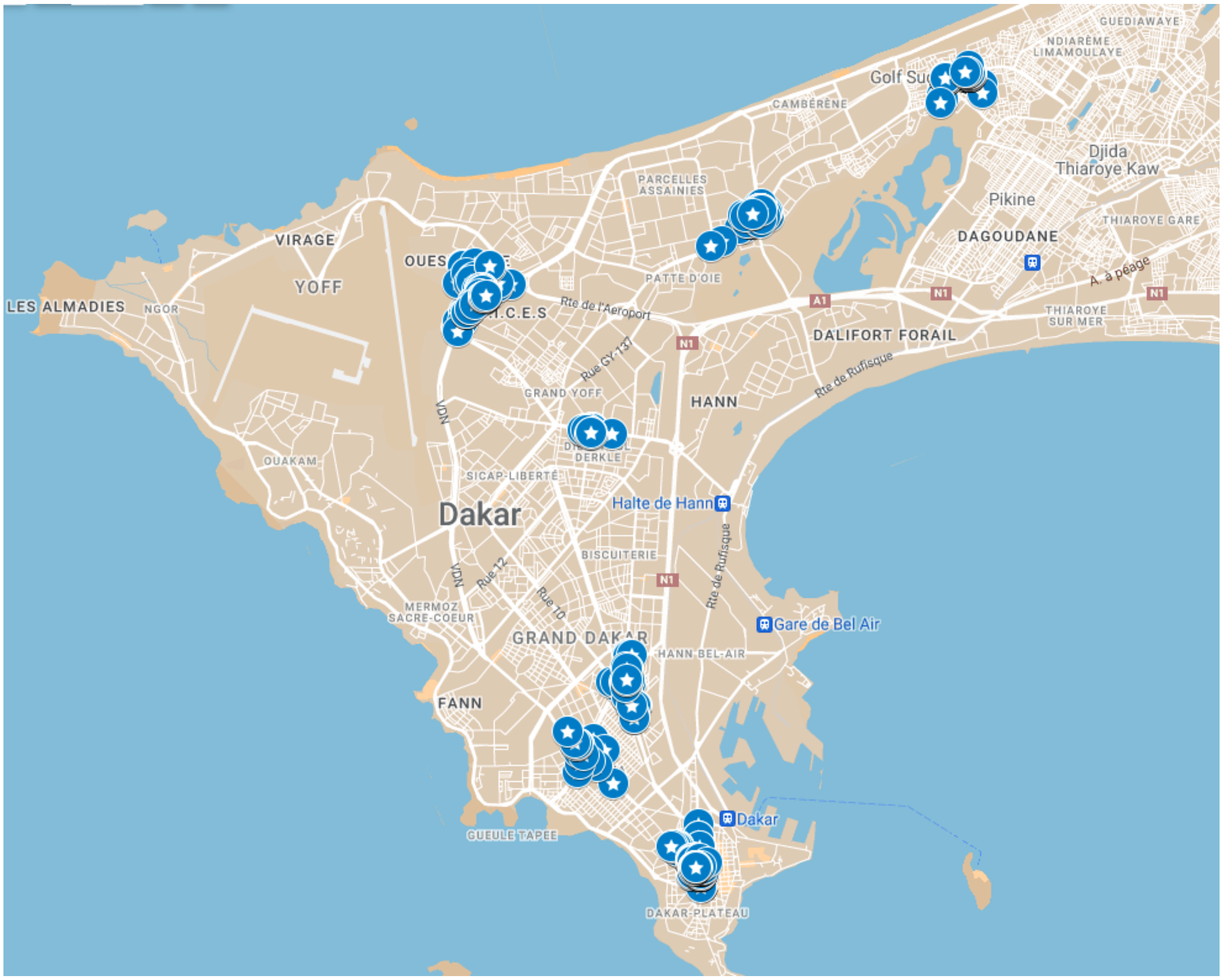


Figure A2: Locations of the Mystery Passenger Audits - August 2022

Notes: The figure displays a map of Dakar, the capital city of Senegal. Each blue dot is the GPS location of the mystery passenger audit activity for a random subset of data points. The goal was to measure (i) drivers' behavior related to digital payments and pricing, and (ii) drivers' effort based on their presence on the road. In particular, in August 2022, I trained twenty mystery passengers to hail taxis throughout Dakar, following a strict procedure to mimic typical price bargaining. Over two weeks, they systematically rotated across six high-traffic locations each day, capturing a broad sample of taxis and driver behavior over a meaningful timeframe. Surveyors asked questions and secretly recorded taxi license numbers. They pretended they had to leave after a pre-set bargaining process—primarily to increase the sample size and reduce field costs. The activity was repeated a sufficient number of times to match taxi drivers with their license numbers in the experimental sample. Specifically, mystery passengers adhered to the following steps: (1) Memorize the randomized destination and pre-specified price on their data collection application, (2) Stop a taxi, (3) Ask the driver's initial price, (4) Suggest the pre-specified low price, (5) Listen to the driver's counteroffer and ask their last price, (6) Suggest a non-rounded price, (7) Ask to use digital payments. Detailed data was recorded on a tablet once the taxi left about each step of the process.

A.2 Illustrations of Digital Observability via Payment Technology

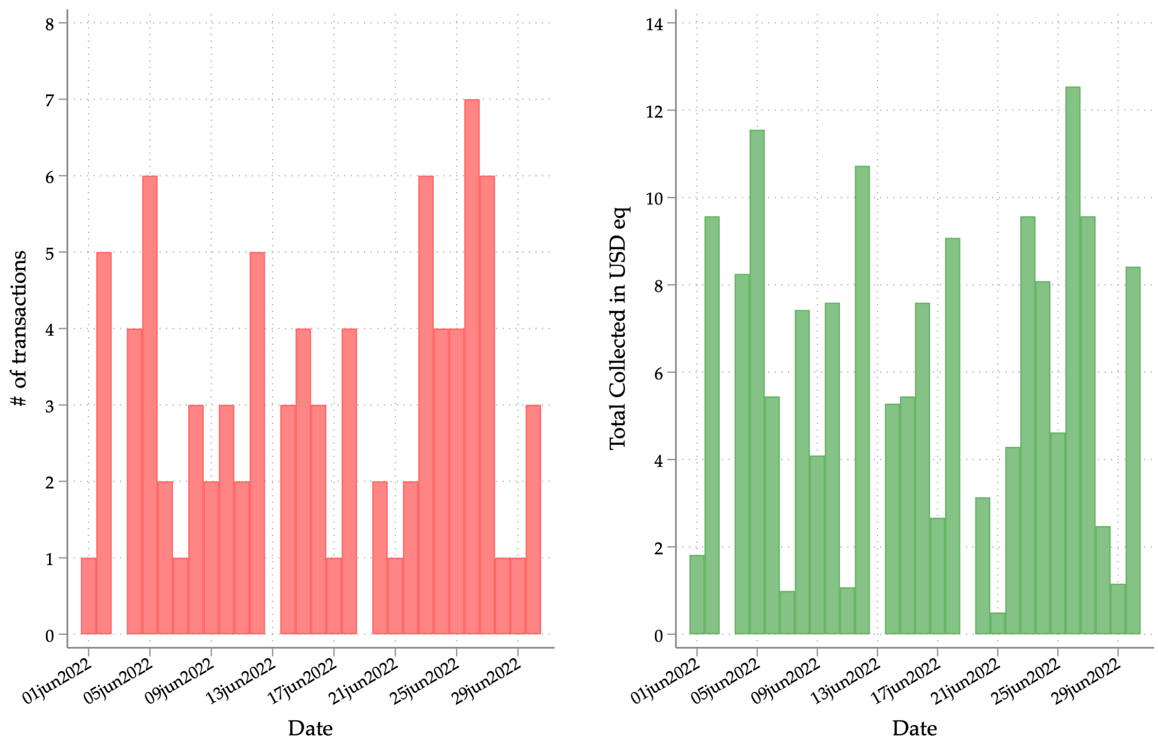


Figure A3: Illustration of a Driver's Transaction History with Observability

Notes: This figure shows an example of the transaction count and the total collected across days for one driver on the digital payment app. This figure illustrates the information available to taxi owners under *Granular Observability*, i.e., the number of transactions and the amount collected. The owners can also better evaluate time and hour of the transaction, as shown in Figure A4.

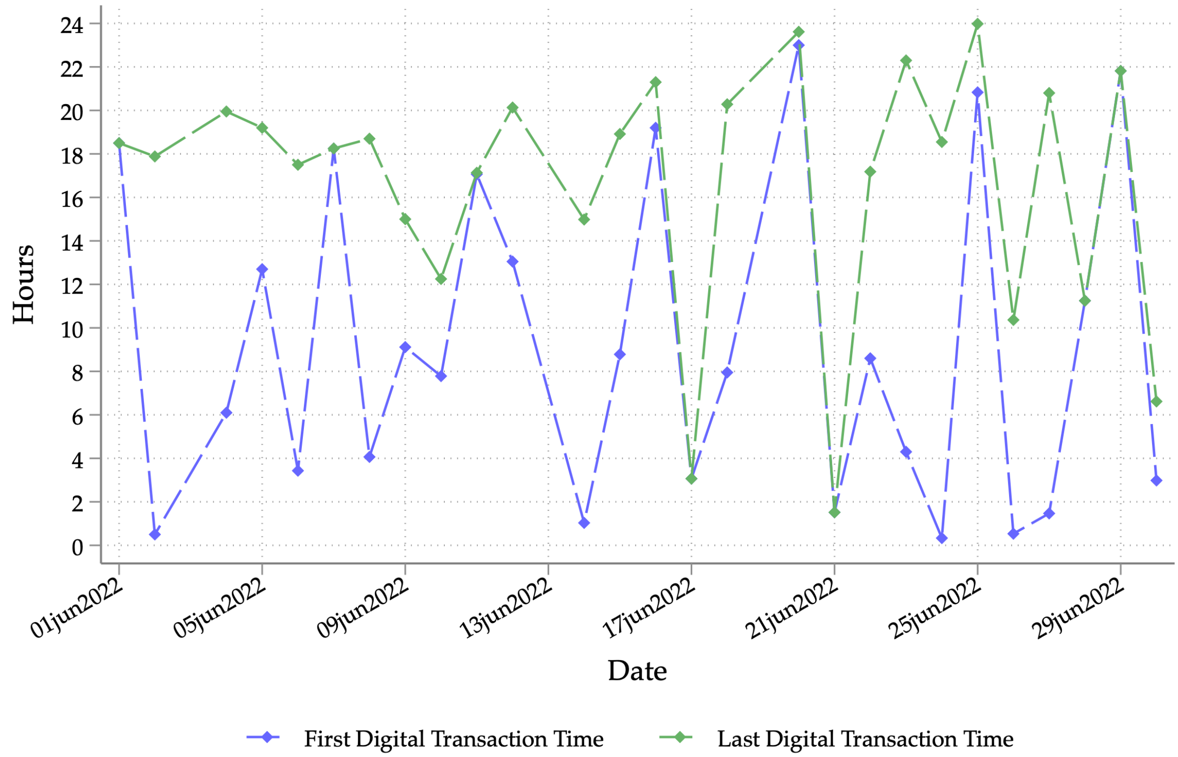


Figure A4: Illustration of start and end hours worked from the transaction history of a driver with observability

Notes: This figure shows an example of the transaction count and the total collected across days for one driver on the digital payment app. It conveys the information available to taxi owners under *Granular Observability*, i.e., an estimate of the start and end working hours, providing a signal for the employee effort.

A.3 Profit Measures in the Taxi Industry

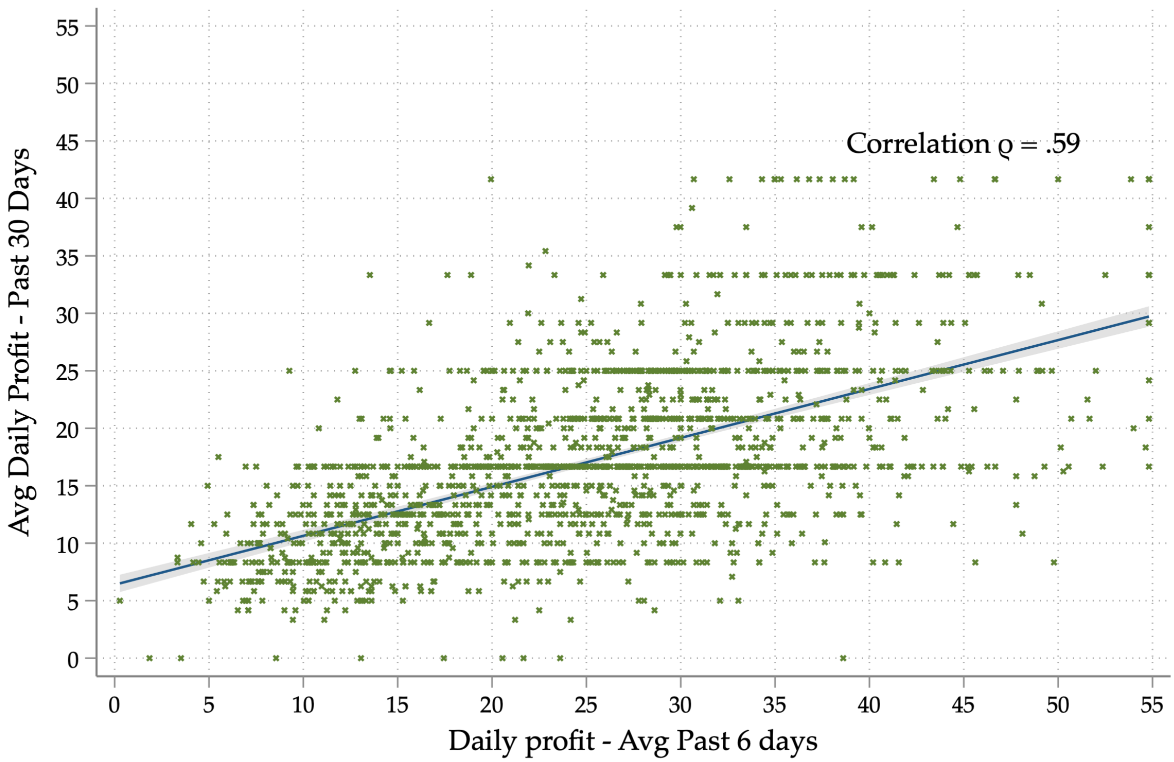


Figure A5: Positive Correlation Between the Two Profit Measures

Notes: This figure plots the average daily profits over the last 6 days and the average self-reported daily profit over the past 30 days (measured twice in the short- and mid-term). The two variables come from two different survey questions, following a similar methodology as described in De Mel et al. (2009): (1) "On [DATE], how much money did you get to keep, once you paid all your work-related expenses, including fuel, payment, repair, and police, contribution, and food?" one question for each of the past 6 days, and then (2) "Over the last 30 days, what is the average income that you managed to keep per worked day, once you paid all your work-related expenses, including fuel, payment, repair, and police, contribution, and food?"

The two measures are winsorized at the 99 percentile and the value in CFA is converted into USD (CFA 600 = USD 1). This figure documents a strong positive correlation between the two measures of profits, but also the large variance in profit across and within drivers, as the scatter plot shows. Driver's past 6 days profit only partly predicts the past 30 days profit and tends to be higher.

A.4 No Evidence of Manipulation of the Signal by Drivers

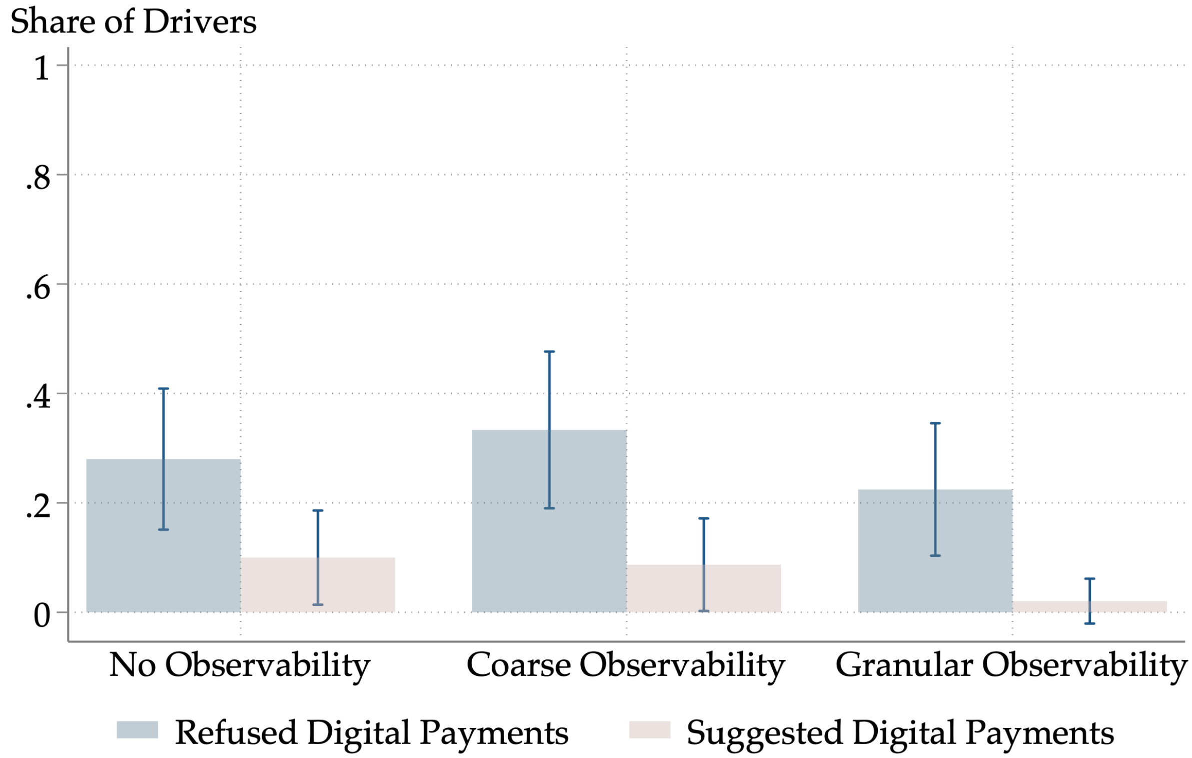


Figure A6: No Manipulation of the Effort and Output Signals - Mystery Passenger Audits

Notes: This figure shows no evidence of manipulation of the output of effort signals by the taxi drivers during the mystery passenger audits. This figure shows the share of drivers who refused digital payments or suggested paying digitally to the mystery passengers during the audit activity. In general, the agent can mostly manipulate the share of digital transactions downward (i.e., processing more cash than digital payments) and not the reverse in this setting. In most cases, passengers have the choice of whether to pay digitally or in cash. I do not find any case where drivers *demanded* digital payments from customers, which is understandable in an economy where cash remains the dominant form of transaction and it is practically difficult to compel customers to pay digitally. This figure shows no differential propensity to engage in manipulating the effort or output signal in the data from mystery passenger audits, neither upward nor downward. The absence of manipulation may be explained by the competitive pressure drivers face to secure passengers, which discourages them to manipulate digital payment usage. Drivers may perceive the short-term loss of passengers—resulting from pressuring or discouraging them to pay digitally to signal something to their owners—as too costly to justify such actions.

A.5 Additional Reduced-Form Impacts of Observability

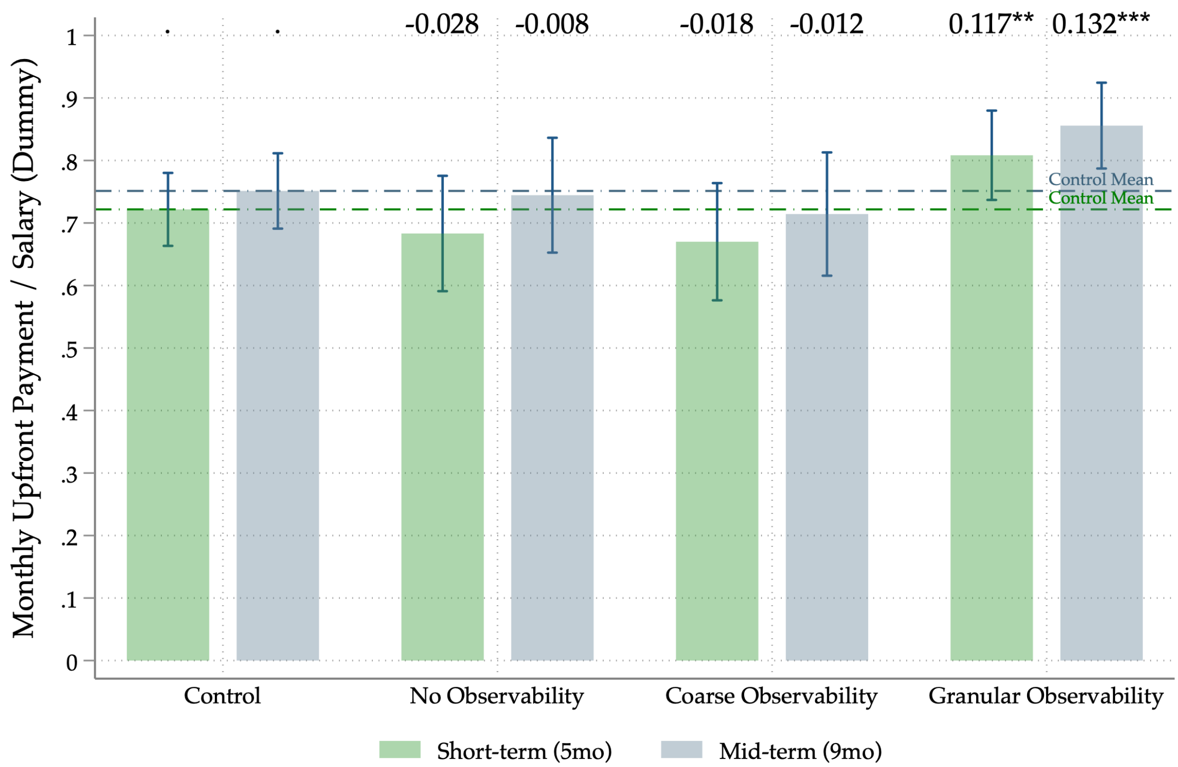
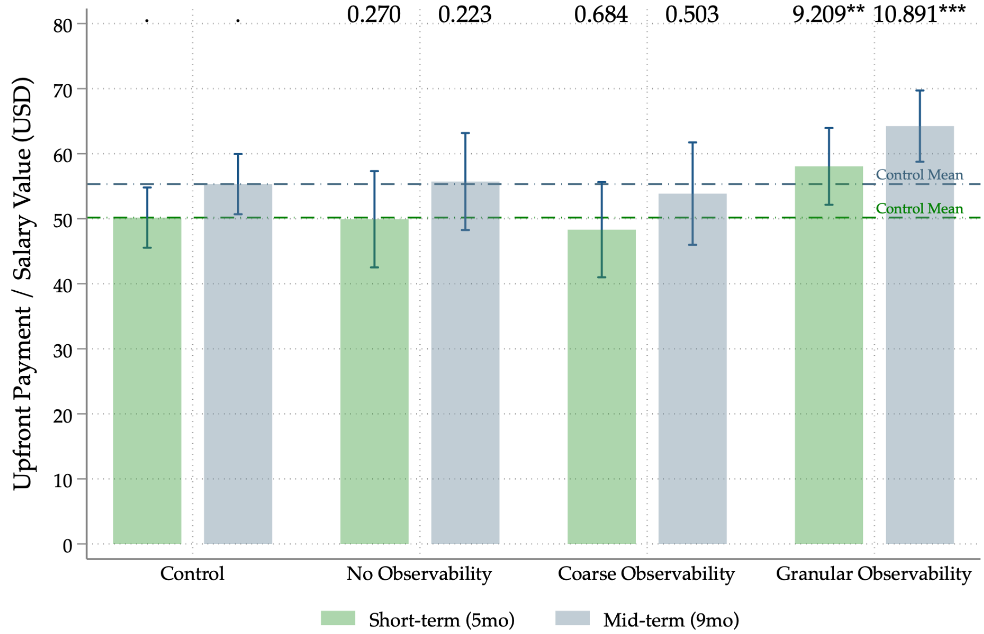
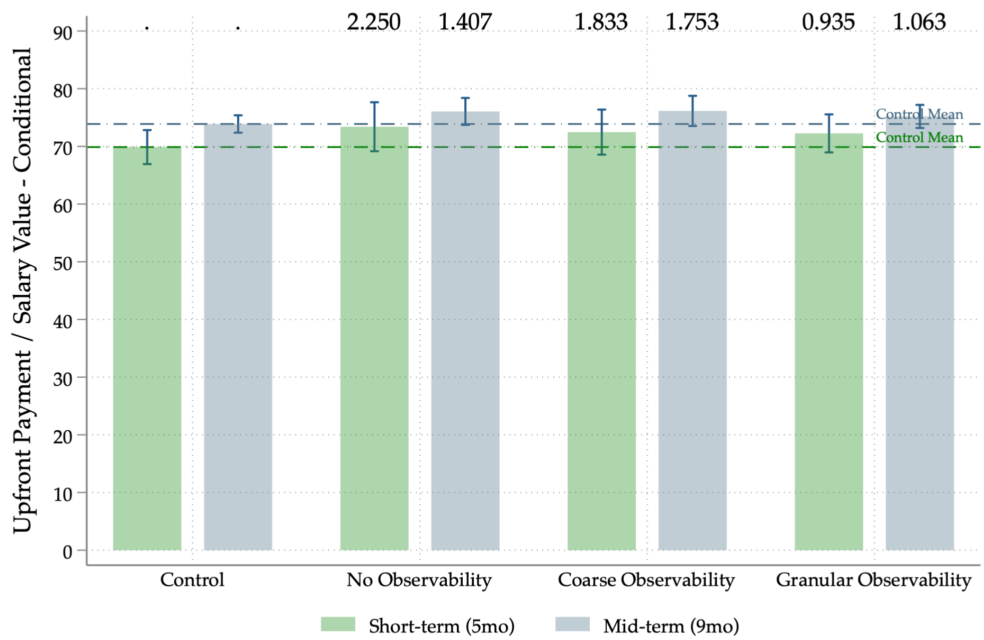


Figure A7: Impact of Observability on the Upfront Payment (Dummy) W

Notes: Business-level regression analyses assess the impact of varying levels of observability on the monthly upfront payment outcome, y_j. The model is defined as $y_j = \beta_0+\beta_1 T_j^{GranularObs}+\beta_2 T_j^{CoarseObs}+\beta_3 T_j^{NoObs}+\alpha_s+\epsilon_j$, where y_j indicates whether the business owner provides an upfront payment to their driver. These regressions are conducted separately for short-term (approximately 5 months, indicated in green) and mid-term periods (approximately 9 months, indicated in blue). The coefficients β_1, β_2, and β_3 are displayed above each respective bar, for each respective treatment arm. Each regression includes controls for strata fixed effects α_s. These results include heteroskedasticity-robust standard errors and the significance of each coefficient is denoted with its significance level: $p < 0.1$ *, $p < 0.05$ **, $p < 0.01$ ***. A horizontal dotted line shows the control group's mean outcome, serving as the baseline comparison.



(a) Unconditional



(b) Conditional

Figure A8: Impact of Observability on Upfront Payment (Salary) Value

Notes: This figure plots the impact of observability levels on the upfront payment value, unconditional, imputing 0 if no upfront payment (top) and conditional on getting an upfront payment (bottom). Business-level regression analyses are run to assess the impact of observability on the monthly upfront payment in value, y_j. The model is defined as $y_j = \beta_0 + \beta_1 T_j^{GranularObs} + \beta_2 T_j^{CoarseObs} + \beta_3 T_j^{NoObs} + \alpha_s + \epsilon_j$, where y_j indicates the value of the monthly upfront payment, converted in USD (USD 1 = CFA 600), conditional on offering an upfront payment. These regressions are conducted separately for short-term (approximately 5 months, indicated in green) and mid-term periods (approximately 9 months, indicated in blue). The coefficients β_1, β_2, and β_3 are displayed above each respective bar, for each respective treatment arm. Each regression includes controls for strata fixed effects α_s. These results include heteroskedasticity-robust standard errors and the significance of each coefficient is denoted by asterisks: $p < 0.1$ *, $p < 0.05$ **, $p < 0.01$ ***. A horizontal dotted line shows the control group's mean outcome, serving as the baseline comparison.

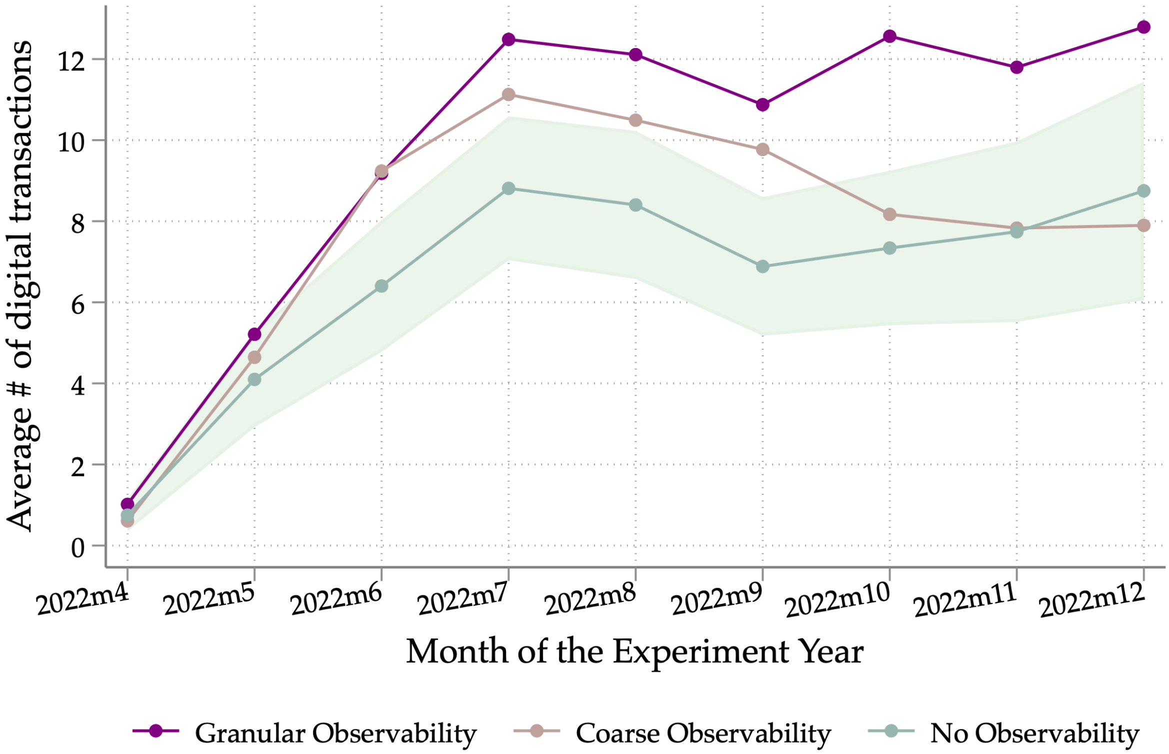


Figure A9: Transaction History of a Driver Under Different Levels of Observability

Notes: This figure shows the total number of digital transactions made by a driver in each of the treatment groups. It includes drivers not using the app (approximately 14%), with zeros retained to allow for better comparison. The data is administrative data provided by the mobile money company among each user. It shows both the increase in digital transactions over time as the technology gets more adopted and the difference in digital usage across groups. 95% confidence intervals use the standard error of the mean.

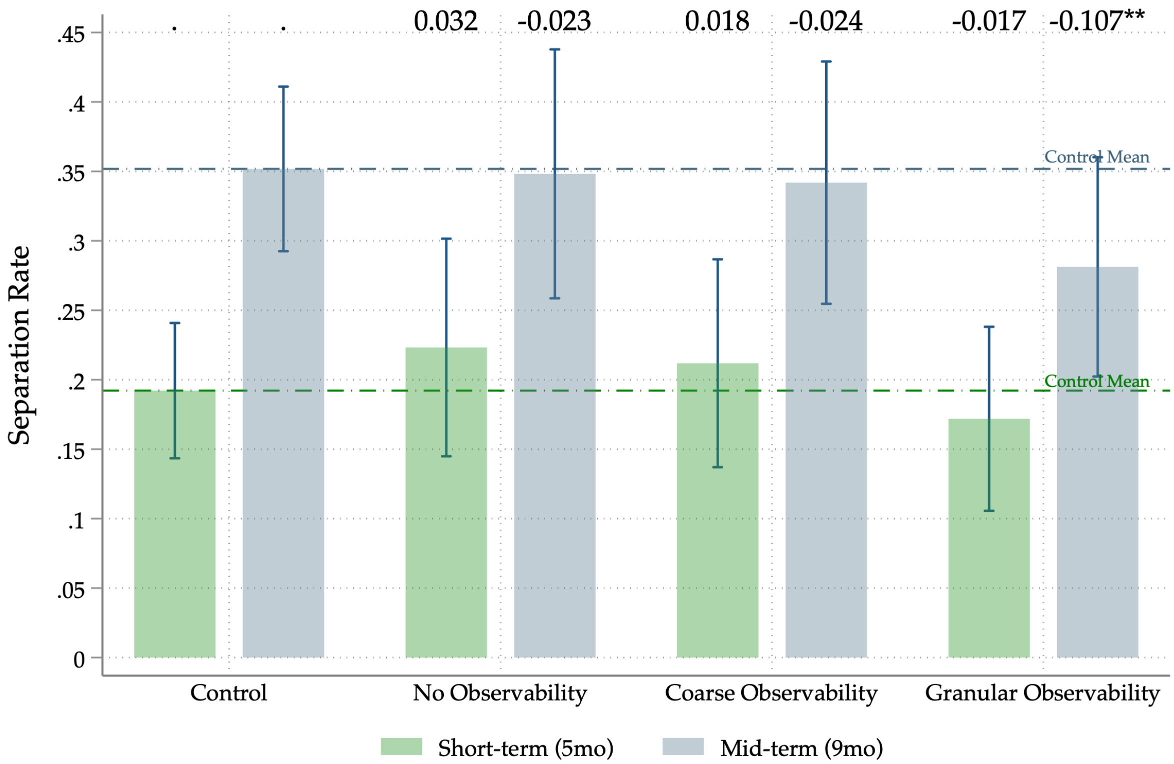


Figure A10: Impact of Observability on Owner-Driver Separation p

Notes: Business-level regression analyses assess the impact of varying levels of observability on the separation rate, y_j. The model is defined as $y_j = \beta_0 + \beta_1 T_j^{GranularObs} + \beta_2 T_j^{CoarseObs} + \beta_3 T_j^{NoObs} + \alpha_s + \epsilon_j$, where y_j indicates whether the owner and the driver work together at the time of the survey. These regressions are conducted separately for short-term (approximately 5 months, indicated in green) and mid-term periods (approximately 9 months, indicated in blue). The coefficients β_1, β_2, and β_3 are displayed above each respective bar, for each respective treatment arm. Each regression includes controls for strata fixed effects. These results include heteroskedasticity-robust standard errors and the significance of each coefficient is denoted with its significance level: $p < 0.1$ *, $p < 0.05$ **, $p < 0.01$ ***. A horizontal dotted line shows the control group's mean outcome, serving as the baseline comparison.

A.6 Long-Term Driver Turnover Rates

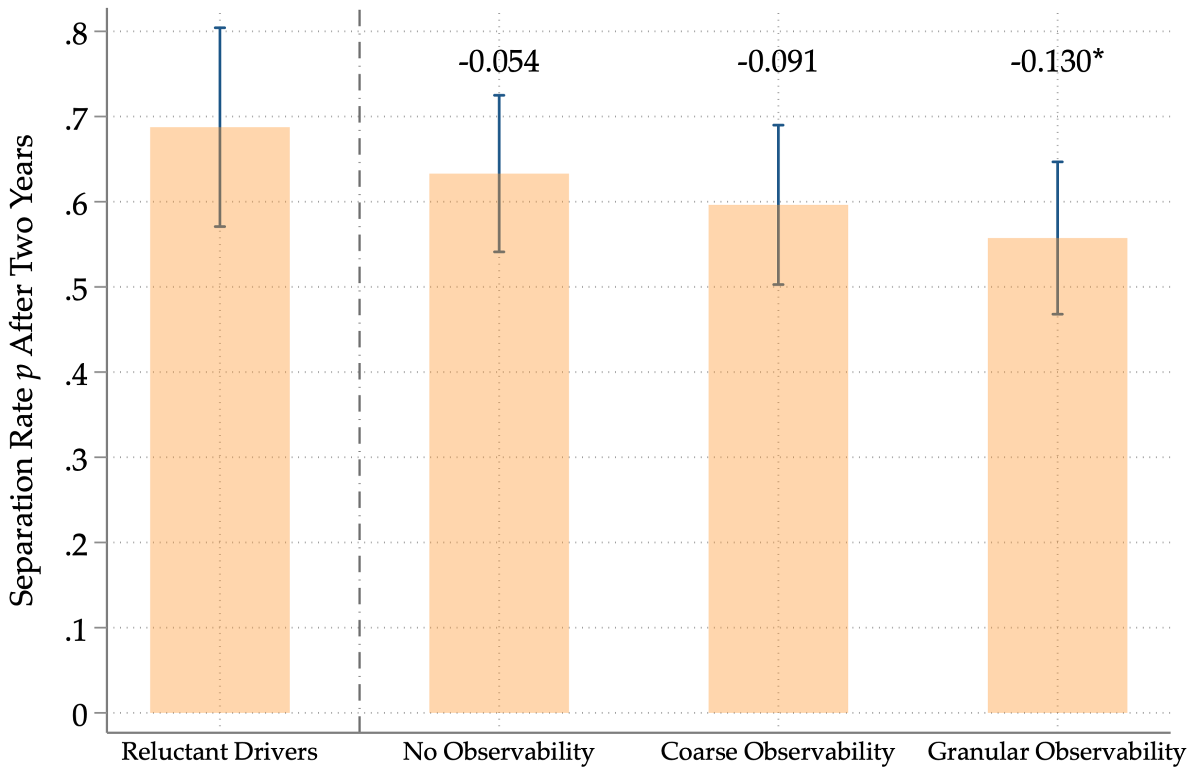


Figure A11: Separation Rates After Nearly Two Years Across Experimental Groups

Notes: This figure shows the separation rates or shares of employment turnover across the different experimental groups. 95% confidence intervals are displayed. From the left bar in order, the first group defines the "low-types" drivers who refused to give their employers' contact in the first place due to concerns regarding transaction observability. The intuition is as follows: the low-type drivers should be the ones that refused to give the owner's contact due to fears of observability (self-reported by drivers), while the ones who refused for other reasons (self-reported by drivers) are less likely to be low-types. Drivers who refused due to observability concerns tend to have shorter relationships (63% vs. 56%). In addition, owner-driver pairs randomized into *Granular Observability* tend to stay together significantly more. The model is defined as $y_j = \beta_0 + \beta_1 T_j^{GranularObs} + \beta_2 T_j^{CoarseObs} + \beta_3 T_j^{NoObs} + \alpha_s + \epsilon_j$, where y_j indicates the separation rate after almost two years, and the omitted category being the reluctant drivers. The coefficients displayed above each respective bar are $\beta_1, \beta_2, \beta_3$, for each respective treatment arm.

A.7 Structural Input – Weekly Collection

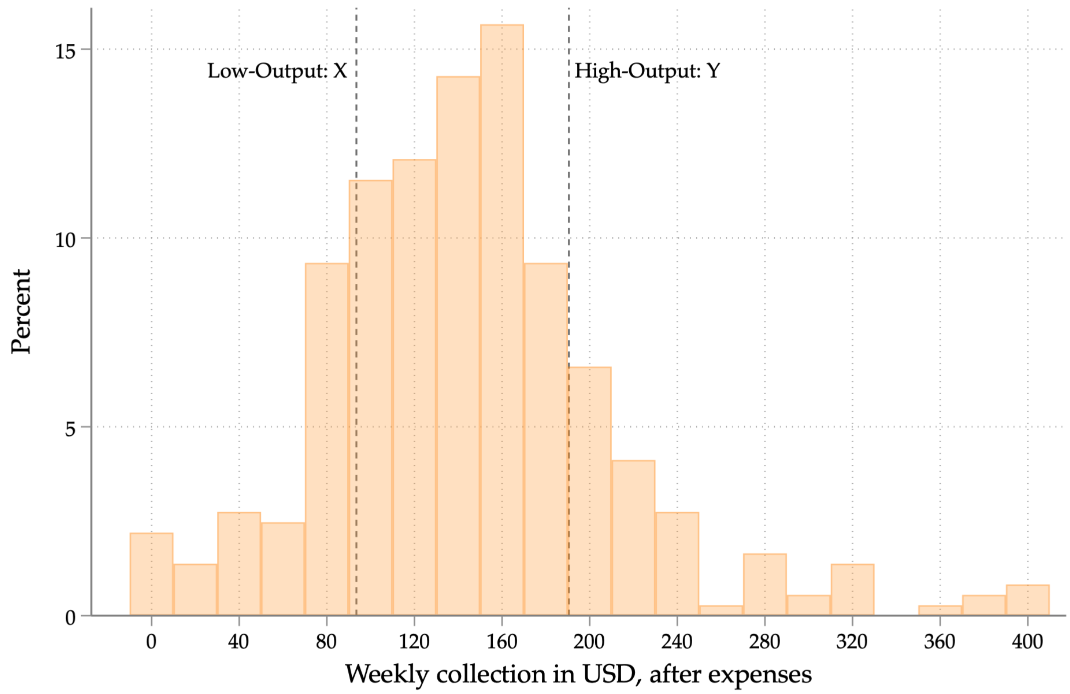


Figure A12: Structural Inputs: Weekly Collection

Notes: This figure shows the distribution of the weekly collection of drivers after removing running expenses such as fuel, maintenance cost paid by the drivers, food and beverages during the day, and any other work expenses (association fees, etc.). To increase precision, I consider the revenue and expenses from the last day of work and multiplied by the median number of days worked in a week. I define high and low output, Y and X, as the average output above or below median output when working more than the median hours: these are displayed as vertical dotted lines.

A.8 Sensitivity of Parameter Estimates to Estimation Moments

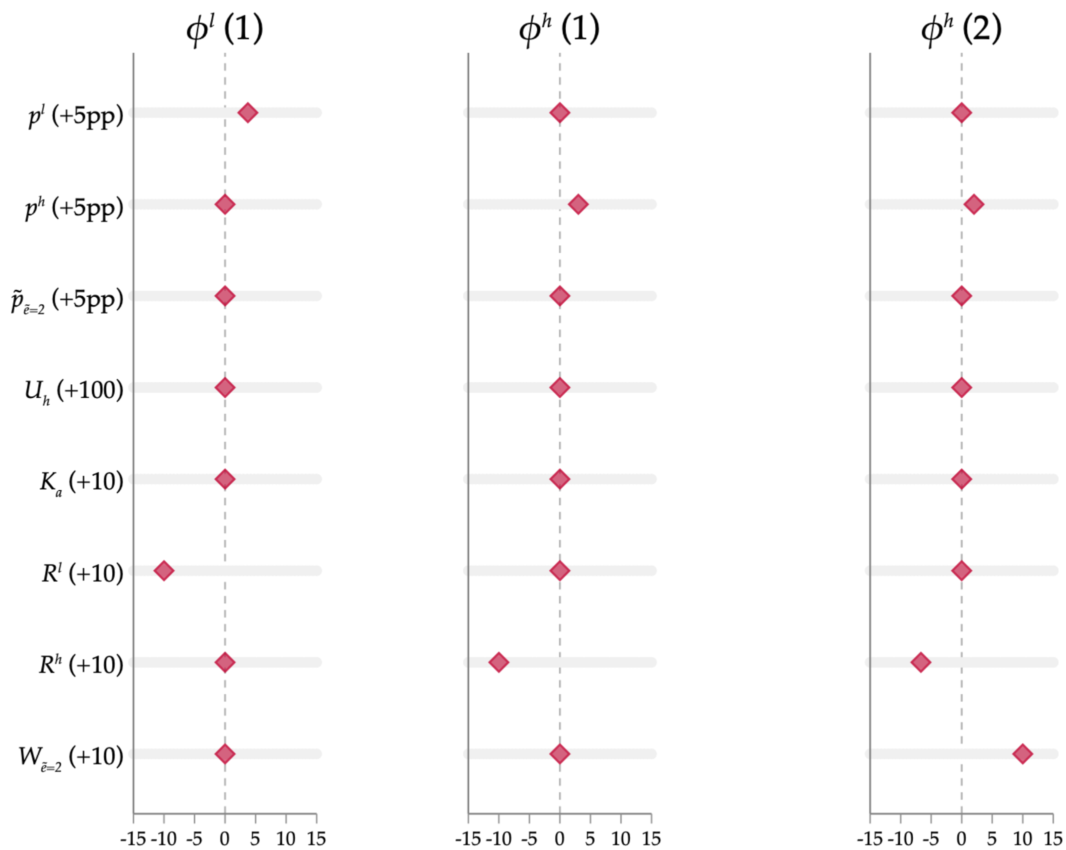


Figure A13: Sensitivity Λ in USD of Parameter Estimates to Estimation Moments

Notes: This figure plots estimated values of $\Lambda = (\mathbf{J}'\mathbf{W}\mathbf{J})^{-1}\mathbf{J}'\mathbf{W}$, where $\mathbf{J}$ is the 8×3 Jacobian matrix of derivatives of the 8 moments with respect to each of the 3 theoretical parameters ϕ_1^l, ϕ_1^h, and ϕ_2^h, each represented in a different panel, and $\mathbf{W}$ is the weighting matrix. It follows the methodology proposed by Andrews et al. (2017) to measure the sensitivity of parameter estimates to moments. Columns of Λ show the sensitivity, in dollars, of each parameter estimate to a unit change in each moment (the rows of Λ), displayed on the left y-axis. The values are transformed to facilitate interpretation of changes in moment values, as indicated in parentheses (e.g., +5pp).

A.9 Counterfactual Analyses in a Single Graph

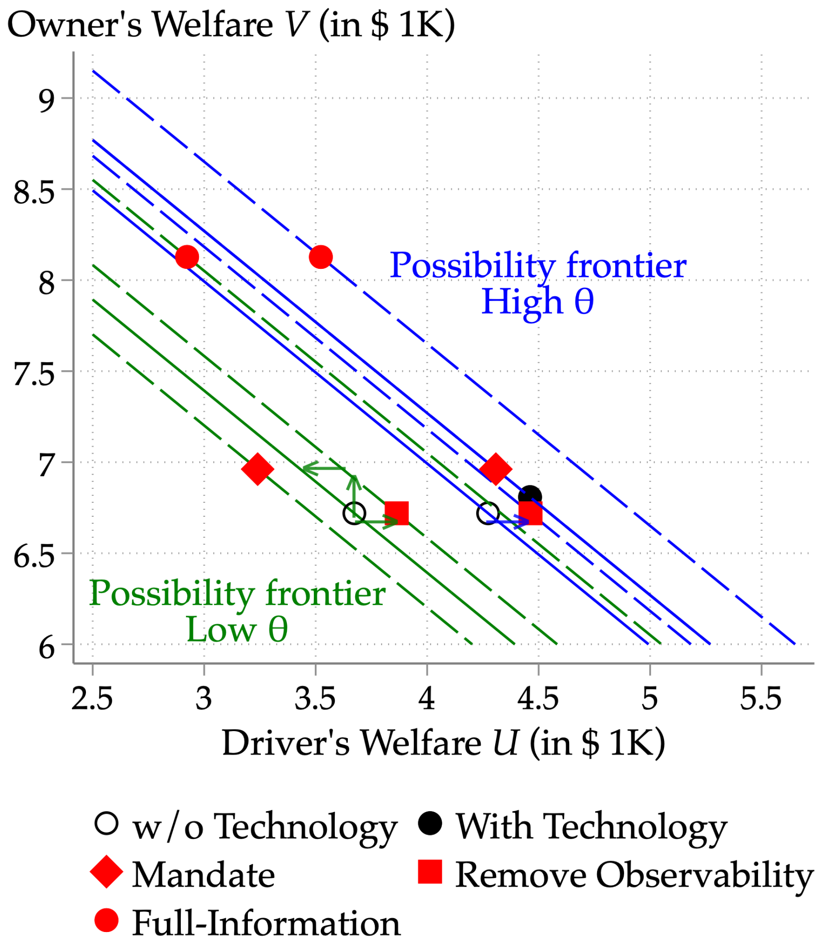


Figure A14: Owner's and Driver's Welfare Under Different Counterfactuals—All

Notes: This figure presents counterfactual analyses by plotting the contract valuations (welfare) for both the owner and the driver on the same graph. All counterfactual analyses are plotted on the same graph. The solid lines depict the utility possibility frontiers for low- and high-type drivers, before and after the introduction of the technology. Dashed lines represent the utility frontiers under various counterfactual scenarios. The graph is re-scaled to zoom into the area of interest. In constructing these graphs, I assume a social planner maximizing total welfare, defined as the sum of the owner's and driver's welfare.
Panel (a) represents the baseline contract (without digital payments) with the *Granular Observability* group (with digital payments). In Panel (b), I analyze the effect of mandating digital payment adoption, which requires both low- and high-type drivers to adopt the technology. Panel (c) examines a counterfactual where the technology is redesigned to remove the observability feature, allowing both types of drivers to adopt. Panel (d) explores a full-information benchmark, where the technology remains unchanged, is universally adopted, and fully reveals the driver's effort and output level.

B Additional Tables

B.1 Randomization Balance Tables

Table B1: Balance Table - Experimental Sample of Taxi Businesses

	Control (1)	Treatment (2)	t-stat (3)	N (4)
Attrition Between Listing and Baseline	0.24 (0.43)	0.26 (0.44)	(1.02)	3026
Panel A. Taxi Businesses				
Owners Not Driving Their Taxi	0.17 (0.38)	0.16 (0.37)	(-0.51)	2269
Owners Driving Their Taxi	0.49 (0.50)	0.52 (0.50)	(1.32)	2269
Taxi Drivers (Non-Owners)	0.29 (0.45)	0.28 (0.45)	(-0.79)	2269
Part of a Taxi Association	0.38 (0.49)	0.41 (0.49)	(1.05)	2269
Daily Hours Worked	10.61 (2.56)	10.61 (2.65)	(0.03)	2269
Panel B. Individual Characteristics				
Male Respondent	1.00 (0.07)	1.00 (0.05)	(0.84)	2269
Education Level: Less Than Primary	0.67 (0.47)	0.68 (0.47)	(0.55)	2269
Literacy (Read And Write)	0.70 (0.46)	0.70 (0.46)	(-0.11)	2269
Senegalese National	0.93 (0.26)	0.95 (0.22)	(1.70*)	2269
Wealth Index - PPI Poverty Line 2011	62.83 (18.77)	63.94 (16.88)	(1.39)	2269
Saved Money In The Past 3 Months	0.57 (0.50)	0.56 (0.50)	(-0.40)	2269
Household Head	0.88 (0.33)	0.87 (0.34)	(-0.67)	2269
Panel C. Mobile Money Characteristics				
# of Mobile Money Connections (In & Out)	58.85 (75.42)	53.74 (57.64)	(-1.68*)	2269
# of Personal Mobile Money Transfers Per Connection	2.57 (1.45)	2.53 (1.68)	(-0.51)	2269
Number of Obs	920	1349		2269

Notes: Treatment: The following regression is run: $Y_i = \alpha + \beta T_i^{Access} + \epsilon_i$. Heteroskedasticity-robust standard errors are clustered at the business level. The t-Test is reported with the following significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All variables are collected during the baseline survey, and missing responses (refused to answer or don't know) are dummied out from the regression. The PPI Index is an aggregated wealth index specific to Senegal, as described in Poverty Probability Index (PPI).

Table B2: Balance Table - Experimental Sample of Owner-Driver Pairs

	G-O (1)	O-O (2)	N-O (3)	Control (4)	F-test p-value (5)	N (6)
Attrition Between Listing and Baseline	0.30 (0.46)	0.31 (0.46)	0.34 (0.48)	0.29 (0.45)	(0.67)	881
Panel A. Business Setup and Contract						
Owners Not Driving *	0.65 (0.48)	0.55 (0.50)	0.58 (0.50)	0.61 (0.49)	(0.43)	613
Owns Only One Taxi *	0.94 (0.23)	0.88 (0.32)	0.93 (0.26)	0.93 (0.26)	(0.28)	613
Long Relationship (> 2 Years) *	0.45 (0.50)	0.47 (0.50)	0.39 (0.49)	0.46 (0.50)	(0.59)	613
Age of the Relationship	3.53 (4.28)	3.25 (3.60)	2.92 (3.64)	3.62 (4.25)	(0.46)	613
Proxy for Risk Aversion (Lining in Garages) *	0.11 (0.32)	0.09 (0.28)	0.11 (0.32)	0.10 (0.30)	(0.90)	613
Upfront Payment / Salary W	0.54 (0.50)	0.50 (0.50)	0.49 (0.50)	0.55 (0.50)	(0.64)	613
Upfront Payment / Salary Value W - Unconditional	39.82 (38.26)	40.88 (47.67)	36.36 (38.52)	43.35 (45.30)	(0.55)	613
Weekly Rent Target Value R	102.92 (11.40)	105.64 (11.18)	105.02 (9.34)	104.21 (11.71)	(0.26)	613
Family Business	0.52 (0.50)	0.53 (0.50)	0.54 (0.50)	0.47 (0.50)	(0.52)	613
Panel B. Driver's Effort						
End-Start Work Time	12.05 (2.85)	12.59 (2.84)	12.53 (3.08)	12.63 (2.93)	(0.34)	613
Daily Hours Worked	10.38 (2.63)	10.79 (2.59)	10.87 (2.78)	10.94 (2.58)	(0.30)	613
Driver Defaulted at Least Once in the Past Month	0.40 (0.49)	0.46 (0.50)	0.50 (0.50)	0.51 (0.50)	(0.22)	613
Avg Daily Revenue Collected	48.14 (9.52)	47.98 (9.71)	48.07 (9.97)	48.17 (9.60)	(1.00)	613
Avg # of Daily Customers	15.77 (5.56)	15.76 (5.07)	16.86 (6.52)	16.58 (6.12)	(0.36)	613
Panel C. Driver's Characteristics						
Education level: less than primary	0.69 (0.46)	0.70 (0.46)	0.74 (0.44)	0.68 (0.47)	(0.79)	613
Wealth Index - PPI Poverty Line 2011	64.56 (16.94)	65.20 (15.76)	63.70 (16.10)	63.50 (17.24)	(0.81)	613
High Digital Users (> 6 Taxi-like Transactions) *	0.52 (0.50)	0.53 (0.50)	0.43 (0.50)	0.47 (0.50)	(0.32)	613
Number of Obs	128	118	112	255		613

Notes: This table compares the business and driver's baseline characteristics across observability treatment groups. The analysis shows no significant imbalances between the groups. The treatment consisted of providing a digital payment technology to the taxi owner and their driver(s). Among the treated pairs, the following observability treatments were randomized. G-O: Granular Observability // C-O: Coarse Observability // N-O: No Observability of the owner on their driver's transactions. p-value from the joint F-Test ANOVA are reported with the following significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
The top row shows the attrition rate among owner-driver pairs formed either during the listing or at the baseline survey.
Variables used to stratify are described in Section 3.4, indicated with a star * in the Table. Digital personal transactions are observed in the administrative data. All other variables are collected during the baseline survey and averaged, and missing responses (refused to answer or don't know) are dummied out from the joint F-test. The PPI Index is an aggregated wealth index, specific to Senegal, to measure poverty likelihood, as described in Poverty Probability Index (PPI). All values are converted to USD (USD 1 = CFA 600).

B.2 Transaction Cost of Using Cash and Additional Impacts of Digital Payments

Table B3: Baseline Costs of Using Cash

	Statistic
	(1)
Had any time lost due to small change in a typical week	0.86
	(0.35)
Number of times significant time was lost due to small change in a typical week	1.30
	(1.93)
Had any customers refused due to small change in a typical week	0.60
	(0.49)
Number of customers refused due to small change in a typical week	2.17
	(2.68)
Had any instances of reduced price due to small change in a typical week	0.92
	(0.26)
Number of times price was reduced due to small change in a typical week	1.34
	(2.10)
Lost any money due to change mistakes in a typical week	0.41
	(0.49)
Imputed loss due to small change in a typical week (USD)	6.66
	(4.64)
Average weekly profit (USD)	97.81
	(47.75)
Percentage of profit lost due to small change (imputed)	0.09
	(0.08)
Observations	1891

Notes: Baseline survey data collected from March to May 2022. Survey data collected about a typical week: at baseline, we did not collect data for the past 7 days due to the specific timing of the baseline survey administration, i.e., during the month of Ramadan, where taxi activity tends to be lower. Values are converted to USD (USD 1 = CFA 600). The following questions were asked, in order displayed in the Table:

- *How many times have you wasted time (more than 10 minutes) looking for small-change during your work?*
- *How many customers have you turned down because they only wanted to pay with electronic money, and not cash?*
- *How many times have you reduced the price of the ride because of the small change problem?*
- *Have you lost part of your earnings with customers due to giving change (e.g., through miscalculation)?*
- *Over the past 3 months, on average, how much money do you keep each day during a normal workday, after paying all your expenses (fuel, transfer, repairs, police, contributions, food, etc.)?*
- Imputed loss is calculated by assigning monetary values to each reported cost, based on fieldwork conducted with a subset of drivers before the experiment. Time lost and mistakes in giving change are estimated at CFA 500 (USD 0.8), refused customers at CFA 1,500 (USD 2.5), and reduced price at CFA 800 (USD 1.3). The share is obtained by dividing this imputed loss by the average profit.

B.3 Stylized Facts about Taxi Owner-Driver Relationships

Table B4: Baseline Owner-Driver Relational Contracts

	Share (1)
Panel A. Owner-Driver Contract	
Agreed Rental Target Transfer	0.96
	(0.20)
Upfront Payment / Salary Provided at Baseline	0.53
	(0.50)
Upfront Payment / Salary Provided After Nine Months	0.73
	(0.45)
Owners' Reasons For Providing The Upfront Payment / Salary:	
Ensure Minimum Income to Driver for Their Work	0.84
	(0.37)
Encourage Driver to Take Fewer Risks on the Road	0.41
	(0.49)
Upfront Payment / Salary Adjusted Down if Insufficient Transfer is Provid	0.15
	(0.36)
Retain Driver - Poaching Considerations	0.11
	(0.31)
Owner Responsible for Maintenance Costs	0.87
	(0.33)
Panel B. Default on Transfer	
Default at Least Once a Month	0.48
	(0.50)
Any Payment Default in Past 3 Months	0.70
	(0.46)
Owner Sees Transfer Default as Conflict Cause	0.65
	(0.48)
Driver Stressed Over Transfers in Past 3 Months	0.48
	(0.50)
Panel C. Limited Liability	
Savings: Drivers Able to Save in Past 3 Months	0.45
	(0.50)
Loan: Drivers Obtained a Loan in Past 3 Months	0.08
	(0.27)
Observations	613

Notes: Survey data collected at baseline, except for the reasons for providing a salary / upfront payment, collected at mid-term and supplemented by short-term data when missing, focusing on the subset of owners who provide a salary to their drivers. This particular question allowed for multiple responses and was framed as follows: 'We know that some taxi owners pay a salary, while others do not. We want to understand, in your case, what are the reasons why you pay a salary to your driver?' The share of respondents selecting each response is displayed in the table. All the questions on contract terms were asked to both owners and drivers separately, and we checked the consistency. While discrepancies between owner and driver reports were limited, any inconsistencies prompted a third review by a senior field coordinator to determine its cause. For savings and loans, only substantial savings and loans are included, defined as those above CFA 50,000 (USD 80).

Table B5: Owner's Beliefs About Driver's Work

	Accurate (1)	Overestimate (2)	Underestimate (3)
Number of Days Worked in a Week	0.46	0.21	0.33
	(0.50)	(0.41)	(0.47)
Avg Working Hours (+/- 2)	0.39	0.33	0.49
	(0.49)	(0.47)	(0.50)
Avg Daily Revenue Collected (+/- USD 5)	0.26	0.34	0.41
	(0.44)	(0.48)	(0.49)
Avg # of Customers (+/- 3)	0.27	0.38	0.44
	(0.45)	(0.49)	(0.50)
Taxi Parked Away From Owner's Home	0.68		
	(0.47)		
Use a GPS Tracker	0.01		
	(0.08)		
Observations	404		

Notes: Survey data was collected at baseline, except for the question on the GPS tracker, which was collected at endline. All belief questions followed this structure: *'Over the past 3 months, how many [days] has your driver worked on average per week?'* These belief questions were answered only by a subset of owners to limit survey length and increase response rates among owners, hence the reduced number of observations. The owner's answers were then matched to their driver's self-reported performance and compared.
In particular, the first column indicates whether the owner was accurate about their driver's work performance (within some reasonable specified ranges for discrepancy). Columns 2 and 3 report whether the owner over-estimated (believed the driver worked more days than reported) or under-estimated (believed the driver worked fewer days than reported) the driver's work. For these two columns, no ranges for discrepancy are considered, but simply compare whether the owner's report is above or below their drivers.

Table B6: Mapping From Effort to Output

	Daily Output			
	(1)	(2)	(3)	(4)
Driver's Hours			2.448***	2.381***
Worked That Day			(0.063)	(0.074)
# of Taxis	1726	1726	1726	1726
Observations - Days	9521	9521	9498	9498
Adjusted R2	0.008	0.015	0.215	0.453
Days of the Week / Days of the Month FEs	Yes	Yes	Yes	Yes
Calendar Date FEs	No	Yes	Yes	Yes
Hours Worked	No	No	Yes	Yes
Driver FEs	No	No	No	Yes

Notes: This tables displays different regressions to predict daily output based on driver's characteristics and effort on a given day. Heteroskedasticity-robust standard errors are provided in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Values are converted into USD (USD 1 = CFA 600). 'Days of Week/Month FEs' specifications control for the day of the week (e.g., Monday) and the day of the month (e.g., the 10th), as these are sometimes considered to influence part of the demand variation in this setting. The location is the same for all drivers as they all operate in Dakar, Senegal.

Table B7: Owner-Driver Separations at Mid-Term

	Share
	(1)
Owner-driver Separated at Mid-term	0.34
	(0.47)
Reasons for Owner-Driver Separation:	
Driver Decided to Leave	0.29
	(0.46)
Driver Was Fired	0.21
	(0.41)
Owner Sold the Taxi	0.20
	(0.40)
Taxi in Maintenance for at Least 2 Months	0.13
	(0.33)
Other Reasons Not Listed Above/Refused to say	0.15
	(0.36)
Total Number of Owner-Driver Pairs Surveyed Mid-Term	563

Notes: Survey data was collected at the time of separation. The question asked to the owner and the driver was as follows: *Could you explain why this person X is no longer your driver/owner?* Enumerators selected answers from a pre-defined list, but open-ended reasons were also collected. In cases of disagreement between the owner and the driver, the owner's response was considered. If one of the two responses was missing, the non-missing response was used.

B.4 Non-Differential Attrition Rates Across Observability Treatments

Table B8: Non-Differential Attrition Rates Across Observability Treatments

	G-O (1)	O-O (2)	N-O (3)	Control (4)	F-test p-value (5)	N (6)
Short-Term Survey	0.01	0.00	0.00	0.00	(0.29)	613
	(0.09)	(0.00)	(0.00)	(0.00)		
Mid-Term Survey	0.07	0.07	0.05	0.05	(0.85)	613
	(0.26)	(0.25)	(0.23)	(0.22)		
Long-Term Survey	0.05	0.08	0.03	0.04	(0.20)	613
	(0.21)	(0.28)	(0.16)	(0.20)		
Number of Obs	128	118	112	255		613

Notes: This table compares the attrition rate within each survey round across observability treatment groups. The analysis shows no differential attrition across groups for all survey rounds. The treatment involved providing a digital payment technology to the taxi owner and their driver(s). Among the treated pairs, the following observability treatments were randomized: G-O (Granular Observability), C-O (Coarse Observability), and N-O (No Observability of the owner on their driver's transactions). p-value from the joint F-Test ANOVA are reported with the following significance level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B9: Impact of Digital Payments on Driver's Own Revenue and Profit

	Avg Profit (1)	Daily Profit (2)	Revenue (3)	# Customers (4)	Avg Price (5)	Customer/Hour (6)	Revenue/Hour (7)
Panel A. Short-Term 5-Month Survey							
Technology Access	-0.107	0.320	-0.278	-0.449*	0.038	0.045	0.036
	(0.385)	(0.725)	(0.605)	(0.253)	(0.193)	(0.067)	(0.082)
Observations	1714	1714	1714	1714	1714	1714	1714
Control Mean at Short-Term	14.37	17.77	51.38	15.93	3.94	1.65	5.37
% Change T at Short-Term	-0.75	1.80	-0.54	-2.82	0.97	2.74	0.66
Panel B. Mid-Term 9-Month Survey							
Technology Access	-0.395	-0.690	-0.343	-0.045	-0.266	-0.035	-0.133*
	(0.391)	(0.725)	(0.572)	(0.200)	(0.196)	(0.024)	(0.075)
Observations	1674	1674	1674	1674	1674	1674	1674
Control Mean at Mid-Term	19.31	22.42	50.72	14.51	4.24	1.52	5.29
% Change T at Mid-Term	-2.04	-3.08	-0.68	-0.31	-6.28	-2.34	-2.51

Notes: Baseline survey data collected from March to June 2022. Short-term survey data collected from July-September 2022. Mid-term survey data collected from October-November 2022 (after 9 months). Driver-level regressions of the outcomes on the treatment (access to the technology), with the omitted category the pure control group. The model is the following: $y_{ij} = \beta_0 + \beta T_{ij}^{Access} + \epsilon_{ij}$, where y_{ij} is the outcome variable displayed at the top of the column, with i the driver in taxi business j. Cluster-robust standard errors at the business level j, and the significance of each coefficient is denoted by asterisks: p < 0.1 *, p< 0.05 **, p< 0.01 ***. The percent change between dividing the coefficient by the control mean is also reported. The sample includes all drivers that were surveyed at least once at short and/or mid-term — the case for about 90% of the sample — and replaced with missing and dummied out otherwise. Each regression includes controls for strata fixed effects and baseline controls when available. The outcomes, except in Column 1, are averaged from the last 3 days prior to the survey date. These are constructed from the following questions:
- *Avg Profit*: Over the past 30 days, what is the average amount of money you earn per working day, after paying all your expenses (fuel, payments, repairs, police, contributions, food, etc.)? — a la De Mel et al. (2009).
-*Daily Profit*: For each day, profit is computed from the taxi drivers' production function, which is composed of the revenue collected, coming from a unique source (passengers), minus each cost category, listed as follows: (i) fuel, (ii), transfer to the owner, (iii) side expenses (small repairs, police, washing, toll, association contribution excluding food), and (vi) food/drinks/stimulants throughout the working days.
- *Revenue*: For each day: On that day, over working hours, how much money did you collect (from passengers' payment)?
-*Customers*: For each day: On that day, over working hours, how many customers did you have?
- *Average Price*: Daily collection divided by the number of customers.
- *Customer/Hour*: Drivers' productivity in terms of number of customers: number of customers divided by the working hours (start-end).
- *Revenue/Hour*: Drivers' productivity in terms of revenue: total revenue collected divided by the working hours (start-end).
All values are converted to USD when relevant (USD 1 = CFA 600).

Table B10: Impact of Digital Payments on Revenue and Profit - Power Calculations and Minimum Detectable Effects (MDE)

	Avg Profit (1)	Daily Profit (2)	Revenue (3)	# Customers (4)	Avg Price (5)	Customer/Hour (6)	Revenue/Hour (7)
Panel A. Short-Term 5-Month Survey							
MDE (in %)	9.59	12.05	3.45	4.80	17.36	11.45	4.45
Observations	1714	1714	1714	1714	1714	1714	1714
Control Mean$_{Avg}$	14.35	17.71	30855.29	16.00	3.91	1.65	5.36
Panel B. Mid-Term 9-Month Survey							
MDE (in %)	6.39	9.59	3.48	4.38	16.21	5.02	4.25
Observations	1674	1674	1674	1674	1674	1674	1674
Control Mean$_{Avg}$	19.33	22.44	30452.81	14.53	4.28	1.52	5.31

Notes: This table computes the minimal detectable effects (MDEs) for the profit outcomes of interest for both survey rounds. It performs a simple two-sample means test between the control and treatment group to quantify the power of the experiment on the profit dimensions, using the mean and standard deviation of the control and treatment groups. All variables, except in column 1, are averaged at the daily level from the last 3 days prior to the survey, as in the main results. MDEs were computed using analytical methods for a power of 0.8 and a significance level of 0.05.
The outcomes are constructed from the following questions:
- *Avg Profit*: Over the past 30 days, what is the average amount of money you earn per working day, after paying all your expenses (fuel, payments, repairs, police, contributions, food, etc.)? — a la De Mel et al. (2009).
-*Daily Profit*: For each day, profit is computed from the taxi drivers' production function is composed of the revenue collected, coming from a unique source (passengers), minus each cost category can be listed as follows: (i) fuel, (ii), transfer to the owner, (iii) side expenses (small repairs, police, washing, toll, association contribution excluding food), and (vi) food/drinks/stimulants throughout the working days.
- *Revenue*: For each day: On that day, over working hours, how much money did you collect (from passengers' payment)?
-*Customers*: For each day: On that day, over working hours, how many customers did you have?
- *Average Price*: Daily collection divided by the number of customers.
- *Customer/Hour*: Drivers' productivity in terms of number of customers: number of customers divided by the working hours (start-end).
- *Revenue/Hour*: Drivers' productivity in terms of number of revenue: total revenue collected divided by the working hours (start-end).
All values are converted to USD when relevant (USD 1 = CFA 600).

Table B11: Impact of Digital Payments on Additional Outcomes

	Theft Anxiety (1)	Keep Records (2)	Luxury Purchases (3)	Able to Save (4)
Panel A. Short-Term 5-Month Survey				
Technology Access	-0.042*	0.051***	-0.060**	-0.013
	(0.023)	(0.009)	(0.024)	(0.023)
Observations	1714	1714	1714	1714
Control Mean at Short-Term	0.76	0.02	0.45	0.60
% Change T at Short-Term	-5.51	216.17	-13.26	-2.12
Panel B. Mid-Term 9-Month Survey				
Technology Access	-0.047*	0.034***		-0.019
	(0.026)	(0.009)		(0.022)
Observations	1674	1674		1674
Control Mean at Mid-Term	0.57	0.02		0.65
% Change T at Mid-Term	-8.18	137.48		-2.93

Notes: Baseline survey data collected from March to June 2022. Short-term survey data collected from July-September 2022. Mid-term survey data collected from October-November 2022 (after 9 months). Driver-level regressions of the outcomes on the treatment (access to the technology), with the pure control group as the omitted category. The model is the following: $y_{ij} = \beta_0 + \beta T_{ij}^{Access} + \epsilon_{ij}$, where y_{ij} is the outcome variable displayed at the top of the column, with i the driver in taxi business j. Cluster-robust standard errors at the business level j, and the significance of each coefficient is denoted by asterisks: $p < 0.1$ *, $p < 0.05$ **, $p < 0.01$ ***. The percent change between dividing the coefficient by the control mean is also reported. The sample includes all drivers that were surveyed at least once at short and/or mid-term — the case for about 90% of the sample — and replaced with missing and dummied out otherwise. Each regression includes controls for strata fixed effects and baseline controls when available.

All outcomes are dummies (0-1) constructed from the following questions:

- *Theft Anxiety*: In the last 3 months, how often are you worried that part of your earnings might be stolen? — Dummy equal to 1 from Always (every day) to sometimes. 0 if almost never and never.
- *Keep Records*: During the last 3 months, have you kept a written or digital record of your transactions for accounting purposes?
- *Luxury Purchases*: Consider the past 7 days: did you buy perfumes, deodorants, clothes, or pillows while driving or working?
- *Able to Save*: During the last 3 months, how much have you been able to save? — A dummy is set to 1 if more than CFA 100,000 (USD 170) was saved, which is considered a significant amount. Robustness checks also performed with other values show no effect.

Table B12: Baseline Transaction Costs Predict Willingness-To-Pay Outcome: Willingness-To-Pay

	log(Baseline Willingness-To-Pay - WTP)				
	(1)	(2)	(3)	(4)	(5)
TC	0.175***	0.109**	0.268***	0.179***	0.054**
	(0.043)	(0.045)	(0.043)	(0.044)	(0.026)
Obs	1702	1702	1702	1702	1702
Benchmark Good Control	Yes	Yes	Yes	Yes	Yes
TC Mean	0.480	0.602	0.533	0.412	1.601
$TC =$	*Any Time Lost*	*Refused Customers*	*Reduced Price*	*Mistakes Giving Change*	*log(Self-Reported Loss)*

Notes: Baseline survey data collected from March to May 2022 about the past 7 days. Standard errors (SE) clustered at the business level. In all regressions, we control for WTP for a benchmark good, i.e., a bottle of water in this study, as recommended in Dizon-Ross and Jayachandran (2022). The outcome is the willingness-to-pay for the technology. Driver's willingness-to-pay was elicited at baseline following the BDM procedure in an incentivized manner (Becker et al., 1964). To preserve the randomization treatment arms, the lottery was run on 5% of the treated sample only drawing a random number from an uniform distribution such that most of the 5% of treated drivers actually were given the product below their WTP. For interpretability, the WTP outcome is converted into log, to be interpreted as percentage change for one unit increase in the independent variable, approximately. The last column (5) is log-log, hence the percent increase in the dependent variable for every 1% increase in the independent variable. Missing values are dummied out.

The regressors TC (Transaction Costs), displayed at the bottom of the table, are dummies (0-1) are coming from survey data about the past 7 days prior to the survey date. They are constructed from the following questions:

- *Any Time Lost*: How many times have you wasted time (more than 10 minutes) looking for small-change during your work?
- *Refused Customers*: How many customers have you turned down because they only wanted to pay with electronic money, and not cash?
- *Reduced Price*: How many times have you reduced the price of the ride because of the small change problem?
- *Mistakes Giving Change*: How many times have you lost part of your earnings with customers due to giving change?
- *Self-Reported Loss*: Because of all these cash-related problems, how much money do you estimate you lost? Converted into a log scale for interpretability.

All values are converted in USD when relevant (USD 1 = CFA 600).

B.5 First-Stage: Observability Impact on Information Frictions

Table B13: First-Stage: Observability Impact on Information Frictions

	Owner's Knowledge				
	Work Days (Cross-Checked)	Hours Worked	Digital Collection	Use SMS To Observe Effort	Technology Monitoring Daily/Weekly
	(1)	(2)	(3)	(4)	(5)
Granular Observability	0.083	0.067	0.148**	0.341***	0.435***
	(0.065)	(0.062)	(0.070)	(0.047)	(0.074)
Coarse Observability	0.047	-0.009	0.062	0.205***	
	(0.063)	(0.064)	(0.067)	(0.042)	
Observations	292	292	292	292	46
Baseline Control	YES	YES	YES		
No Observability Mean	0.29	0.59	0.21	0.00	0.00
% Change Granular Observability	28.7	11.4	71.9		

Notes: Sample of owners surveyed at mid-term (around 9 months), all in the groups of drivers with access to the technology. The group of control drivers is not included since their owners were not asked about their (nonexistent) digital collection. Robust standard errors are shown in parentheses with the following significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Owners' knowledge in the 7 days prior to the interview was elicited for each of the outcomes in Columns 1 to 3. Respectively, these columns represent whether the owner claimed to know the driver's effort (work days and hours) and the driver's total revenue collected digitally. Since 92% of owners in the 'no observability' group claimed to know their drivers' days worked, all reports were cross-checked by comparing the days reported by the owners to the actual work days of the drivers (as reported by the drivers in a separate survey). A substantial share of owners were incorrect. I report a dummy for whether the two reports match (i.e., both owners and drivers report the same number of days worked in the past 7 days). Refusals to answer were limited but are dummied out in these regressions.
The last two columns display the share of owners reporting the use of the technology to observe their driver's working hours and days (Column 4), with a dummy equal to 0 for the 'no observability' group due to treatment compliance. The last column shows the share of owners reporting that they check the driver's transactions daily or weekly (and 0 for the 'no observability' and 'coarse observability' groups). It is important to note that the sample size for this particular question is smaller because it was only added at the end of the mid-term data collection survey.

Table B14: Positive Correlation Between Self-Reported Effort and Output And Digital Payments Usage

	Day Worked		# Hours Worked		Total Revenue Collected		# Of Passengers	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
# of Digital Transactions	0.104***		0.213***		2.568***		0.824***	
	(0.004)		(0.080)		(0.487)		(0.181)	
Revenue Collected Digitally		0.028***		0.047**		0.711***		0.129***
		(0.001)		(0.021)		(0.121)		(0.047)
# of Taxis	1041	1041	1041	1041	1041	1041	1041	1041
Mean No Digital Activity	0.75	0.75	9.91	9.91	50.27	50.27	14.89	14.89
Sample of Days Considered	All		Used Digital		Used Digital		Used Digital	

Notes: Regressions are performed at the day level, controlling for day fixed effects and survey timing (mid-term or long-term) for treated taxis. The outcome variables are self-reported by the drivers for the past 3 days prior to the survey, at both mid-term and long-term. Drivers are asked about specific dates, and those self-reports are compared to the actual digital transactions recorded by the mobile money partner company. The two dependent variables are the number of digital transactions on a given day and the total money collected. The first two columns leverage the fact that, by asking drivers to report on the past 3 days worked, the responses also reveal information about days not worked. In Columns 3 to 8, the sample size is restricted to days where at least one digital transaction was made. Robust standard errors are provided in parentheses, with the following significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
The 'mean no digital activity' represents the mean outcome for days with no digital activity (i.e., the number of transactions was zero on those days). Taxi drivers who refused to respond to these questions are excluded. Values are converted from CFA to USD, with CFA 600 = USD 1.

B.6 Impacts of Observability on Pairs Remaining Together

Table B15: Impact of Observability on Contracts — Pairs Remained Together at Mid-Term

	Upfront Payment 'Salary' Dummy	Upfront Payment 'Salary' Value (USD)	Weekly Rent Target Value	Default Rate
	(1)	(2)	(3)	(4)
Granular Observability	0.097**	8.978**	0.119	-0.083
	(0.045)	(3.556)	(1.235)	(0.060)
Coarse Observability	-0.040	-1.430	1.528	0.048
	(0.055)	(4.332)	(1.501)	(0.068)
No Observability	-0.010	0.072	-0.209	-0.005
	(0.052)	(4.133)	(1.520)	(0.066)
Observations	406	406	406	406
Control Mean	0.77	57.11	100.79	0.30
% Change Granular Observability	12.49	15.72	0.12	-27.23
F-test Granular O = No O (p-value)	0.05	0.04	0.83	0.28

Notes: Tables restricted to pairs that remained together at mid-term (endogenous restriction). Business-level regressions of the contract outcomes on the three treatment arms, with the omitted category the pure control group. The model is defined as $y_j = \alpha + \beta_1 T_j^1 + \beta_2 T_j^2 + \beta_3 T_j^3 + \epsilon_j$, where y_j is the outcome variable displayed at the top of the column. These regressions are conducted for mid-term period (approximately 9 months after the baseline survey). Heteroskedasticity-robust standard errors are used and the significance of each coefficient is denoted by asterisks: $p < 0.1$ *, $p < 0.05$ **, $p < 0.01$ ***. Each regression includes strata fixed effects.

- *Upfront Payment 'Salary' Dummy*: Whether the owner provides a monthly upfront fixed payment to their driver (also referred to as a 'salary' in the industry).
- *Upfront Payment 'Salary' Value (USD)*: Value of the monthly upfront fixed payment to the driver, including 0 if no payment is provided. Values converted in USD (USD 1 = CFA 600).
- *Weekly Rent Target Value*: Weekly rental target amount from driver to owner, as agreed between owner and driver. Values converted in USD (USD 1 = CFA 600).
- *Default Rate*: The proportion of drivers who default at least once a month (in the past three months), according to either the owner or their driver's reports.

B.7 Long-Term Impacts of Observability on Relationships

Table B16: Impact of Observability on Contracts and Relationships - Long-Term (Nearly 2 Years)

	Upfront Payment 'Salary' Dummy (1)	Upfront Payment 'Salary' Value (USD) (2)	Weekly Rent Target Value (3)	Separation (4)
Granular Observability	0.125*	8.886	-0.140	-0.061
	(0.075)	(5.953)	(2.037)	(0.067)
Coarse Observability	-0.032	-3.033	3.090	-0.038
	(0.074)	(5.844)	(3.439)	(0.069)
Observations	211	211	211	338
Control Mean	0.73	55.85	99.50	0.63
% Change Granular Observability	17.20	15.91	-0.14	-9.63

Notes: Business-level OLS regressions of the contract outcomes on the three treatment arms, with the omitted category the No Observability group (the control group received the technology after nine months). The model is defined as $y_j = \beta_0 + \beta_1 T_j^{GranularObs} + \beta_2 T_j^{CoarseObs} + \alpha_s + \epsilon_j$, where y_j is the outcome variable displayed at the top of the column and $alpha_s$ are the strata fixed effects. These regressions are conducted at long-term (nearly 2 years after the baseline survey). Heteroskedasticity-robust standard errors are used and the significance of each coefficient is denoted by asterisks: p < 0.1 *, p< 0.05 **, p< 0.01 ***.

- *Upfront Payment 'Salary' Dummy*: Whether the owner provides a monthly upfront fixed payment to their driver (also referred to as a 'salary' in the industry).
- *Upfront Payment 'Salary' Value (USD)*: Value of the monthly upfront fixed payment to the driver, including 0 if no payment is provided. Values converted in USD (USD 1 = CFA 600).
- *Weekly Rent Target Value*: Weekly rental target amount from driver to owner, as agreed between owner and driver. Values converted in USD (USD 1 = CFA 600).
- *Separation*: Owner and driver are not working together at the time of the survey.

B.8 Observability Impacts Robustness: No Contamination Bias With Multiple Treatments

Table B17: Impact of Observability on Contracts — Contamination Bias With Multiple Treatments (Robustness)

	Upfront Payment 'Salary' Dummy (1)	Upfront Payment 'Salary' Value (USD) (2)	Weekly Rent Target Value (3)	Separation (4)
Granular Observability	0.131***	10.771***	0.312	-0.098**
	(0.042)	(3.292)	(1.054)	(0.049)
Coarse Observability	-0.017	0.074	1.717	-0.024
	(0.050)	(3.885)	(1.250)	(0.051)
No Observability	-0.009	0.095	0.271	-0.021
	(0.049)	(3.843)	(1.232)	(0.051)
Observations	479	479	479	577
Control Mean	0.75	55.34	100.60	0.32
% Change Granular Observability	17.39	19.46	0.31	-30.47
Chi-squared test Granular O = No O (p-value)	0.01	0.01	0.98	0.19

Notes: This table estimates the effect of a single as-good-as-randomly assigned treatment using a partially linear model, following Goldsmith-Pinkham et al. (2024), so that β identifies a convex average of heterogeneous treatment effects. This analysis accounts for potential contamination bias in randomized experiments with multiple treatments and strata fixed effects. Results remain unchanged. The table shows business-level regressions of the contract outcomes on the three treatment arms, with the omitted category the pure control group. The outcome variable is displayed at the top of the column. These regressions are conducted for mid-term period (approximately 9 months after the baseline survey). Heteroskedasticity-robust standard errors and the significance of each coefficient is denoted by asterisks: $p < 0.1$ *, $p < 0.05$ **, $p < 0.01$ ***. Each regression considers strata fixed effects.

- *Upfront Payment 'Salary' Dummy*: Whether the owner provides a monthly upfront fixed payment to their driver (also referred to as a 'salary' in the industry).
- *Upfront Payment 'Salary' Value (USD)*: Value of the monthly upfront fixed payment to the driver, including 0 if no payment is provided. Values converted in USD (USD 1 = CFA 600).
- *Weekly Rent Target Value*: Weekly rental target amount from driver to owner, as agreed between owner and driver. Values converted in USD (USD 1 = CFA 600).
- *Separation*: Owner and driver are not working together at the time of the survey.

B.9 Additional Impacts of Observability and Heterogeneity Analyses

Table B18: Drivers Under Observability Have Higher Digital Usage

	# of transactions	# of transactions	Amount (USD)	# of active weeks	# of active days
	(1)	(2)	(3)	(4)	(5)
Granular Observability	0.397***	0.346**	79.054**	0.162*	0.240**
	(0.146)	(0.142)	(38.418)	(0.088)	(0.113)
Coarse Observability	0.165	0.089	7.197	0.085	0.062
	(0.143)	(0.142)	(29.752)	(0.088)	(0.112)
Obs	358	358	358	358	358
Mean No Observability	59.18	59.18	221.20	15.29	38.54
Baseline P2P Control	NO	YES	YES	YES	YES

Notes: Administrative data provided by the mobile money partner company, collected from April to December 2022, the end of the mid-term survey. Zeros (drivers not using the technology at all) are included and account for approximately 15% of drivers. Robust standard errors (HC3) are used. The significance of each coefficient is denoted as follows: $p < 0.1$ *, $p < 0.05$ **, $p < 0.01$ ***. Columns 1, 2, 4, and 5 display coefficients from Poisson regressions (their outcomes are counts), while Column 3 represents an OLS regression. The pure control group sample is excluded because they did not have access to the technology.
The variable # of transactions represents the total number of transactions, while the amount corresponds to the total value of those transactions (in USD). The number of active weeks/days refers to the count of time periods with at least one transaction received.
Baseline P2P Control denotes the number of P2P transactions received before the experiment launch, categorized as within or outside the taxi value ranges. This control is included to mitigate variations caused by potential pre-trends in digital usage.
The taxi value range is defined as CFA 1,000–3,500 (USD 1.5–6), representing the 5th and 95th percentiles of taxi P2B transactions received.

Table B19: Impact of Observability on Self-Reported Effort

	Hours Worked	End-Start Time	Revenue	# Customers
	(1)	(2)	(3)	(4)
Granular Observability	0.662**	0.642**	0.042	0.377
	(0.291)	(0.314)	(1.447)	(0.686)
Coarse Observability	0.200	0.125	-1.956	0.264
	(0.313)	(0.344)	(1.445)	(0.651)
No Observability	0.323	0.422	-0.225	-0.461
	(0.341)	(0.393)	(1.326)	(0.799)
Observations	598	598	598	598
Control Mean	10.20	11.76	50.60	14.88
% Change Granular Observability	6.5	5.5	0.1	2.5
F-test Granular O = No O (p-value)	0.38	0.61	0.86	0.33

Notes: Short-term survey data were collected from July to September 2022, and mid-term survey data from October to November 2022 (after 9 months). The outcome is the self-reported effort by the driver in the past 3 days, averaged before the survey. Difference-in-difference regressions were conducted, controlling for individual fixed effects. Baseline controls include the same variable, but instead of using the past 3 days, they average over the last 3 days worked to avoid specific variations during Ramadan (which occurred during part of the baseline phase). Robust standard errors are used. The sample includes all drivers surveyed at least once at short and/or mid-term, with missing values replaced and dummied out otherwise. The F-stat testing for the difference between the estimates of No Observability and Granular Observability is shown at the bottom of the table.

- *Hours Worked*: On day X, please indicate what time you started driving, what time you finished driving, and how many hours of break you took in between.
- *End-Start Time*: On day X, please indicate what time you started driving and what time you finished driving.
- *Customers*: On day X, between Xh and Yh, how many customers did you have?
- *Revenue*: On day X, between Xh and Yh, how much money did you collect in total? — converted to USD, using USD 1 = CFA 600.

Table B20: Heterogeneous Impact of Observability on Contracts Based on Driver's Digital Usage

Top 50% Number of Daily Digital Transactions

	Upfront Payment 'Salary' Dummy	Upfront Payment 'Salary' Value (USD)	Weekly Rent Target Value	Separation
	(1)	(2)	(3)	(4)
Granular Observability	0.064	-2.892	-2.357	-0.059
	(0.073)	(9.134)	(1.946)	(0.089)
Granular Observability $\times$ X	0.131	14.903	4.254	-0.035
	(0.117)	(12.214)	(2.642)	(0.122)
Coarse Observability	-0.020	-6.502	3.192	-0.007
	(0.088)	(9.016)	(2.809)	(0.086)
Coarse Observability $\times$ X	0.056	12.693	-2.908	0.072
	(0.137)	(14.098)	(3.225)	(0.123)
X = Proxy for Digital Intensity: 50p	-0.072	-7.701	0.644	-0.142*
	(0.094)	(9.905)	(2.130)	(0.085)
Observations	278	278	278	335
Mean No Observability	0.74	58.95	101.01	0.35

Notes: Business-level regressions of the contract outcomes on the three treatment arms, with the pure control group as the omitted category. The outcome variable is displayed at the top of each column. These regressions are conducted for the mid-term period until December 2022 (approximately 9 months after the baseline survey). They include interaction terms to study heterogeneity based on digital usage. The proxy for Digital Intensity is defined as the 50th percentile across drivers based on the average number of transactions received daily throughout the experiment, for days with at least one transaction. Although digital usage is also an endogenous variable, this table aims to provide suggestive evidence linking digital usage to contract changes.

Robust heteroskedasticity-consistent standard errors are reported, and the significance of each coefficient is denoted by asterisks: $p < 0.1$ *, $p < 0.05$ **, $p < 0.01$ ***. The pure control group is not included because their digital usage is zero. Each regression includes controls for strata fixed effects.

- *Upfront Payment 'Salary' Dummy*: Whether the owner provides a monthly upfront fixed payment to their driver (also referred to as a 'salary' in the industry).
- *Upfront Payment 'Salary' Value (USD)*: Value of the monthly upfront fixed payment to the driver, including 0 if no payment is provided. Values converted to USD (USD 1 = CFA 600).
- *Weekly Rent Target Value*: Weekly rental target amount from driver to owner, as agreed between owner and driver. Values converted to USD (USD 1 = CFA 600).
- *Separation*: Owner and driver are not working together at the time of the survey.

Table B21: Similar Observability Impact on Peer-to-peer Transactions Received by Drivers

	# of transactions	# of transactions	Amount (USD)	# of active weeks	# of active days
	(1)	(2)	(3)	(4)	(5)
Panel A. P2P transactions Within Taxi Value Ranges					
Granular Observability	0.206**	0.140	8.216	0.033	0.071
	(0.097)	(0.088)	(6.579)	(0.059)	(0.076)
Coarse Observability	0.272***	0.163*	13.911**	0.044	0.108
	(0.096)	(0.088)	(7.044)	(0.061)	(0.078)
Obs	358	358	358	358	358
Mean No Observability	22.35	22.35	69.87	10.30	18.21
Baseline P2P Taxi Control	NO	YES	YES	YES	YES
Panel B. P2P transactions Outside Taxi Value Ranges					
Granular Observability	0.205	0.132	14.554	-0.058	0.040
	(0.171)	(0.152)	(121.297)	(0.070)	(0.122)
Coarse Observability	0.192	0.081	100.244	0.070	0.114
	(0.140)	(0.136)	(130.590)	(0.070)	(0.117)
Obs	358	358	358	358	358
Mean No Observability	28.99	28.99	789.05	12.07	22.73
Baseline P2P Non-Taxi Control	NO	YES	YES	YES	YES

Notes: Administrative data provided by the mobile money partner company, collected from April to December 2022, the end of the mid-term survey. Zeros (drivers not using the technology at all) are included and account for approximately 15% of drivers. Robust standard errors are used. The significance of each coefficient is denoted as follows: $p < 0.1$ *, $p < 0.05$ **, $p < 0.01$ ***. Columns 1, 2, 4, and 5 display coefficients from Poisson regressions (their outcomes are counts), while Column 3 represents an OLS regression. The pure control group sample is excluded because these drivers did not access the technology.

The variable # of transactions represents the total number of transactions, while the amount corresponds to the total value of those transactions (in USD). The number of active weeks/days refers to the count of time periods with at least one transaction received.

Baseline P2P Control denotes the number of P2P transactions received before the experiment launch, categorized as within or outside the taxi value ranges. This control is included to mitigate variations caused by potential pre-trends in digital usage.

The taxi value range is defined as CFA 1,000–3,500 (USD 1.5–6), representing the 5th and 95th percentiles of taxi P2B transactions received.

Table B22: Impact of Observability On Driver's Profit

	Driver's Profit (1)	Driver's Profit (2)	Driver's Profit (3)
Granular Observability	10.146***	10.727***	9.957***
	(3.401)	(3.280)	(3.312)
Coarse Observability	4.138	4.791	4.859
	(4.415)	(4.204)	(4.162)
No Observability	2.925	4.597	4.164
	(4.908)	(4.714)	(4.654)
Observations	529	525	525
Control Mean	440.57	440.57	440.57
% Change Granular Observability	2.30	2.43	2.26
F-test Granular O = No O (p-value)	0.15	0.21	0.23
Baseline Control	NO	YES	YES +

Notes: Outcomes are regressed on the three treatment arms, with the control group as the omitted category. Robust heteroskedasticity-consistent standard errors (SE) are used, and the significance of each coefficient is denoted by asterisks: $p < 0.1$ *, $p < 0.05$ **, $p < 0.01$ ***. The regressions control for strata fixed effects, and the outcome is measured at the mid-term point, 7 to 9 months after the baseline survey. For respondents who refused to answer or indicated uncertainty, their responses were replaced with dummy variables. All monetary values are converted to USD (USD 1 = CFA 600). Driver profits are self-reported average earnings over the past 30 days, following De Mel et al. (2009). Specifically, drivers were asked: 'Over the last 30 days, what is the average income that you managed to keep per worked day, after paying all your work-related expenses, including fuel, rental payment, repair, police, contributions, and food?' This figure was then multiplied by the number of days worked in a typical month, with the addition of the monthly salary value paid by the taxi owner, if applicable. Baseline controls in Column (3) include not only average profit at baseline but also a dummy variable for whether the driver had a salary at baseline and the number of days worked per week at baseline.

B.10 Observability and Owner's Hiring Decision

Table B23: Impact of Observability on Owner's Hiring Decision

	Intention (9mo)		*Actual hiring decision*	
	Intent to hire	Intent to participate in a hiring meeting	Mid-term (9mo)	Long-term (2y)
	(1)	(2)	(3)	(4)
Granular Observability	0.029	0.010	0.066**	0.043
	(0.040)	(0.042)	(0.030)	(0.035)
Coarse Observability	0.013	0.034	0.012	0.004
	(0.041)	(0.044)	(0.030)	(0.035)
Observations	820	820	820	746
No Observability Mean	0.322	0.489	0.153	0.297
% Change Granular Observability	9	2	43	14

Notes: Survey data includes all taxi owners, including those driving their taxi alone (i.e., with no employed driver) at baseline. Robust standard errors are used. p-values are displayed in brackets, with the significance of each coefficient denoted as follows: $p < 0.1$ *, $p < 0.05$ **, $p < 0.01$ ***. The pure control group is excluded because it was treated after 9 months. The mean reported in the bottom row is the mean in the 'No Observability' treatment arm. The regression controls for the owner's baseline characteristics, such as whether they were driving the taxi themselves or had an employed driver.
- *Intent to hire*: Given the difficulty of the driver's job, do you intend to hire a driver to help you or leave the taxi to them in the near future?
- *Intent to participate in a hiring meeting*: Would you be interested in participating in a meeting between owners and drivers to facilitate the hiring of a driver for your taxi?
- *Hired a driver*: Whether the owner had effectively hired a driver at the time of the follow-up survey (after 9 months and about 2 years, i.e., mid- and long-term).
Single firm owners under the observability treatment were told the following: "Your option with this technology is 'Granular Observability.' As a taxi owner, you will have access to the transactions of your drivers, be able to observe the driver's transaction history and trips, and receive an SMS indicating the total daily transactions. If you decide to hire a driver in the future, you will also have this option." Owners in the 'No Observability' treatment were explicitly told they would not have access to the driver's transaction data.

B.11 Heterogeneous Impact of Observability on Separation Rate

Table B24: Heterogeneous Treatment Effect of Observability on Separation Rate

	Separation Rate							
	Mid-term	Long-term	Mid-term	Long-term	Mid-term	Long-term	Mid-term	Long-term
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$X =$			*Recent Relationship*		*Non-Family Business*		*Non-Risk-Averse Agent*	
X			0.078	0.004	0.127	0.243***	0.066	0.028
			(0.089)	(0.103)	(0.086)	(0.093)	(0.093)	(0.107)
Granular Observability	-0.100*	-0.061	-0.079	-0.041	-0.089	0.005	-0.058	-0.074
	(0.060)	(0.067)	(0.093)	(0.107)	(0.079)	(0.094)	(0.093)	(0.107)
Granular Observability $\times X$			-0.032	-0.033	-0.030	-0.144	-0.081	-0.124
			(0.124)	(0.141)	(0.122)	(0.130)	(0.130)	(0.148)
Coarse Observability	0.011	-0.038	0.031	-0.062	-0.026	-0.020	0.120	0.035
	(0.062)	(0.069)	(0.095)	(0.105)	(0.080)	(0.095)	(0.099)	(0.109)
Coarse Observability $\times X$			-0.024	0.050	0.075	-0.047	-0.158	-0.178
			(0.125)	(0.147)	(0.124)	(0.138)	(0.132)	(0.154)
Observations	335	338	335	338	335	338	335	338
No Observability Mean	.35	.63	.35	.63	.35	.63	.35	.63
Mean X			.56	.56	.47	.47	.49	.49

Notes: Business-level regressions estimate the effect of observability on separation rates at mid-term (9 months) and long-term (about two years). The omitted category is the 'No Observability' group, and the analysis includes interaction terms to study heterogeneity. The pure control group is excluded to ensure comparability across columns, as all pairs were treated after the mid-term survey. Robust heteroskedasticity-consistent standard errors (SE) are used, and regressions control for strata fixed effects. The heterogeneity analysis includes the following variables:
- *Recent Relationship:* Relationship lasted less than or equal to 2 years at baseline, before the start of the experiment.
- *Non-Family Business:* Relationship identified as non-family (e.g., close friends, friends, or neutral relations), as opposed to family members.
- *Non-Risk-Averse Agent:* Driver with a coefficient of relative risk aversion above 1 (CRRA utility function), as elicited in the field using an incentivized game.

B.12 Driver's Adoption Decision and Preference for Observability

Table B25: Reasons for Driver Refusals to Share Owner Contact Information

	Share
	(1)
Concerns over Owner's Privacy / Need to Discuss with Owners First	0.48
	(0.50)
Digital Observability of Transactions	0.20
	(0.40)
Owner Not Available or Not Interested	0.15
	(0.35)
Lack of Trust in Research Team	0.01
	(0.12)
Other Reasons: Don't Have Phone Number, Not in Country, etc.	0.11
	(0.31)
Refused To Answer	0.08
	(0.27)
Observations	433

Notes: Survey data were collected from drivers during the listing survey conducted from March to May 2022. Drivers were asked, "Why do you refuse to share your owner's contact information?" and their responses were recorded. Multiple answers were possible, and the share of drivers in each category is reported. Refusing to share the owner's contact information prevented drivers from adopting the digital payment technology.

Table B26: Impact of Transaction Observability on Technology Adoption and Privacy Concerns

	Technology Adoption (Willing to Share Owner's Information)	
	(1)	(2)
Removing Observability	0.281***	0.073*
	(0.090)	(0.042)
Observations	87	346
Control Mean	0.119	0.149
Observability Concerns Cited	YES	NO
% Change Removing Observability	236	49

Notes: Survey data were collected from June 15 to July 7, 2022, on drivers who refused to provide their owner's contact information during the listing. Driver-level regressions are performed. The outcome *Adoption* is whether the driver provided the owner's contact information to the surveyor, thus enabling them to adopt the digital payment technology. Heteroskedasticity-robust standard errors are reported, and the significance of each coefficient is denoted by asterisks: $p < 0.1$ *, $p < 0.05$ **, $p < 0.01$ ***. The random assignment of removing the owner's observability of drivers' digital transactions was not stratified.

I run two separate regressions for two distinct groups of individuals: (1) those who cited transaction observability as the reason for not providing their owner's contact information, and (2) those who cited other reasons during the listing process (see Table B25). These regressions aim to demonstrate two key points: (a) the treatment effect is significant for both groups, including those who did not raise observability as an issue, and (b) even for the group where observability was a concern, adoption remains incomplete when observability is removed. This suggests that the treatment effect may be underestimated, as some drivers might be uncertain whether observability will actually be removed.

Table B27: Robustness: Non-Adopters Are Low-Types and Poorest Drivers

	Willing To Adopt	Reluctant Drivers	Difference $\beta_{ReluctantDrivers}$
	(1)	(2)	(3)
Panel A. Driver's Performance			
Performance Index (Z-Score)	0.094	-0.046	(-0.140)
	(0.746)	(0.885)	
Number of Passengers (3 Days)	43.482	39.444	(-4.038**)
	(14.339)	(15.194)	
Total Collection (3 Days, USD)	152.515	151.264	(-1.251)
	(28.898)	(41.065)	
Effective Hours Worked (Avg 3 Days)	10.298	9.938	(-0.360)
	(2.098)	(2.194)	
Total Work Time (End to Start - Avg 3 Days)	11.930	11.413	(-0.517)
	(2.495)	(2.616)	
Panel B. Relationship and Contracts			
Monthly Upfront Payment 'Salary' (Dummy) W	0.750	0.901	(0.151***)
	(0.435)	(0.300)	
Weekly Rent Value R (USD)	101.389	101.760	(0.372)
	(11.055)	(7.787)	
Separation Rate p	0.348	0.363	(0.015)
	(0.479)	(0.482)	
Owner-Driver Relationship $>$ Two Years	0.449	0.420	(-0.028)
	(0.500)	(0.495)	
Panel C. Driver's Characteristics			
Risk-Averse Agents (CRRA>1)	0.400	0.658	(0.258***)
	(0.492)	(0.476)	
Often Stressed About Rent Transfers	0.077	0.147	(0.070*)
	(0.268)	(0.355)	
Panel D. Demographics and Poverty			
Education (At Least Primary)	0.264	0.230	(-0.035)
	(0.443)	(0.422)	
Literacy (Reading and Writing)	0.575	0.609	(0.033)
	(0.497)	(0.489)	
Wealth Index (PPI-IPA)	63.700	58.742	(-4.958**)
	(16.099)	(16.267)	
Additional Revenue Source	0.124	0.222	(0.099*)
	(0.331)	(0.418)	
Observations	112	190	

Notes: This robustness table complements Table 9 and summarizes the characteristics of drivers who were willing to adopt the digital payment technology compared to those who did not adopt because they refused to share their owner's contact information, where drivers in both groups eventually were offered the technology with *No Observability*. Data, except for demographics and risk-aversion coefficients, were collected during the mid-term survey with drivers from October to December 2022. The first two columns present the mean values of each variable for drivers willing to adopt and non-adopters, with standard deviations in brackets below. More precisely, to characterize selection, I compare drivers in the impact experiment—who agree to potential observability but whose owners are randomly assigned to *No Observability*—to reluctant drivers assigned to *No Observability*. This approach is useful since only mid-term performance data could be collected as many reluctant drivers refused to participate in the one-month follow-up survey due to their unwillingness to provide their owner's contact, knowing it would prevent adoption. I compare performance between drivers 'willing to adopt'—assigned to *No Observability* or *Control*—and 'non-adopters'—the reluctant drivers randomly assigned to *No Observability*. The third column shows the estimate from the regression of the variable on being a non-adopter, that is $Y_i = \beta ReluctantDrivers + \epsilon_i$. Heteroskedasticity-robust standard errors and the significance of each coefficient is denoted by asterisks: p $<$ 0.1 *, p $<$ 0.05 **, p$<$ 0.01 ***.
Values are converted to USD (USD 1 = CFA 600). In particular: The Z-Score Productivity Index is a combination of mean of the z-scores of the number of passengers, the total collected, and the hours worked. Risk-averse agents are defined as driver with a coefficient of relative risk aversion above 1 (CRRA utility function), as determined in the field in an incentivized game. The Wealth Index is defined using the methodology developed by IPA in Senegal to measure household wealth, referred to as the Poverty Probability Index (PPI) based on the poverty survey (ESPS-II) developed in 2011 in Senegal.

Table B28: Predicting Driver's Preferences for Observability At Baseline

	Preference For Observability (1)
Panel A. Driver's Characteristics	
Avg Daily Revenue (Past 3 Days, Tens of USD)	0.097***
	(0.020)
Avg Number of Days Worked in a Week	0.045***
	(0.017)
Default at Least Once a Month	-0.055
	(0.034)
Productivity - Value/Hour	0.017
	(0.016)
Hours Worked in a Day	-0.004
	(0.006)
High-Types: Above Median Top Performers Among Drivers	0.101***
	(0.035)
Low-Types: Above Median Low Performers Among Drivers	-0.082**
	(0.035)
Owner-Driver Relationship > Two Years	0.084**
	(0.036)
Panel B. Owner's Beliefs	
Underestimates Number of Days Worked in a Week	0.028
	(0.047)
Underestimates Avg Hours Worked in a Day	0.042
	(0.046)
Underestimates Avg Daily Revenue	0.188***
	(0.051)
Observations	599
Mean of Drivers Preferring Observability	0.229

Notes: Baseline survey data from March to June 2022. Drivers were asked what level of transaction observability they would prefer, if they had to choose (before the treatment was randomized). The outcome variable is defined as 'Preference for Observability' if 'Granular Observability' was the driver's top choice. This variable is regressed on different X baseline variables, each in separate regressions. The regression model is specified as $PreferenceForObservability_i = \alpha + \beta X_i + \epsilon_i$, with significance levels denoted as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Heteroskedasticity-robust standard errors are used. Drivers who refused to respond to this question are excluded. Note that belief questions were only answered by a subset of owners to limit survey length and increase response rates among owners. Questions regarding driver types were framed as follows:
High-Types: Considering a typical 30-day period in the last 3 months, how many days did you collect more than 40,000 FCFA (USD 66)? The median across drivers was *2.0*, and drivers above this median were classified as high types, or top performers.
Low-Types: Considering a typical 30-day period in the last 3 months, how many days did you collect less than 25,000 FCFA (USD 41)? The median across drivers was *8.0*, and drivers above this median were classified as low types, or below-median performers.
An owner is defined as underestimating their driver's work if the value they provide is below the actual value reported by the driver.

B.13 Framework Inputs

Table B29: Framework Inputs

Input	Value (USD)	Source
From the literature:		
Discount factor δ	0.99	Calibrated from relevant setting, Ethiopia, Yesuf and Bluffstone (2019).
Survey Calibration:		
Owner's replacement cost K_p	268.28	Survey question - about 33 days of profit lost.
Share of high-θ μ	0.42	Survey question to taxi owners.
Baseline upfront payment W	9.40	Average salary in control and N-O groups.
Target transfer R^l and R^h	100.29	Median transfer - 91% have this exact target.
High, low output: (Y, X)	(190.50, 93.51)	Avg output above/below median output with high effort in control and N-O groups.
Production function q_0 at $e = 0$	0.27	Likelihood of above median output with low effort.
Production function q_1 at $e = 1$	0.57	Likelihood of above median output with high effort in control and N-O groups.
Driver's surplus π_1 at $e = 1$	51.42	Calibrated from above estimates.
Owner's Maintenance Cost MC	19.03	Average maintenance cost self-reported by owners at baseline.
Agent's Outside Options $(\bar{u}^l, \bar{u}^h)$	(27.33, 33.33)	Average profit of a small merchant in Dakar from a representative survey I conducted.
Reduced-Form Estimates:		
Technological Gains for Drivers G	1.90	Treatment effect on imputed loss associated with costs of using cash at short-term.
Production function q_2 at $e = 2$	0.67	Likelihood of above median output with high effort in Granular Observability group.
Upfront Payment with Observability $W_{\tilde{e}=2}$	10.94	Unconditional average in the Granular Observability group.
Probability to keep with Observability $p_{\tilde{e}=2}$	0.97	Share remaining together under Granular Observability group.

Notes: All parameters, except the discount rate, are calibrated from the survey data. Specifically, to calibrate q_1, I use the proportion of times drivers achieved upper median output when working more than the median hours in the 'No Observability' (N-O) and control groups at mid-term. Similarly, to calibrate q_0, I use the proportion of times drivers in the entire sample achieved upper median output when working less than the 25^{th} percentile of hours worked. High and low outputs (Y and X) were determined based on drivers' revenue minus expenses, which include weekly food and beverages, all measured in the survey. I define high and low outputs, Y and X, as the average output above or below the median output when working more than the median hours, respectively, as illustrated in Figure A12.

I infer the proportion of high-type drivers (μ) by directly asking owners for their estimates and averaging their responses. The exact survey question was: 'We are trying to understand your perception of drivers. There are good and bad taxi drivers in terms of work. Imagine that 10 drivers present themselves to you at random. Out of 10 drivers, how many do you consider to be 'good' drivers?' The owner's replacement cost K_p was calibrated using the response to the question: 'If you were to lose your current taxi driver, how long do you estimate it would take for you to find another similar taxi driver?'

To estimate outside options, I conducted a representative survey of merchants—common outside options for drivers—in September 2022. From this survey, I calculated the median profit for merchants with 0 or 1 employee. To distinguish between high-type and low-type outside options, I used the difference in poverty likelihood, based on the 200% National Poverty Line as computed by IPA. This measure allows me to estimate variation in outside options: I assume high-type drivers can earn the median profit level observed among merchants, while low-type drivers can earn a proportion of this median, adjusted to reflect the wealth index gap between the two empirical samples.

Then, I use reduced-form estimates to obtain the following parameters: the private technology gains for drivers (G), which are assumed to equal the total transaction cost of using cash in the control group at mid-term, reflecting anticipated technological gains as described in Section 5.2; the production function for drivers under 'Granular Observability'. Specifically, to calibrate q_2, I use the proportion of times drivers achieved upper median output when working more than the median hours in the 'Granular Observability' treatment group at mid-term; and the contract characteristics under 'Granular Observability', particularly the observed salary $W_{\tilde{e}=2}$ and the probability of retaining the driver $p_{\tilde{e}=2}$.

Table B30: Threshold Values of the Discount Rate in the Structural Estimation

Discount Rate and Thresholds	Value (1)
$\underline{\delta}$	0.23
δ	0.99
$\bar{\delta}$	1.04
$\bar{\delta}^{tech}$	1.03
$\bar{\delta}^{FI}$	1.03

Notes: This table presents the structural threshold values for the discount rate from the no-deviation conditions, under which the various results hold in the structural estimation.

C Theoretical Framework: Proofs and Derivations

C.1 Baseline Contract and Agents' Informational Rents

This section elaborates on the proof of the baseline contract and investigates the source of rents for drivers at baseline.

The principal seeks to maximize expected transfers and the future discounted value of the relationship. The objective functions of the principal V^θ matched with the agent type θ at a given time can thus be written:

$$V^\theta = \max_{t,p,e} \mathbb{E}[t(\tilde{y}) + \delta[p(\tilde{y})V^\theta + (1-p(\tilde{y}))(-K_p + \mu V^h + (1-\mu)V^l)]|e] \tag{C1}$$

This optimization is subject to the following constraints:

$$\begin{cases} \mathbb{E}[(y(e) - t(\tilde{y})) - \phi^\theta(e) + \delta U^\theta - \delta K_a(1-p(\tilde{y}))|e] \geq \max\{-\delta K_a + \delta U^\theta; \bar{u}\} & \text{Participation Constraint } \textit{(IR)} \\ e \in \arg\max_{\tilde{e}\in\{0,1,2\}} \mathbb{E}[y(e) - t(\tilde{y}) + \delta U^\theta - \delta K_a(1-p(\tilde{y}))|\tilde{e}] - \phi^\theta(\tilde{e}) & \text{Incentive Compatibility } (IC) \\ t(\tilde{y}) \leq \tilde{y} \leq y(e) & \text{Limited Liability } \textit{(LL)} \\ Y - t(Y) + \delta(U^\theta - K_a(1-p(Y))) \geq Y - t(X) + \delta(U^\theta - K_a(1-p(X))) & \text{Truth-Telling } \textit{(TT)} \\ y(e) - t(\tilde{y}) + \delta(U^\theta - (1-p(\tilde{y}))K_a) \geq y(e) + \delta(U^\theta - K_a) & \text{Dynamic Enforceability } \textit{(DE)} \end{cases}$$

Result 1. (Baseline Contract Without Digital Payments) *Under Assumptions 1–4, $\exists\ \underline{K}_p < \bar{K}_p$, $\underline{\delta} < \bar{\delta}$, s.t. when $K_p \in (\underline{K}_p, \bar{K}_p)$ and $\delta \in (\underline{\delta}, \bar{\delta})$, the principal's best type-dependent stationary contract is:*

$$\bar{t}^\theta = \begin{pmatrix} t(Y) = R^\theta \\ t(X) = X \end{pmatrix} \quad \textit{and} \quad \bar{p}^\theta = \begin{pmatrix} p(Y) = 1 \\ p(X) = \bar{p}^\theta \end{pmatrix}$$

where the continuation probability for a low-output outcome, $\bar{p}^\theta$, is given by:

$$\bar{p}^\theta = \min \begin{Bmatrix} \bar{p}^{\theta,IC} = 1 + \dfrac{\bar{u}}{\delta K_a} - \dfrac{q_0\phi^\theta(1)}{\delta K_a(q_1 - q_0)}, \\ \bar{p}^{\theta,TT} = 1 - \dfrac{1}{\delta K_a}\left[q_1(Y-X) - \phi^\theta(1) - \bar{u}\right] \end{Bmatrix}$$

and the rental transfer for a high-output outcome, R^θ, is given by:

$$R^\theta = \min \begin{Bmatrix} R^{\theta,IC} = Y - \bar{u} - \dfrac{(1-q_0)\phi^\theta(1)}{q_1 - q_0}, \\ R^{\theta,TT} = q_1(Y-X) + X - \phi^\theta(1) - \bar{u} \end{Bmatrix}$$

with $\bar{p}^h > \bar{p}^l$, $R^h > R^l$. The agent induced effort is $e^l = e^h = 1 < 2$.

The contract values $(R^\theta, \bar{p}^\theta)$ will ultimately depend on which of the incentive compatibility constraint (IC) or the truth-telling constraint on output (TT) binds for agent θ.

Proof. The proof proceeds in two parts. First, I derive the optimal terms of the best stationary contract whether the incentive compatibility constraint or the truth-telling constraint binds. Second, I derive the conditions under which such contract is preferable for the principal.

Best stationary contract terms I first demonstrate that $p(Y) = 1$. I proceed by contradiction. Assume, for contradiction, that $p(Y) < 1$. First, consider the principal's objective to maximize the continuation value. If $p(Y) < 1$, the principal can increase $p(Y)$ to 1 and obtain a higher continuation value. Now, consider the agent's incentive compatibility. The agent's effort and truth-telling are affected by the expected punishment $p(Y)$. Increasing $p(Y)$ up to 1 enhances the agent's incentives for effort or truth-telling, given the structure of the framework. Since the principal incurs a cost K_p for punishing the agent with $K_p > 0$, it is therefore less preferable for the principal to set $p(Y)$ to any value less than 1. Hence, the assumption that $p(Y) < 1$ leads to a contradiction. Therefore, it must be that $p(Y) = 1$.

Second, I show that $t(X) = X$. Assume, for contradiction, that $t(X) < X$. The truth-telling constraint implies $p(X) \leq 1 - \frac{t(Y)-t(X)}{\delta K_a}$. Given that punishing with probability $p(X)$ incurs a cost $K_p > 0$ for the principal, reducing $t(X)$ unnecessarily imposing additional constraints on $p(X)$ and reduces transfers without enhancing the agent's reporting incentives. Hence, the owner is strictly better off setting $t(X) = X$.[53] Note that $t(X)$ cannot be more than X given the agent's limited liability constraint.

Let's define $t(Y) = R^\theta$ and $p(X) = \bar{p}^\theta$, where $R^\theta \in [0, Y]$ and $\bar{p}^\theta \in [0, 1]$. R^θ is the net transfer from the agent θ to the principal in a high reported output period. It is assumed that both parties observe a public randomization device at the end of the stage game, as commonly used in the literature—see Mailath and Samuelson (2006), Chapter 7. The principal can use this public randomization device for p, which would determine whether the agent is terminated for outcome $y = X$. The deviation, where the principal does not follow through with the randomization device, is assumed to lead to an equilibrium in which the agent always misreports output and exerts no effort. Thus, the principal is always better off following through with the public randomization device.

Upon agent reporting low output, the principal must terminate on path, with $\bar{p}^\theta < 1$. This costly punishment arises because of the limited liability constraint (Assumption 2): following low output, the principal cannot extract money so the only way to discipline incentives requires inefficient punishment (Fuchs, 2007). Specifically, this constraint implies $t(X) = X$, as the principal cannot demand more than what the agent collects.

To incentivize effort $e = 1$ by agent θ, the principal sets $\bar{p}$ such that the IC constraint is satisfied:

[53] One can enrich the model and include a minimum payment to be provided to the agent in the low-output period, so he would receive more than 0 due to limited liability.

$$\frac{q_1(Y-t(Y))+\delta(-K_a(1-\bar{p})(1-q_1))-\phi^\theta(1)}{1-\delta} \geq \frac{q_0(Y-t(Y))+\delta(-K_a(1-\bar{p})(1-q_0))}{1-\delta} \iff$$

$$\bar{p} \leq 1+\frac{Y-t(Y)}{\delta K_a}-\frac{\phi^\theta(1)}{\delta K_a(q_1-q_0)} \equiv \bar{p}^\theta \tag{C2}$$

From the truth-telling constraint, we get: $\bar{p}^\theta \leq 1-\dfrac{R^\theta - X}{\delta K_a} \ \forall \ \theta$

We also need to make sure that the dynamic constraint of the driver does not bind in a low-output period, meaning: $\bar{p}^\theta \geq \dfrac{X}{\delta K_a} \ \forall \ \theta$.

To summarize:

$$\begin{cases} \bar{p}^\theta & \leq 1+\dfrac{Y-R^\theta}{\delta K_a}-\dfrac{\phi^l(1)}{\delta K_a(q_1-q_0)} \quad \text{implied from } (IC^\theta) \\ \bar{p}^\theta & \leq 1-\dfrac{R^\theta-X}{\delta K_a} \quad \text{implied from } (TT) \\ \bar{p}^\theta & \geq \dfrac{X}{\delta K_a} \quad \text{implied from } (DE) \end{cases} \tag{C3}$$

Either the incentive compatibility (IC) or the truth-telling (TT) constraint binds because terminating the agent is costly to the principal.

Let's first assume that the principal sets the transfer in high-output periods to the maximum such that the agent remains in the relationship, i.e., if the following is true:

$$\begin{gathered} \frac{\partial V^\theta}{\partial t(Y)} > 0 \iff \\ K_p < \frac{q_1 K_a}{(1-q_1)} \equiv \bar{K}_p \end{gathered} \tag{C4}$$

with $V^\theta = \dfrac{q_1 t(Y)+(1-q_1)X-\delta[K_p(1-\bar{p}^\theta)(1-q_1)]}{1-\delta}$.

Intuitively, this suggests that the replacement cost for the principal should be sufficiently low, ensuring that termination is an effective tool for the principal. Otherwise, the principal never fires the agent ($\bar{p}=1$) and the agent always reports low-output, which I rule out.

Case 1. Incentive compatibility constraints binds When (IC) binds, then:

$$\bar{p}^{\theta,IC} = 1+\frac{Y-R^{\theta,IC}}{\delta K_a}-\frac{\phi^\theta(1)}{\delta K_a(q_1-q_0)}$$

The limited liability constraint implies that the participation constraint is slack and both agents have an informational rent at baseline. In particular, the principal selects R^θ such that the agent's present discounted utility of the agent binds, with K_a the cost of being rehired, which is paid only when the agent gets rematched to a principal.

$$\frac{q_1(Y - t(Y)) - \phi^\theta(1) - \delta(K_a(1-\bar{p}^\theta)(1-q_1))}{1-\delta} = \frac{\bar{u}}{1-\delta}$$

From the (IC) constraint binding, I replace $\bar{p}^{\theta,IC} = 1 + \dfrac{Y - t(Y)}{\delta K_a} - \dfrac{\phi^\theta(1)}{\delta K_a(q_1 - q_0)}$

To obtain:

$$\begin{cases} t(Y) = R^{\theta,IC} = Y - \bar{u} - \dfrac{(1-q_0)\phi^\theta(1)}{q_1 - q_0} \\ \bar{p}^{\theta,IC} = 1 + \dfrac{\bar{u}}{\delta K_a} - \dfrac{q_0 \phi^\theta(1)}{\delta K_a(q_1 - q_0)} \end{cases}$$

Informational Rent: Hence, in each period, the agent captures some surplus above their outside option $\bar{u}$ of walking away and thus receive the following *informational rent*:

$$q_1(Y - R^{\theta,IC}) - \phi^\theta(1) - \bar{u} = \frac{\phi^\theta(1) q_0(1-q_1)}{q_1 - q_0} - \bar{u}(1-q_1)$$

In addition, the following condition needs to hold for the principal to be better off incentivizing $e = 1$ instead of $e = 2$ for the two agents. Specifically, to incentivize $e = 2$, the principal would require $\bar{p}_2^{\theta,IC} = 1 + \dfrac{Y - R^{\theta,IC}}{\delta K_a} - \dfrac{\phi^\theta(2) - \phi^\theta(1)}{\delta K_a(q_2 - q_1)}$, with $\bar{p}_2^{\theta,IC} < \bar{p}^{\theta,IC} \quad \forall \quad \theta$.

And using the same reasoning as before, we would get:

$$\begin{cases} R^{\theta,IC}_{e=2} = Y - \bar{u} - \dfrac{\phi^\theta(2)(1-q_1) - \phi^\theta(1)(1-q_2)}{q_2 - q_1} \\ \bar{p}^{\theta,IC}_{e=2} = 1 + \dfrac{\bar{u}}{\delta K_a} - \dfrac{q_1\phi^\theta(2) - q_2\phi^\theta(1)}{\delta K_a(q_2 - q_1)} \end{cases}$$

The principal is better off incentivizing $e = 1$ if:

$$\begin{aligned} V^\theta_{e=1}(R^{\theta,IC}_{e=1}, \bar{p}^{\theta,IC}_{e=1}) > V^\theta_{e=2}(R^{\theta,IC}_{e=2}, \bar{p}^{\theta,IC}_{e=2}) \iff \\ K_p > K_p^{\theta,IC} \end{aligned} \tag{C5}$$

Case 2. Truth-telling constraints binds When the truth-telling constraint binds, then the following holds:

$$\bar{p}^{\theta,TT} = 1 - \frac{R^{\theta,TT} - X}{\delta K_a} \tag{C6}$$

With the same reasoning as before, we can recover R^{TT} such that:

$$\begin{cases} R^{\theta,TT} = q_1(Y - X) + X - \phi^\theta(1) - \bar{u} \\ \bar{p}^{\theta,TT} = 1 - \dfrac{1}{\delta K_a}[q_1(Y - X) - \phi^\theta(1) - \bar{u}] \end{cases} \tag{C7}$$

The agent gets the following *informational rent*:

$$q_1(Y - R^{\theta,TT}) - \phi^\theta(1) - \bar{u} = q_1(Y - q_1(Y - X)) - (\bar{u} + \phi^\theta(1))(1 - q_1)$$

Similarly as in Case 1, we can check that $V^{\theta}_{e=1}(R^{\theta,TT}, \bar{p}^{\theta,TT}) > V^{\theta}_{e=2}(R^{\theta,IC}_{e=2}, \bar{p}^{\theta,IC}_{e=2}) \iff K_p > K_p^{\theta,TT}$

In particular, $\underline{K}_p = max(K_p^{h,IC}, K_p^{h,TT})$ such that the principal is better off incentivizing both agent, high and low-ability, to exert $e = 1$ at baseline, when either condition binds. Intuitively, if the principal's replacement cost, K_p, is above a certain level, it becomes unprofitable for the principal to incentivize effort level $e = 2$ at baseline. Achieving this higher level of effort would require a punishment that is too severe to be cost-effective for the principal.

I derive the condition under which the truth-telling constraint binds:

$$\begin{aligned} \bar{p}^{\theta,TT} < \bar{p}^{\theta,IC} \iff \\ (Y - X) - \frac{\phi^{\theta}(1)}{q_1 - q_0} > 0 \end{aligned} \tag{C8}$$

This condition depends on the agent's type. Intuitively, it may be that the (TT) constraint binds for the high-type agents, but not for the low-type agents, because it is "easier" (lower punishment needed) to incentivize effort from a high-type agent than a low-type one. In other words, a low-type agent is more likely to have his (IC) constraint binds.

Conditions for the discount factor δ I check whether the agent's dynamic enforceability constraint holds, such that both agents value the future enough to come back to the owner at the end of a low-output period, in particular:

$$\delta > \max\left(\frac{(q_1 - q_0)(X - \bar{u}) + q_0\phi^l(1)}{(q_1 - q_0)K_a}, \frac{X + q_1(Y - X) - \phi^h(1) - \bar{u}}{K_a}\right) \equiv \underline{\delta} \tag{C9}$$

On the other hand, I derive $\bar{\delta}$ such that Result 1 holds only when the principal is not too patient, i.e., $\underline{\delta} < \delta < \bar{\delta}$, or when the share of low-types is sufficiently high. The following "no-deviation" condition needs to hold:

- At equilibrium: $\mu < \dfrac{V_l(\frac{1}{\delta} - 1) + K_p}{V_h - V_l}$ or $\delta < \dfrac{V_l}{-K_p + \mu V_h + (1 - \mu)V_l} \equiv \bar{\delta}$ with V_h and V_l equilibrium values. That is, it must be that the share of high-type agents is sufficiently low, such that it is not profitable for the principal to fire low-type in the hope of being rematched with a high-type agent. This condition is then verified in the structural estimation.

This completes the proof of Result 1.

□

C.2 Full Information Benchmark and Fixed Wage Contract

Lemma 1. (Full Information Benchmark) *Under Assumptions 2-4 and $\delta < \bar{\delta}^{FI}$, the principal's best stationary equilibrium is a wage contract that guarantees re-hiring when the agent exerts the optimal effort* *level, with no termination occurring on the equilibrium path.*

The wages are set such that the high and low-type agents' participation constraints bind.[54] In the full information benchmark, the owner must sufficiently compensate the agent for his optimal level of effort so that the agent is indifferent between working and his outside option. This optimal level of effort will ultimately depend on the disutility of work of the agent, in particular the first-best level of effort will be $e = 1$ for agent θ if:

$$\begin{aligned} V^{FI}_{e=1} > V^{FI}_{e=2} &\iff q_1 Y + (1-q_1)X - \left(\phi^\theta(1) + \bar{u}\right) > q_2 Y + (1-q_2)X - \left(\phi^\theta(2) + \bar{u}\right) \\ &\iff \phi^\theta(2) > (q_2 - q_1)(Y - X) + \phi^\theta(1) \end{aligned}$$

This effort level would guarantee re-hiring with $\bar{p}^\theta = 1$, as terminating the agent is costly for the principal. Various compensation schemes could theoretically achieve this goal. Here, I provide the formal argument demonstrating that a fixed wage is weakly preferable from the principal's perspective compared to a bonus payment.

Proof. I proceed by showing why a fixed wage contract is preferable to a bonus payment. Suppose the principal opts to use a bonus system, where $\bar{W}$ is given in high-output periods and $\underline{W}$ in low-output periods, ensuring the agent provides $e = 2$ and is always re-hired. The objective function of the principal can be written as:

$$V^\theta = \max_{W}(qY + (1-q)X) - W + \delta V^\theta \tag{C10}$$

Given the relational nature of the contract, the principal must not renege on the agreed bonus in high-output periods. The principal's non-reneging constraint can be expressed as:

$$\begin{gathered} Y - \bar{W} + \delta V > Y - \underline{W} + \delta(V - h), \\ \delta > \frac{\bar{W} - \underline{W}}{K_p}. \end{gathered}$$

This condition indicates that the discount factor δ must be sufficiently large to deter the principal from paying a lower bonus $\underline{W}$ in high-output periods (instead of the promised large bonus). This constraint does not apply if a fixed wage is paid directly to the agent before knowing the output, as it eliminates the incentive for the principal to renege.

Therefore, a fixed wage is preferred over a bonus system because it is feasible over a wider range of δ, ensuring the principal's dynamic enforceability constraint and maintaining the agent's effort.

[54] Similar full-information benchmarks are discussed in a range of papers with different environments, such as Shetty (1988) with limited liability; in *Proposition 5* of Shavell (1979) with risk averse agents; and in Ghatak and Pandey (2000) with joint moral hazard in effort and risk.

Second, because the principal can be re-matched to any types of agents, we must verify the condition under which the principal is better off keeping the low-type drivers in the full-information benchmark, than firing them in the hope of being re-match to a high-type with probability μ. Let's define W^l and W^h the fixed wage of the low- and high-type agents in the full information benchmark. In the full-information benchmark, when $e = 1$ is the optimal level of effort, $W_{e=1} = \phi^\theta(1) + \bar{u}$, as the agent are paid at their outside option when leaving the taxi industry. For both types to remain in the market, the following "no-deviation" condition needs to be satisfied, meaning the share of low-types $(1 - \mu)$ should be high enough.

$$\delta < \frac{V_l^{FI}}{-K_p + \mu V_h^{FI} + (1-\mu)V_l^{FI}} \equiv \bar{\delta}^{FI}$$

$\square$

C.3 Observability of Agent's Output Only

Consider a scenario where only the employee's output y is verifiable, not his effort e.

Lemma 2. (Information on Output) *Under Assumptions 1–4, complete output information for the principal relaxes the truth-telling constraint (TT). When (TT) binds, this information leads to an increase in the continuation probability $\bar{p}$ and R to satisfy the incentive compatibility constraint (IC).*

This lemma indicates how information on output y can affect the principal-agent relationship. It enables an increase in p and R without revealing the agent's effort. This contract adjustment emerges because observability of output eliminates the need to impose penalties on the agent to ensure truthful reporting. Revealing output allows the principal to not punish the agent to tell the truth on the output collected.

C.4 Stage 2: Imperfect Information on Agent's Effort and Output

Digital payments provide the principal with *imperfect* information on their agent for at least two reasons empirically. First, cash is still being used, meaning the digitalization of payments is not complete. Second, while digital transactions include timestamps and transaction values, they only partially reflect the agent's effort and output. This section shows that such imperfect information leads to similar predictions to the benchmarks detailed above, using the informativeness principle (Holmström, 1979). Let's define s the high-effort signal and $\kappa = P(s|e = 2)$ and assume $P(s|e = 1) = P(s|e = 0) = 0$.

Result 2. (Imperfect Information on Effort) *Under Assumptions 1–4, when (IC) binds, for $K_p \in (\underline{K}_p, \bar{K}_p)$ and $\delta \in (\underline{\delta}, \bar{\delta})$, $\kappa > \bar{\kappa}$ for $\bar{\kappa} < 1$, and $\phi^\theta(2) < \tilde{\phi}\ \forall\ \theta$, the principal's best type-dependent stationary contract is:*

$$\vec{t}^\theta = \begin{pmatrix} t(Y) = R^\theta - W^\theta_{\tilde{e}=2} \\ t(X) = X - W^\theta_{\tilde{e}=2} \end{pmatrix} \quad \textit{and}$$

$$p(\tilde{y}, s) = \begin{cases} 1 & \textit{if } \tilde{y} = Y, \\ \bar{p}_{TT} & \textit{if } \tilde{y} = X \textit{ and } s \textit{ is observed} \\ \bar{p}^{\theta\prime} < \bar{p}^{\theta} & \textit{if } \tilde{y} = X \textit{ and } s \textit{ is not observed} \end{cases}$$

The agent θ induced effort is $e^{\theta} = 2$.

Each period, the principal uses $W^{\theta}_{\tilde{e}=2}$ to provide incentives for adoption. To mitigate concerns about reneging in this relational contract framework, the principal incentivizes the agent to adopt the technology by offering an upfront payment, $W^{\theta}_{\tilde{e}=2}$, rather than reducing the target rental payment. More formally, the upfront payment is preferred over a promised reduction in rent because otherwise the principal would have incentives to renege on his promise. A similar logic applies when the principal obtains *imperfect* information on output, relaxing the truth-telling constraint (TT) when the latter is binding at baseline. Consequently, imperfect information can be advantageously incorporated into the contract under minimal conditions.

The proof that follows proceeds in two steps. I first show why the principal's payment is preferably made *upfront* in this relational contract framework. Second, I rationalize the new contract structure and resulting effort.

Proof. Consider the contract structure where the transfer $t(Y)$ is defined as

$$t(Y) = R^{\theta} - W^{\theta}_{\tilde{e}=2},$$

For contradiction, assume $W^{\theta}_{\tilde{e}=2}$ represents the *promised* reduction in rental transfer at the end of the period, based on the high-effort signal.

At the end of the work period, if high output Y is achieved, the principal has an incentive to deviate by increasing the transfer by ϵ:

$$t(Y) = R^{\theta} - W^{\theta}_{\tilde{e}=2} + \epsilon,$$

where $\epsilon > 0$ is the marginal increase in the transfer. The agent would still be willing to accept this adjusted transfer, as it remains above their outside option. A similar reneging concern exists with low output X, allowing the principal to extract

$$t(X) = X - W^{\theta}_{\tilde{e}=2} + \epsilon$$

without breaking the agent's participation constraint.

This leads to a contradiction, as there exists a profitable one-shot deviation for the principal. To prevent such ex-post reneging concerns, the principal would prefer to provide $W^{\theta}_{\tilde{e}=2}$ upfront, eliminating the temptation to deviate in this relational contract framework. □

The second part of the proof is to rationalize the new contract structure and resulting incentives for effort on the agent side.

Proof. Given that $\kappa = P(s|e = 2)$ and $P(s|e = 1) = P(s|e = 0) = 0$, with s the high-effort signal. The higher κ, the more valuable the information. Consider a scenario where the incentive compatibility constraint (IC^θ) binds at baseline for agent θ, making imperfect information on effort valuable to relax this constraint. Specifically, the probability of retaining the agent becomes $\bar{p}_{TT}$ in a low-output period if the owner observes high effort $e = 2$, and $\bar{p}^{\theta\prime} < \bar{p}^\theta$ otherwise. This can be expressed as follows:

$$p(\tilde{y}, s) = \begin{cases} 1 & \text{if } \tilde{y} = Y, \\ \bar{p}_{TT} & \text{if } \tilde{y} = X \text{ and } s \text{ is observed} \\ \bar{p}^{\theta\prime} < \bar{p}^\theta & \text{if } \tilde{y} = X \text{ and } s \text{ is not observed} \end{cases}$$

For the sake of the argument, let's assume $\bar{p}_{TT} = 1$ meaning the truth-telling constraint does not bind in this setting. This assumption will simplify the following derivations.

The upfront payment or 'salary' must be sufficiently high to ensure that:

$$U_2^\theta + W_{\tilde{e}=2}^\theta \geq U_{notech}^\theta \tag{C11}$$

Note that $W_{\tilde{e}=2}^\theta > 0 \iff$
$\phi^\theta(2) > (q_2 - q_1)(Y - R^\theta) + \phi^\theta(1) - \delta K_a((1 - \bar{p}^\theta)(1 - q_1) - (1 - \kappa)(1 - \bar{p}^{\theta\prime})(1 - q_2))$

The IC^θ constraint requires that the agent θ must then be incentivized to exert high effort ($e = 2$). When the principal does not observe the signal, with probability $1 - \kappa$, he must terminate the relationship with some sufficiently low probability $\bar{p}^{\theta\prime}$ to incentivize effort. Intuitively, the high-effort signal enables the principal to incentivize effort at a lower cost (punishment) when the signal is valuable enough.

Mathematically, the IC^θ constraint can be written as:

$$\frac{W_{\tilde{e}=2}^\theta + q_2(Y - R^\theta) - \phi^\theta(2) - \delta K_a(1 - q_2)(1 - \kappa)(1 - \bar{p}^{\theta\prime})}{1 - \delta} \quad \text{(Agent exerts } e = 2\text{)}$$
$$> \frac{W_{\tilde{e}=2}^\theta + q_1(Y - R^\theta) - \phi^\theta(1) - \delta K_a(1 - q_1)(1 - \bar{p}^{\theta\prime})}{1 - \delta} \quad \text{(Agent exerts } e = 1\text{)}$$
$$\iff \bar{p}^{\theta\prime} \leq 1 - \frac{\phi^\theta(2) - \phi^\theta(1) - (q_2 - q_1)(Y - R^\theta)}{((1 - q_1) - (1 - q_2)(1 - \kappa))\delta K_a} \tag{C12}$$

Now I turn to the equilibrium and examine the conditions under which the owner is better off with this new contract.

$$V_{e=2}^\theta = \quad q_2 R^\theta + (1 - q_2)X - W_{\tilde{e}=2}^\theta + \delta(V - K_p(1 - \kappa)(1 - q_2)(1 - \bar{p}^{\theta\prime}))$$

The principal will be better off offering this upfront payment if $V^{\theta}_{e=2} > V^{\theta}_{baseline}$. Let's have $\bar{p}^{\theta\prime}$ such that the IC binds. One can show that

$$\frac{\partial \bar{p}^{\theta\prime}}{\partial \kappa} > 0 \tag{C13}$$

And since we know that:

$$\frac{\partial V}{\partial \bar{p}^{\theta\prime}} > 0, \tag{C14}$$

There exists a threshold $\exists\ \kappa \leq \bar{\kappa}$ such that the principal would not want to incentivize the agent and would instead retain the initial contract upon adoption. In other words, the precision of information κ must be sufficiently high to induce a profitable change in the contract.

Using the same reasoning,

$$\frac{\partial \bar{p}^{\theta\prime}}{\partial \phi^{\theta}(2)} > 0 \tag{C15}$$

There exists a threshold $\exists\ \phi^{\theta}(2) < \tilde{\phi}\ \forall\ \theta$ such that the principal would want to incentivize both agent's types to adopt. In other words, the disutility of high-effort $\phi^{\theta}(2)$ must be sufficiently low to allow a profitable change in the type-dependent contract.

□

Then, I turn to examine how imperfect information on output impacts the principal-agent relationship. Empirically, the digital payment technology provides only a low-output signal to business owners because only a fraction of driver's transactions were digital during the experiment. Therefore, I define s' the low-output signal and $\xi = P(s'|y = X)$ and assume that $P(s'|y = Y) = 0$.

Lemma 3. (Imperfect Information on Output) *Under Assumptions 1–4, when (TT) binds, for $K_p \in (\underline{K}_p, \bar{K}_p)$ and $\delta \in (\underline{\delta}, \bar{\delta})$, and $\xi > \bar{\xi}$ for $\bar{\xi} < 1$, the principal's best type-dependent stationary contract is:*

$$\bar{t}^{\theta} = \begin{pmatrix} t(Y) = R^{\theta} \\ t(X) = X \end{pmatrix} \quad \textit{and}$$

$$p(\tilde{y}, s') = \begin{cases} 1 & \textit{if } \tilde{y} = Y, \\ \bar{p}_{IC} & \textit{if } \tilde{y} = X \textit{ and } s' \textit{ is observed} \\ \bar{p}^{\theta} & \textit{if } \tilde{y} = X \textit{ and } s' \textit{ is not observed} \end{cases}$$

The imperfect information on output enables the principal to relax the truth-telling constraint and not punish the agent for misreporting when the signal indicates low output. With a similar reasoning as before, since $\frac{\partial V}{\partial \bar{p}^{\theta}} > 0$, the principal is better off using the low-output signal when the information is accurate enough—that is, when the probability of low output given the signal is sufficiently high, $\xi > \bar{\xi}$. Specifically, the principal benefits from increasing the continuation probability following a low-output signal. The target rental transfer R^{θ} remains unaffected, as the absence of signal does not perfectly reveal the high output: increasing R^θ would lead the agent to misreport the revenue.

The experiment specifically examines the importance of this output channel under the *Coarse Observability* treatment arm. In particular, *Coarse Observability* limits owners to seeing only low digital output levels, which means the agent can only signal *low output* to owners through the app. This treatment is expected to reduce the firing punishment for misreporting revenue when the driver fails to pay the rent, while preserving part of informational rent for the driver by not substantially revealing their effort or high-output levels.

C.5 Agent's Manipulation of Effort and Output Signals

Lemma 4. (No Manipulation) *Under Assumptions 1–4, the agent has no incentive to manipulate the imperfect information on effort or output provided to owners since inaccurate reporting or distorted information would limit the overall impact of the technology.*

Proof. The proof proceeds separately for the information on effort and output. In this context, passengers generally decide between digital and cash payments, yet I examine the potential for manipulation by drivers. Here, I assume that drivers can only lower the proportion of digital transactions (i.e., favoring cash over digital payments) without the ability to increase it, aligning with the observed characteristics of this setting. Indeed, there was no empirical evidence of drivers enforcing digital payments during the mystery passenger audits. Cash remains the primary method of transaction in Dakar, and passengers always expect the option to pay with cash.

1. *Manipulating the effort information.* Consider a strategy in which the agent manipulates the observable effort metric by under-utilizing digital payments. Suppose, for contradiction, that the agent reduces the share of digital transactions, signaling low effort to the principal. By Result 2, this manipulation reduces the likelihood of observing the high-effort signal s, thereby increasing the agent's termination probability. As a result, the agent's expected payoff is reduced. Consequently, the agent strictly prefers not to manipulate his effort level downward (and cannot realistically manipulate it upward, as discussed above).

2. *Manipulate the output information.* I show that with sufficiently high termination costs for the agent (i.e., large K_a), the agent has no incentive to manipulate the output information downward because that always leads to a higher termination probability. I focus on the high-output period (as the low-output period cannot be manipulated by the agent given the limited liability assumption, see above). By manipulating the low-output signal s' to signal low-output when actually $y = Y$, the agent keeps a greater surplus but will be punished with probability $(1 - \bar{p}_{IC})$ according to Lemma 3. Therefore, the agent will not manipulate if K_a is sufficiently high, that is formally:

$$\underbrace{Y - X + \delta[U - K_a(1 - \bar{p}_{IC})]}_{\text{Agent's present value manipulating output information}} \leq \underbrace{Y - R^{\theta} + \delta U}_{\text{Not manipulating the output information}} \tag{C16}$$
$$R^{\theta} - X \leq \delta K_a(1 - \bar{p}_{IC})$$

For Condition C16 to be satisfied, the principal, anticipating manipulation, has to set the termination probability sufficiently high to incentivize no manipulation. In other words, if anything, manipulation concerns would lower the positive treatment effect of the output signal on retention.

□

The experiment is designed to empirically test manipulation by comparing drivers' digital usage under *Coarse Observability* and *No Observability* treatment arms. If drivers were manipulating, we would expect them to use digital payments less under *Coarse Observability*. However, the findings show no evidence of manipulation; in fact, drivers tend to use digital payments slightly more under *Coarse Observability* in the administrative data (see Figure A9). Additionally, mystery passenger audits confirm this absence of manipulation, as drivers in both treatment groups respond similarly when passengers request to pay digitally (see Figure A6). The absence of manipulation may be explained by the competitive pressure drivers face to secure passengers, which discourages them to manipulate digital payment usage. Drivers may perceive the short-term loss of passengers—resulting from pressuring or discouraging them to pay digitally to signal something to their owners—as too costly to justify such actions.

C.6 Stage 1: Differential Digital Technology Adoption

After having established that the technology, once adopted, may increase owner's utility from the (imperfect) information on either the agent's output or effort, this section shows how agents' type influences who adopts the technology in the first place, see the following Result 3. This demonstration provides bounds for the disutility of work of the low-type agent, used in the structural simulations.

Result 3. (Differential Adoption) *Under Assumptions 1–4, $\exists\, \underline{K}_p < \bar{K}_p, \underline{\delta} < \bar{\delta}^{tech}$ and $\bar{\phi}^h < \bar{\phi}^l$ s.t. if $K_p \in (\underline{K}_p, \bar{K}_p)$ and $\delta \in (\underline{\delta}, \bar{\delta}^{tech})$, $\phi^l(2) > \bar{\phi}^l$, and $\phi^h(2) < \bar{\phi}^h$, then only high-ability agents adopt the technology, while low-ability agents opt not to adopt it.*

Proof. The proof proceeds by solving the two-stage game using backward induction. In Stage 1 ("adoption"), *the agent* decides whether to adopt the new technology based on expected utility in Stage 2 ("impact"). In Stage 2, as shown in Result 2, *the principal* would offer a new contract to the agent. The goal of this proof is to show which type of agents would adopt the technology in Stage 1 and under which conditions. For simplicity, we assume throughout this proof that $\kappa = 1$, meaning the high-effort signal perfectly reveals high effort $e = 2$. While the comparative statics still hold for $\bar{\kappa} < \kappa < 1$, the derivations become more complex.

There exists values of $\phi^h(2)$ and $\phi^l(2)$ such that the agents can be better off once they adopt the technology. To make the analysis more interesting, let's focus on the case where $\phi^h(2)$ and $\phi^l(2)$ are both too high such that no agents would adopt the technology absent changes to the contract. The principal can incentivize agents to adopt by offering a compensation upfront. As described before, this minimum upfront payment or 'salary' is denoted by w. In particular, the 'salary' should be sufficiently high such that, when IC binds:

$$\begin{aligned} U_2^\theta + w \geq U_{notech}^\theta \iff \\ w \geq \phi^\theta(2) + \bar{u} - q_2(\bar{u} + \frac{(1-q_0)\phi^\theta(1)}{q_1 - q_0}) \equiv W_{\tilde{e}=2}^\theta \end{aligned} \tag{C17}$$

Note that $W_{\tilde{e}=2}^\theta > 0 \iff \phi^\theta(2) > q_2(\bar{u} + \frac{(1-q_0)\phi^\theta(1)}{q_1 - q_0}) - \bar{u} \equiv \underline{\phi}^\theta$

The principal is better off offering a positive payment $W_{\tilde{e}=2}^\theta$ if the present discounted value to be matched with an agent adopting the technology $V_{e=2}^\theta$ minus this upfront payment is greater than the principal's objective function at baseline $V_{e=1,notech}^\theta$.

$$\frac{V_{e=2}^\theta - W_{\tilde{e}=2}^\theta}{1-\delta} \geq \frac{V_{e=1,notech}^\theta}{1-\delta} \tag{C18}$$

In Stage 1, there exists a range of low enough of disutility of high-effort for high-type $\phi^h(2) < \bar{\phi}^h$ and high enough disutility of high-effort for low-type $\phi^l(2) > \bar{\phi}^l$ such that the principal would prefer only the high-type to adopt and exert $e = 2$. The principal matched with low-types would be worse off offering $W_{\tilde{e}=2}^l$, so they have no incentive to offer it in the first place, given the low-type agent's cost of high effort. The reason is that the low-type agent is already exerting the "first-best" effort level, $e = 1$, from the social planner's point of view. Adopting the technology, which only reveals "high-effort", would push the agent to exert high effort $e = 2$. In other words, the principal would find it unprofitable to fully compensate the agent given his high disutility. Due to the lack of formal commitment, the principal cannot credibly commit to not demanding high effort once the agent adopts the technology, and would fire the agent if the effort signal is not provided.

As before, the following "no-deviation" condition needs to hold in equilibrium for the principal of a low-type agent to keep working with him in equilibrium once the technology is introduced:

$$\delta < \frac{V_{tech}^l}{-K_p + \mu V_{tech}^h + (1-\mu)V_{tech}^l} \equiv \bar{\delta}^{tech} \tag{C19}$$

with V_{tech}^l the new value function upon introduction of the technology for owners matched with a low-type not adopting the technology (where the outside option of the principal increases as they can now match with a high-type adopting the technology). This will always be true for sufficiently high K_p. This "no-deviation" condition is ultimately verified in the structural estimation, see Table B30.

Note that I assume that the agent can keep the technology with them upon termination, meaning the principal would still need to incentivize the next matched low-type agent for adoption.

This completes the proof of Result 3. □

D Structural Estimation

D.1 Framework Inputs

D.1.1 From Survey Data

The survey data collection was carefully designed to provide the necessary components for estimating key parameters and conducting welfare analysis. All parameters used in the framework are calibrated from the survey data or used as moments (except the discount rate). Specifically, for the production function, I calibrate q_1, the likelihood of a high output when effort is one $e = 1$, by using the proportion of times drivers achieved upper median output when working more than the median hours in the *No Observability* (N-O) and *Control* groups at mid-term. Similarly, to calibrate q_0, the likelihood of a high output when effort is zero $e = 0$, I use the proportion of times drivers in the entire sample achieved upper median output when working less than the 25^{th} percentile of hours worked. This survey-based calibration aims to map the discrete effort-output production function, and the sensitivity of the estimation is discussed in Section D.4.

High and low outputs (Y and X) are calibrated based on drivers' revenue minus expenses, which include weekly food and beverages of drivers, all measured in the survey. To be consistent with the calibration of q_1, I define high and low output, Y and X, as the average output above or below median output when working more than the median hours, respectively, as illustrated in Figure A12.

The proportion of high-type drivers (μ) is directly inferred by asking the owners for their estimates and averaging their responses. The exact survey question was: "We are trying to understand your perception of drivers. There are good and bad taxi drivers in terms of work. Imagine that 10 drivers present themselves to you at random. Out of 10 drivers, how many do you consider to be 'good' drivers?".

To estimate driver's outside options $\bar{u}^l, \bar{u}^h$, which I here allow to vary across agents to enrich the framework, I conducted a representative survey of traders in September 2022. Being a trader is a common outside option for drivers, based on empirical observations of drivers who left the taxi industry during the two-year experiment. From this survey, I calculated the average profit of merchants with 0 or 1 employee. Then, to distinguish between high-type and low-type driver's outside options, I used the difference in poverty likelihood, based on the 200% National Poverty Line, as measured by IPA. This poverty likelihood difference allows me to infer the difference in outside options between high-type and low-type drivers.

The owner's replacement cost K_p is calibrated using the owner's response to the survey question: "If you were to lose your current taxi driver, how long do you estimate it would take for you to find another similar taxi driver?".

D.1.2 Reduced-form Estimates

I use the reduced-form estimates to calibrate the following parameters:

- The private technological gains for the drivers, G, which is the treatment effect on imputed loss associated with costs of using cash at short-term, as described in Section 5.2.

- The production function for high-type drivers relies on the driver's effort and production under *Granular Observability*. Specifically, to calibrate q_2, I use the proportion of times drivers achieved an upper median output when they work more than the median hours in the *Granular Observability* group at mid-term.

- The contract characteristics rely on the owner-driver pairs under *Granular Observability*, in particular the observed upfront payment $W_{\tilde{e}=2}$ and the probability of retaining the driver $p_{\tilde{e}=2}$.

To test the sensitivity of the estimation to the framework inputs, I estimate the standard errors for each parameter using a bootstrap procedure, resampling with 1,000 bootstrap replications of the survey data.

D.2 Matched Moments: Description

Enriching the Framework: Main changes To better align the framework with the empirical evidence, I introduce three main modifications. First, I incorporate the fact that the technology provides benefits to drivers through reduced cash-related costs, denoted by G, which drivers capture directly. Second, I account for different outside options for each type of driver, $\theta \in l, h$. Third, I include an upfront payment, $W > 0$, given to drivers at baseline and received in both states of the world, such that this payment W does not affect the drivers' incentive compatibility or truth-telling constraints but influences their utility. These modifications are further discussed in relation to the specific moments they affect below.

The following empirical moments are matched to the theoretical moments in order to recover the disutility parameters: $\phi^h(1), \phi^l(1), \phi^h(2)$.

Moment 1: Baseline Continuation Probabilities $\bar{p}^l$

$$\bar{p}^l = \min \left\{ \begin{aligned} &\bar{p}^{l,IC} = 1 + \frac{\bar{u}}{\delta K_a} - \frac{q_0 \phi^l(1)}{\delta K_a (q_1 - q_0)}, \\ &\bar{p}^{l,TT} = 1 - \frac{1}{\delta K_a} \left[q_1 (Y - X) - \phi^l(1) - \bar{u} \right] \end{aligned} \right\}$$

The principal selects the minimum between the two cutoffs to incentivize both effort and truth-telling. The low-type $\theta = l$ empirical moment is considered to be the re-hiring probabilities from the reluctant drivers, as they were not influenced by the technology, and by not adopting, reveal their type to the researcher.

Moment 2: Baseline Continuation Probability $\bar{p}^h$

$$\bar{p}^h = \min \begin{Bmatrix} \bar{p}^{l,IC} = 1 + \dfrac{\bar{u}}{\delta K_a} - \dfrac{q_0 \phi^h(1)}{\delta K_a(q_1 - q_0)}, \\ \bar{p}^{l,TT} = 1 - \dfrac{1}{\delta K_a}\left[q_1(Y - X) - \phi^h(1) - \bar{u}\right] \end{Bmatrix}$$

I use the re-hiring rate in the control group of the impact experiment. This treatment is composed of high-type workers as only high-types adopt the technology in the lens of the framework. However, these workers are not affected by the technology, as they are part of the control group. Therefore, this treatment arm allows me to recover $\bar{p}^h$ for agents who are willing to adopt the technology but do not ultimately receive it.

Moment 3: Continuation Probability under Granular Observability $\tilde{p}_{\tilde{e}=2}$ The theoretical moment is derived in Appendix C.6, see Equation C12, and reported below.

$$\bar{p}^{\theta\prime} = \bar{p}^{\theta\prime} \leq 1 - \frac{\phi^\theta(2) - \phi^\theta(1) - (q_2 - q_1)(Y - R^\theta)}{((1 - q_1) - (1 - q_2)(1 - \kappa))\delta K_a} \tag{D1}$$

In the lens of the framework, it is implied that $\tilde{p}_{\tilde{e}=2} = (1 - (1 - \kappa) * (1 - q_2) * (1 - \bar{p}^{\theta\prime}))$. To simplify, I assume that $\kappa = 1$ such that the owner receives a perfect signal of high-effort, and that the owner also receives a perfect signal of low-output under *Granular Observability*, with no change in the rental payment R (which closely maps what is observed empirically). The empirical moment is the reduced-form re-hiring rate in the *Granular Observability* group when the high-type agent is induced to exert $e = 2$. This treatment arm is expected to mitigate moral hazard in effort and in output reporting, as discussed in Section 6.4.

Moment 4: High-Type Agent's Contract Valuation U^h

$$U^h = \frac{W + q_1(Y - R^h) - \phi^h(1) - \delta K_a(1 - \bar{p}^h)(1 - q_1)}{1 - \delta} \tag{D2}$$

The empirical moment is the driver's self-reported contract valuation. Specifically, I asked drivers who adopted the technology the following question to approximate the driver's value of the relationship: "Imagine the following scenario: how much would another taxi owner need to pay you to leave your current relationship with your owner and work for them in their taxi?" Here, I enrich the framework by considering that this contract valuation incorporates the calibrated baseline upfront payment "salary" W.

Moment 5: Agent's replacement cost K_a

The empirical moment is the agent's perceived replacement cost, derived from the following survey question asked at baseline: "If you were to lose your current job as a taxi driver, how long do you estimate it would take for you to find another similar taxi driver's job?" The number of days estimated is then multiplied by the daily driver profit to calculate the empirical value of the perceived replacement cost.[55]

Moments 6 and 7: Agent's Transfers in High-Output Periods R^l and R^h

These empirical moments represent the agents' median transfers in high-output periods, as derived from the baseline survey. Empirically, 91% of pairs have the same target transfer R^θ at baseline, resulting in a similar median across the two types of drivers. Within the framework, this can be explained by the higher disutility of work and the lower outside option for low-types, which may offset each other and lead to a similar R^θ across types. The formula for each agent type $\theta \in l, h$ is as follows:

$$R^\theta = \min \begin{Bmatrix} R^{\theta,IC} = Y - \bar{u} - \dfrac{(1-q_0)\phi^\theta(1)}{q_1 - q_0}, \\ R^{\theta,TT} = q_1(Y - X) + X - \phi^\theta(1) - \bar{u} \end{Bmatrix}$$

Moment 8: Upfront payment/Salary under Granular Observability $W_{\tilde{e}=2}$ The empirical moment is the upfront payment offered to drivers in the *Granular Observability* group when the high-type agent is induced to exert $e = 2$. The rationale for this upfront payment is to incentivize adoption and compensate the agent for exerting a higher level of effort $e = 2$. The estimation takes into account that some drivers were already receiving an upfront payment "salary" $W > 0$ at baseline.

Intuitions Behind Each Moment These eight empirical moments are matched to the theoretical moments to recover the following parameters: $\phi^h(1)$, $\phi^l(1)$, $\phi^h(2)$. The model is thus over-identified.

The disutility of work of the low-types for $e = 1$, $\phi^l(1)$, is identified using **Moment 1** and **Moment 6**. Intuitively, if the job becomes tougher, then the continuation probability in a low-output period or the transfer in a high-output period would need to be lower to incentivize effort at baseline. Using the same intuition, I estimate the disutility of work for $e = 1$ for the high-types, $\phi^h(1)$, using **Moment 2**, **Moment 4**, and **Moment 7**.

The disutility of work of the high-types with $e = 2$, $\phi^h(2)$, is estimated using **Moment 3** and **Moment 8**. The intuition is that the owner must compensate the driver for increased effort under *Granular Observability* by offering both a higher upfront payment, $W_{\tilde{e}=2} > W$, and a higher continuation probability, $\tilde{p}_{\tilde{e}=2}$. **Moment 5** is used to identify all the parameters.

On the other hand, I estimate the low-type driver's disutility of work with $e = 2$, $\phi^l(2)$. Since the low-type drivers did not adopt the technology, I can only obtain a lower bound for this parameter using the following theoretical intuition: a low-type θ driver would require a sufficiently high

[55] I choose to estimate K_a rather than calibrate it because this parameter may be noisy. This approach allows for a better match of other key parameters, leading to more accurate estimation of the dis-utilities of work. Calibrating it does not qualitatively change the results.

upfront payment, $W^l_{\tilde{e}=2} > W$, (for a given $\tilde{p}_{\tilde{e}=2}$) to exert high effort $e = 2$. I compute the minimum value for $\phi^l(2)$ such that, at $W^l_{\tilde{e}=2}$, the owner would actually prefer to maintain the (baseline) status quo. Appendix C.6 details the derivations used to compute the lower bound for $\phi^l(2)$.[56]

Untargeted Moment: Low-Type Driver's Contract Valuation U^l. To test the framework's fit, I consider an untargeted moment: the present-discounted contract valuation for low-type drivers at baseline. This moment was not included in the estimation procedure because I did not collect baseline contract valuation data from drivers who refused the technology, as their survey was intentionally made shorter, as discussed before. However, I use the empirical contract valuation from the group of adopting drivers who initially preferred their employer not to observe their transactions at baseline. I discuss in the main text (Section 7.3) how well the framework's predicted structural components align with both the targeted and untargeted empirical moments.

Finally, the owner's present-discounted contract valuation is defined with the following formula. The owner receives the driver's rental transfer, while the owner's costs include the upfront payment W offered at baseline to some drivers and the maintenance cost MC, along with the calibrated replacement cost K_p.

$$V^h = \frac{q_1 R + (1-q_1)X - W - MC - \delta K_p (1-\bar{p}^h)(1-q_1)}{1-\delta} \tag{D3}$$

D.3 Parameter Estimation Details

I estimate the parameters of interest using a GMM approach, which minimizes the distance between the structural and reduced-form components. My data $\mathbf{X}_i$ comprises eight empirical moments as described above. The inputs form the structural component. The GMM estimator minimizes the following objective function:

$$\hat{\beta} = \arg\min_{\beta \in \Theta} \left(\frac{1}{T} \sum_{i=1}^{T} g(\mathbf{X}_i, \beta) \right)^{\top} \hat{\mathbf{W}} \left(\frac{1}{T} \sum_{i=1}^{T} g(\mathbf{X}_i, \beta) \right)$$

Here, $g(\mathbf{X}_i, \beta)$ represents the difference between the vector of empirical moments ($\bar{p}$, p^h, $\tilde{p}$, U^h, K, R^l, R^h, $W_{\tilde{e}=2}$) and the vector of structural moments described above. Each empirical moment corresponds to a structural moment predicted by the model, allowing the GMM estimator to match observed and theoretical behavior. The weighting matrix $\mathbf{W}$ consists of the inverse variance of the estimation moments.

[56] I also specify and check in the data the "no-deviation" condition, which states that, upon adoption of the technology by high-type agents, the owners matched with low-types should have no incentive to deviate by terminating the low-type agents, incurring the replacement cost K_p, and recruiting a new agent, with probability μ of being matched with a high-type who accepts the technology. See Table B30.

I also verify the "no-deviation" conditions specified in the theoretical framework for the results to hold (see Table B30) and that the replacement cost ensures the principal is better off requiring $e = 1$ at baseline with no information on effort and output (and not $e = 2$) and setting the transfer to the maximum in high-output periods.

D.4 Sensitivity of Parameter Estimates to Estimation Moments

I assess the sensitivity of the parameter estimates derived above to the matched moments, following Andrews et al. (2017). In particular, I derive the following sensitivity parameter:

$$\Lambda = (\mathbf{J}'\mathbf{W}\mathbf{J})^{-1}\mathbf{J}'\mathbf{W}$$

where $\mathbf{J}$ is the 5×4 Jacobian matrix of derivatives of the 8 moments with respect to the 3 parameters ϕ_1^l, ϕ_1^h, and ϕ_2^h; and $\mathbf{W}$ is the weighting matrix, as described above. The sensitivity measures the asymptotic bias of the parameter estimates under local perturbations when all other parameters are held fixed. More specifically, the columns of Λ represent the sensitivity in dollars of a given parameter estimate to a one-unit change in each of the moments; the rows of Λ represent the moments.

To simplify interpretation, I convert the sensitivity values as follows: a 5-percentage-point change in the probability by the end of the experiment (after 28 weeks), the contract valuation U_h as a USD 100 change, and the other parameters—the replacement cost K_a, the transfers in high-outcome periods R^h and R^l, and the upfront payment $W_{\tilde{e}=2}$—as USD 10 changes.

Figure A13 displays the sensitivity matrix Λ in four panels, each corresponding to one parameter. I observe relative differences in parameters' sensitivity to the estimated moments. I find limited sensitivity to most moments for the disutility of work for low- and high-types. The three disutilities of work are primarily determined by the continuation probabilities p and the transfers R, consistent with the economic intuition that these primarily set the incentive compatibility constraints of the agent. The disutilities are not very sensitive to the contract valuation or replacement cost. Overall, the sensitivity is within reasonable dollar ranges, supporting the robustness of the structural results.