Misconduct and Reputation under Imperfect Information

Francis Annan[♣]
Georgia State University

v: MARCH 2021

Abstract

Misconduct – market actions that are unethical and indicative of fraud or wrongdoing – is a significant yet poorly understood issue that underlies many economic and financial transactions. Does misconduct in markets matter? When and how does reputation acts as a discipline against seller misconduct? We design a field experiment to study the impact of two-sided anti-misconduct information programs on markets, which we deploy on the local markets for mobile money (Human ATMs) in Ghana. We show that, at baseline, these markets are characterized by substantial imperfect information, consumer mistrust, and vendor misconduct. The information programs led to a large reduction in misconduct (-21 pp = -72%) and as a result, an increase in overall market activity, firm sales and consumer welfare. We develop a simple sanctioning framework between vendors and consumers that shows the treatment effect is due to a combination of more accurate consumers' beliefs about misconduct and increased reputation concerns. Together, our results indicate a potentially significant source of local financial market frictions, where market activities are underprovided due to misconduct and difficulty in building reputation. Social sanctions through reputational impacts can promote formal local markets when formal sanctions are weak.

Keywords: *forensics and information* (D18, **D83**), *reputation* (**L14**, Z13), *household finance* (D14, **O12**)

♣ Georgia State University, J. Mack Robinson College of Business, 35 Broad St NW, Atlanta, GA 30303. e-mail: `fannan@gsu.edu`. For helpful discussions and suggestions, we thank Arun Chandrasekhar, Xavier Gine, Paul Goldsmith-Pinkham, Glenn Harrison, Erzo F.P. Luttmer, Sandip Sukhtankar, Dean Karlan, Supreet Kaur, Pascaline Dupas, Isaac Mbiti, Sean Higgins, and seminar participants: Household Finance Seminar, Georgia State University CEAR/RMI, University of Virginia Economics, PEDL Conference on Firms in Low-Income Countries, and BREAD 2021. We thank the NBER Development Economics Program Fall meeting participants for comments on an earlier draft of this paper. Field support from Samuel Adotevi and Kwamena Arkafra (of Ghana Statistical Service) is acknowledged. We are grateful to the World Bank for funding. The World Bank had no involvement in any substantive aspect of the research project. Institutional Review Board (IRB) approvals for research data collection were obtained from Georgia State University. The project was registered in the AEA RCT Registry, ID-0003812.

I Introduction

Misconduct – market actions that are unethical and indicative of fraud or wrongdoing – is a common and partially observed phenomenon that underlies many economic and financial transactions. Recent studies have begun to illuminate the nature and potential welfare consequences of misconduct in transactional markets (Egan, Matvos and Seru 2019a, Annan 2020; Egan, Matvos and Seru 2019b). In theory, concerns for reputation (by profit maximizing firms or businesses) deter misconduct and encourage quality provision in markets (Karpoff 2012; Shapiro [1982, 1983]). In practice, however, it might be difficult to establish and maintain reputational capital in a market environment with much imperfect information (see, Bai 2019), as it is difficult to demonstrate the quality of services or transactions between transacting parties.

Reputation itself becomes effective and disciplinary if there is a high probability of detecting misconduct (Burkhardt 2018) and thus, the presence of imperfect information might exacerbate misconduct, with potential impacts on market efficiency. For example, misconduct can raise the marginal cost of transactions and increase uncertainty in prices (Olken and Pande 2012), which may reduce market demand (Shapiro 1983; Coppejans et al. 2007) and overall firm growth (Jensen and Miller 2018). In emerging and developing financial market contexts, misconduct is likely to be particularly significant as transacting parties are poorly informed and institutions are weak.

This paper asks two related questions. First, does misconduct in markets matter and if so how? Second, when and how does reputation acts as a discipline against seller (or vendor) misconduct? We use a field experiment to address these questions, examining the impact of providing low-cost market-level anti-misconduct information sets on vendor misconduct, market activity, and consumer welfare in eastern Ghana. If imperfect information exacerbates vendor misconduct which in turn reduces consumer demand and overall firm growth, then anti-misconduct information programs should reduce vendor misconduct and improve outcomes for both consumers and enterprises. If the information programs increase the probability of detecting and punishing misconduct, then this should raise vendor concerns for reputation.

We conduct our experiment on the market for mobile money (M-Money), an economically important financial market innovation which has been shown to improve welfare and reduce poverty (Jack and Suri 2014; Suri and Jack 2016). These markets are, however, characterized by much imperfect information about the official transaction tariffs (poor consumer knowledge), substantial vendor misconduct (market vendors overcharge over 22% of transactions), consumer mistrust and misperceived beliefs about misconduct. These features make M-Money an ideal setting to study misconduct and reputation under imperfect information.

We construct a unique census of local markets for M-Money between February-March 2019 and then perform our experiment by randomly assigning these markets (villages) to three anti-misconduct information programs about either price transparency (PT), monitor and report (MR), or both (PT+MR, their interaction). In the PT treatment, consumers receive relevant information and training about the official transaction charges. In the MR treatment, consumers were given a toll-free number to report suspected misconduct to providers or authorities. The joint treatment combines PT and MR information sets. In all cases, vendors were informed that customers have received such information earlier and the same information sets are then given to the vendors, making our interventions two-sided.

M-Money provides financial services which are delivered on digital mobile networks to potential consumers and comprises market vendors – who are small business outlets that provide cash-in and cash-out services to the consumers (Human ATMs), earn transactional commissions as their profit, and exchange cash for so-called e-money. A typical market (village) is made up of about 3 vendors. One distinguishing feature of M-Money is that the official charges on transactions are ex-ante set by providers that the market vendors work for. We use this to cleanly define misconduct as all transactions at the vendor point that are over-charged, which can be derived by comparing observed transaction charges to the provider-approved prices (Egan, Matvos and Seru 2019b; Annan 2020).

The experiment involves 130 independent local markets in 130 different localities across 9 districts. The large number of markets allows for randomization at the market-level. Markets designate reconstructed pairs of randomly selected vendors and their nearby customers, randomized into the 2×2 information design. The intervention lasted over twenty-two weeks. We tracked several outcomes at endline: households or consumers usage of M-Money, shocks exposure and mitigation (experiences of household shocks that they could not financially remedy), poverty, and collected vendors sales records of mobile money and other goods to examine the supply side effects and to directly validate the household transaction data.

We propose an audit study to measure vendor misconduct: trained auditors visited vendor points to make actual transactions, whose charges are compared to the official tariffs to infer misconduct. By using real transactions that span different transaction types, we recover rich information about market behavior and avoid major criticisms of standard audit studies within economics: deception and its subsequent effect on the market (see e.g., Ortmann and Hertwig 2002; Kessler, Low and Sullivan 2019). Our dataset is unique due to its size (130 random vendors and 990 customers), the expansive set of outcomes from both sides of the market, the administrative audit measures of misconduct, market census and surveys, and the 2×2 random information variation at market-level. We find four set of results.

First, as a first stage, the intervention reduced vendor misconduct dramatically. Overall, the incidence of vendor misconduct decreased by -21 pp = -72%, while the severity of misconduct decreased by -GHS0.68 (-$0.14) = -86%. With a control mean of GHS0.78, the latter means the intervention led the total fee (official charge + misconduct) to fall from about 1.80% to about 1.10%, implying about 40% reduction of typical M-Money transactions fees. The joint intervention and the MR intervention show economically larger reduction in market vendors' misconduct, however the PT program also had a negative impact on misconduct. Next, we find significant spillover effects: non-treated vendors located in treated villages reduced their misconduct (-15 pp overall), suggesting a large market-wide impact of our information programs on overall local market behavior. This dramatic reduction in vendor misconduct due to the information sets impacted various real consumer and market outcomes.

Consumer (or household) outcomes improved except for overall poverty. Customers meaningfully increase their uptake of transactional services (+11.2% to +40%) and savings likelihood (7.6 pp =+12.6%) at vendor points to levels that enable them to better mitigate unexpected household shocks (-6.8 pp =-7.6%). That is, consumers in treated markets were about 7.6% less likely to experience shocks that they couldn't financially remedy. We do not find evidence for an impact on overall poverty levels. The joint program shows larger impacts across the various consumer outcomes, compared to the alternative individual information, suggesting that the two individual information sets are (informationally) complements.

Vendors transactional sales volume increased. Overall, the information programs significantly increased vendors sales volume (+GHS557 per day =+46.5%). This reaffirms the estimated impacts on consumers, and shows that reducing vendor misconduct can enhance the efficiency of local financial markets by increasing the provision of market activity. For context, the 40% increase in consumer demand (47% increase in vendor sales) in response to a 40% total fee (official charge + misconduct) reduction is reasonable; it is an elasticity of about 1.0 (1.2). In additional tests, we find extended large positive impacts on vendors' non M-Money transactions, suggesting positive spillover effects of the information program on overall local market activities.

Finally, we present evidence on consumers' beliefs about misconduct and reputational concerns. The information programs caused consumers' perception about honest vendor behavior to increase (+6.7 pp = +30% overall), and importantly made such beliefs more positively correlated with the objective audit measure of misconduct (accurate and updated beliefs). The effects appear to be much larger for the joint program. Thus, thinking well of the vendors and trusting that they won't be cheated by vendors, customers increased their usage of M-Money. Vendors are also reinforced to reduce their misconduct behavior.

We show robustness of the various findings to several inference procedures, post-double-selection LASSO estimation procedure (Belloni et al. [2014]), including adjustments for multiple testing (List, Shaikh and Xu [2019]) and attrition (Lee [2009], Behaghel et al. [2015]).

What explains the estimated anti-misconduct information impacts? Our underlying hypothesis is that of reputational concerns, which we also provide evidence for. We set up a simple sanctioning and reputation framework which guides our information programs and illustrate a reputation interpretation of the results. By providing symmetric market information about vendor misconduct to (potentially uninformed) consumers, it raises vendor concerns for reputation as customers now have the technologies to infer irresponsible vendors and directly report misconduct behavior, and then go to vendors who don't engage in misconduct (reputational revenue). If vendors care about such negative or positive perceptions, then misconduct will fall, which has market-wide implications for the outcomes of our study. The model generates testable implications and allows us to make progress towards the measurement of reputational concerns. Our model is an instance of standard microeconomic analysis as applied to misconduct and market behavior, yet our empirical work is innovative: reducing vendor misconduct using two-sided symmetric market information programs, and measuring reputational concerns based on how customers are able to infer vendors misconduct and then recognizing that, vendors dramatically reduced their misconduct behavior.

We make three main contributions to the literature. First, is the literature on forensic economics (see e.g., Olken and Pande 2012; Zitzewitz 2012 detail reviews). Misconduct underlies many economic and financial transactions (Egan, Matvos and Seru 2019a, Annan 2020; Egan, Matvos and Seru 2019b), yet the sources of such concealed behavior remain less understood. We emphasize how the presence of imperfect information might exacerbate misconduct in markets, showing in an experiment that providing symmetric information to transacting parties raises concerns for reputation. Very little is known about how reputational losses act as a discipline against business misconduct (Karpoff 2012 provides a review indicating ambiguous effects). In addition, this result speaks to the broader notion on the use of local sanctions via reputation-building to promote rural financial institutions and development in low-income settings (see, Munshi 2014 for a review).

Second, is the literature on household finance and Fintech adoption (see e.g., Higgins 2020 and references therein). Much research exist on the consumer effects of Fintech, but very little is known about supply side behavior. Here, we emphasize seller misconduct as a key barrier to both sides of the market and that reducing it has meaningful impacts on consumers and vendors. Third, is the literature on information, firm behavior and growth in developing countries (see e.g., Jensen and Miller 2018; Bai 2019). We show that promoting market transparency and monitoring induce firms to seek desirable choices or quality with meaningful impacts on market activity and consumer welfare in transactional markets. Firms' actions are rewarded via increased consumer demand if customers are informed. This increases the size or sales of M-Money firms (that are small to medium sized). Overall, our findings shed lights on why small to medium firms may not grow because they fail to provide quality (by engaging in market misconduct) as quality provision is under-rewarded due to imperfect information.

From a policy perspective, our results highlight how the simple provision of low-cost two-sided information might influence misconduct and consumers trust that they won't be cheated by sellers, and how this might eventually facilitate efficient market behavior, particularly in vulnerable market environments. This is important for setting relevant consumer protection policies. Evaluating how uninformed local market buyers are and providing information about price transparency and monitoring to both sides of the market could potentially be used to build trust and increase the benefits of emerging financial markets, particularly Fintech.

II Experiment: Design

II.1 Background: Mobile Money, Market Census, and Market Facts

II.1.1 Mobile Money

The market for M-Money comprises (i) vendors, (ii) customers, and (iii) service providers. Market vendors correspond to an outlet, shop, premises or local banking channels where M-Money transactions can be carried out on behalf of the providers – which are joint partnerships between mobile network operators (MNOs) and commercial banks. Particularly, vendors register accounts for customers and act as cash-in and cash-out transaction points for customers (i.e., Human ATMs). These vendors generically earn commissions on transactions by acting on behalf of the financial service operators. The introduction and significant penetration of digital mobile telecommunications has provided a cheap infrastructure to make M-Money services accessible even to the poor and low-income societies. In these poor environments, formal financial institutions are shallow and largely absent (see, Banerjee and Duflo [2006; 2011] for authoritative surveys about this), making M-Money a competitive financial option in low-income environments.

Similar to other banking and financial services, the business of M-Money likely faces fraud and misconduct, which could take different forms. In the policy circles, regulators from Bank of Ghana, for example, have expressed concerns about such potential market misconduct. There are ongoing regulator and stakeholder discussions about eliminating emerging risks and recognizable fraud on this market and proving ultimate consumer confidence in mobile financial services. In Ghana, the MNOs and their commercial partners have been charged to build more risk and fraud-resilient financial infrastructures.[1] Our present study is designed

[1] "We also want you [Mobile Network Operators] to make your service affordable, we also want you [Mobile Network Operators] to put in place systems to minimize or eliminate fraud if possible and we also want you [Mobile Network Operators] to give wonderful customer service to your customers as they come to your premises to transact business. We want your system to have what it takes, to give very good audit trail of transactions." – **Bank of Ghana's payments oversight office head Clarence Blay, speaking at a stakeholder conference titled Expanding Cashless Payments**

to understand financial misconduct at vendor points (see, Figure 10 in Appendix G) and examine its potential efficiency costs. We do this in a rural context where the business of M-Money could have larger impacts, if well designed.

II.1.2 Market Census

Detail vendor × customer data on M-Money is unavailable. So, between February-March 2019, we carried out a unique census of the market for M-Money in Eastern Ghana, spanning 9 districts. Districts are made up of sub-administrative units called "localities" or villages. Eastern Ghana was chosen for its two attractive features: (i) it covers an expansive number of villages, with potentially mobile money vendor sites, and (ii) our initial pilot works in other parts of this region suggest substantial levels of misconduct in this market (Annan 2017). Our census exercise documents the universe of all vendor points (both formal and informal), and other surrounding households (within 5 houses radius around a given vendor) successfully across 130 localities. We focus on nearby households in order to maximize our chances of studying households that might make transactions with select vendors, while minimizing costs. We define a local market as the pair: vendor × the set of all nearby customers (see, Annan 2020 for details).

II.1.3 Market Facts

Our baseline census solicited information from all market participants: both vendors and customers. We asked information on their basic demographics, poverty and assets, detail market records on M-Money and non M-Money services, including general to specific knowledge about M-Money transactions. Additional household information on personal finance, debts, savings, shocks and investments were obtained from customers. Here, we will focus on data that are relevant to our study of market impacts of misconduct. Detail summaries, and other patterns about the market are available in Annan (2020) and upon request.

Table 14 shows the summary statistics for the market. To facilitate comparisons between both sides of the market, the relevant statistics for vendors and customers are displayed

---

Through Mobile Wallet Transactions, 2015.

next to it each other. Female vendorship is 39% – meaning that these local markets are disproportionately made up of more male vendors. 62% of the potential customers are females, and customers are more likely to be self-employed, married and older relative to vendors. Approximately and strikingly, half of the vendorship have received formal training about the market for M-Money before joining the business. The overwhelmingly majority (90% [SD=0.29]) of customers, their close family and friends networks have registered for a M-Money account (also called "wallet"), indicating that it is likely a popular financial technology.

We turn next to specific features of the market. With an average experience of 2 years in doing M-Money business, a vast majority (75% [SD=0.43]) of vendors operate as a joint venture – bundling this with other services.[2] The average daily sales per vendor is about GHS2,260 [US$442]. Thus, most of these vendors operate relatively small to medium size enterprises. The majority of households or customers use M-Money services than other alternative commercial financial services: 95% of customers are M-Money users, 80% are past formal bank users, while just 9% are post-office users. This can be explained by the convenient access and lower charges of M-Money, difficulty in access and distance to nearby services: we estimate an average distance of approximately 61 meters to the closest mobile money vendor site, while this distance is about 383 meters for post-offices.

II.1.4 Motivating Features: Asymmetric Information, Misconduct, Perceptions about Misconduct, Reputation

The presence of asymmetric information and regard for reputation are key ingredients of our study: information programs, their effects and interpretations. In the Theory section, we show evidence that our (baseline) setting reflects a market environment where (i) consumers are objectively uninformed (less sophisticated), (ii) vendor misconduct incidence is high

[2]We identified joint venture services like: groceries and provisions, local medicine, multi TV installation, registration of SIM cards, phones and accessories, airtime recharge cards, mini-credit transfers, acting as agents for land and house sales, electricals and accessories, photocopying and typesetting, educational/online results checking, electric prepaid credit, among others.

(vendors overcharge over 22% of transactions measured using administrative audit exercises versus 19-59% measured using survey elicitations), (iii) consumers misperceive the level of misconduct, and (iv) vendors value their reputation as there is a positive return to good reputation but difficult to establish such reputation (i.e., with scope to build reputational capital).

II.2 Intervention and Timetable

We evaluate the impacts of different information sets that reduce market misconduct on both customers and vendors. As we discuss later in the Theory section, the provision of relevant market information about vendor misconduct to (potentially uninformed) consumers raises vendor concerns for reputation as customers are likely able to infer (ir)responsible vendors and then assign reputational payoffs to the vendors. If vendors care about such (negative or positive) perceptions, then misconduct will fall, which has market-wide implications for the outcomes of our study. This provides theoretical basis to fix our ideas and motivate the information programs.

All local markets (vendor $\times$ customers) receive a physical research visit, and markets assigned to treatment receive additional information about misconduct. For all markets, we show subjects the reconstructed market rosters, ask them to indicate where their last financial transactions were conducted, and provide contact information of our research team for further assistance. Markets assigned to treatments additionally receive either of the following:

- Treatment program I: Price Transparency (PT) – Addresses the question of "what to ask vendors while at vendor points". It informs and educates consumers about the true tariffs for common local transactions, and thus improves consumer sophistication about detecting misconduct.
- Treatment program II: Monitor and Report (MR) – Addresses the question of "how to

report seller misconduct". It informs customers by providing a toll-free number to report suspected misconduct to authorities, and thus raises the potential cost of misconduct to vendors if caught.

- Treatment program III: joint PT+MR – A joint program that tests the interaction of programs I and II. (see, Exhibits in Appendix F for the specific information sets).
- Control program: no additional information.

To ensure meaningful treatment effects, we visit the assigned local markets 3 consecutive times over a 3 months period (once per month) to first deliver and then repeat the information programs to subjects. Visits are concluded with the subjects summarizing the information they received and keeping hard copies of the treatment program. More uniquely, we ensured that vendors are equally aware of the interventions by communicating the same information set to the vendors right after seeding the information with the nearby households – yielding a two-sided information design. Together, our treatment programs aim to reduce potential information frictions and increase the social cost of vendors' misconduct.

To roughly gauge the likely significance of the information programs, the recipients were *ex-ante* asked to rate the usefulness of the information we provided for their financial decision-making (i.e., customers) and businesses (i.e., vendors) on a 5-point scale: `1 (Not useful), 2 (Quite useful), 3 (Useful), 4 (Very useful), 5 (Extremely useful)`. Overall, the median value = 3 (mean=3.38, [SD=0.82]), suggesting that subjects view our information interventions as useful, and thus likely to be *ex-post* effective.[3] Program I is a popular consumer protection policy instrument. By benchmarking this with programs II and III,

[3]In practice, there were instances where the experimental subjects (specifically the customers) took "costly actions" to call our research team to discuss their M-Money 2-3 months after the provision of the information programs. This will suggest that subjects are willing to pay for our information programs, perhaps because they find the information credible. In addition, this will suggest that subjects' rating of the usefulness of the information provided is less likely affected by potential experimenter demand (pleasing) effects (see, de Quidt, Haushofer and Roth 2018).

we can evaluate program I's relative effectiveness in reducing market misconduct committed against consumers, and assess whether program I is compatible with other information programs or whether it only becomes effective when combined with an alternative that increases the cost of misconduct to firms.

Table 1: **Study timeline**

	DATE	ACTIVITY
Part 1	February 2017	**Pilots:** Misconduct– incidence, correlates (Annan 2017)
Part 2	Feb 15-Mar 20, 2019	**Baseline:** Market census– detail market records, demographics, main outcomes, misconduct beliefs
	Sep 01-Oct 15, 2019	**Audit study I:** Estimate misconduct, $\geq$ 1 in 5 transactions (22%)
Part 3	Oct 15-Dec 15, 2019	**Intervention:** Information assignment Control: no information Treatment I: price transparency Treatment II: monitor and report Treatment III: price transparency + monitor and report
		Transaction networks data: family vs friends vs strangers
Part 4	May 15-May 30, 2020	**Endline:** Phone survey + manual tracing supplement main outcomes, misconduct beliefs
	Aug 15-Sep 01, 2020	**Audit study II:** Re-estimate misconduct (12% of transactions)
	> Sep 15, 2020	**Administrative data (vendors):** transaction record volumes

II.3 Data Collected

We gather information from multiple rounds of data collection (i) combined listing and baseline market census (process discussed in "Market Census" above), (ii) baseline audit study (process discussed below), (iii) transaction networks data, (iv) 22-weeks follow-up (phone) market survey, and 33-weeks administrative audit study, which we call an endline.

II.3.1 Administrative Audit Data

To objectively measure "true" misconduct, we employ an audit study where auditors (experimental customers) were given cash to make actual transactions on M-Money, as credible data on misconduct is directly unavailable. The transactions span multiple transaction types which are common in the market (12 different transactions in total): sending *versus* receiving transactions. Tariffs on transactions are ex-ante set by the providers. To mimic the local market context and properly capture misconduct, we recruit and use local residents,[4] who were trained to trained to follow the same approach on how to interact with the vendors, particularly use uniform language at visits to vendors (see, Annan 2020 for details). This approach has the strengths of measuring the true incidence of misconduct (unlike other survey-based measures of misconduct; DeLiema et al. 2018), while avoiding deception and its later effect on the market (unlike other standard audit studies; Kessler, Low and Sullivan 2019).

In our market setting, (and as in Egan, Matvos and Seru 2019b; Annan 2020), we define misconduct to entail transactions that are over-charged when compared to the regulator and provider-approved tariff rates. Table 15 and Figure 5 in Appendix C show the baseline results across the various transactions. We estimate that 22% of transactions are overcharged (reflects the incidence of misconduct) and GHS3.3 (= 82% of the official tariffs) overpaid to the vendor as a result of misconduct (reflects the severity of misconduct). There is heterogeneity in misconduct levels across the different types of transactions or groups. More importantly, misconduct is concentrated in over-the-counter (OTC) transactions, which by construct involve little to no automation or active verification from the side of the customer, and thus more vulnerable to vendor misconduct. This is reassuring and alleviates potential concerns that the auditors might be over- or under- measuring misconduct.

[4]A potential concern is that vendors cheat strangers (like the auditors) but not the local repeat customers that they know. This is not a major concern here for two reasons. First, it might be more risky to cheat strangers because they might be more informed. This reduces that possibility of systematically cheating strangers. Second, in our market environment, we estimate that a very large share of the market transactions are conducted with customers who have no family and/or close relations: customers from our study area were shown the locality-level roster of all vendors and then asked to indicate where they last transacted at and how they are related to that vendor: 8.0% of transactions were between participants who are blood-related, 22.0% were between participants who are friends, and 70.0% are not related at all.

II.3.2 Market Survey Data

We measure several repeated outcomes at different stages of the study. For customers, we restrict attention to 4 relevant outcomes: (i) adoption and usage of money services: ask whether households use the money services, and the transaction amount involved per week, (ii) savings on M-Money: ask whether households save on their money wallets, (iii) specific shock experiences (health, revenue, household expenditure shocks) and risk mitigation: ask whether customers experienced unexpected shocks that they could not financially remedy or pay for – providing an objective proxy for insurance (Dupas and Robinson [2013]; Breza and Chandrasekhar [2019]), and (iv) poverty. Since our study focuses on mobile money in low-income and poor environments, we fielded questions that allow us to directly examine poverty. We adapted a recently develop short-cut—yet rigorous, inexpensive, simple and transparent—measure of poverty called the "Simple Poverty Scorecard". Details about this poverty scoring methodology can be found in (Schreiner 2015).

For vendors, we measure sales volume: solicit transaction records for their M-Money business and non M-Money services (if the vendor operates a joint venture). Together, we gather data from both sides of the market, which allows us to cross-validate accuracy of the records. For example, one will expect increases in households money transactions to (positively) correlate with increases in nearby vendor sales, all else equal.

II.4 Treatment Assignment

We use a 2×2 factorial design, randomizing the total 130 "representative" markets (as defined below) into 4 experimental anti-misconduct programs: PT-only (31 markets $\equiv$ 31 representative vendors $\times$ 272 nearby customers), MR-only (32 markets $\equiv$ 32 representative vendors $\times$ 257 nearby customers), joint program (35 markets $\equiv$ 35 representative vendors $\times$ 276 nearby customers), and control program (32 markets $\equiv$ 32 representative vendors $\times$ 185 nearby customers). We stratified based on districts, and all misfits are resolved and randomly assigned.

II.5 Balance and Validity of Design

II.5.1 Balance I

We focus our study on selected "representative" markets (villages) drawn from a listing of the baseline market census. Each of the 130 localities has one or more vendor(s) (range=1-12, average=3.3) with their surrounding customers or households (range=5-47, average=20.8). To maximize statistical power, we "randomly" select one vendor and his/her nearby customers per locality for our study. We call this combination (representative vendor × nearby households) a "representative" market. Sample representativeness requires that being a "representative" market is independent of any relevant market-level statistics. To test that these samples are comparable to the market population, we run the regression

$$y_{mv} = \alpha + \beta S_{mv} + \epsilon_{mv}$$

on the baseline census data, where $S_{mv} = 1$ if market pair m from the pairs in village v is selected to be a representative in the *pre* intervention period. We consider a number of different relevant outcomes, and show that both sides of the market show no observable differences across the two groups. Tables 9 and 10 report the results, where we find no difference across markets selected and those not-selected to be representative.

II.5.2 Balance II

We base our treatment analysis on a comparison of "representative" local markets ($m = v$ now) that received the information treatments with those that did not receive the treatments. Successful randomization of treatments, and thus identification requires that the assignments to treatments (i.e., price transparency-only, monitor and report-only, and joint information sets) are independent of any relevant household or market-level statistics. Similarly, to test that these markets are comparable, we run the regression

$$y_{iv} = \alpha + \beta \mathbf{I}_v + \epsilon_{iv}$$

on the baseline data, where $\mathbf{I}_v = 1$ if local market v in district d received an information treatment, 0 otherwise. We consider the various treatments separately and together (i.e., meta) for a number of different outcomes, and show that both sides of the market show no observable differences across the two groups. Tables 11 and 12 report the results, and provides strong evidence in favor of balance with no difference across subjects i (households or vendors) in assigned (treated) and non-assigned (control) markets.

II.5.3 Attrition

Our randomization is based on the selected "representative" markets that draws on the baseline market census. Table 13 displays the breakdown of response rates and attrition between baseline and endline. Here, attrition may be linked to subjects non-response, migration to outside the locality, and inability to reach the participants either because their phone numbers are inactive or out of network coverage area. To maximize response rates at endline, trained field officers conducted multiple phone calls (see, Figure 4) at different time horizons of the day, varying either weekdays or weekends, combined with manual contact tracing for subjects with inactive phone numbers. We record an overall attrition rate of 18%, which is low given that the business of M-Money is subject to high degree of migration and operator turnovers. Attrition looks non-differential. For our endline audit transactional exercises, 129 out of the 130 representative vendors were reached, implying an attrition rate of just 0.8%. In our empirical estimations, we evaluate and formally show robustness to attrition by treatment status.

III Experiment: Results

We present and discuss the treatment effects. Since all our treatments are about information provision, we first report the (combined) meta effect of information assignment, and then the separate effects for the different treatments.

III.1 Empirical Specifications

We estimate treatment effects using the model:

$$y_{ivd} = \beta \mathbf{I}_{vd} + \eta_d + \beta_0 y_{base,ivd} + \mathbf{X}'_{ivd}\xi + \epsilon_{ivd}$$

which links various endline outcome(s) y_{ivd} of subject (customer or vendor) i in locality (village) v in district d to the random treatment variable(s) $\mathbf{I}_{vd}$, district-level (stratification unit) dummies η_d, baseline outcomes $y_{base,ivd}$ and additional vector of controls $\mathbf{X}_{ivd}$. We include baseline outcomes to primarily increase precision but these control for potential confounds (*if* any). For the meta effects, $\mathbf{I}_{vd}$ is a 0-1 indicator for whether a locality received any of the information programs, and thus β captures the (meta) treatment effect. For the separate effects, $\mathbf{I}_{vd}$ is a 0-1 indicator for whether a locality received a specific information program. We denote by β_1, β_2 and δ the separate treatment effects for PT-only, MR-only, and joint information sets, respectively (i.e., $\beta = (\beta_1, \beta_2, \delta)'$).

For inference and robustness, we report various standard errors including, the wild bootstrap cluster-t and randomization inference both clustered at the (village) market level. To address the potential issue of multiple testing, we adjust p-values for multiple testing across family of outcomes following the procedure presented in List, Shaikh and Xu (2019). To evaluate and show robustness for "potential" attrition bias, we report Lee (2009) attrition bounds (trimming based on observed attrition rates; see, Table 13), Imbens and Manski (2004) confidence sets, and Behaghel et al. (2015) attrition bounds (trimming based on the number of times subjects were called before answering the phone survey; see, Figure 4). In alternative models, we choose $\mathbf{X}_{ivd}$ using post-double-selection LASSO (for good estimation performance, in addition to minimizing researcher degrees of freedom and the possibility for p-hacking; Belloni et al. [2014]). We will sometimes discuss effects that contain useful economic information (i.e., looking at effect sign and effect size)–whether significant or not (Abadie 2020).

III.2 Treatment Effects of Information Sets on Misconduct

As a first stage, we ask "whether the information programs are anti-misconduct?" Table 2 reports the meta and separate treatment effects, and shows that the intervention meaningfully reduced vendor misconduct (measured using actual audit transactions). We estimate a meta effect of -21 pp (-72%+ of control mean) for misconduct incidence and -GHS0.55 for misconduct amount (-63% of control mean). The effects are economically much larger for the joint and MR programs, however the differences across the programs are barely distinguishable statistically.

In additional tests (see, Table 16), we find significant spillover effects: non-treated vendors located in treated localities (or markets) reduced their misconduct (-15 pp meta effect). Motivated by previous theoretical and applied research (Matsa 2011; Annan 2020), we examine heterogeneity in effects along two dimensions: market competition and vendors' gender. Baseline data on vendor sales is used to construct a Herfindahl-Hirschman index, where a lower index reflects higher levels of market competition. The estimates (Tables 17 and 18) show that the reduction in misconduct is much larger in localities with more competition, particularly for the joint information program. The effects are similar across gender, which means female vendors might have responded more to the information programs because at baseline (pre-treatment), the female vendors were significantly more likely to commit misconduct relative to male vendors. This suggests that both the underlying market structure and vendors' gender matter for the impact of anti-misconduct information programs. In this case, corrective policies to influence misconduct committed against consumers can include schemes that facilitate competition in financial services for the poor, and /or bear on the gender distribution of market vendors.

These results strongly confirm that the information programs are indeed anti-misconduct, yielding economically very large and statistically significant decrease in both incidence (the occurrence) and intensity (shift in the distribution) of misconduct. We next evaluate how this dramatic reduction in misconduct due to the information sets impacted the various consumer and market outcomes.

III.3 Real Effects: Graphical Evidence of Treatment Effects

We provide graphical illustration of the treatment effects. Figure 1 plots the empirical cumulative distributions of endline log(transaction amounts per week) by treatment status. The effects are displayed for the various treatments together (meta) and separately (in keeping with the approach of reporting the meta versus separate information effects). There is strong visual evidence of positive effects of the information programs on customers transactional outcomes. This implies increased uptake of the M-Money financial services as a result of the information program. What is more striking is that the effects does not seem to be driven by specific parts of the distribution. A Kolmogorov–Smirnov (KS) test for equality of distributions rejects the null that the distributional pairs are equal in all cases (p-values<0.080) except for the PT-only program (p-value=0.288). Thus, there is a considerable difference between the distribution of treated versus control local markets as we reject the null hypothesis of no distributional effects. We proceed to quantify the impacts for the various economic outcomes. Our estimates are robust to alternative controls, inference procedures, and adjustment for attrition.

III.4 Information Assignment – Meta Estimates

III.4.1 Effects on M-Money usage and Savings

Tables 3 and 4 report the estimated meta effects on usage of services and savings, respectively. There is increased transaction amount per week (see, Table 3), with a semi-elasticity of 0.402 (=+11.2% of control mean, p-value=0.048). In Appendix Table 20, we report the effects on the probability of using financial services, showing increased transaction likelihood of usage per week (7.1 pp =+9.8% of control mean, p-value=0.049). For savings, there is evidence of increased savings rate (see, Table 4) by 7.6 pp (=+12.6% of control mean, p-value=0.099).

III.4.2 Effects on Mitigation of shocks: revenue, health, and expenditure

Did customers' (or households) increase their transactional services and savings likelihood in meaningful enough levels that they are better able to mitigate unexpected shocks? Table 5 shows the estimated meta effects on customers experiences to unmitigated shocks. We report this for general shocks (any experience), and then individually for shocks related to household revenue, health, and household expenditures, respectively.

There is reduced instance(s) of general unexpected shocks that consumers cannot financially remedy or pay for (i.e., when resource limits bind) (-6.8 pp =-7.6% of control mean, p-value=0.068). This effect is mainly driven by household expenditures, which has the largest significant reduction of 10.7 pp. However, both the health and revenue sources are equally meaningful looking at their effect sizes (7.2 pp and 5.6 pp, respectively). These estimates provide a large and objective proxy for resilience and insurance value of reducing market misconduct to consumers.

III.4.3 Effects on Poverty

In Table 7, we test whether the information program impacted poverty (Schneider 2005)–which is an objective measure of consumer welfare. We do not find evidence for an impact on overall poverty levels. This is less surprising given that poverty is a structural and composite outcome, which can take a while to see effects. We hypothesize, however, that increased adoption of the financial services, savings and continued resilience of households to unexpected shocks (reported earlier), might translate into longer-term changes in poverty. Moreover, there could be significant distributional impacts on poverty.

III.5 Information Sets –What's necessary, What's sufficient?

We now report the separate impacts by the different information programs.

III.5.1 Effects on M-Money usage and Savings

Table 6 shows the estimated effects of the various information sets for the uptake of services and savings. For uptake of services, the effects are positively much larger for the joint program (semi-elasticity of 0.506 =+14.1% of control mean, p-value=0.035), compared with the other individual information sets. The results are similar for savings behavior at vendor points. Customers are significantly more likely to save on M-Money with much larger impacts for the joint program (semi-elasticity of 0.123 = +20.2% of control mean, p-value=0.024), compared with the other individual information sets. A Wald test rejects the null that the savings effect from the joint program is equal to effect from the monitor and report-only information set (p-value=0.066)

We combine all the usage and savings outcomes (via PCA) (see, column 5 of Table 6), finding that the effects are consistently larger for the joint program. This is followed by the MR-only, and then the PT-only information sets. These results indicate that the MR-only and PT-only programs are (informationally) complements, and that the PT alone (a popular consumer protection instrument) may not be sufficient except when combined with random information assignment about MR.

III.5.2 Effects on Mitigation of shocks and Poverty

The estimated impacts for the various information sets on both shock mitigation and welfare are reported in Table 7. For shock mitigation, the joint information program show significantly (negative) larger impacts, compared to the alternative individual information counterparts. As in the meta estimate, this effect is mainly driven by mitigation of unexpected shocks related to household expenses. Effects from the MR-only program are rather smaller and insignificant. For poverty, we also do not find evidence for an impact on overall poverty levels across the various programs. These results agree with our earlier findings that the two individual information sets are (informationally) complements and that the impact on poverty as a structural or composite outcome may distributional.

III.6 Treatment effects on vendor transactions

Did market vendors' experience an increase in sales? If the consumers records are accurate, and hence the estimated treatment effects, then one might expect direct increases in vendors transactional volumes (all else equal). We test for this and in doing so, evaluate the impact of the information intervention on vendors market activity.

Table 8 reports the estimated impacts on vendors. As expected, the information programs significantly increased vendors sales volume (+GHS557 per day = +46.5% of control mean) overall. The estimated +GHS557 per day seems reasonable: it translates to about GHS3,899 per week (for vendors). If we divide this weekly estimate by +GHS95 (as customers transact about +GHS95 per week in response to the information intervention; see, Table 19), we get approximately +41 more customers per week at a representative vendor point. This may reflect both repeat and/ or distinct customers.

For context, the typical transaction here is about GHS100 (based on the experimental transactions of GHS50, GHS160 and GHS1100 which were chosen to be typical of the market setting, Table 15). The regular and official fee will be 1% of the transaction value, which implies a fee value of GHS1. The experiment led the total fee (regular fee + misconduct) to fall from about 1.75% to about 1.10% (Table 2). This is about a 40% reduction of the transaction fee. The 40% increase in consumer demand (46% increase in vendor sales) in response to a 40% fee reduction is reasonable; it is an elasticity of about 1.0 (1.2).

In additional tests (not reported), we find extended large positive impacts on non M-Money transactions, suggesting positive spillover effects of the information program on overall local market activities...

IV Framework: Interpreting the Results

We present a framework to guide the interpretation of our results. We seek to understand what happens when we give relevant seller misconduct information to both a (potentially dishonest and informed) vendor and (potentially uninformed) consumers in a local finance context. One could tell several stories about how the information intervention might act to affect misconduct, and thus the market outcomes. Our underlying hypothesis, however, is that vendors expect that they are more likely to be perceived (by potential customers) as *irresponsible* if they commit misconduct in our experiment. Such negative perceptions trigger direct punishments and affect vendors' reputation (e.g., via a reduction in vendors sales in other joint lines of business, customer referrals, including other future market and social relations akin to relational contracting, Gibbons and Roberts 2012). This yields a misconduct sanctioning vs. reputation-type interpretation.

Our goal is not to develop a general theory of either misconduct (e.g., Banerjee et al. 2012 for corruption) or reputation and moral hazard (e.g., Board and Meyer-ter-Vehn 2013). We rather provide a parsimonious model of moral hazard under revelation that embeds misconduct and sanctioning to deliver highly stylized predictions which guide the interpretation of our results. We turn first to relevant features of our setting to motivate the modeling framework and subsequent interpretation. These features provide an empirical analog and building blocks of the model and empirical tests.

IV.1 Baseline Setting

We document relevant features of our empirical setting by providing 3 pieces of descriptive evidence: the presence of asymmetric (imperfect) information about the true transactional prices between vendors and customers, difficulty of vendors in establishing market reputation – amplified by the limited trust of customers in transacting, and misperceptions about misconduct – making it difficult for customers to infer otherwise (ir)responsible vendors.

IV.1.1 Feature 1: Asymmetric (Imperfect) Information

`Customers are less knowledgeable about true prices relative to market vendors (at baseline)?`

We draw on data from the baseline market census to examine if vendors have superior knowledge about the true transactional prices compared to customers. In a series of tests, both vendors and customers were asked to indicate the true charges for two randomly chosen transactions of sizes GHS200 (small to medium) and GHS1200 (large). For the vendors, we were careful to inform them at the beginning that we were not there to perform any actual transactions, but to rather assess their overall knowledge about the market (to alleviate any potential incentives for misconduct). Knowledge tests were taken towards the end of the surveys for both subjects. With reference to the official charges, this provides us an estimate of their knowledge about the true charges, specifically the % of subjects whose answers were correct across the markets.

Results are displayed in Figure 6, showing strong evidence of asymmetric information: vendors have superior knowledge about the true transactional charges relative to customers. While market vendors are relatively more knowledgeable, their knowledge is also imperfect. This noise can also limit the ability of vendors to build reputation as it exacerbates the incidence of misconduct. These results are expected because (unlike customers) vendors receive formal training about the market for M-Money before they start their businesses. Although a universal requirement, approximately and strikingly, only 50% of the vendors indicated they received formal training at the baseline (see, Table 14).

IV.1.2 Feature 2: Reputation

`I. Vendors: importance of good market reputation to vendors?`

We asked a random sample of vendors in the control group of our experiment *post*-endline about how important is it to show a high degree of good market reputation (or image and responsibility) to potential customers through their market transactions. As shown in Figure 7, the vast majority of vendors (81% [SD=0.391]) consider good market reputation or image as important. This descriptive suggests that the vendors value their reputation as there is likely a positive return or reward to good market reputation.

`II. Customers: (mis)trust for carrying out m-money transactions (at baseline)?`

Our baseline census solicited information about customers' level of trust in carrying out their transactions in the market. Figure 8 reports the results, suggesting limited level of trust. About 62% [SD=0.48, n=1275] of the customers indicate distrust in transacting at vendor points, while the rest (i.e., 38% [SD=0.48, n=779]) have trust making their money transactions. This suggests that vendors have low reputation in the market, perhaps because either (i) some vendors find it difficult to establish such reputation, or (ii) some customers are unable to infer the responsible vendors, both consistent with our earlier evidence of imperfect information about transactional tariffs.

IV.1.3 Feature 3: Perceptions about Misconduct

`Misperceived beliefs about misconduct (at baseline)?`

Figure 9 compares the true versus subjective beliefs of misconduct. Our actual audit transactions provide an objective (true) misconduct incidence of 22% [SD=0.41, n=663] at vendor points. We denote this by $(1-\pi)$, implying that π is the % of honest transactions (i.e., transactions not overcharged). Details about the objective estimates of misconduct (dishonest transactions) across a range of market audit transactions are illustrated in Table 15 and Figure 5, as noted. Next, we also asked customers views, at baseline, about the incidence of misconduct, yielding an overall subjective incidence of 59% [SD=0.49, n=1921] (denote that by $(1-\hat{\pi})$; implying a subjective incidence of honest transactions $\hat{\pi}$=41%). Of course, the subjective belief estimate about honest transactions $\hat{\pi}$ could be much higher, depending on how it is elicited. For our analysis, we thus assume consumers misperceive the level of honest vendor behavior (and hence will allow for misperceived beliefs $\hat{\pi}$), and that the measured $\hat{\pi}$ is a (good) proxy for the relevant $\hat{\pi}$, which is lower. This assumption agrees with the observed departure of $\hat{\pi}$ from π and why misconduct is "prevalent" in the market at baseline.

Thus, this reflect an empirical setting where (i) consumers are objectively less sophisticated (uninformed), (ii) market vendors value their reputation in the market but such reputation is difficult to establish because consumers cannot observe whether they are cheated because they don't know the official price that they are supposed to be charged. This suggest that there is a positive return for good market reputation (e.g., extended sales, borrowing, referrals) if viewed by customers as responsible, and (iii) at baseline, consumers under or misperceive the level of honest transactions. Our setup and information programs work to reduce vendors misconduct and enhance consumers subjective belief or perception about the level of honest vendor behavior. Moreover, consumers might transact more if misconduct (equivalently, the marginal cost of transactions) is low.

IV.2 Model: Misconduct, Punishment and Reputation

IV.2.1 Environment

Assume a continuum of local markets, defined by the pair: {representative vendor i, potential customer(s) j}. This is akin to our experiment's design, whereby we construct a local market using a randomly selected representative vendor and the nearby households as customers per locality to maximize statistical power. In each locality, the other vendors and customers have no designated role; our model will inherit the same design. We present a simple model of moral hazard under revelation with reputational effects and direct punishment.

The vendor chooses an action $s \in \{0, 1\}$, where $s = 0$ refers to a dishonest action (does overcharge market transaction) and $s = 1$ refers to an honest action (does not overcharge market transaction and thus responsible). Taking an honest action $s = 1$ generates a cost C. Customers imperfectly observe the vendor's action, but learn about the transaction through public signals σ, giving rise to a moral hazard problem (Board and Meyer-ter-Vehn 2013). Denote by π the % of honest transactions (or probability that the vendor will be honest), so $\Pr(s = 1) = \pi$. We allow customers to hold imperfect belief about the probability that the vendor will be honest, which we denote by $\hat{\pi}$. $\hat{\pi}$ is assumed to be common knowledge to avoid instances of higher-order beliefs.

Let p_j denote the probability that potential customer(s) j visit the vendor point and that a financial transaction takes place. So, in every local market {representative vendor; nearby customer(s)}, potential customer(s) j visit each representative vendor, independently, with probability p_j. If the transaction takes place, the vendor receives a revenue in two ways: reputation (from honest behavior) and "uncertain" direct benefits (from dishonest behavior). First, given the public information σ, consumers' willingness to pay is $\mathbb{E}^{\hat{\pi}}[s = 1|\sigma, s]$; this equals the vendor's reputational payoff given the signal. We call this reputational payoff as the vendor cares about $\mathbb{E}^{\hat{\pi}}[s = 1|\sigma, s]$ that customers' compute (i.e., posterior that the vendor is honest) and assigns, which takes place immediately (as in Shapiro 1983). As a practical foundation: if the customer thinks well of the vendor, the vendor will have access to valuable future opportunities e.g., extended sales, borrowing, referrals. Second, if the vendor chooses $s = 0$ (being dishonest), s/he receives an additional benefit $Y > 0$ corresponding to the overcharged transaction amount. However, with probability q, consumers can directly punish the vendor by reporting the dishonest behavior; the vendor gets $Y^s \mathbf{I}_{s=0} < Y \mathbf{I}_{s=0}$ if reported. Given the vendor's action s and market consumers' belief about this action $\hat{\pi}$, the vendor's profits $\mathbf{\Pi}(s)$ equal

$$-C\mathbf{I}_{s\neq 0} + [qY^s + (1-q)Y]\,\mathbf{I}_{s=0} + \hat{\pi}\mathbb{E}^{\hat{\pi}}[s = 1|\sigma, s](\sum_j p_j) + (1-\hat{\pi})(1 - \mathbb{E}^{\hat{\pi}}[s = 1|\sigma, s])(\sum_j p_j)$$

IV.2.2 Mapping Model to Experiment

Before analyzing the framework, it is useful to discuss how our model and analysis map to our experimental set-up. Market vendor(s) decide whether to commit misconduct ($s = 1$) or not ($s = 0$). Consumers (uniformed vs informed) learn about the transactional action through public signals σ. Based on consumers' inference about the vendor's action given the available signal, customers either assign a reputational payoff ($\mathbb{E}^{\hat{\pi}}[s = 1|\sigma, s]$) to the vendor (via PT information program) or report the vendor's dishonest behavior as a direct punishment (via MR information program). If customers believe or perceive (via $\hat{\pi}$) that the vendor is honest, then the vendor receives higher revenue (e.g., through repeated visits to transact, not reporting the vendor) and vice versa.

Our goal is to compare market information sets about misconduct: one "without" information and another "with" information assignment about misconduct. For the information assignment, we vary the information sets: one with the technology to detect and reward misconduct behavior (reputation; where $\sigma = s$), another with the technology to directly report and punish misconduct behavior (punishment), and both. We model assignment of the anti-misconduct market information as either a shift in the distribution of $\hat{\pi}$ or $\mathbb{E}^{\hat{\pi}}[s=1|\sigma,s]$. As we show (and as implied by the model), the information assignment (i) increased customers beliefs about % honest transactions $\hat{\pi}$, (ii) caused customers to update their beliefs about honest vendor behavior (thus to assign $\mathbb{E}^{\hat{\pi}}[s=1|\sigma,s]$), and (iii) caused vendors themselves to update their beliefs about how informed consumers are and the likelihood of direct punishment. Together, these increased honest market vendor actions ($s=1$) and improved market outcomes by increasing consumers' demand for services and vendor's transaction volumes.

IV.2.3 Analysis

In the game, we are interested in Perfect Bayesian Equilibria. Let us define $\sum_j p_j = \eta$: an intensity factor on customers' belief about the vendor's responsibility or honesty (In Appendix A, we provide detail foundations for η). Denote $\hat{\pi}^* = \frac{qY^s+(1-q)Y+C}{2\sum_j p_j} + 1/2 = \frac{qY^s+(1-q)Y+C}{2\eta} + 1/2$ (assume $\hat{\pi}^* < 1$).

Proposition 1. Equilibrium: *Consider the model and stated assumptions. There is a Perfect Bayesian equilibrium (PBE) which is a cutoff such that*

$$s = \begin{cases} 1 & if\ \hat{\pi} \geq \hat{\pi}^* \\ 0 & otherwise \end{cases}$$

This PBE is supported by the following beliefs:

- $\Pr(s=1) = \Pr(s=1|p_j>0) = \Pr(s=1|p_j=0) = \hat{\pi}$
- $\Pr(s=1|\sigma=s=1, p_j>0, \hat{\pi} \geq \hat{\pi}^*) = 1$ *and* $\Pr(s=1|\sigma=s=0, p_j>0, \hat{\pi} \geq \hat{\pi}^*) = 0$
- $\Pr(s=1|\sigma=s=1, p_j>0, \hat{\pi} < \hat{\pi}^*) = \underbrace{x \in (0,1)}$ *and* $\Pr(s=1|\sigma=s=0, p_j>0, \hat{\pi} < \hat{\pi}^*) = \hat{\pi}$

Proof. See Appendix A. ∎

When $p_j = 0$, there is no customer visit to the vendor and as such there is no updating: the posterior equals the prior $\hat{\pi}$. The maximal extent of reputation gain is given by the difference: $\triangle\mathbb{E}^{\hat{\pi}}[s = 1|\sigma, s] = \mathbb{E}^{\hat{\pi}}[s = 1|\sigma = s = 1] - \mathbb{E}^{\hat{\pi}}[s = 0|\sigma = s = 0]$ which depends on the available signal about the vendor's action σ and the posterior payoff the customer computes and assigns.

Proposition 2. Information Intervention Effect: *(i) Changing subjective belief: $\hat{\pi}' > \hat{\pi}$ i.e., $\hat{\pi}' \in (\hat{\pi}, \hat{\pi} + \epsilon; \epsilon > 0)$. By shifting beliefs $\hat{\pi}' > \hat{\pi}$, it increases the number of $s = 1$. (ii) Changing the number of informed (sophisticated) customers. Denote by θ the number of informed customers. By shifting θ: $\theta' > \theta$ i.e., $\theta' \in (\theta, \theta' + \epsilon; \epsilon > 0)$, it (weakly) increases the number of customers visits to the vendor, η, making equilibrium honest behavior $s = 1$ more likely. The informed consumers thus exert a positive externality on the uninformed ones by driving up honest vendor behavior. (iii) Increasing either $\triangle\mu$ (PT information program) or q (MR information program) increases the number of $s = 1$.*

Proof. See Appendix A. ■

IV.3 Effects – Subjective Beliefs, Reputation, and Belief updates

Subjective Beliefs: From the assumed lower $\hat{\pi}$, Proposition 2 indicates that an upward shift in $\hat{\pi}$ (as well as the number of informed customers θ) should increase the $\Pr(s = 1)$. Thus, a necessary requirement for our information program to reduce misconduct (with impacts on the allied market outcomes) is to check whether $\hat{\pi}$ increased. Did our information intervention actually increased $\hat{\pi}$? First, in Figure 2, we plot the distribution of $\hat{\pi}$ at endline – reflecting subjective beliefs about customers experiences of honest transactions by treatment status. These are displayed for the various treatments together (meta) and separately. Second, we estimate

$$\hat{\pi}_{jvd} = \gamma\mathbf{I}_{jvd} + \gamma_0\hat{\pi}_{base,jvd} + \mathbf{X}'_{jvd}\xi + \zeta_{jvd}$$

controlling for consumers baseline beliefs about the likelihood of honest transactions. Table

21 reports the estimated effects of the information program on $\hat{\pi}$. There is strong evidence (both visual and formal) that the intervention shifted $\hat{\pi}$ upward in meaningful levels. We estimate a meta effect of +6.7 pp (+30% of control mean) increase in customers subjective beliefs about honest vendor behavior. The effect appears to be much larger for the joint program as expected.

Reputation: We measure reputation based on either how customers are able to infer vendor misconduct, or how vendors themselves are able to detect informed-customers who might reward honest vendor actions (or report dishonest behavior). These are two major ingredients for reputation and its concerns for market vendors based on our simplified setup. Indeed, by providing market information about misconduct, it becomes more likely to detect misconduct and thus raising the importance of reputation.

Empirical Test I: Consumers updated their beliefs about vendors misconduct? We define this as the probability of a customer guessing (or inferring) correctly misconduct and vendor irresponsibility (or responsibility) given the information treatment $\Pr(s|\mathbf{I}_{jvd}) \equiv \mathbb{E}^{\hat{\pi}}[s = 1|\sigma, s]$. We compute this as an indicator that equals 1 whenever the customer guessed the presence of misconduct in the locality (or market) and the audit exercise objectively revealed misconduct in that locality. As we defined, this corresponds to the reputation payoff that customers assign to vendors. We estimate[5]

$$\Pr(s|\mathbf{I}_{jvd}) = \gamma \mathbf{I}_{jvd} + \gamma_0 \Pr(s)_{base,jvd} + \mathbf{X}'_{jvd}\xi + \zeta_{jvd}$$

We estimate a meta effect of +8.8 pp (87% of control mean) increase in customers' ability to guess misconduct behavior, and the effects are economically larger for the joint information

[5] Alternatively, we estimate:

$$\hat{\pi}_{jvd} = \gamma_1 Misconduct_{ivd} + \gamma_2 \mathbf{I}_{jvd} + \gamma_3 Misconduct_{ivd} \times \mathbf{I}_{jvd} + \gamma_0 \hat{\pi}_{base,jvd} + \mathbf{X}'_{jvd}\xi + \zeta_{jvd}$$

This specification has perception of misconduct as the dependent variable and examines how the intervention caused consumers perceptions to more closely correlate with the audit measure of misconduct ($\gamma_3 > 0$). We estimate a meta effect of $\hat{\gamma}_3$=36 pp (p-value=0.075).

program (see, Table 22). These results (i) provide evidence of consumers' belief update - i.e., increased ability of customers to predict irresponsible vendors, and (ii) shows increased sophistication of consumers. If vendors recognize this, then they might update their beliefs about consumers sophistication or likelihood of reporting dishonest behavior by reducing misconduct. This leads to the second empirical test:

Empirical Test II: Vendors updated their beliefs about customers sophistication (i.e., θ) and likelihood of reporting dishonest behavior? We define this as the reduction in vendors misconduct as a result of the anti-misconduct information programs. This trivially coincides with our first-stage results, Table 2, where we document significant and robust reduction in misconduct due to the information sets. Results from follow-up surveys (not reported) also provide corroborative evidence that the vendors' updated their beliefs about consumer sophistication. Overall, these results are strongly consistent with our proposed reputation-based interpretation: by providing symmetric information about misconduct to both parties (uninformed customers and informed vendors), it attenuates misbeliefs about misconduct, and raises vendors concerns for market reputation. In response, vendors reduce their misconduct which has market-wide impacts.

IV.4 Corroborative Evidence

We present further heterogeneity results that lend corroborative support for the model, direct punishment and reputational effects. First, concerns for reputation is likely to be more relevant for market vendors that operate joint ventures (i.e., M-Money bundled with other services) due to relational contracting: vendors can leverage their ongoing customer relationships or goodwill with M-Money transaction services for the other non-M-Money services they provide (Gibbons and Roberts 2012). Thus, we expect the information effects to be larger for vendors that bundle M-Money with other services, relative to market vendors that operate only M-Money services. This is consistent with our earlier evidence indicating large positive spillover impacts of the information program on vendors' non-M-Money sales volume. Table 23 shows additional robust evidence that the information effects on misconduct are particularly concentrated on vendors that bundle services at baseline.

Second, under much asymmetric information about the true transactional tariffs (Figure 6), consumers might find it difficult to detect, report and thus reward good vendor behavior, which would also be especially true for customers that were vulnerable (illiterate, marginalized) at baseline. We show consistent evidence that the information intervention benefited female customers more, who also performed poorly in our baseline knowledge tests about the true transactional tariffs.

Finally, we have shown that more accurate consumers' beliefs about misconduct and reputation drive the estimated impacts of our anti-misconduct information programs. However, since the market (particularly consumers) became more informed about the true tariffs, this might also turn on two interesting alternative mechanisms: price effect or bargaining effect for the real consumer and market outcomes. For price effects, this can considered as a by-product of reputation: vendors took honest actions because of concerns that they might be perceived by consumers as irresponsible, which lead to lower prices and as a result, a price response for consumer demand and other market outcomes. Such price effects are consistent with and re-affirms reputation. For bargaining effects, mobile money is not a market where participants negotiate over transactions. The price is fixed for a given market transaction and consumers take this as given. Misconduct arises when a vendor decides to overcharge or not the market transaction. We therefore believe that bargaining is not driving our findings.

V Cost-Effectiveness of the Information Program

How cost-effective is our information intervention? Does this compare well with other financial education interventions? When computing cost-effectiveness, we focus on usage of services-only measure for customers and sales volume-only measure for vendors. This is a very conservative approach in the sense that it does not consider the significant treatment effects on savings, risk mitigation, poverty outcomes, and other positive externalities of the information program. For example, we find additional improvements on non M-Money sales volume for vendors (not reported).

We estimate costs based on the # of trained field officers utilized (3 officers to minimize cross-officer heterogeneity), the # of times or rounds the experimental subjects were visited (3$\times$) to deliver the interventions over the period (October 15-December 15, 2020), transportation costs (GHS385 per officer $\times$ 3 officers $\times$ 3 rounds = GHS3,465), remuneration and allowance for officers (GHS1,200 per officer $\times$ 3 officers $\times$ 3 rounds = GHS10,800), and occasional accommodation for officers during field visits (GHS100 per officer $\times$ 3 officers $\times$ 3 rounds = GHS900). The total cost equals to GHS15,165. About 632 panel of treated customers were reached.[6] Similarly, about 97 panel of treated vendors were reached ($\frac{98+96+98}{3} = 97$), bringing the total # of subjects to 730. We then estimate $\frac{\text{GHS15,165}}{730} =$ GHS20.8 per subject, or US$4.0 per person at an exchange rate of US$1=GHS5.12. The opportunity cost of time-use for the subjects is very limited: it took roughly 7 minutes per visit to deliver the information intervention. When compared with the minimum wage in Ghana (GHS10.65 per day), the time-use and thus its cost on subjects is very negligible. Thus, the information sets cost approximately US$4.0 per subject.

Overall, our cost-effectiveness ratio is 1:5 – a per subject cost of US$4.0 for about +US$19.3 increase in the usage of financial services for customers (see, Table 19), with sizable implications for consumer welfare (risk mitigation, poverty; see, Table 5). For vendors, the treatment effect (+GHS591; see, Table 8) implies a ratio of 1:28 improvement in vendor outcomes. These rough cost-effectiveness estimates compare favorably with other financial information programs. For example, Frisancho (2018) reports a cost per pupil of US$4.80 and a US$1 increase in financial education program's expense for a 3.3 point improvement in financial literacy. In a recent meta-analyses about financial education interventions, Kaiser et al. (2020) reports a cost-effectiveness ratio of $60.40 per person for one-fifth of a standard deviation improvement in outcomes. Our findings suggest that providing market-level infor-

[6] $\frac{1}{3}\sum_{r=1}^{3}$ # of subjects reachable per round$_r$ $= \frac{629+617+642}{3} = 632$.

mation that reduce misconduct in local market transactions could be a cost-effective way to improve market efficiency or activity.

VI Conclusion

Misconduct in markets matters: in terms of efficiency. By providing information sets that deter and reduce market vendors misconduct – customers meaningfully increase their take-up of transactional services and savings behavior at vendor points, enabling customers to better mitigate unexpected shocks. Market vendors experience meaningful increases in their sales volume, suggesting improved market efficiency.

Reputation does matter for misconduct. In rural financial environments, where markets are subject to a high degree of information imperfections, the importance of reputation as a discipline device against market misconduct is limited. Misconduct may be high because consumers cannot observe whether they are cheated since they don't know the official price, making it difficult to establish meaningful reputation. Reputation however becomes effective and disciplinary if there is a high probability of inferring misconduct (Shapiro 1982, Burkhardt 2018) and vendors can easily demonstrate the quality of their market services. Such reputation-driven misconduct is illuminated drawing on features of our empirical setting and the provision of relevant market information that improves subjects' ability to report misconduct and make inferences about misconduct.

Our field experiment is carefully designed to: (i) reduce market misconduct using information programs about misconduct, (ii) quantify the impacts on important economic outcomes on both sides of the market: uptake of transactional services, savings, risk mitigation, poverty and sales, and (iii) show that these effects are driven by a combination of more accurate consumers' beliefs about misconduct and increased reputation concerns. We do this by constructing a unique census of markets for mobile money in rural Ghana, showing baseline evidence of significance misconduct, information imperfections and misbeliefs about vendor misconduct, and assigning these local markets to information programs about misconduct. Overall, our results emphasize the significance of local sanctions to support the growth of rural financial institutions (Karpoff 2012; Munshi 2014) and provide a proof-of-concept of a potentially significant source of local financial market friction, where market activities (i.e., adoption, transactions, savings, sales) are underprovided (Jensen and Miller 2018; Bai 2019) due to misconduct, with implications for market efficiency in transactional markets. Reputation (or trust) in markets might be difficult to build and thus low, likely due to imperfect information.

References

[1] Abadie, Alberto. 2020. "Statistical Nonsignificance in Empirical Economics." American Economic Review: Insights, 2 (2): 193-208.

[2] Annan, Francis. 2017. "Fraud on Mobile Financial Markets: Evidence from a Pilot Audit Study." NET Institute Working Paper No. 17-16.

[3] Annan, Francis. 2020. "Gender and Financial Misconduct: A Field Experiment on Mobile Money." Georgia State University - Working Paper.

[4] Bai, Jie. 2019. "Melons as Lemons: Asymmetric Information, Consumer Learning and Quality Provision." Revise and Resubmit, Review of Economic Studies.

[5] Banerjee, Abhijit and Esther Duflo. 2007. "The Economic Lives of the Poor." Journal of Economic Perspectives, 21(1): 141-68.

[6] Banerjee, Abhijit and Esther Duflo. 2011. *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*. New York: Public Affairs. 15th Edition.

[7] Banerjee, A, Hanna R, and Mullainathan S. 2012. "Corruption". Ch. 27, pages 1109-1147 of: Handbook of Organizational Economics.

[8] Belloni Alexandre, Victor Chernozhukov, and Christian Hansen. 2014. "Inference on Treatment Effects after Selection among High-Dimensional Controls." Review of Economic Studies 81(2): 608-650.

[9] Breza, E., and A. Chandrasekhar. 2019. "Social Networks, Reputation, and Commitment: Evidence from a Savings Monitors Experiment." Econometrica, 87(1): 175–216.

[10] Board, Simon and Moritz Meyer-ter-Vehn. 2013. "Reputation for Quality." Econometrica, 81 (6), 2381-2462.

[11] Burkhardt, Kirsten. 2018. *Private Equity Firms: Their Role in the Formation of Strategic Alliances*. John Wiley & Sons, Inc.

[12] Coppejans, M., Gilleskie, D., Sieg, H. and K. 2007. Strumpf. "Consumer Demand under Price Uncertainty: Empirical Evidence from the Market for Cigarettes." Review of Economics and Statistics, 89 (3): 510-521.

[13] de Quidt, Jonathan, Johannes Haushofer, and Christopher Roth. 2018. "Measuring and Bounding Experimenter Demand." American Economic Review, 108 (11): 3266-3302.

[14] Dupas, P., and J. Robinson. 2013. "Why Don't the Poor Save More? Evidence From Health Savings Experiments." American Economic Review, 103 (4), 1138–1171.

[15] Egan, Mark, Gregor Matvos and Amit Seru. 2019a. "The Market for Financial Adviser Misconduct." Journal of Political Economy, 127 (1): 233-295.

[16] Egan, Mark, Gregor Matvos and Amit Seru. 2019b. "When Harry fired Sally: The Double Standard in Punishing Misconduct." Revise and Resubmit, Journal of Political Economy.

[17] Frisancho, Veronica. 2018. "The Impact of School-Based Financial Education on High School Students and Their Teachers: Experimental Evidence from Peru". IDB Working Paper No. IDB-WP-871.

[18] Gibbons, R and J. Roberts. 2012. "The Handbook of Organizational Economics." Princeton University Press.

[19] Higgins, Sean. 2020. "Financial Technology Adoption." Northwestern University - Working paper.

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

[21] Jensen, Robert, and Nolan H. Miller. 2018. "Market Integration, Demand, and the Growth of Firms: Evidence from a Natural Experiment in India." American Economic Review, 108 (12): 3583-3625.

[22] Lee, David. 2009. "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects." Review of Economics Studies, 76(3): 1071-1102.

[23] Kaiser, Tim and Annamaria Lusardi, Lukas Menkhoff, and Carly Urban. 2020. "Financial Education Affects Financial Knowledge and Downstream Behaviors". Wharton Pension Research Council Working Paper No. 2020-07.

[24] Karpoff, M. Jonathan. 2012. "Does Reputation Work to Discipline Corporate Misconduct?" Ch. 18, pages 361-382 of: Oxford Handbook of Corporate Reputation. https://faculty.washington.edu/karpoff/Karpoff%20reputation%20survey%202012%20copy.pdf

[25] Kessler, Judd B., Corinne Low, and Colin D. Sullivan. 2019. "Incentivized Resume Rating: Eliciting Employer Preferences without Deception." American Economic Review, 109 (11): 3713-44.

[26] List, John A., Azeem M. Shaikh and Yang Xu. 2016. "Multiple Hypothesis Testing in Experimental Economics". Experimental Economics, 22(4): 773-793.

[27] Matsa, A. David. 2011. “Competition and Product Quality in the Supermarket Industry”. Quarterly Journal of Economics, 126 (3): 1539-1591.

[28] Munshi, Kevin 2014. “Community Networks and the Process of Development.” Journal of Economic Perspectives, 28 (4), 49-76.

[29] Olken, Benjamin and Rohini Pande. 2012. “Corruption in developing countries.” Annual Review of Economics, 4(): 479-509.

[30] Ortmann, Andreas and Ralph Hertwig. 2002. “The Costs of Deception: Evidence from Psychology.” Experimental Economics, 5(): 111–131.

[31] Schreiner, Mark. 2015. “Simple poverty scorecard--Poverty-assessment tool for Ghana.” Available here: http://www.simplepovertyscorecard.com/GHA_2012_ENG.pdf

[32] Shapiro, C. 1982. “Consumer information, Product Quality, and Seller reputation.” The Bell Journal of Economics, 13(1), 20-35.

[33] Shapiro, C. 1983. “Premiums for High Quality Products as Returns to Reputations.” Quarterly Journal of Economics, XCVIII, 658–679.

[34] Suri, Tavneet and William Jack. 2016. “The long-run Poverty and Gender Impacts of Mobile Money”. Science, 354 (6317): 1288-1292.

[35] Zitzewitz, Eric. 2012. “Forensic Economics.” Journal of Economic Literature, 50 (3): 731-69.

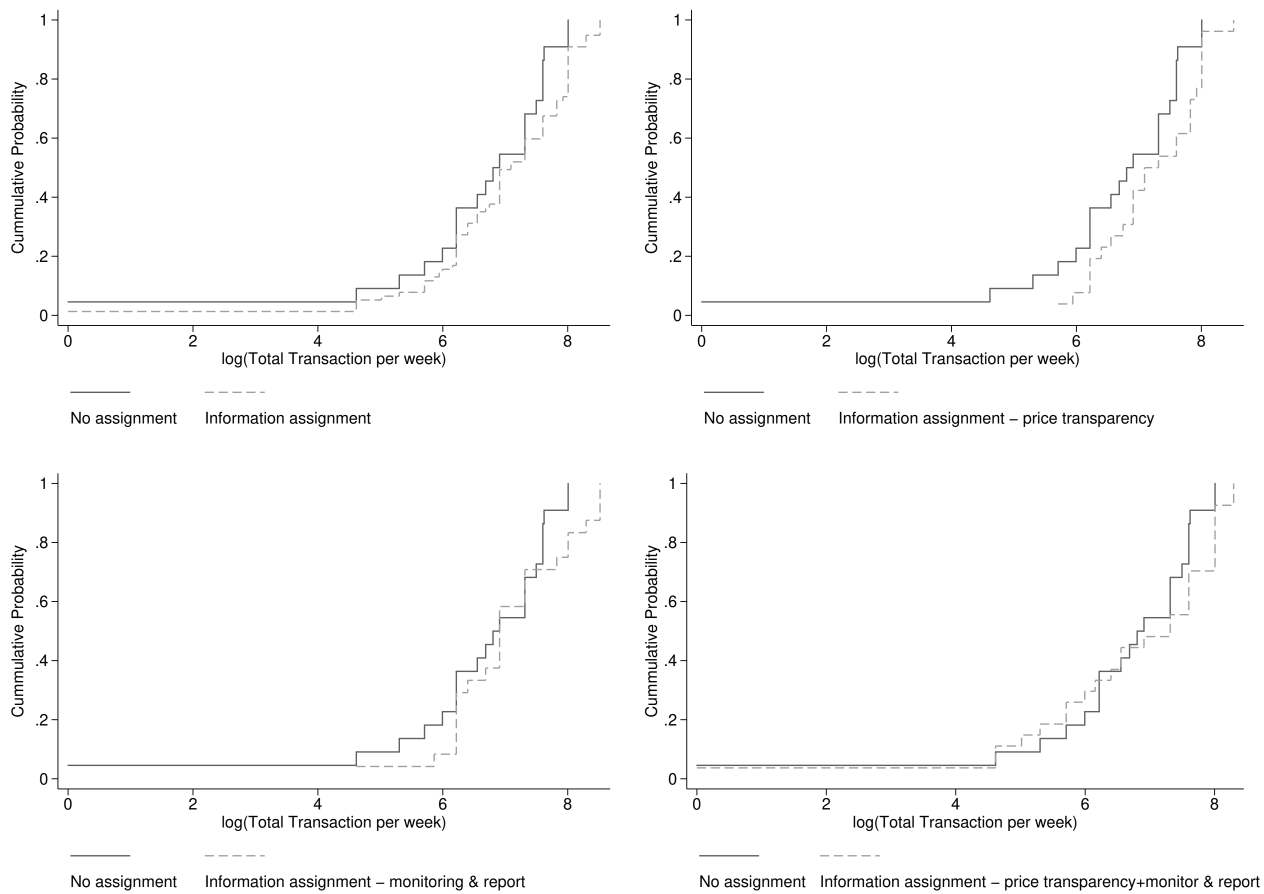
Main Results for Text

Table 2: **Effect of information sets on vendors misconduct**

	1(Misconduct=Yes)		Amount-Misconduct, GHS	
PANEL A				
Treatment: Information Assignment (β)	-0.231	-0.211	-0.675	-0.551
	(0.055)	(0.086)	(0.185)	(0.255)
	[-0.324, -0.138]	[-0.354, -0.067]	[-0.984, -0.367]	[-0.975, -0.125]
Baseline misconduct	X	X	X	X
Market District F.E.	X		X	
Market District × Transaction × Date F.E.		X		X
Observations	335	335	335	335
Mean of dependent variable (control)	0.294	0.294	0.778	0.778
Lee (2009) Attrition Bounds	<-0.174, -0.164>		<-0.484, -0.1435>	
Imbens and Manski (2004) CS	[-0.225, -0.094]		[-0.642, -0.085]	
PANEL B				
Price Transparency (β_1)	-0.177	-0.184	-0.550	-0.439
	(0.065)	(0.094)	(0.199)	(0.276)
	[-0.285, -0.069]	[-0.342, -0.027]	[-1.881, -0.219]	[-0.898, 0.020]
Monitor and Report (β_2)	-0.257	-0.217	-0.687	-0.574
	(0.063)	(0.093)	(0.222)	(0.275)
	[-0.363 0.151]	[-0.373, -0.061]	[-1.057, -0.317]	[-1.032, -0.117]
Joint program: PT + MR (δ)	-0.233	-0.212	-0.718	-0.555
	(0.064)	(0.089)	(0.198)	(0.279)
	[-0.340, -0.127]	[-0.360, -0.062]	[-1.048, -0.388]	[-1.019, -0.089]
Baseline misconduct	X	X	X	X
Market District F.E.	X		X	
Market District × Transaction × Date F.E.		X		X
Observations	335	335	335	335
Mean of dependent variable (control)	0.294	0.294	0.778	0.778
p-value (test: $\beta_1 = \delta$)	0.327	0.670	0.280	0.553
p-value (test: $\beta_2 = \delta$)	0.660	0.921	0.860	0.923
p-value (test: $\beta_1 = \beta_2$)	0.104	0.563	0.347	0.411
p-value (test: $\beta_1 + \beta_2 = \delta$)	0.027	0.108	0.074	0.204

Note: **1**(.) is a logical indicator that equals 1 if the argument in the parenthesis is true, 0 otherwise. Includes randomization strata (market district) dummies, baseline outcomes, and additional controls. Observations are at the vendor x transaction x date level. Clustered standard errors (at local market level) reported in parenthesis. 90% confidence intervals (CI) and confidence sets (CS) are reported in brackets. Panel A reports meta estimate of treatment effects, while panel B shows effects separately for each information program. Results similar to post-double-selection LASSO estimates clustered at the (local) market level and to alternative inference procedures (Wild cluster bootstrap and permutation test clustered at the market level).

Figure 1: **Distribution of log (total transactions) at endline by treatment status**



Note: Figure plots the distributions (CDFs) of log(Total Transactions per week+1) at endline for different experimental subsamples. From a Kolmogorov–Smirnov test for the equality of distributions, *p*-values equal 0.058, 0.288, 0.043 and 0.074, respectively. Observations are at the customer level.

Table 3: **Effect of treatment on usage**

Model: Linear
DV: Log Total Transaction per week

	(1)	(2)	(3)	(4)
Treatment: Information Assignment (β)	0.561	0.470	0.416	0.402
	(0.225)	(0.217)	(0.220)	(0.213)
	[0.189, 0.932]	[0.111, 0.828]	[0.052, 0.779]	[0.050, 0.755]
Inference Robustness (β)				
CI: Clustered S.E.	[0.059,1.062]	[0.096, 0.843]	[0.096, 0.843]	[0.082, 0.723]
CI: Wild Bootstrap	[0.191, 0.922]	[0.113, 0.821]	[0.113, 0.821]	[0.024, 0.789]
p-value: Permutation Test	0.015	0.032	0.041	0.048
p-value: L-S-X MHT Corr (2019)	0.012			
Market District F.E.		X	X	X
Baseline usage			X	X
Controls				X
Observations	763	763	723	723
R-squared	0.009	0.064	0.076	0.108
Mean of dependent variable (control)	3.583	3.583	3.583	3.583
Lee (2009) Attrition Bounds				
Lower Bound:	0.432			
	(0.271)			
	[-0.013, 0.878]			
Upper Bound:	0.805			
	(0.299)			
	[0.313, 1.297]			
Imbens and Manski (2004) CS	[0.076, 1.197]			
Behaghel et al. (2015) Attrition Bounds				
Lower Bound:	0.430			
	(0.226)			
	[0.059, 0.806]			
Upper Bound:	0.738			
	(0.225)			
	[0.366, 1.110]			

Note: Market district is the randomization strata. Observations are at the customer level. Robust standard errors reported in parenthesis. 90% confidence intervals (CI) and confidence sets (CS) are reported in brackets. Reported confidence CIs for wild bootstrap and permutation tests derived from running 1000 replications in each case. Permutation test (for randomization inference) clustered at the (local) market level. L-S-X MHT Corr (2019) refers to the multiple hypothesis testing procedure presented in List, Shaikh and Xu (2019) for transactions outcomes family (services usage; savings). 0-1 indicators for baseline migration motives (desire to migrate, plan to migrate, and permanent migration) used as predictors of attrition to tighten attrition bounds. Results similar to post-double-selection LASSO estimates clustered at the market level.

Table 4: **Effect of treatment on savings**

Model: Linear
DV: 0-1 Indicator for whether consumer is "saving" on M-Money

	(1)	(2)	(3)	(4)
Treatment: Information Assignment (β)	0.089	0.078	0.080	0.076
	(0.045)	(0.045)	(0.046)	(0.045)
	[0.013, 0.164]	[-0.003, 0.147]	[0.004, 0.157]	[0.001, 0.151]
Inference Robustness (β)				
CI: Clustered S.E.	[-0.001, 0.178]	[-0.002, 0.146]	[0.007, 0.153]	[0.004, 0.148]
CI: Wild Bootstrap	[0.013, 0.165]	[-0.005, 0.149]	[0.004, 0.156]	[0.004, 0.149]
p-value: Permutation Test	0.059	0.108	0.080	0.099
p-value: L-S-X MHT Corr (2019)	0.048			
Market District F.E.		X	X	X
Baseline savings			X	X
Controls				X
Observations	763	763	689	689
R-squared	0.005	0.027	0.075	0.105
Mean of dependent variable (control)	0.605	0.605	0.605	0.605
Lee Attrition Bounds				
Lower Bound:	0.070			
	(0.050)			
	[-0.014, 0.152]			
Upper Bound:	0.125			
	(0.056)			
	[0.031, 0.218]			
Imbens and Manski (2004) CS	[0.001, 0.201]			
Behaghel et al. (2015) Attrition Bounds				
Lower Bound:	0.078			
	(0.046)			
	[0.003, 0.154]			
Upper Bound:	0.120			
	(0.045)			
	[0.045, 0.196]			

Note: Market district is the randomization strata. Observations are at the customer level. Robust standard errors reported in parenthesis. 90% confidence intervals (CI) and confidence sets (CS) are reported in brackets. Reported confidence CIs for wild bootstrap and permutation tests derived from running 1000 replications in each case. Permutation test (for randomization inference) clustered at the (local) market level. L-S-X MHT Corr (2019) refers to the multiple hypothesis testing procedure presented in List, Shaikh and Xu (2019) for transactions outcomes family (services usage; savings). 0-1 indicators for baseline migration motives (desire to migrate, plan to migrate, and permanent migration) used as predictors of attrition to tighten attrition bounds. Results similar to post-double-selection LASSO estimates clustered at the market level.

Table 5: **Effect of treatment on shock mitigation and poverty**

Model: Linear

	u-Shocks Experience (1)	u-Shocks HH Revenue (2)	u-Shocks Health (3)	u-Shocks HH Expenditure (4)	Poverty Likelihood (5)
Treatment: Information	-0.068	-0.072	-0.056	-0.107	1.033
Assignment (β)	(0.030)	(0.040)	(0.044)	(0.044)	(1.254)
	[-0.117, -0.019]	[-0.140, -0.005]	[-.0130, 0.016]	[-0.180, -0.034]	[-1.033, 3.099]
Inference Robustness (β)					
CI: Clustered S.E.	[-0.128, -0.008]	[-0.159, 0.013]	[-0.163, 0.05]	[-0.206, -0.008]	[-1.306, 3.373]
CI: Wild Bootstrap	[-0.117, -0.020]	[-0.141, -0.007]	[-.1319, .018]	[-0.182, -0.033]	[-0.984, 3.107]
p-value: Permutation Test	0.068	0.176	0.332	0.091	0.451
p-value: L-S-X MHT Corr (2019)	0.027	0.057	0.601	0.161	0.140
Observations	763	763	763	763	763
R-squared	0.095	0.059	0.179	0.152	0.121
Mean of dependent variable (control)	0.890	0.773	0.525	0.416	10.18
Lee (2009) Attrition Bounds	[-0.089, -0.043]	[-0.103, -0.050	[-0.055, 0.003]	[-0.112, -0.053]	[-0.361, 3.286]
Imbens and Manski (2004) CS	[-0.134, 0.024]	[-0.164, 0.020]	[-0.128, 0.078]	[-0.190, 0.015]	[-2.761, 5.248]
Behaghel et al. (2015) Attrition Bounds	[-0.089, -0.045]	[-0.101, -0.058]	[-.058, -0.018]	[-0.099, -0.059]	[-0.178, 2.371]

Note: u denotes unmitigated. Includes randomization strata (market district) dummies, baseline outcomes, and additional controls. Observations are at the customer level. Robust standard errors reported in parenthesis. 90% confidence intervals (CI) and confidence sets (CS) are reported in brackets. Reported confidence CIs for wild bootstrap and permutation tests derived from running 1000 replications in each case. Permutation test (for randomization inference) clustered at the (local) market level. L-S-X MHT Corr (2019) refers to the multiple hypothesis testing procedure presented in List, Shaikh and Xu (2019) for welfare outcomes family (shocks mitigation; poverty). 0-1 indicators for baseline migration motives (desire to migrate, plan to migrate, and permanent migration) used as predictors of attrition to tighten attrition bounds. Results similar to post-double-selection LASSO estimates clustered at the market level.

Table 6: **Effect of information sets on usage and savings**

	Log Total Transaction per week	Total Transaction per week	Using M-Money	Saving on M-Money	PCA Index (1, 3, 4)
	(1) Linear-OLS	(2) Tobit	(3) Linear-OLS	(4) Linear-OLS	(5) Linear-OLS
Price Transparency (β_1)	0.280	39.684	0.059	0.064	0.088
	(0.247)	(54.369)	(0.044)	(0.053)	(0.110)
Robust S.E.	[-0.127, 0.687]	[-49.863, 129.231]	[-0.014, 0.133]	[-0.022, 0.152]	[-0.093, 0.270]
Clustered S.E.	[-0.103, 0.664]	[-38.782, 118.151]	[-0.011, 0.130]	[-0.071, 0.247]	[-0.069, 0.247]
Wild Bootstrap	[-0.124, 0.688]		[-0.014, 0.135]	[-0.021, 0.150]	[-0.097, 0.273]
p-value: Permutation Test	0.281	0.583	0.171	0.260	0.413
p-value: L-S-X MHT Corr (2019)	0.188		0.163	0.336	
Lee (2009) Attrition Bounds	<0.151, 0.767>		<0.051, 0.142>	<0.024, 0.122>	<0.060, 0.207>
Monitor and Report (β_2)	0.431	173.007	0.0705	0.036	0.188
	(0.253)	(83.049)	(0.044)	(0.054)	(0.110)
Robust S.E.	[0.014, 0.849]	[36.222, 309.792]	[-0.002, 0.143]	[-0.054, 0.126]	[0.007, 0.369]
Clustered S.E.	[0.031, 0.831]	[33.908, 312.106]	[-0.001, 0.142]	[-0.056, 0.128]	[0.026, 0.350]
Wild Bootstrap	[0.021, 0.842]		[-0.003, 0.143]	[-0.054, 0.125]	[0.001, 0.372]
p-value: Permutation Test	0.091	0.013	0.119	0.549	0.080
p-value: L-S-X MHT Corr (2019)	0.003		0.007	0.257	
Lee (2009) Attrition Bounds	<0.605, 0.790>		<0.106, 0.134>	<0.035, 0.072>	<0.262, 0.334>
Joint program: PT + MR (δ)	0.506	83.276	0.080	0.123	0.220
	(0.248)	(53.138)	(0.044)	(0.052)	(0.108)
Robust S.E.	[0.097, 0.915]	[-4.243, 170.797]	[0.008, 0.153]	[0.037, 0.208]	[0.042, 0.398]
Clustered S.E.	[0.129, 0.883]	[5.898, 160.655]	[0.012, 0.148]	[0.038, 0.207]	[0.067, 0.372]
Wild Bootstrap	[0.108, 0.907]		[0.007, 0.152]	[0.035, 0.211]	[0.036, 0.406]
p-value: Permutation Test	0.035	0.244	0.073	0.024	0.034
p-value: L-S-X MHT Corr (2019)	0.009		0.021	0.002	
Lee (2009) Attrition Bounds	<0.451, 0.877>		<0.096, 0.152>	<0.134, 0.191>	<0.198, 0.626>
Observations	723	723	723	689	689
Mean of dependent variable (control)	3.583	198.956	0.722	0.605	-0.201
p-value (test: $\beta_1 = \delta$)	0.298	0.336	0.583	0.203	0.175
p-value (test: $\beta_2 = \delta$)	0.739	0.204	0.786	0.066	0.745
p-value (test: $\beta_1 = \beta_2$)	0.502	0.077	0.780	0.562	0.315
p-value (test: $\beta_1 + \beta_2 = \delta$)	0.536	0.158	0.397	0.753	0.696

Note: Includes randomization strata (market district) dummies, baseline outcomes, and additional controls. Observations are at the customer level. Robust standard errors reported in parenthesis. 90% confidence intervals (CI) are reported in brackets. Reported confidence CIs for wild bootstrap and permutation tests derived from running 1000 replications in each case. Permutation test (for randomization inference) clustered at the (local) market level. L-S-X MHT Corr (2019) refers to the multiple hypothesis testing procedure presented in List, Shaikh and Xu (2019) for welfare outcomes family (shocks mitigation; poverty). 0-1 indicators for baseline migration motives (desire to migrate, plan to migrate, and permanent migration) used as predictors of attrition to tighten attrition bounds. Results similar to post-double-selection LASSO estimates clustered at the market level.

Table 7: **Effect of information sets on shock mitigation and poverty**

Model: Linear

	u-Shocks Experience (1)	u-Shocks HH Revenue (2)	u-Shocks Health (3)	u-Shocks HH Expenditure (4)	Poverty Likelihood (5)
Price Transparency (β_1)	-0.090	-0.110	-0.073	-0.128	1.680
	(0.036)	(0.047)	(0.052)	(0.051)	(1.509)
Robust S.E.	[-0.150, -0.029]	[-0.188, -0.031]	[-0.159, 0.012]	[-0.212, -0.044]	[-0.806, 4.167]
Clustered S.E.	[-0.159, -0.021]	[-0.214, -0.006]	[-0.194, 0.047]	[-0.244, -0.012]	[-1.077, 4.438]
Wild Bootstrap	[-0.151, -0.028]	[-0.188, -0.033]	[-0.161, 0.014]	[-0.214, -0.042]	[-0.712, 4.102]
p-value: Permutation Test	0.053	0.103	0.327	0.107	0.335
p-value: L-S-X MHT Corr (2019)	0.024	0.038	0.328	0.048	0.046
Lee (2009) Attrition Bounds	<-0.103, -0.004>	<-0.130, -0.031>	<-0.104, -0.005>	<-0.173, -0.074>	<-0.613, 4.974>
Monitor and Report (β_2)	-0.019	-0.001	-0.001	-0.041	1.439
	(0.036)	(0.049)	(0.052)	(0.049)	(1.552)
Robust S.E.	[-0.080, 0.041]	[-0.082, 0.079]	[-0.087, 0.084]	[-0.128, 0.045]	[-1.117, 3.997]
Clustered S.E.	[-0.088, 0.050]	[-0.105, 0.102]	[-0.126, 0.124]	[-0.168, 0.085]	[-1.231, 4.111]
Wild Bootstrap	[-0.081, 0.042]	[-0.080, 0.081]	[-0.086, 0.083]	[-0.132, 0.050]	[-1.202, 4.055]
p-value: Permutation Test	0.684	0.986	0.985	0.597	0.416
p-value: L-S-X MHT Corr (2019)	0.410	0.621	0.302	0.637	0.107
Lee (2009) Attrition Bounds	<-0.036, 0.0003>	<-0.032, 0.003>	<0.042, 0.079>	<0.006, 0.042>	<0.862, 3.716>
Joint program: PT + MR (δ)	-0.089	-0.096	-0.089	-0.143	0.022
	(0.036)	(0.048)	(0.051)	(0.049)	(1.456)
Robust S.E.	[-0.150, -0.029]	[-0.176, -0.016]	[-0.174, -0.005]	[-0.226, -0.061]	[-2.377, 2.421]
Clustered S.E.	[-0.167, -0.011]	[-0.195, 0.003]	[-0.207, 0.028]	[-0.250, -0.036]	[-2.895, 2.939]
Wild Bootstrap	[-0.150, -0.029]	[-0.176, -0.014]	[-0.176, -0.002]	[-0.229, -0.061]	[-2.492, 2.529]
p-value: Permutation Test	0.057	0.142	0.215	0.067	0.989
p-value: L-S-X MHT Corr (2019)	0.018	0.030	0.204	0.034	0.904
Lee (2009) Attrition Bounds	<-0.103, -0.029>	<-0.128, -0.054>	<-0.107, -0.034>	<-0.160, -0.086>	<-2.809, 2.336>
Observations	763	763	763	763	763
Mean of dependent variable (control)	0.890	0.773	0.525	0.416	10.18
p-value (test: $\beta_1 = \delta$)	0.983	0.751	0.714	0.718	0.235
p-value (test: $\beta_2 = \delta$)	0.052	0.034	0.050	0.021	0.326
p-value (test: $\beta_1 = \beta_2$)	0.057	0.015	0.123	0.059	0.870
p-value (test: $\beta_1 + \beta_2 = \delta$)	0.698	0.813	0.825	0.701	0.140

Note: u denotes unmitigated. Includes randomization strata (market district) dummies, baseline outcomes, and additional controls. Observations are at the customer level. Robust standard errors reported in parenthesis. 90% confidence intervals (CI) are reported in brackets. Reported confidence CIs for wild bootstrap and permutation tests derived from running 1000 replications in each case. Permutation test (for randomization inference) clustered at the (local) market level. 0-1 indicators for baseline migration motives (desire to migrate, plan to migrate, and permanent migration) used as predictors of attrition to tighten attrition bounds. L-S-X MHT Corr (2019) refers to the multiple hypothesis testing procedure presented in List, Shaikh and Xu (2019) for welfare outcomes family (shocks mitigation; poverty). Results similar to post-double-selection LASSO regression estimates clustered at the market level.

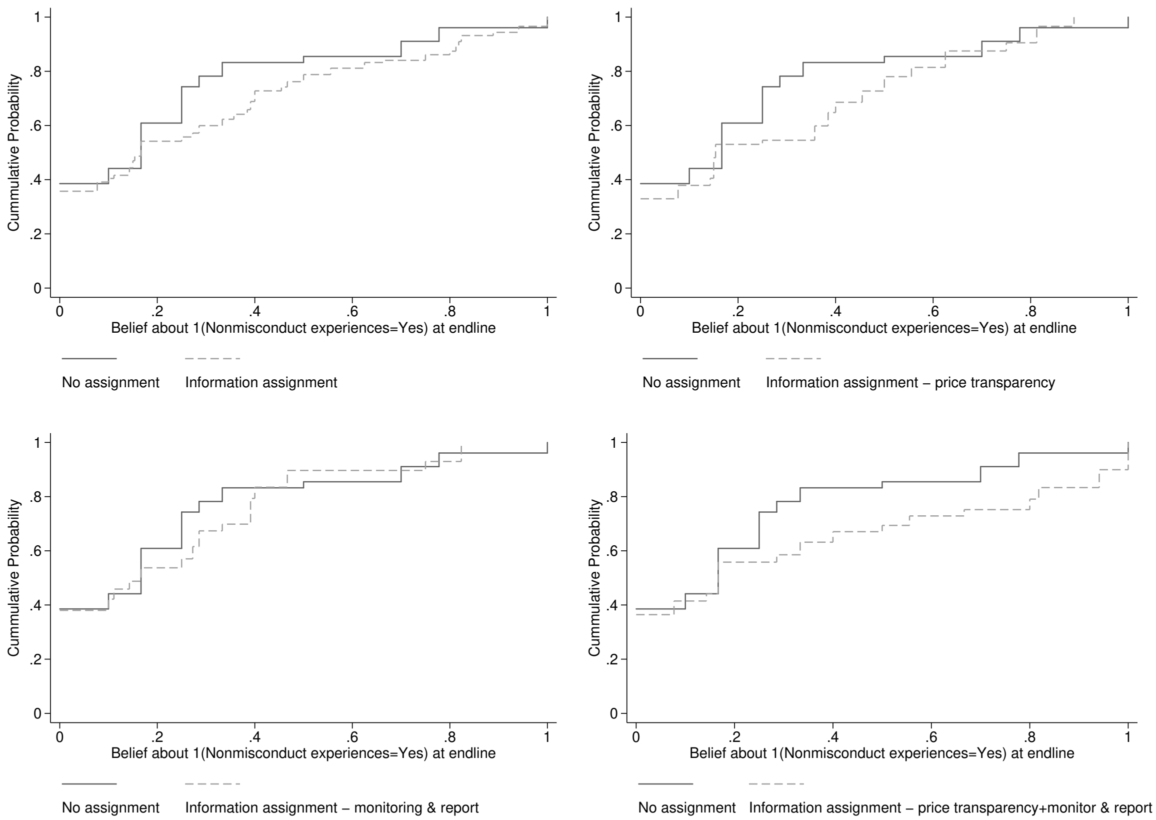
Table 8: **Effect of treatment on vendors market sales**

DV: Total sales per day (GHS)

	(1) Linear-OLS	(2) Tobit
PANEL A		
Treatment: Information	557.342	591.568
Assignment (β)	(278.916)	(274.918)
Robust S.E.	[93.386, 1021.298]	[44.862, 1138.273]
Clustered S.E.	[94.1867 102.498]	[44.862, 1138.273]
Wild Bootstrap	[92.150, 1026.000]	
Lee (2009) Attrition Bounds	<211.479, 622.647>	
Behaghel et al. (2015) Attrition Bounds	[432.877, 574.305]	
Observations	99	99
Mean of dependent variable (control)	1192.727	1192.727
PANEL B		
Price Transparency (β_1)	663.086	699.722
	(311.063)	(299.543)
Robust S.E.	[145.512, 1180.659]	[201.387, 1198.057]
Clustered S.E.	[146.549, 1179.623]	[201.387, 1198.057]
Wild Bootstrap	[169.900, 1169.000]	
Lee (2009) Attrition Bounds	<339.936, 702.517>	
Behaghel et al. (2015) Attrition Bounds	[443.286, 776.013]	
Monitor and Report (β_2)	569.105	648.863
	(438.705)	(439.512)
Robust S.E.	[-160.850, 1299.061]	[-82.330, 1380.058]
Clustered S.E.	[-159.388, 1297.599]	[-82.330, 1380.058]
Wild Bootstrap	[-177.400, 1290.000]	
Lee (2009) Attrition Bounds	<104.594, 714.902>	
Behaghel et al. (2015) Attrition Bounds	[157.272, 612.034]	
Joint program: PT + MR (δ)	421.780	422.605
	(310.394)	(296.952)
Robust S.E.	[-94.680, 938.241]	[-71.419, 916.630]
Clustered S.E.	[-93.646, 937.207]	[-71.419, 916.630]
Wild Bootstrap	[-195.700, 1361.000]	
Lee (2009) Attrition Bounds	<249.224, 435.588>	
Behaghel et al. (2015) Attrition Bounds	[261.620, 460.001]	
Observations	99	99
Mean of dependent variable (control)	1192.727	1192.727
p-value (test: $\beta_1 = \delta$)	0.408	0.321
p-value (test: $\beta_2 = \delta$)	0.707	0.562
p-value (test: $\beta_1 = \beta_2$)	0.808	0.894
p-value (test: $\beta_1 + \beta_2 = \delta$)	0.130	0.082

Note: Includes randomization strata (market district) dummies, baseline outcomes, and additional controls. Observations are at the vendor level. Robust standard errors reported in parenthesis. 90% confidence intervals (CI) are reported in brackets. Reported confidence CIs for wild bootstrap derived from running 1000 replications. Differential attrition bounds are reported. 0-1 indicators for baseline migration motives (desire to migrate, plan to migrate, and permanent migration) used as predictors of attrition to tighten attrition bounds. Results similar to post-double-selection LASSO estimates clustered at the (local) market level.

Figure 2: **Distributions of customers belief about vendor responsibility (non-misconduct) at endline by treatment status**



Note: Figure plots the distributions (CDFs) of belief about (Non-misconduct experiences=Yes) or honest vendor experiences at endline for different experimental subsamples. **1**(.) is a logical indicator that equals 1 if the argument in the parenthesis is true, 0 otherwise. Observations are at the customer level. In each market, we compute the share of experimental customers that indicate they believe not experiencing misconduct (i.e., non-misconduct) at endline. From a Kolmogorov–Smirnov test for the equality of distributions, p-value = 0.000 for all cases.

Supplementary Appendix
(Not For Publication)

A Proofs

A.1 Proof of Proposition 1

Proof. $s = 1$ IFF

$$\mathbf{\Pi}(s=1) > \mathbf{\Pi}(s=0)$$

$$-\frac{C}{(\sum_j p_j)} + \hat{\pi}\mathbb{E}^{\hat{\pi}}[s=1|\sigma, s=1] + (1-\hat{\pi})(1-\mathbb{E}^{\hat{\pi}}[s=1|\sigma, s=1])$$
$$>$$
$$\frac{qY^s + (1-q)Y}{(\sum_j p_j)} + \hat{\pi}\mathbb{E}^{\hat{\pi}}[s=1|\sigma, s=0] + (1-\hat{\pi})(1-\mathbb{E}^{\hat{\pi}}[s=1|\sigma, s=0])$$

$\mathbb{E}^{\hat{\pi}}[s=1|\sigma,s] = \Pr[s=1|\sigma,s]$ and $\sum_j p_j = \eta$, so we write:

$$-\frac{C}{\eta} + \hat{\pi}\underbrace{\Pr[s=1|\sigma, s=1]}_{\mu(1,1)} + (1-\hat{\pi})(1-\Pr[s=1|\sigma, s=1])$$
$$>$$
$$\frac{qY^s + (1-q)Y}{\eta} + \hat{\pi}\underbrace{\Pr[s=1|\sigma, s=0]}_{\mu(1,0)} + (1-\hat{\pi})(1-\Pr[s=1|\sigma, s=0])$$

We get

$$2\hat{\pi}\mu(1,1) - 2\hat{\pi}\mu(1,0) - \mu(1,1) + \mu(1,0) > \frac{qY^s + (1-q)Y + C}{\eta}$$

$$\hat{\pi} > \frac{qY^s + (1-q)Y + C}{2\eta\triangle\mu} + 1/2$$

where $\triangle\mu = \mu(1,1) - \mu(1,0)$. In this PBE:
If $\hat{\pi} > \hat{\pi}^*$, then $\mu(1,1) = \Pr(s=1|\sigma, s=1) = 1$ and $\mu(1,0) = \Pr(s=1|\sigma, s=0) = 0$. Since $\hat{\pi}$ is common knowledge, consumers calculate that if $\hat{\pi} > \hat{\pi}^*$, then $\triangle\mu = 1$ which assigns the maximum reputational revenue. Thus, $\triangle\mu = 1$, implying $\hat{\pi} > \frac{qY^s+(1-q)Y+C}{2\eta(1-0)} + 1/2 \geq \hat{\pi}^*$. If $\hat{\pi} < \hat{\pi}^*$, then $\mu(1,1) = \Pr(s=1|\sigma, s=1) = x \in (0,1)$ (it can be anything), $\mu(1,0) = \Pr(s=1|\sigma, s=0) = \hat{\pi}$, $\Delta\mu < 1$ and

$$\hat{\pi} < \hat{\pi}^* = \frac{qY^s + (1-q)Y + C}{2\eta(1-0)} + 1/2$$

The vendor does not find it worthwhile to choose an honest action $s_i = 1$ to seek for any reputation; not even the maximum reputation gain $\Delta\mu = (1-0) = 1$ makes it worthwhile to choose an honest action $s = 1$. The cost C is too high. In our experiment, by providing symmetric two-sided information about the official prices of transactions, consumers' signal σ is the same as the s action chosen by the vendor. There is revelation of the imperfectly observed vendor's actions and beliefs are updated to the posterior $\Pr[s = 1|\sigma = s = 1] = 1$ and $\Pr[s = 1|\sigma = s = 0] = 0$. ■

A.2 Proof of Proposition 2

Proof. For (i), it follows directly by noting that $\Pr(s = 1|\hat{\pi})$ *is increasing in* $\hat{\pi}$. To prove (ii), we first provide foundations for η.

Foundations: Computing η: Denote by θ the fraction of informed customers, v_G the value of ethical transactions to the customer, v_B the value of unethical transactions to the customer, where $v_G > v_B$. For simplicity, we assume that customers have the same willingness to pay for ethical transactions. The expected value of transacting (for customers) is: $v(\Pr[s = 1|\sigma, s]) = \Pr[s = 1|\sigma, s]v_G + (1 - \Pr[s = 1|\sigma, s])v_B$, with a reduced form demand function: $D_i(\Pr[s = 1|\sigma, s] = 1) = v(\Pr[s = 1|\sigma, s] = 1) = \theta v_G$ for the informed customers versus $D_u(\Pr[s = 1|\sigma, s]) = v(\Pr[s = 1|\sigma, s]) = (1-\theta)v(\Pr[s = 1|\sigma, s])$ for the uninformed customers. Thus, the aggregate market demand for honest transactions is

$$D_{s=1}(\Pr[s = 1|\sigma, s], .) = \underbrace{\theta v_G}_{D_i} + \underbrace{(1-\theta)v(\Pr[s = 1|\sigma, s])}_{D_u}$$

Similarly, the aggregate demand is $D_{s=0}(\Pr[s = 1|\sigma, s], .) = \theta v_B + (1-\theta)v(\Pr[s = 1|\sigma, s])$ for dishonest transactions.

Effects: Letting η equal the aggregate demand D_s, and observing that $\frac{\partial D_{s=1}}{\partial \theta} = v_G - v(\Pr[s = 1|\sigma, s]) = v_G - \Pr[s = 1|\sigma, s]v_G - (1 - \Pr[s = 1|\sigma, s])v_B \geq 0|_{\Pr[s=1|\sigma,s]=1}$ in equilibrium. For dishonest transactions, $\frac{\partial D_{s=0}}{\partial \theta} = v_B - v(\Pr[s = 1|\sigma, s]) \leq 0|_{\Pr[s=1|\sigma,s]=1}$. We thus have the following result: For (ii), $\eta(\theta)$ *is weakly-increasing in* θ*. Since* $\hat{\pi}^*$ *is decreasing in* η*, noting that* $\lim_{\eta\to+\infty} \hat{\pi}^* = 0$*, it follows that* $\Pr(s = 1)$ *is more likely.*

To prove (iii), it suffice to show that $\frac{\partial \hat{\pi}}{\partial q}|_{\hat{\pi}=\hat{\pi}^{**}} < 0$ and $\frac{\partial \hat{\pi}}{\partial \triangle\mu}|_{\hat{\pi}=\hat{\pi}^{**}} < 0$ where $\hat{\pi}^{**} = \frac{qY^s+(1-q)Y+C}{2\eta\triangle\mu} + 1/2$ since both make $\Pr(s = 1)$ more likely. We have that $\frac{\partial \hat{\pi}}{\partial q}|_{\hat{\pi}=\hat{\pi}^{**}} = \frac{Y^s-Y}{2\eta\triangle\mu} < 0$ because $Y^s < Y$. Similarly, $\frac{\partial \hat{\pi}}{\partial \triangle\mu}|_{\hat{\pi}=\hat{\pi}^{**}} = -\frac{2\eta(qY^s+(1-q)Y+C)}{(2\eta\triangle\mu)^2} < 0$. ■

B Balance and Attrition

Table 9: **Balance test I: representativeness of select-sample with market population (vendors)**

Supply side: Vendors

	Constant	Select
Demographic Characteristics		
Female	0.398***	0.021
	(0.049)	(0.076)
Married	0.205***	0.083
	(0.043)	(0.065)
Akan ethnic	0.571***	8.96*e*-04
	(0.054)	(0.076)
Age	26.456***	0.716
	(0.585)	(1.117)
Education (any)	0.725***	-0.040
	(0.050)	(0.076)
Self employment	0.552***	-0.126*
	(0.058)	(0.075)
M-Money training	0.493***	0.043
	0.050	(0.070)
Poverty Indicators		
Household head read English	4.104***	0.102
	(0.163)	(0.223)
Outer wall used cement	3.909***	-0.306
	(0.222)	(0.342)
Toilet facility	4.617***	-0.349
	(0.140)	(0.268)
Number working mobile phones	8.466***	0.366
	(0.208)	(0.261)
Own working bicycle/ motor bicycle / car	1.554***	0.715
	(0.287)	(0.499)
Market: Size + Sales		
M-Money: Total volume [GHS] (daily)	2296.046***	24.611
	(129.932)	(178.263)
Non M-Money: Number customers (daily)	32.829***	-0.023
	(1.796)	(2.520)
Non M-Money: Total volume [GHS] (daily)	156.404***	-0.726
	(6.272)	(8.799)
Joint F-test (linear), *p*-value	0.375	
Chi-squared test (probit), *p*-value	0.460	

Note: Observations are at the vendor or market level. Each row is a separate regression. The F and Chi-squared tests are conducted excluding all market outcomes. Robust standard errors are in parentheses *** p<0.01, ** p<0.05, * p<0.10.

Table 10: **Balance test i: representativeness of select-sample with market population (customers)**

Demand side: Customers

	Constant	Select
Demographic Characteristics		
Female	0.628***	-2.0e-3
	(0.022)	(0.026)
Married	0.517***	0.021
	(0.019)	(0.024)
Akan ethnic	0.623***	-2.7e-3
	(0.036)	(0.039)
Age	38.635***	1.688*
	(0.737)	(0.891)
Education (any)	0.890***	9.7e-3
	(0.015)	(0.016)
Self employment	0.665***	0.025
	(0.029)	(0.029)
M-Money registered	0.905***	1.2e-3
	(0.014)	(0.017)
Poverty Indicators		
Household head read English	3.428***	-0.124
	(0.114)	(0.152)
Outer wall used cement	3.664***	-0.272
	(0.196)	(0.195)
Toilet facility	4.372***	-0.584
	(0.137)	(0.182)
Number working mobile phones	7.151***	-0.159
	(0.123)	(0.159)
Own working bicycle/ motor bicycle / car	1.180***	0.238
	(0.143)	(0.176)
Subjective Assessment: Fraud or Misconduct		
Attempted fraud experience (any)	0.611***	-0.041
	(0.040)	(0.039)
Ever over-charged/ unauthorized account use	0.292***	0.013
	(0.024)	(0.028)
Market: Features + Transactions		
Distance to closest formal bank (meters)	286.079***	147.891
	(73.105)	(107.315)
Distance to closest M-Money (meters)	66.295***	-10.758
	12.787	(13.021)
M-Money: Total use volume [GHS] (weekly)	129.227***	29.280
	(12.982)	(19.406)
Non M-Money: Number use (weekly)	2.062***	0.430
	(0.531)	(0.782)
Non M-Money: Total use volume [GHS] (weekly)	46.149*	-0.449
	(24.141)	(25.959)
Borrowing + Savings		
Likelihood to borrow via M-Money (1-5 scale)	1.515***	-0.065
	(0.073)	(0.069)
Likelihood to save via M-Money (1-5 scale)	2.126***	4.55e-3
	(0.095)	(0.104)
Joint F-test (linear), p-value	0.181	
Chi-squared test (probit), p-value	0.206	

Note: Observations are at the vendor or market level. Each row is a separate regression. The F and Chi-squared tests are conducted excluding all market outcomes. Robust standard errors are in parentheses *** p<0.01, ** p<0.05, * p<0.10.

Table 11: **Balance test II: pre-intervention treatment-control differences (vendors)**

Supply side: Vendors	Constant	**PT**	**MR**	**Joint: PT + MR**
Demographic Characteristics				
Female	0.551***	-0.180	-0.255*	-0.058
	(0.118)	(0.159)	(0.153)	(0.159)
Married	0.389***	-0.037	-0.202	-0.131
	(0.117)	(0.160)	(0.145)	(0.153)
Akan ethnic	0.491***	0.218	-0.118	0.189
	(0.119)	(0.156)	(0.161)	(0.151)
Age	27.097***	-0.413	2.163	-1.358
	(1.955)	(2.973)	2.845	(2.454)
Education (any)	0.697	-0.044	0.042	-0.041
	(0.126)	0.169	(0.165)	(0.163)
Self employment	0.443	0.058	0.008	-0.124
	(0.118)	(0.163)	(0.163)	(0.151)
M-Money training	0.340	0.265	0.293	0.170
	(0.119)	(0.163)	(0.159)	(0.160)
Poverty Indicators				
Household head read English	4.248***	-0.213	-0.093	0.139
	(0.295)	(0.506)	(0.480)	(0.4178)
Outer wall used cement	3.783***	0.038	-0.204	-0.486
	(0.591)	(0.790)	(0.794)	(0.784)
Toilet facility	4.464***	0.400	-0.581	-0.530
	(0.370)	(0.561)	(0.679)	(0.560)
Number working mobile phones	8.854***	-0.089	0.383	-0.346
	(0.276)	(0.490)	(0.490)	(0.449)
Own working bicycle/ motor bicycle / car	2.037***	0.004	0.359	0.483
	(0.642)	(1.072)	(1.002)	(1.052)
Poverty rate (Schneider 2015)	5.326	5.299	2.299	4.821
	(3.270)	(6.184)	(4.116)	(4.219)
Market: Size + Sales				
M-Money: Total volume [GHS] (daily)	1925.800***	305.049	478.480	665.939
	(555.950)	(789.582)	(902.508)	(1654.237)
Non M-Money: Number customers (daily)	32.473***	-2.080	-8.057	10.789
	(6.788)	(9.202)	(8.859)	(14.256)
Non M-Money: Total volume [GHS] (daily)	163.750***	-30.789	-14.096	14.986
	(61.630)	(66.831)	(69.562)	(73.869)
Joint F-test (linear), *p*-value			0.711	
Chi-squared test (probit), *p*-value			0.534	

Note: Observations are at the vendor or market level. Each row is a separate regression. The F and Chi-squared tests are conducted using the meta indicator **1(Information Assignment)** as the outcome and excluding all market outcomes. Robust standard errors are in parentheses *** p<0.01, ** p<0.05, * p<0.10.

Table 12: **Balance test II: pre-intervention treatment-control differences (customers)**

Demand side: Customers

	Constant	**PT**	**MR**	**Joint: PT + MR**
Demographic Characteristics				
Female	0.635***	0.003	-0.001	-0.034
	(0.053)	(0.061)	(0.069)	(0.064)
Married	0.505***	0.038	0.004	0.077
	(0.039)	(0.048)	(0.051)	(0.056)
Akan ethnic	.548***	0.101	0.077	0.092
	(0.072)	(0.092)	(0.102)	(0.090)
Age	39.380***	2.189	0.436	0.818
	(1.370)	(1.987)	(1.932)	(1.754)
Education (any)	0.891***	0.035	-0.027	0.021
	(0.025)	(0.029)	(0.042)	(0.033)
Self employment	0.668***	0.015	0.039	0.030
	(0.041)	(0.054)	(0.067)	(0.060)
M-Money registered	0.896***	-0.010	0.017	0.019
	(0.029)	(0.044)	(0.037)	(0.036)
Poverty Indicators				
Household head read English	3.353***	-0.081	-0.345	0.226
	(0.212)	(0.321)	(0.347)	(0.305)
Outer wall used cement	3.315***	-0.263	0.245	0.307
	(0.456)	(0.551)	(0.520)	(0.560)
Toilet facility	4.206***	-0.427	-0.478	-0.634*
	(0.169)	(0.377)	(0.405)	(0.327)
Number working mobile phones	7.086***	-0.415	-0.005	0.072
	(0.204)	(0.298)	(0.315)	(0.300)
Own working bicycle/ motor bicycle / car	1.141***	0.124	0.395	0.503
	(0.284)	(0.372)	(0.372)	(0.414)
Poverty rate (Schneider 2015)	11.280***	2.772	1.704	0.046
	(1.478)	(2.420)	(2.191)	(1.976)
Subjective Assessment: Fraud or Misconduct				
Attempted fraud experience (any)	0.565***	-0.000	0.018	$2.41e$-16
	(0.044)	(0.070)	(0.065)	(0.067)
Ever over-charged/ unauthorized account use	0.336***	-0.067	-0.037	-0.010
	(0.041)	(0.057)	(0.056)	(0.056)
Market: Features + Transactions				
Distance to closest formal bank (meters)	249.470**	-33.832	242.196	447.365*
	(96.807)	(127.385)	(255.640)	240.233
Distance to closest M-Money (meters)	45.623***	28.577	5.426	2.920
	(15.154)	(22.952)	(19.682)	(17.788)
M-Money: Total use volume [GHS] (weekly)	158.005***	-28.246	-9.495	37.712
	(35.465)	(40.296)	(41.623)	(55.060)
Non M-Money: Number use (weekly)	2.141***	-.255	1.049	0.532
	(0.606)	(0.748)	(1.972)	(1.230)
Non M-Money: Total use volume [GHS] (weekly)	26.706**	31.607	20.569	17.800
	(12.093)	(28.309)	(19.784)	(20.181)
Borrowing + Savings				
Likelihood to borrow via M-Money (1-5 scale)	1.391***	-0.011	0.098	0.130
	(0.120)	(0.141)	(0.171)	(0.174)
Likelihood to save via M-Money (1-5 scale)	2.103***	-0.070	0.087	0.085
	(0.177)	(0.248)	(0.246)	(0.264)
Joint F-test (linear), p-value			0.850	
Chi-squared test (probit), p-value			0.846	

Note: Observations are at the vendor or market level. Each row is a separate regression. The F and Chi-squared tests are conducted using the meta indicator **1(Information Assignment)** as the outcome and excluding all market outcomes. Robust standard errors are in parentheses *** p<0.01, ** p<0.05, * p<0.10.

Table 13: **Attrition**

	PT	**MR**	**Joint: PT + MR**	**Control**	Total	Attrition
CENSUS (*Joint baseline*)						
Vendors					333	
Customers					1,921	
Markets (vendor×customers)					333	
SELECT SAMPLE (*Randomized*)						
Vendors	31	32	35	32	130	
Customers	272	257	276	185	990	
Markets (vendor×customers)	31	32	35	32	130	
ENDLINE (*Follow-up*)						
Vendors	26	28	28	25	107	23
	(84%)	(88%)	(80%)	(78%)	(82%)	(18%)
	(SD=37%)	(SD=33%)	(SD=40%)	(SD=42%)	(SD=38%)	(SD=38%)
Customers	230	207	230	143	810	180
	(85%)	(81%)	(83%)	(77%)	(82%)	(18%)
	(SD=36%)	(SD=39%)	SD=37%)	(SD=42%)	(SD=39%)	(SD=39%)
Markets (vendor×customers)	26	28	28	25	107	23
	(84%)	(88%)	(80%)	(78%)	(82%)	(18%)
	(SD=37%)	(SD=33%)	SD=40%)	(SD=42%)	(SD=38%)	(SD=38%)

Note: Table reports the summary statistics for the subsample that was successfully reached for a follow-up and for the subsample that was not successfully reached in endline phone surveys and manual contact tracing. Shown for both sides of the market (vendors versus customers). Attrition for endline audit exercises is 0.8%: 129 out of the 130 representative vendors were reached. There was only one unreachable vendor in the joint PT + MR program.

C Descriptive Statistics

Table 14: **Summary statistics of relevant variables from the market census**

	Vendors		**Customers**	
	Mean	SD	Mean	SD
Demographic Characteristics				
Female	0.398	0.489	0.623	0.484
Self employment	0.479	0.499	0.681	0.466
Self income -- monthly [GHS]	2.014	1.483	1.376	0.868
Married	0.249	0.432	0.535	0.498
Akan ethnic	0.572	0.494	0.621	0.485
Age (years)	26.291	8.242	39.545	15.021
Education (any)	0.691	0.461	0.896	0.304
M-Money training	0.508	0.500		
M-Money registered (self + any close person)			0.905	0.293
Poverty Indicators				
Household size (above 5)	0.223	0.416	0.244	0.430
Household head read English	0.769	0.421	0.606	0.488
Outer wall used cement	0.749	0.433	0.705	0.456
Toilet facility	0.891	0.311	0.849	0.357
Working mobile phone(s)	0.976	0.152	0.976	0.151
Own working bicycle/ motor bicycle/ car	0.280	0.449	0.214	0.410
Market: Access + Transactions + Sales				
Doing business experience (years)	2.051	2.12		
Joint venture: M-Money + other services	0.752	0.431		
M-Money: Total volume [GHS] (daily)	2260.569	3775.947		
Non M-Money: Number customers (daily)	32.791	47.067		
Non M-Money: Total volume [GHS](daily)	155.156	164.574		
Distance to closest formal bank (meters)			338.577	751.370
Distance to closest post office (meters)			382.932	250.737
Distance to closest M-Money (meters)			61.288	94.928
Formal bank user (of nearby banks)			0.806	0.395
Post-office user (of nearby offices)			0.092	0.290
M-Money user (of nearby vendors)			0.946	0.224
M-Money: Total use volume [GHS] (weekly)			144.199	396.283
Non M-Money: Number use (weekly)			2.272	14.766
Non M-Money: Total use volume [GHS] (weekly)			44.700	505.107
Borrowing + Savings				
Likelihood to borrow via M-Money (1-5 scale)			1.477	0.877
Likelihood to save via M-Money (1-5 scale)			2.112	1.213
Subject Assessment: Fraud or Misconduct				
Attempted fraud experience (any)			0.589	0.492
Ever over-charged			0.191	0.403
Ever over-charged + unauthorized account use			0.293	0.455
Number of observations	333		1,921	

Note: Table reports the summary statistics of relevant variables from our market census separately for both sides of the market: vendors *versus* customers. This include information about demographics, poverty indicators, and market outcomes, respectively. Customers' borrowing and savings behavior and their subjective assessment of market misconduct on M-Money are also shown. The census cover 333 vendors and 1,921 customers or households across a space of 137 villages. The exchange rate during the market census period is US$ 1.0 = GHS 5.12.

Figure 3: **Distribution (histogram) of total transactions at endline**

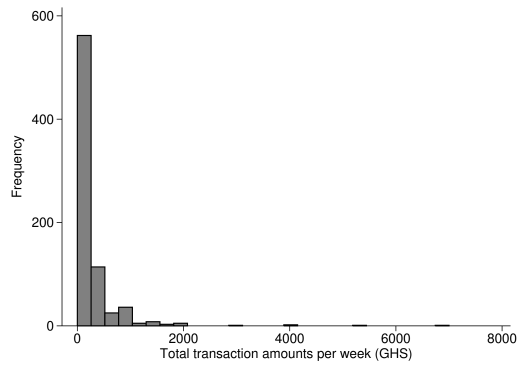


Figure 4: **Phone calls and reachability of subjects**

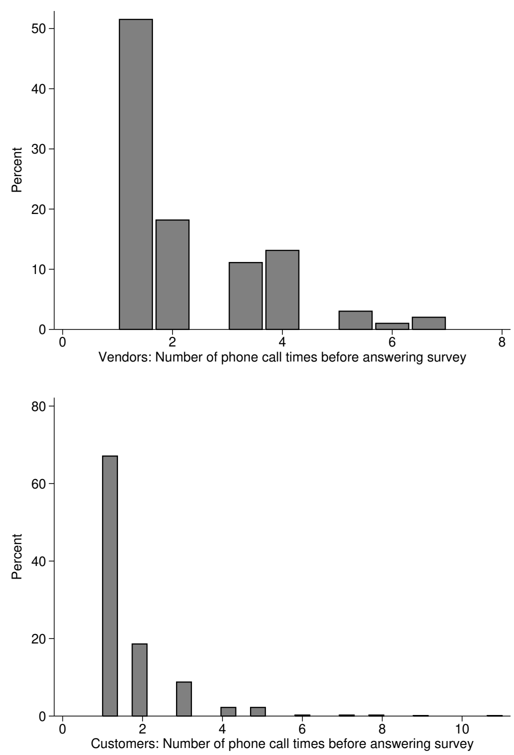
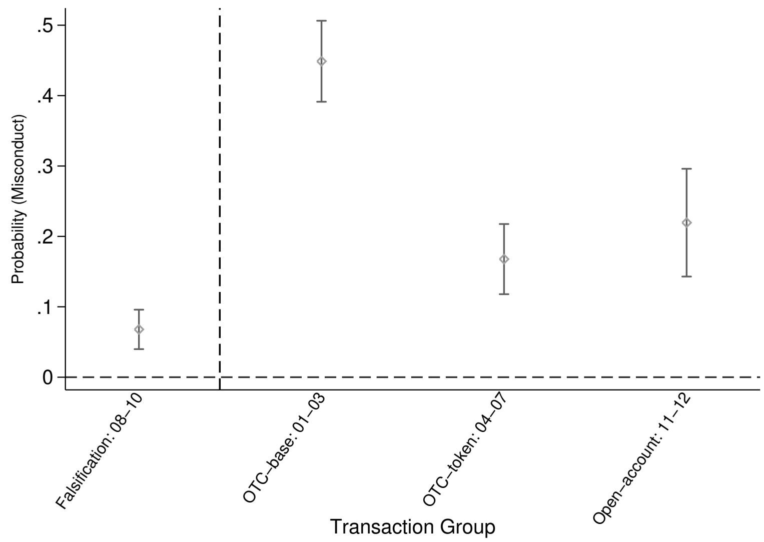


Table 15: **Misconduct at baseline: descriptive statistics based on transactional audit exercise, details**

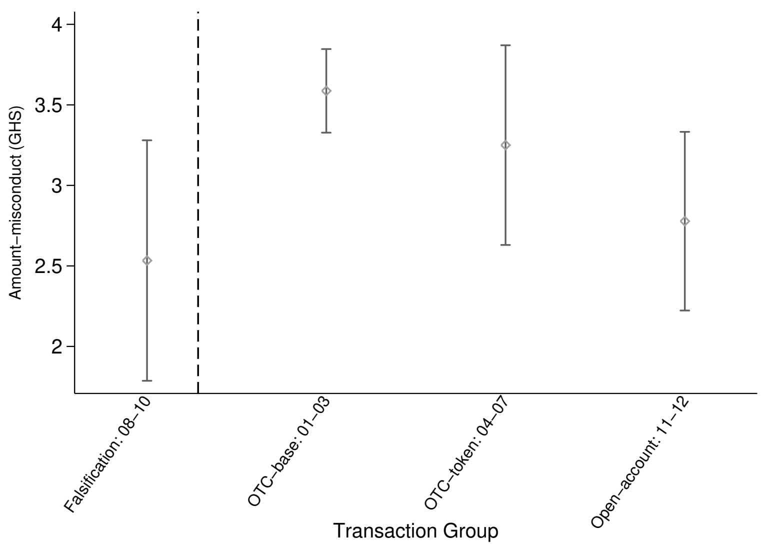
# Transaction type (description)	Outcome variable	Mean	SD	Transaction group	Mean	SD
01 Cash-in GHS50 - to others wallet	**1**[Misconduct=Yes]	0.35	0.480		0.44	0.498
	Overcharged [GHS]	4.65	1.093		3.58	1.498
02 Cash-in GHS160 - to others wallet	**1**[Misconduct=Yes]	0.52	0.502	$\{= \quad OTC - base$		
	Overcharged [GHS]	4.07	0.269			
03 Cash-in GHS1100 - to others wallet	**1**[Misconduct=Yes]	0.48	0.504			
	Overcharged [GHS]	1.85	1.406			
04 Send GHS50 token - to others	**1**[Misconduct=Yes]	0.18	0.390		0.16	0.374
	Overcharged [GHS]	3.68	1.624		3.25	1.850
05 Send GHS1100 token - to others	**1**[Misconduct=Yes]	0.19	0.397			
	Overcharged [GHS]	3.25	1.982	$\{= \quad OTC - token$		
06 Receive GHS50 token - from others	**1**[Misconduct=Yes]	0.20	0.405			
	Overcharged [GHS]	2.71	2.138			
07 Receive GHS1100 token-from others	**1**[Misconduct=Yes]	0.08	0.287			
	Overcharged [GHS]	3.33	2.081			
08 Cash-in GHS50 - to own wallet	**1**[Misconduct=Yes]	0.07	0.259		0.06	0.252
	Overcharged [GHS]	3.20	2.049		2.53	1.641
09 Cash-in GHS160 - to own wallet	**1**[Misconduct=Yes]	0.08	0.274	$\{= \quad Falsification$		
	Overcharged [GHS]	2.00	1.549			
10 Cash-out GHS50 - from own wallet	**1**[Misconduct=Yes]	0.05	0.223			
	Overcharged [GHS]	2.50	1.290			
11 Purchase new SIM card	**1**[Misconduct=Yes]	0.32	0.473		0.21	0.416
	Overcharged [GHS]	2.73	1.099	$\{= \quad Open - account$	2.77	1.352
12 Register new M-Money wallet	**1**[Misconduct=Yes]	0.08	0.280			
	Overcharged [GHS]	3.00	2.645			
Overall	**1**[Misconduct=Yes]	0.22	0.419		0.22	0.419
	Overcharged [GHS]	3.32	1.591		3.32	1.591
Number of transactions		663-1,548			663-1,548	

Note: Table reports the specific transactions used for the actual transactional exercises and shows the descriptive statistics of misconduct. These misconduct outcomes are based on the transactional exercises. Transactions are categorized into four groups, namely: OTC-base, OTC-token, Falsification, and Open-account. OTC denotes over-the-counter and captures transactions that involve little to no automation or active verification from the side of the customer (i.e., more room for vendors to overcharge OTCs). **1**[.] is a logical indicator that takes the value 1 whenever the argument in the bracket is true, and zero otherwise. Overall, the incidence of misconduct is 22% [SD=0.419] and the average overcharged-amount due to misconduct is GHS3.32 [SD=1.591], which represents $\frac{3.32}{4.03} \times 100 = 82\%$ of the average "official charge" for the transactional amounts used in the audit exercises. Our field market transactions are allowed to vary in sizes of GHS50 (small), GHS160 (medium) and GHS1,100 (large). Their official charges are GHS0.50, GHS1.60 and GHS10.00 respectively. Thus, the average official charge, pooling all the 3 varying transaction sizes, is approximately GHS4.03.

Figure 5: **Misconduct at baseline: descriptive statistics based on transactional audit exercise, graphical**



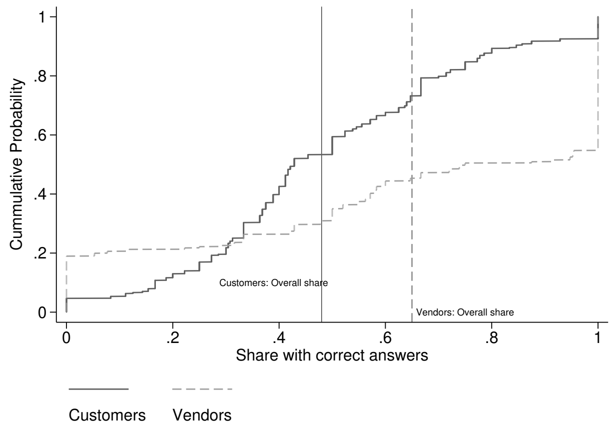
(a) **MISCONDUCT INCIDENCE × TRANSACTION GROUP**



(b) **MISCONDUCT SEVERITY × TRANSACTION GROUP**

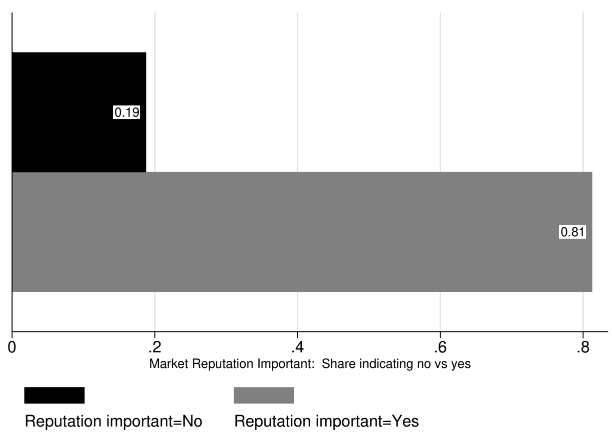
Note: Figures display the distribution of misconduct, measured as either the probability of the vendor committing a misconduct "incidence" (Figure (a)) or the amount overcharged as result of misconduct "severity" (Figure (b)) using actual transactional exercises at baseline. Transactions are categorized into four groups, namely: OTC-base, OTC-token, Falsification, and Open-account. OTC denotes over-the-counter and captures transactions that involve little to no automation from the side of the customer. The specific transactions in each transaction group are reported in the Table 15. 90% confidence intervals (CI) are displayed around the estimates. As expected, misconduct is much higher in the OTC-type transactions (i.e., little to no automation/verification required from the customer) compared to the Falsification group (automation and active verification required from the customer).

Figure 6: **Asymmetric information about transactional prices**



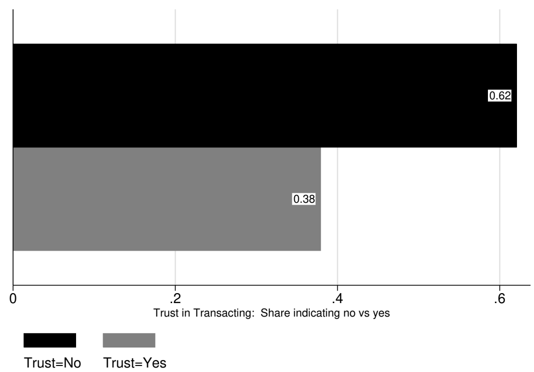
Note: Figure plots the distributions (CDFs) of the share of subjects with accurate answers for charges on randomly selected popular transactions (GHS200; GHS1200) derived with reference to their official or mandated rates (2GHS; 10GHS, respectively). A subject is correct if his/her answer matches the mandated rate. Shown separately for customers and vendors. From a Kolmogorov–Smirnov test for the equality of distributions, p-value < 0.01.

Figure 7: **Importance of reputation to vendors**



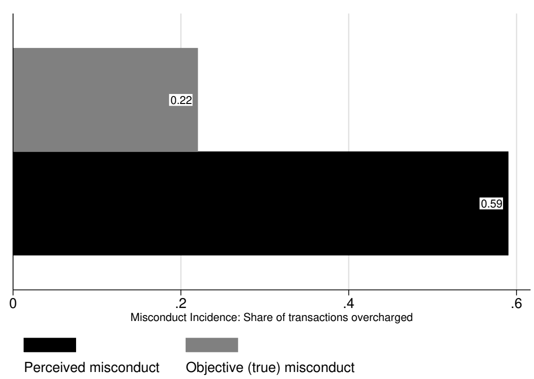
Note: Figure plots the share of vendors that value good market reputation through their money market transactions. Subjects were asked to indicate how important it is to show a high degree of good market image and responsibility to potential customers when carrying out M-Money transactions on a scale of 1 (not important) to 5 (very important). To ease the exposition, we first obtain the frequency distribution of the 1-5 value data and then find the median value (i.e., 4). All values above the median are recoded to be "yes" (reputation important), and those below are recorded as "no" (reputation not important). From an unpaired t-test for equality of vendors proportions of reputation-important and reputation-not important, p-value $= 0.000$.

Figure 8: **Consumer trust in performing money transactions at vendor points**



Note: Figure plots the share of customers, at baseline, that trust or not the money transactions they make at vendor banking points. Subjects were asked to indicate their level of trust for carrying out M-Money transactions at the vendor points from a scale of 1 (low) to 5 (high). To ease the exposition, we first obtain the frequency distribution of the 1-5 value data and then find the median value (i.e., 3). All values strictly above the median are recoded to be "yes" for trust in transacting (trust), and those below are recorded as "no" (distrust). From an unpaired t-test for equality of customers proportions of distrust and trust, p-value $= 0.000$.

Figure 9: **Misperceived beliefs about misconduct**



Note: Figure plots the share of transactions that are actually overcharged (truth) versus customers estimate of the share that are overcharged (perceived). From an unpaired t-test for equality of true misconduct $(1-\pi)$ and perceived misconduct $(1-\hat{\pi})$, p-value $=0.000$. $\pi =$ the share of non-misconduct or transactions not-overcharged.

D Further Results: Treatment Effects

Table 16: **Spillover effect of information sets on vendors misconduct**

	1(Misconduct=Yes)		Amount-Misconduct, GHS	
PANEL A				
Treatment: Information Assignment (β)	-0.153	-0.218	-0.473	-0.648
	(0.055)	(0.065)	(0.173)	(0.206)
	[-0.245, -0.060]	[-0.326, -0.109]	[-0.763, -0.184]	[-0.992, -0.303]
Baseline misconduct				
Market District F.E.	X		X	
Market District × Transaction × Date F.E.		X		X
Observations	411	411	411	411
Mean of dependent variable (control)	0.278	0.278	0.783	0.783
Lee (2009) Attrition Bounds	<-0.170, -0.155>		<-0.569, -0.479>	
Imbens and Manski (2004) CS	[-0.305, -0.076]		[-1.211, -0.220]	
PANEL B				
Price Transparency (β_1)	-0.163	-0.232	-0.567	-0.720
	(0.058)	(0.070)	(0.172)	(0.196)
	[-0.260, -0.065]	[-0.351 -0.114]	[-0.856, -0.279]	[-1.048, 0.391]
Monitor and Report (β_2)	-0.182	-0.239	-0.470	-0.693
	(0.056)	(0.075)	(0.191)	(0.242)
	[-0.277 0.087]	[-0.364, -0.113]	[-0.789, -0.151]	[-1.098, -0.287]
Joint program: PT + MR (δ)	-0.122	-0.178	-0.409	-0.524
	(0.069)	(0.070)	(0.211)	(0.224)
	[-0.238, -0.006]	[-0.296, -0.060]	[-0.762, -0.055]	[-0.900 -0.149]
Baseline misconduct				
Market District F.E.	X		X	
Market District × Transaction × Date F.E.		X		X
Observations	405	405	405	405
Mean of dependent variable (control)	0.278	0.278	0.783	0.783
p-value (test: $\beta_1 = \delta$)	0.512	0.315	0.353	0.179
p-value (test: $\beta_2 = \delta$)	0.235	0.235	0.712	0.323
p-value (test: $\beta_1 = \beta_2$)	0.640	0.915	0.482	0.859
p-value (test: $\beta_1 + \beta_2 = \delta$)	0.007	0.001	0.011	0.001

Note: For spillover effects, estimations compare non-treated vendors located in treated localities (or markets) to the pure control localities. **1**(.) is a logical indicator that equals 1 if the argument in the parenthesis is true, 0 otherwise. Includes randomization strata (market district) dummies. Observations are at the vendor x transaction x date level. Clustered standard errors (at local market level) reported in parenthesis. 90% confidence intervals (CI) and confidence sets (CS) are reported in brackets. Panel A reports meta estimate of treatment effects, while panel B shows effects separately for each information program. Results similar to post-double-selection LASSO estimates clustered at the (local) market level and to alternative inference procedures (Wild cluster bootstrap and permutation test clustered at the market level).

Table 17: **Heterogeneity in effect of information sets on vendor misconduct**

		MARKET COMPETITION			VENDORS' GENDER	
		1(Misconduct=Yes)	Misconduct, GHS		**1**(Misconduct=Yes)	Misconduct, GHS
Treatment: Information Assignment (β)		-0.905	-2.796		-0.254	-0.658
		(0.271)	(1.271)		0.097	(0.295)
		[-1.362, -0.448]	[-4.937, -0.656]		[-0.417, -0.092]	[-1.150, -0.166]
	x Competition	-1.237	-4.303	**x** Female	0.129	0.320
		(0.658)	(2.730)		(0.143)	(0.448)
		[-2.345, -0.128]	[-8.898, 0.292]		[-0.109, 0.368]	[-0.424, 1.065]
	Competition	1.164	3.885	Female	-0.161	-0.396
		(0.655)	(2.817)		(0.131)	(0.434)
		[0.061, 2.267]	[-0.855, 8.625]		[-0.381, 0.057]	[-1.118, 0.324]
Baseline misconduct		X	X		X	X
Market District $\times$ Transaction $\times$ Date F.E.		X	X		X	X
Observations		159	159		335	335
Mean of dep var (control)		0.294	0.778		0.294	0.778

Note: **1**(.) is a logical indicator that equals 1 if the argument in the parenthesis is true, 0 otherwise. Includes randomization strata (market district) dummies. Observations are at the vendor x transaction x date level. Clustered standard errors (at local market level) reported in parenthesis. 90% confidence intervals (CI) are reported in brackets. Results similar to post-double-selection LASSO estimates clustered at the (local) market level and to alternative inference procedures (Wild cluster bootstrap and permutation test clustered at the market level). Market competition index is defined as negative of the Herfindahl-Hirschman (HH) index trimmed to the closed interval $(0, 1)$ to minimize extreme influences.

Table 18: **Heterogeneity in effect of information sets on vendor misconduct**

		MARKET COMPETITION			VENDORS' GENDER	
		1(Misconduct=Yes)	Misconduct, GHS		**1**(Misconduct=Yes)	Misconduct, GHS
Price Transparency (β_1)		-0.652	-2.094		-0.224	-0.549
		(0.321)	(1.565)		(0.109)	(0.326)
		[-1.193, -0.110]	[-4.729, 0.540]		[-0.407, -0.042]	[-1.091, -0.007]
	x Competition	-0.728	-2.802	**x** Female	0.155	0.358
	(b_1)	(0.731)	(3.202)	(b_1)	(0.166)	(0.528)
		[-1.960, 0.502]	[-8.191, 2.587]		[-0.120, 0.432]	[-0.519, 1.237]
Monitor and Report (β_2)		-0.713	-2.111		-0.237	-0.680
		(0.340)	(1.471)		(0.109)	(0.337)
		[-1.286, -0.139]	[-4.587, 0.364]		[-0.419, -0.054]	[-1.241, -0.120]
	x Competition	-0.742	-2.410	**x** Female	0.086	0.324
	(b_2)	(0.786)	(3.059)	(b_2)	(0.164)	(0.513)
		[-2.065, 0.580]	[-7.559, 2.737]		[-0.186, 0.359]	[-0.528, 1.177]
Joint program: PT+MR (δ)		-0.965	-2.880		-0.278	-0.673
		(0.291)	(1.333)		(0.104)	(0.317)
		[-1.456, -0.473]	[-5.124, -0.637]		[-0.452, -0.104]	[-1.200, -0.145]
	x Competition	-1.502	-5.028	**x** Female	0.197	0.350
	(d)	(0.702)	(2.953)	(d)	(0.165)	(0.548)
		[-2.684, -0.320]	[-9.998, -0.057]		[-0.076, 0.472]	[-0.561, 1.262]
	Competition	0.834	2.681	Female	-0.170	-0.407
		(0.704)	(3.068)		(0.134)	(0.440)
		[-0.351, 2.019]	[-2.481, 7.845]		[-0.393, 0.052]	[-1.139, 0.323]
Baseline misconduct		X	X		X	X
Market District ×		X	X		X	X
Transaction × Date F.E.						
Observations		159	159		335	335
Mean of dep var (control)		0.294	0.778		0.294	0.778
p-value (test: $b_1 = d$)		0.116	0.376		0.787	0.987
p-value (test: $b_2 = d$)		0.118	0.698		0.366	0.950
p-value (test: $b_1 = b_2$)		0.965	0.535		0.612	0.942
p-value (test: $b_1 + b_2 = d$)		0.974	0.074		0.838	0.628

Note: **1**(.) is a logical indicator that equals 1 if the argument in the parenthesis is true, 0 otherwise. Includes randomization strata (market district) dummies. Observations are at the vendor x transaction x date level. Clustered standard errors (at local market level) reported in parenthesis. 90% confidence intervals (CI) are reported in brackets. Results similar to post-double-selection LASSO estimates clustered at the (local) market level and to alternative inference procedures (Wild cluster bootstrap and permutation test clustered at the market level). Market competition index is defined as negative of the Herfindahl-Hirschman (HH) index trimmed to the closed interval $(0, 1)$ to minimize extreme influences.

Table 19: **Effect of treatment on usage**

Model: Tobit
DV: Total Transaction per week (GHS)

	(1)	(2)	(3)	(4)
Treatment: Information Assignment (β)	116.628	106.077	99.402	95.292
	(52.439)	(52.149)	(53.718)	(52.489)
	[30.267, 202.989]	[20.194, 191.960]	[10.928, 187.875]	[8.840, 81.743]
sigma (σ)	581.695	576.667	571.064	563.983
	(83.946)	(83.240)	(83.464)	(82.838)
	[443.447, 719.942]	[439.580, 713.754]	[433.598, 708.529]	[427.547, 700.418]
Inference Robustness (β)				
Clustered S.E.	[18.033, 215.222]	[15.901, 196.253]	15.97649 182.828	[15.380, 175.203]
p-value: Permutation Test	0.069	0.085	0.085	0.091
Market District F.E.		X	X	X
Baseline usage			X	X
Controls				X
Observations	763	763	723	723
Mean of dependent variable (control)	198.956	198.956	198.956	198.956

Note: Market district is the randomization strata. Observations are at the customer level. Robust standard errors reported in parenthesis. 90% confidence intervals (CI) and confidence sets (CS) are reported in brackets. Reported confidence CIs for wild bootstrap and permutation tests derived from running 1000 replications in each case. Permutation test (for randomization inference) clustered at the (local) market level.

Table 20: **Effect of treatment on usage**

Linear Model
DV: 0-1 Indicator for whether consumer is "using" M-Money

	(1)	(2)	(3)	(4)
Treatment: Information Assignment (β)	0.096	0.078	0.071	0.071
	(0.041)	(0.039)	(0.039)	(0.038)
	[0.028, 0.164]	[0.013, 0.143]	[0.006, 0.136]	[0.007, 0.133]
Inference Robustness (β)				
CI: Clustered S.E.	[-0.003, 0.197]	[0.010, 0.146]	[0.006, 0.136]	[0.011, 0.129]
CI: Wild Bootstrap	[0.028, 0.164]	[0.008, 0.135]	[0.002, 0.139]	[0.008, 0.132]
p-value: Permutation Test	0.017	0.028	0.045	0.049
p-value: L-S-X MHT Corr (2019)	0.022			
Market District F.E.		X	X	X
Baseline adoption			X	X
Controls				X
Observations	763	763	723	723
R-squared	0.008	0.074	0.075	0.105
Mean of dependent variable (control)	0.722	0.722	0.722	0.722
Lee (2009) Attrition Bounds				
Lower Bound:	0.083			
	(0.043)			
	[0.011, 0.154]			
Upper Bound:	0.142			
	(0.056)			
	[0.048, 0.234]			
Imbens and Manski (2004) CS	[0.025, 0.217]			
Behaghel et al. (2015) Attrition Bounds				
Lower Bound:	0.086			
	(0.041)			
	[0.005, 0.168]			
Upper Bound:	0.128			
	(0.041)			
	[0.047, 0.209]			

Note: Market district is the randomization strata. Observations are at the customer level. Robust standard errors reported in parenthesis. 90% confidence intervals (CI) and confidence sets (CS) are reported in brackets. Reported confidence CIs for wild bootstrap and permutation tests derived from running 1000 replications in each case. Permutation test (for randomization inference) clustered at the (local) market level. L-S-X MHT Corr (2019) refers to the multiple hypothesis testing procedure presented in List, Shaikh and Xu (2019) for outcomes family (services usage; savings). 0-1 indicators for baseline migration motives (desire to migrate, plan to migrate, and permanent migration) used as predictors of attrition to tighten attrition bounds.

E Further Results: Belief Updates and Heterogeneity

Table 21: **Customers subjective belief about vendor responsibility (non-misconduct) increase at endline**

BELIEFS $\hat{\pi}$			
PANEL A			
Treatment: Information Assignment ($\gamma = \beta$)	0.065	0.069	0.067
	(0.035)	(0.035)	(0.035)
	[0.006, 0.122]	[0.011, 0.127]	[0.009, 0.126]
Baseline belief about $\mathbf{1}$(Non-misconduct=Yes)		X	X
Controls			X
Observations	943	941	941
Mean of dependent variable (control)	0.223	0.223	0.223
PANEL B			
Price Transparency ($\gamma \equiv \beta_1$)	0.064	0.067	0.066
	(0.042)	(0.042)	(0.042)
	[-0.005, 0.133]	[-0.001, 0.137]	[-0.004, 0.137]
Monitor and Report ($\gamma \equiv \beta_2$)	0.010	0.010	0.002
	(0.041)	(0.041)	(0.041)
	[-0.060, 0.076]	[-0.062, 0.073]	[-0.066, 0.069]
Joint program: PT + MR ($\gamma \equiv \delta$)	0.117	0.133	0.132
	(0.043)	(0.042)	(0.043)
	[0.047, 0.188]	[0.062, 0.203]	[0.061, 0.203]
Baseline belief about $\mathbf{1}$(Non-misconduct=Yes)		X	X
Controls			X
Observations	943	941	941
Mean of dependent variable (control)		0.223	
p-value (test: $\beta_1 = \delta$)	0.191	0.109	0.109
p-value (test: $\beta_2 = \delta$)	0.006	0.001	0.001
p-value (test: $\beta_1 = \beta_2$)	0.147	0.109	0.101
p-value (test: $\beta_1 + \beta_2 = \delta$)	0.435	0.301	0.272

Note: $\mathbf{1}(.)$ is a logical indicator that equals 1 if the argument in the parenthesis is true, 0 otherwise. In each market, we compute the baseline outcome as the share of experimental customers that indicate no experiences of misconduct (i.e., non-misconduct). Includes baseline outcomes, and additional controls. Observations are at the customer level. Robust standard errors reported in parenthesis. 90% confidence intervals (CI) are reported in brackets. Panel A reports the meta estimate of treatment effects, while panel B shows the effects separately for each information program.

Table 22: **Customers belief update: effect of information sets on customers correct guess of vendor irresponsibility (or misconduct)**

$\mathbf{1}(\text{GUESS}_i\text{=AUDIT TRIAL}_j)$			
PANEL A			
Treatment: Information Assignment ($\gamma \equiv \beta$)	0.084	0.085	0.088
	(0.030)	(0.030)	(0.029)
	[0.033, 0.134]	[0.034, 0.136]	[0.039, 0.137]
Baseline belief about $\mathbf{1}$(Misconduct=Yes) $\approx \Pr(t_i)_{base,j}$		X	X
Controls			X
Observations	763	763	763
Mean of dependent variable (control)	0.101	0.101	0.101
PANEL B			
Price Transparency ($\gamma \equiv \beta_1$)	0.067	0.068	0.072
	(0.037)	(0.037)	(0.036)
	[0.007, 0.128]	[0.007, 0.129]	[0.012, 0.131]
Monitor and Report ($\gamma \equiv \beta_2$)	0.012	0.013	0.013
	(0.036)	(0.036)	(0.035)
	[-0.047, 0.071]	[-0.046, 0.072]	[-0.044, 0.071]
Joint program: PT + MR ($\gamma \equiv \delta$)	0.164	0.167	0.173
	(0.039)	(0.040)	(0.039)
	[0.098, 0.229]	[0.101, 0.233]	[0.109, 0.238]
Baseline belief about $\mathbf{1}$(Misconduct=Yes) $\approx \Pr(t_i)_{base,j}$		X	X
Controls			X
Observations	763	763	763
Mean of dependent variable (control)	0.101	0.101	0.101
p-value (test: $\beta_1 = \delta$)	0.011	0.010	0.009
p-value (test: $\beta_2 = \delta$)	0.000	0.000	0.000
p-value (test: $\beta_1 = \beta_2$)	0.112	0.114	0.093
p-value (test: $\beta_1 + \beta_2 = \delta$)	0.111	0.105	0.092

Note: $\mathbf{1}(.)$ is a logical indicator that equals 1 if the argument in the parenthesis is true, 0 otherwise. Includes randomization strata (market district) dummies, baseline outcomes, and additional controls. Observations are at the customer level. Robust standard errors reported in parenthesis. 90% confidence intervals (CI) are reported in brackets. Panel A reports the meta estimate of treatment effects, while panel B shows the effects separately for each information program.

Table 23: **Corroborative evidence for reputation**

	1(Misconduct=Yes)		Amount-Misconduct, GHS	
Treatment: Information Assignment (β)	0.197	0.400	0.318	1.221
	(0.177)	(0.260)	(0.626)	(0.884)
	[-0.096, 0.492]	[-0.032, 0.834]	[-0.722, 1.358]	[-0.248, 2.691]
x Bundle	-0.318	-0.346	-0.722	-0.903
	(0.124)	(0.169)	(0.369)	(0.426)
	[-0.525, -0.111]	[-0.627, -0.064]	[-1.336, -0.108]	[-1.611, -0.195]
x Married	-0.159	-0.075	-0.312	-0.291
	(0.152)	(0.233)	(0.604)	(0.870)
	[-0.412, 0.093]	[-0.464, 0.312]	[-1.317, 0.692]	[-1.737, 1.155]
x Incorrect (knowledge test)	0.148	-0.010	-0.388	-0.190
	(0.142)	(0.252)	(0.562)	(0.988)
	[-0.087, 0.384]	[-0.429, 0.409]	[-1.323, 0.547]	[-1.832, 1.452]
x Incorrect (knowledge test) and Female	-0.291	-0.243	-0.114	-0.121
	(0.131)	(0.111)	(0.362)	(0.330)
	[-0.510, -0.072]	[-0.428, -0.058]	[-0.716, 0.488]	[-0.670, 0.427]
x Inexperienced	-0.309	-0.480	-0.666	-1.474
	(0.188)	(0.254)	(0.668)	(0.864)
	[-0.623, 0.003]	[-0.903, -0.057]	[-1.778, 0.444]	[-2.91, -0.037]
Bundle	0.235	0.209	0.612	0.593
	(0.122)	(0.167)	(0.373)	(0.399)
	[0.030, 0.439]	[-0.068, 0.488]	[-0.009, 1.233]	[-0.069, 1.257]
Married	0.173	0.143	0.296	0.465
	(0.144)	(0.220)	(0.579)	(0.825)
	[-0.066, 0.413]	[-0.222, 0.510]	[-0.667, 1.259]	[-0.907, 1.837]
Incorrect (knowledge test)	-0.158	-0.030	0.359	0.133
	(0.136)	(0.248)	(0.558)	(0.952)
	[-0.385, 0.067]	[-0.443, 0.382]	[-0.569, 1.288]	[-1.449, 1.716]
Incorrect (knowledge test) and Female	0.197	0.128	-0.157	-0.041
	(0.114)	(0.103)	(0.331)	(0.298)
	[0.007, 0.387]	[-0.042, 0.300]	[-0.707, 0.393]	[-0.538, 0.454]
Inexperienced	0.251	0.364	0.712	1.288
	(0.186)	(0.256)	(0.638)	(0.860)
	[-0.058, 0.561]	[-0.061, 0.791]	[-0.348, 1.773]	[-0.141, 2.718]
Baseline misconduct	X	X	X	X
Controls	X	X	X	X
Market District F.E.	X		X	
Market District $\times$ Transaction $\times$ Date F.E.		X		X
Observations	335	335	335	335
Mean of dependent variable (control)	0.294	0.294	0.778	0.778

Note: **1**(.) is a logical indicator that equals 1 if the argument in the parenthesis is true, 0 otherwise. Includes randomization strata (market district) dummies, baseline outcomes, and additional controls. Observations are at the vendor x transaction x date level. Clustered standard errors (at local market level) reported in parenthesis. 90% confidence intervals (CI) are reported in brackets. Results similar to post-double-selection LASSO estimates clustered at the (local) market level and to alternative inference procedures (Wild cluster bootstrap and permutation test clustered at the market level). Bundle is a 0-1 indicator for whether the vendor bundles M-Money with other business services at the banking point. Married is a 0-1 indicator for whether the subject is married. Incorrect (knowledge test) is a 0-1 indicator for whether the vendor was incorrect in answering the baseline knowledge tests about the true transaction prices. Inexperienced is a 0-1 indicator for whether the subject have been in the M-Money business for less that 12 months.

F Anti-Misconduct Information Programs – Exhibits

F.1 FIRST: VISIT NEARBY CUSTOMERS

PREAMBLE: Greetings Madam/ Sir... My name is....
Please recall we visited your unit in February 2019 to do a survey of (the M-Money business) to find out (how customers, like you, understand the business of M-Money and other services their centers provide). Today, we have come to provide additional education about M-Money for research and to help make the market better and understandable. You may call the research team anytime *if* in any doubt (Phone: XXXXXXXXXX) (omitted to preserve privacy).

F.1.1 T1 - PRICE TRANSPARENCY, PT

Our message is simple. We want to remind you that you should:

- Make sure to ask for official tariff sheets when transacting: e.g., opening new Wallet, OTC, sending. Simply ask.
- When opening a new Wallet don't pay fees – deposit should be credited to your account, check it right away.
- Example of common charges: (i) Pay 0.5GHC if putting 50GHC on someone's account; (ii) 1.6GHC if putting 160GHC on someone's account; 10GHC if putting 1100GHC on someone's account; (iv) it's free to put any money on your own Wallet.
- Research Officer, ask customers: (1) Repeat information provided. (2) Please how would you rate the usefulness of this information we provided for your financial decisions on a 5-point scale? [1=Not useful, 2=Quite useful, 3=Useful, 4=Very useful, 5=Extremely useful]. (3) Leave a hard (paper) copy of this information with subject.

F.1.2 T2 - (MARKET) MONITOR AND REPORT, MR

Our message is simple. We want to remind you that *if* you:

- Suspect any discrepancy or glitches as you make any M-Money transactions, you should call MTN fraud department on NUMBER (Toll-Free number: 100) to report it, right away.

- There is an MTN fraud department; **free** to call. They always help.
- Research Officer, ask customers: (1) Repeat information provided. (2) Please how would you rate the usefulness of this information we provided for your financial decisions on a 5-point scale? [1=Not useful, 2=Quite useful, 3=Useful, 4=Very useful, 5=Extremely useful]. (3) Leave a hard (paper) copy of this information with subject.

F.1.3 T3 - PT+MR

We have two main messages:

- First, we want to remind you that you should: Make sure to ask for official tariff sheets when transacting: e.g., opening new Wallet, OTC, sending. When opening a new Wallet don't pay fees – deposit should be credited to your account, check it right away. Example of common charges: (i) Pay 0.5GHC if putting 50GHC on someone's account; (ii) 1.6GHC if putting 160GHC on someone's account; 10GHC if putting 1100GHC on someone's account; (iv) it's free to put any money on your own Wallet.
- Second, we want to remind you that *if* you: Suspect any discrepancy or glitches as you make any M-Money transactions, you should call MTN fraud department on NUMBER (Toll-Free number: 100) to report it, right away. There is an MTN fraud department; **free** to call. They always help.
- Research Officer, ask customers: (1) Repeat information provided. (2) Please how would you rate the usefulness of this information we provided for your financial decisions on a 5-point scale? [1=Not useful, 2=Quite useful, 3=Useful, 4=Very useful, 5=Extremely useful]. (3) Leave a hard (paper) copy of this information with subject.

F.2 SECOND: VISIT REPRESENTATIVE OR SELECT VENDOR

PREAMBLE: Greetings Madam/ Sir…My name is….
Please recall we visited your unit in February 2019 to do a survey of (the M-Money business) to find out (how merchants, like you, understand the business of M-Money and other services that your centers provide). Today, we have come to provide additional education about M-Money for research and to help make the market better and understandable. You may call the research team anytime *if* in any doubt (Phone: XXXXXXXXXX) (omitted to preserve privacy).
[RESEARCH OFFICER: LET'S BLUFF ABOUT INTERVENTIONS GIVEN TO CUSTOMERS]: We have educated "nearby" customers in this locality about M-Money (since many of them don't understand M-Money's workings well) that:

F.2.1 T1 - PRICE TRANSPARENCY, PT

They should:

- Make sure to ask for official tariff sheets when transacting: e.g., opening new Wallet, OTC, sending.
- When opening a new Wallet don't pay fees – deposit should be credited to their account, check it right away
- Example of common charges: (i) Pay 0.5GHC if putting 50GHC on someone's account; (ii) 1.6GHC if putting 160GHC on someone's account; 10GHC if putting 1100GHC on someone's account; (iv) it's free to put any money on their own Wallet.
- Research Officer, ask vendors: (1) Please how would you rate the usefulness of this information we provided for your financial business on a 5-point scale? [1=Not useful, 2=Quite useful, 3=Useful, 4=Very useful, 5=Extremely useful]. (2) Leave a hard (paper) copy of this information with subject.

F.2.2 T2 - (MARKET) MONITOR AND REPORT, MR

- *If* they: Suspect any discrepancy or glitches as they make any M-Money transactions, they should call MTN fraud department on NUMBER (Toll-Free number: 100) to report it, right away.
- There is an MTN fraud department; **free** to call. They always help.
- Research Officer, ask vendors: (1) Please how would you rate the usefulness of this information we provided for your financial business on a 5-point scale? [1=Not useful, 2=Quite useful, 3=Useful, 4=Very useful, 5=Extremely useful]. (2) Leave a hard (paper) copy of this information with subject.

F.2.3 T3 - PT+MR

Two main messages:

- First, they should: Make sure to ask for official tariff sheets when transacting: e.g., opening new Wallet, OTC, sending. When opening a new Wallet don't pay fees – deposit should be credited to their account, check it right away. Example of common charges: (i) Pay 0.5GHC if putting 50GHC on someone's account; (ii) 1.6GHC if putting 160GHC on someone's account; 10GHC if putting 1100GHC on someone's account; (iv) it's free to put any money on their own Wallet.

- Second, *if* they: Suspect any discrepancy or glitches as they make any M-Money transactions, they should call MTN fraud department on NUMBER (Toll-Free number: 100) to report it, right away. There is an MTN fraud department; **free** to call. They always help.

- Research Officer, ask vendors: (1) Please how would you rate the usefulness of this information we provided for your financial business on a 5-point scale? [1=Not useful, 2=Quite useful, 3=Useful, 4=Very useful, 5=Extremely useful]. (2) Leave a hard (paper) copy of this information with subject.

G Vendor Banking Points – Photos

Figure 10: **Vendor banking points**

Note–Providers: MTN Mobile Money, AirtelTigo Money, Voda Cash