ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/345110909

Misconduct and Reputation under Imperfect Information

Preprint · January 2022

CITATIONS
0

READS
284

1 author:

Francis Annan
Georgia State University

27 PUBLICATIONS 166 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project Climate Change, Epidemic Disease and Human Capital Outcomes View project

All content following this page was uploaded by Francis Annan on 18 January 2022.

The user has requested enhancement of the downloaded file.

Misconduct and Reputation under Imperfect Information

Francis Annan[♣]

Georgia State University

AUGUST 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 act 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 lead to a large reduction in misconduct (-21 pp = -72%) and as a result, an increase in overall market activity, firm sales revenue, and consumer welfare. We develop a simple sanctioning and moral hazard framework between vendors and consumers that shows the treatment effect is due to a combination of more accurate consumer beliefs about misconduct and increased reputation concerns for vendors. 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* (**D83**), *vertical markets and reputation* (**L14**, Z13), *household finance* (D14, **O12**), *consumer protection* (**D18**), *entrepreneurship and firm behavior* (**L26**, **M13**)

♣ 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, Pascaline Dupas, Chiara Farronato, Mariassunta Giannetti (discussant), Xavier Gine, Paul Goldsmith-Pinkham, Brett Green, Glenn Harrison, Sean Higgins, Erzo F.P. Luttmer, Dean Karlan, Supreet Kaur, Isaac Mbiti, Matt Shum, Sandip Sukhtankar, Tavneet Suri (discussant), Catherine Tucker, Iwan Barankay, and several seminar participants at Household Finance Seminar, University of Virginia Economics, University of Pennsylvania-Wharton, Georgia State University, Michigan State University, PEDL 2020 Conference on Firms in Low-Income Countries, CSAE Oxford 2021, BREAD 2021 Conference on Behavioral Economics and Development, NBER 2021 Economics of Digitization Meeting, BREAD 2021 Conference on the Economics of Africa, and NBER 2021 Summer Institute. We thank the NBER Development Economics Program Fall 2020 meeting participants for comments on an earlier draft of this paper. Field support from Samuel Adotevi and Kwamena Arkafra (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 (Bai 2019), as it is difficult to demonstrate and/or reveal the quality of services between businesses and consumers.

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 reflect shrouded prices (Gabaix and Laibson 2006; Brown, Hossain and Morgan 2010), 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; Higgins 2020) and overall firm growth (Jensen and Miller 2018). In emerging and developing financial market contexts, misconduct is likely to be particularly significant as consumers 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 act as a discipline against seller or vendor misconduct? We use a field experiment to address these important questions, examining the impact of 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 a large-scale 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 (upwardly-biased consumer beliefs) about misconduct. These features, which we show at baseline, make M-Money an ideal setting to study misconduct and reputation under imperfect information. Indeed, this form of seller misconduct in payment markets can be found in many other countries. In recent cross-country consumer protection surveys of digital finance users, Blackmon, Mazer, and Warren (2021) document significant rates of vendor misconduct against consumers in Kenya (3%), Uganda (32%), and Nigeria (42%), with corroborative evidence of poor consumer knowledge about official prices and high consumer mistrust.

We construct a unique census of local markets (local communities or villages) for M-Money between February–March 2019 as detailed vendor $\times$ customer data is unavailable and then perform our experiment by randomly assigning these markets to three anti-misconduct information programs about either price transparency (PT), monitoring and reporting (MR), or both (PT+MR, their interaction). In the PT treatment, consumers receive relevant information and training about official transaction charges. In the MR treatment, consumers are 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 are informed that customers have received such information earlier and the same information sets are then given to the vendors, making our interventions two-sided. Thus, the interventions empower consumers with technologies to enforce market vendors' trustworthy behavior relying on social sanctions and/or punishment.

M-Money provides financial services which are delivered on digital mobile networks to potential consumers. M-Money market vendors are small business outlets that provide account opening, cash-in, and cash-out services (Human ATMs), earn transactional commissions as their profit, and exchange cash for so-called e-money. A typical local community is made up of about three 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, so vendors are not allowed to marginalize. We use this feature to cleanly define misconduct as all transactions at the vendor point that are overcharged, which can be derived by comparing observed transaction charges to provider-approved prices (Egan, Matvos, and Seru 2019b; Annan 2020).

The experiment involves 130 independent local markets in 130 different localities across nine 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 lasts over twenty-two weeks. We track several outcomes at endline: household or consumer usage of M-Money; shocks exposure and mitigation (experiences of household shocks that consumers could not financially remedy); poverty; and collected vendor sales revenue records of M-Money and other goods to examine the supply side effects and directly validate the household transaction data. For each locality, while the intervention is applied to one random vendor and their nearby customers, we track additional non-treated vendors at endline to examine spillover effects.

We propose an innovative audit study to measure vendor misconduct: trained auditors visit 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 (Ortmann and Hertwig 2002; Kessler, Low, and Sullivan 2019). Misconduct in markets remains a poorly understood issue due to the empirical difficulties in measuring it objectively. Here, we develop a procedure to cleanly measure misconduct connected to increased transaction costs and shrouded prices. 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 reduces vendor misconduct dramatically. Overall, the incidence of vendor misconduct decreases by -21 pp = -72%, while the severity of misconduct decreases by -GHS0.68 (-\$0.14) = -86%. With a control mean of GHS0.78, the latter means the intervention leads the total fee (official charge + misconduct) to fall from about 1.80% to about 1.10%, implying about 40% reduction of typical M-Money transaction fees. The joint intervention shows an economically larger reduction in market vendors' misconduct; however, the PT and MR programs also have meaningful negative impacts on misconduct. Next, we find significant spillover effects: non-treated vendors located in treated villages reduce 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 disclosure sets impacts various real consumer and business outcomes.

Consumer (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 are about 7.6% less likely to experience shocks that they could not 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 complement one another.

Business (vendor) transactional sales revenue increases. Overall, the information programs significantly increase vendors' total sales volume (+36%). 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 (or 36% increase in total vendor sales) in response to a 40% total fee (official charge + misconduct) reduction is reasonable; it is an elasticity of about 1.0 (or 0.9). In additional tests, we find extended large positive impacts on vendors' non-M-Money business transactions, suggesting positive spillover effects of the information program on overall local market activities. We do not find evidence for an impact on the number of customers or exits of businesses from the local market.

Next, we present evidence on consumer beliefs about seller misconduct and reputational concerns for vendors. The information programs cause consumers' perception of honest vendor behavior to increase (+7.0 pp = +30% overall), and importantly, make such beliefs more positively correlated (+27 pp = +51%) with the objective audit measure of misconduct (accurate and updated beliefs). The effects appear to be much larger for the joint program. Thus, when customers thought well of the vendors and trusted that they would not be cheated, they increase their demand for M-Money and other non-M-Money business items at the vendors premises. Vendors are also reinforced to reduce their misconduct behavior since consumers now have the technologies to enforce vendors' trustworthy behavior using the channels activated – social sanctions and/or punishment.

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

What explains the estimated impacts of anti-misconduct information? Our underlying hypothesis is that of reputational concerns, and our evidence supports this claim. We set up a simple sanctioning and moral hazard framework to guide our information programs and illustrate the reputation interpretation of the results. A vendor's reputation is defined by consumers' perceptions about their tendency to commit misconduct behavior (Macchiavello and Morjaria 2015). When (potentially uninformed) customers are provided with symmetric market information about vendor misconduct, they are able to infer irresponsible vendors, directly report misconduct behavior, and choose to engage with vendors who do not engage in misconduct, generating reputational revenue. This informed consumer base raises vendor concern for reputation. If vendors care about consumers' negative or positive perceptions, then misconduct will fall, with impacts on the various outcomes. 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; measuring reputational concerns based on how customers are able to infer vendor misconduct; and measuring how vendor recognition of customer judgment led to a dramatic reduction in vendor misconduct behavior.

I.1 Related Literature

We make three main distinct contributions to the literature. First, we add to 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 that the use of local sanctions via reputation-building may promote rural financial institutions and development in low-income settings (see Munshi 2014 for a review).

Second, we contribute to the literature on information, business behavior, and growth in developing countries. Previous studies have emphasized several barriers to business growth, including managerial constraints (Bloom et al. 2013), network and interfirm relations (Cai and Szeidl 2017), lack of capital (De Mel, McKenzie, and Woodruff 2008), market access or lack thereof (Atkin, Khandelwal, and Osman 2017), and information (Jensen and Miller 2018; Bai 2019). Here, in addition to imperfect information, we emphasize miscalibrated consumer beliefs about seller misconduct and vertical market structure as other potential barriers to business performance and behavior. Our findings shed light on why small to medium firms may not grow when they fail to provide quality or honest services (by engaging in misconduct), as honesty is under-rewarded in market contexts dominated by imperfect information.

Third, we add to the literature on information disclosure, household finance, and FinTech adoption. There is much existing research on the consumer effects of FinTech (Jack and Suri 2014; Suri and Jack 2016), but there is almost no work about supply side behavior (Higgins 2020 and references therein). Here, we emphasize seller misconduct as a key barrier to both sides of the market and show that reducing it via information disclosure has meaningful impacts on consumers, sellers, and businesses. We show that disclosure – transparency and monitoring – is beneficial to businesses and improves sales revenue as in Brown, Hossain and Morgan (2010) for retail sellers on Yahoo and eBay, specifically in a market setting with low transaction tariffs. Moreover, we document misconduct in payment markets which is an open—and high priority— area of research, particularly in developing countries, where consumers lack experience with FinTech (Garz et al. 2021) and higher transaction fees can act as a barrier to the adoption of payment services (Higgins 2020) and reduce risk sharing across households (Jack and Suri 2014). Our study is the first, to our knowledge, to provide quantitative estimates on both seller misconduct in payment or digital financial markets and the value of anti-misconduct information programs, particularly in environments where M-

Money has the ability to reduce poverty and meaningfully improve the welfare of consumers.

From a policy perspective, our results highlight how the provision of low-cost two-sided information might influence vendor misconduct and consumer trust, 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 payment and digital financial markets.

II Research Setting

II.1 Mobile Money

Market Structure: The market for M-Money comprises (i) service providers, (ii) vendors, and (iii) customers. In Ghana, there are four providers (MTN M-Money, Vodafone Voda-Cash, AirtelTigo Money, and GCB Ltd.'s G-Money), with MTN representing about 90% share of this market. Providers are joint partnerships between mobile network operators (MNOs) and commercial banks. Market vendors (or sellers) correspond to outlets, shops, premises, or local banking channels where M-Money transactions can be carried out on behalf of the providers.

Vendors register new accounts (also called "wallets") for customers and act as cash-in (deposits, transfers) and cash-out (withdrawals) transaction points for customers (i.e., Human ATMs). Vendors can freely enter and exit the market. To start the business of M-Money, vendors need to have the required documentation and meet certain structural and monetary requirements. Vendors should have a permanent space from which to operate and a minimum startup capital of GHS4000 ($US781.25) (MTN Mobile Money 2021)[1], which we observe in practice can be relaxed depending on the environment. All vendors are required to receive official business training about the tariffs, commissions, and other services, and generically

[1]MTN Mobile Money 2021: https://mtn.com.gh/momo/agent/

earn transactional commissions on sales revenue as their profit. In comparison, customers receive little to no information about M-Money's transaction tariffs and services when they sign up. The tariffs on transactions at vendor points are *ex-ante* set by providers, so market vendors are not allowed to marginalize. Thus, the M-Money setup has a vertical market structure: service providers at the upstream set up vendors at the downstream, who work for them and earn commissions on sales. This vertical structure underlies the improvement in business outcomes and baseline vendor misconduct behavior, as we discuss later in the Results section.

The introduction and significant penetration of digital mobile telecommunications has provided a cheap infrastructure to make M-Money services accessible even to 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), making M-Money a competitive financial option. Evidence suggests that M-Money has the potential to reduce poverty and improve the welfare of consumers in Sub-Saharan Africa and Asia through several channels (see e.g., Jack and Suri 2014; Suri and Jack 2016). M-Money is an important market but could be constrained by market misconduct that shrouds prices and increase transaction costs. Providers at the upstream have limited oversight into the behavior of downstream vendors and consumers in poor and low-income environments are poorly informed.

Misconduct Prevalence: Similar to other banking and financial services, the business of M-Money likely faces fraud and misconduct, which could take different forms. Indeed, vendor misconduct is widespread. Recent surveys from Innovations for Poverty Action (IPA) compares market misconduct (overcharging of services) in Kenya, Uganda, and Nigeria (Blackmon, Mazer, and Warren 2021): 33%, 42%, and 3% of consumers reported vendor overcharging in Uganda, Nigeria, and Kenya, respectively. In 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 increasing ultimate consumer confidence in digital financial services. In Ghana, the MNOs and their commercial partners have been charged to build more risk- and fraud-resilient financial infrastructures.[2] Our present study is designed to carefully understand misconduct at vendor points (Figure G.1 in Appendix G), the effect of social sanctions and/or punishment, and evaluate its potential market-wide effects. We do this in a rural context where the business of M-Money could have larger impacts, if well designed.

II.2 Market Census

Detailed vendor $\times$ customer data on M-Money is unavailable. So, between February and March 2019, we carry out a unique census of the market for M-Money in Eastern Ghana, spanning nine districts. Districts are made up of sub-administrative units called "localities" or villages. Eastern Ghana was chosen for two attractive features: (i) it covers an expansive number of villages, with potential M-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 successfully documents the universe of all vendor points (both formal and informal) and other surrounding households (within a five-house radius around a given vendor) across 130 localities. This yields a total of 333 vendors and 1,921 customers or households. 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 $\times$ the set of all nearby customers.

II.3 Market Facts

Our baseline census solicits information from all market participants: both vendors and customers. We ask about their basic demographics, poverty and assets, and detailed market records on M-Money and non-M-Money services, including general to specific knowledge

[2] "*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 Through Mobile Wallet Transactions, 2015.

about M-Money transactions. We also obtain additional household information on personal finance, debts, savings, shocks and investments from customers. Here, we will focus on data that are relevant to our study of market impacts of misconduct. Detailed summaries, and other patterns about the market are available in Annan (2020) and upon request.

Table C.1 shows summary statistics for the market. To facilitate comparisons between both sides of the market, relevant statistics for vendors and customers are displayed next to it each other. Female vendorship is 39%, meaning that these local markets are disproportionately made up of male vendors. Of potential customers, 62% are females, and customers are more likely to be self-employed, married, and older relative to vendors. Strikingly, approximately half the vendorship had received formal training about the market for M-Money before joining the business. The overwhelming majority (90% [SD=0.29]) of customers, as well as their networks of close family and friends, have registered for a M-Money account, indicating that it is likely a popular financial technology.

We turn next to specific features of the market. With an average experience of two years doing M-Money business, a vast majority (75% [SD=0.43]) of vendors operate as a bundled store, bundling M-Money with other services.[3] The average daily sales per vendor for M-Money is about GHS2,260 (US$442). With a sales commission of 1%, the average vendor will earn a daily profit of around GHS23. Thus, most of these vendors operate relatively small to medium size enterprises. The majority of households or customers use M-Money services rather 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, and relative inaccessibility and distance of other nearby services: we estimate an average distance of approximately 61 meters to the closest M-Money vendor site versus about 383 meters to the nearest post-office.

[3]We identified bundled services including 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, electronics and accessories, photocopying and typesetting, educational/online results checking, and electric prepaid credit, among others. Baseline sales revenue from these non-M-Money services represents about 7% of the sales revenue from M-Money (Table C.1).

II.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 of information programs and their effects and interpretations. In the Framework or Theory section, we show evidence that our (baseline) setting reflects a market environment where (i) consumers are objectively uninformed (less sophisticated than vendors); (ii) vendor misconduct incidence is high (vendors overcharge over 22% of transactions objectively measured using administrative audit exercises, versus 59% measured using survey elicitations); (iii) consumers have upwardly-biased beliefs about the level of vendor misconduct and mistrust vendor-based transactions; and (iv) vendors value the positive returns associated with good reputation, but find it difficult to establish good reputation and build reputational capital. Together, we show that information frictions matter and that there is room to build reputation.

III Experiment: Design

III.1 Intervention and Timetable

We evaluate the impacts on both customers and vendors of different information sets that reduce market misconduct. 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 use of 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 treatment 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: Monitoring and Reporting (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 three consecutive times over a two month period (once per every two-three weeks) to first deliver and then repeat the information programs to subjects. We conclude visits by asking subjects to summarize the information they received and giving them hard copies of the treatment program. More uniquely, we ensure that vendors are equally aware of the interventions by communicating the same information set to the vendors right after seeding the information with nearby households, yielding a two-sided information design. Together, our treatment programs aim to reduce potential information frictions and increase the social cost of vendor misconduct.

To roughly gauge the likely significance of the information programs, the recipients are *ex-ante* asked to rate the usefulness of the information we provide for their financial decision-making (i.e., customers) and businesses (i.e., vendors) on a five-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.[4] Program I is a popular consumer protection policy instrument in practice. By benchmarking this with programs II and III, 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 shows the timetable of all field activities.

III.2 Data Collected

We gather information from multiple sources and rounds of data collection (see Table 1): (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, 33-weeks administrative audit study, and market-level transaction data from the largest service provider, which we call an endline.

III.2.1 Administrative Audit Data

To objectively measure true misconduct, we develop an audit study procedure where auditors (experimental customers) are given cash to make actual transactions on M-Money at vendor points, 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): cash-

[4] 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 two to three months after the provision of the information programs. This suggests that subjects are willing to pay for our information programs, perhaps because they find the information credible. In addition, this suggests that subjects' rating of the usefulness of the information provided is less likely affected by potential experimenter demand (pleasing) effects (de Quidt, Haushofer, and Roth 2018).

in, cash-out, and account opening.[5] As mentioned, 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,[6] who are trained to follow a consistent approach to interacting with vendors, including using uniform language, a short and transparent transaction script (see Appendix H for details).

We implement several quality controls for the transactional exercises. First, we set up a computer-adaptive data collection platform (called data HQ), which allows us to track and verify the data in real time and space. Right after every visit, auditors complete a brief questionnaire about the transaction using their Tablets (see Table H.1 in Appendix H) and synchronize the data to our data HQ for immediate access and verification. The GPS coordinates of all transactions are traceable. Second, we pilot the proposed audit approach in February 2017 (as noted in the Market Census section), which yields patterns of misconduct similar to the main experiment. Third, we include transaction types that are either easy or difficult for the seller to overcharge, finding consistent evidence of higher misconduct for the easy to overcharge transactions, as discussed below. Together, these quality controls strengthen our proposed approach by 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).

We define misconduct to entail transactions that are over-charged when compared to the provider-approved tariff rates (as in Egan, Matvos, and Seru 2019b; Annan 2020). Table

[5] Importantly, these include transactions that inherently mimic the fee structure. The typical fee/ tariff structure set by providers is piecewise linear: GHS0.50 for all transaction values $\leq$GHS50, 1% of the value for transaction values between GHS50 and GHS1,000, and GHS10 for all transaction values $\geq$1,000. Similarly, the cost of a new SIM card is GHS2.0 and registering for a new M-Money wallet is free but requires an initial minimum account deposit of GHS5.0. Appendix H and Table C.2 contain details.

[6] A potential concern is that vendors cheat strangers (like the auditors) but not local repeat customers that they know. This is not a major concern here for several reasons. First, it might be more risky to cheat strangers because they might be more informed, which is especially true in this market context with much imperfect information. This reduces the possibility that vendors systematically cheat strangers. Second, in our market environment, we estimate that a very large share of 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 asked to indicate where they last transacted 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% were between unrelated participants. Third, we vary the type of transactions, and auditors conduct multiple or repeat transactions at a vendor point to mimic repeat customers.

C.2 and Figure C.3 in Appendix C show baseline results across the various transactions. We estimate that 22% of transactions are overcharged (which reflects the incidence of misconduct), which results in GHS3.3 (= 82% of the official tariffs) overpaid to the vendor (which reflects the severity of misconduct). There is heterogeneity in misconduct levels across the different types of transactions or groups. Misconduct is concentrated in over-the-counter (OTC) transactions, which involve little to no automation or active verification from the customer, and are thus more vulnerable to vendor misconduct. Non-OTC transactions (e.g., opening a new account) are also overcharged, but at a much lower rate. This is reassuring and alleviates several potential concerns, including the concern that auditors might be over- or under- measuring misconduct.

III.2.2 Market Survey Data

We measure several repeated outcomes at different stages of the study. For customers, we restrict attention to four relevant outcomes: (i) adoption and usage of money services: we ask whether households use money services, and if so, the transaction amount involved per week; (ii) savings on M-Money: we ask whether households saved on their money wallets within the month; (iii) specific shock experiences (such as health, revenue, and household expenditures) and risk mitigation: we ask whether customers experienced unexpected shocks that they could not financially remedy, providing an objective proxy for insurance (Dupas and Robinson 2013; Breza and Chandrasekhar 2019); and (iv) poverty. Since our study focuses on M-Money in low-income and poor environments, we field questions that allow us to directly examine poverty. We adapt a recently developed measure of poverty called the "Simple Poverty Scorecard" that is rigorous, inexpensive, simple, and transparent (for details, see Schreiner 2015).

For vendors, we measure sales revenue by soliciting transaction records for their M-Money business and non-M-Money services (if the vendor operates a bundled store). With these combined measurements, 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 household money transactions to (positively) correlate with increases in nearby vendor sales revenue, all else equal. See Appendix I for definitions of relevant select variables.

III.3 Treatment Assignment

We use a 2×2 factorial design, randomizing the total 130 randomly selected markets (as defined below) into four experimental anti-misconduct programs: PT-only (31 markets ≡ 31 select vendors × 272 nearby customers); MR-only (32 markets ≡ 32 select vendors × 257 nearby customers); joint program (35 markets ≡ 35 select vendors × 276 nearby customers); and control program (32 markets ≡ 32 select vendors × 185 nearby customers). We stratify based on districts, and all misfits are resolved and randomly assigned.

III.4 Balance and Validity of Design

III.4.1 Balance I

We focus our study on randomly selected markets 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 their nearby customers per locality for our study. We call this combination (selected vendor × nearby households) a randomly selected market. Sample representativeness requires that being a randomly selected 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 randomly selected in the pre-intervention period. We consider a number of different relevant outcomes, and show that neither side of the market demonstrates any observable differences across the two groups. Tables B.1 and B.2 report the results, where we find no difference across markets selected and those not selected.

III.4.2 Balance II

We base our treatment analysis on a comparison of randomly selected 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 reporting-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 receives an information treatment, 0 otherwise. We consider the various treatments separately and together (i.e., pooled) for a number of different outcomes, and show that neither side of the market demonstrates any observable differences across the two groups. Tables B.3 and B.4 report the results, providing strong evidence in favor of balance with no difference across subjects i (households or vendors) in assigned (treated) and non-assigned (control) markets.

III.4.3 Attrition

Our randomization is based on randomly selected markets and draws on the baseline market census. Table B.5 displays the breakdown of response rates and attrition between baseline and endline. Here, attrition may be linked to subjects' non-response or migration to outside the locality, and/or our inability to reach participants because their phone numbers are either inactive or out of network coverage area. To maximize response rates at endline, trained field officers conduct multiple phone calls (see Figure C.2) 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 a high degree of migration and operator turnovers. Attrition is non-differential. For our endline audit transactional exercises, 129 out of the 130 randomly selected vendors were reached, implying an attrition rate of just 0.8%.[7] In our empirical estimations, we evaluate and formally show robustness to attrition by treatment status.

IV Experiment: Results

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

IV.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)[8] 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 primarily to increase precision and to control for potential confounds (if any). For the pooled effects, $\mathbf{I}_{vd}$ is a 0-1 indicator for whether a locality received any of the information programs, and thus β captures the (pooled) 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

[7] The interventions did not lead to significant vendor exits from the local market (demonstrating limited adverse selection effects). They rather reduced vendor misconduct behavior, which is consistent with moral hazard effects (similar to Klein, Lambertz, and Stahl 2016).

[8] We have a few continuous outcomes with zero values and likely censored at zero: (i) seller misconduct amount / severity and (ii) consumers' weekly usage of transactional services (Figure C.1). To account for this, we also report results using either a Tobit regression or use an inverse hyperbolic sine transformation `asinh`.

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 B.5), Imbens and Manski (2004) confidence sets, and Behaghel et al. (2015) attrition bounds (improved trimming based on the number of times subjects are called before they answer the phone survey; see Figure C.2). 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).

IV.2 Treatment Effects of Information Sets on Seller Misconduct

As a first stage, we ask whether the information programs are anti-misconduct. Table 2 reports the pooled and separate treatment effects, and shows that the intervention meaningfully reduced vendor misconduct (measured using actual audit transactions). We estimate a pooled 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-only programs, however, the differences across the programs are barely distinguishable statistically.

In additional tests (Table D.1), we find significant spillover effects: non-treated vendors located in treated localities (or markets) reduce their misconduct (-15 pp pooled effect). This broader impacts on vendor misconduct is consistent with misconduct being contagious with externality effects, which is typical of vertical markets (Tirole 1988, Chapter 4). Meaningful vendor reputation is difficult to build in the baseline environment due to externality effects of misconduct, when combined with imperfect information between vendors and consumers. 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 D.2 and D.3) show that the reduction in misconduct is much larger in localities with more competition,[9] particularly for the joint information program. The effects are similar across gender, which means female vendors might respond more to the information programs because at baseline (pre-treatment), female vendors are significantly more likely to commit misconduct relative to male vendors. This suggests that both 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 information disclosure combined with 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 decreases in both incidence (the occurrence) and intensity (shift in the distribution) of seller misconduct. We next evaluate how this dramatic reduction in misconduct due to the information sets impacted various consumer and business outcomes.

Treatment Effects on Consumers

IV.3 Real Effects: Graphical Evidence of Treatment Effects

We provide graphical illustration of the treatment effects on consumers. Figure 1 plots the empirical cumulative distributions of the log of total transaction amounts per week at endline by treatment status. This is a key outcome of interest. The effects are displayed for

[9]At baseline, vendor misconduct was not significantly different between less and more competitive markets, which can be explained by the existence of much imperfect information. This means some vendors were committing misconduct at baseline, even in markets with more vendor competition, which is consistent with several classic papers discussing the possibility that prices can increase in markets with more firms (see e.g., Satterthwaite 1979; Rosenthal 1980). Consumer search costs can be higher in a larger market with more vendors, which will imply that vendors in larger markets are able to exercise more market power and hence engage in higher misconduct.

the various treatments together (pooled) and separately (in keeping with the approach of reporting the pooled 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 M-Money financial services as a result of the information programs. What is more striking is that effects do 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<0.091) except for the PT-only program (p-value=0.481). 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.

IV.4 Information Assignment – Pooled Effects

IV.4.1 Effects on M-Money Usage and Savings

Table 3 reports the estimated pooled effects on usage of services and savings, respectively. There is increased transaction amount per week, with a elasticity of 40.2% (p-value=0.048). In Appendix Table D.5, we report the effects on the probability of using the financial services, showing increased transaction likelihood of usage (7.1 pp =+9.8% of control mean, p-value=0.049). For savings, there is evidence of increased savings rate by 7.6 pp (=+12.6% of control mean, p-value=0.099).

IV.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 4 shows the estimated pooled effects on customers' experiences of unmitigated shocks. We report on general shocks (any experience), and individually on shocks related to household revenue, health, and household expenditures.

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 health and revenue sources are equally meaningful based on their effect sizes (7.2 pp and 5.6 pp, respectively). These estimates provide a large and objective proxy for financial resilience and insurance value of reducing seller misconduct to consumer welfare. In Table D.6, we do not find evidence for an impact on overall poverty levels.

IV.5 Information Sets: What's Necessary, What's Sufficient?

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

IV.5.1 Effects on M-Money Usage and Savings

Table 5 shows the estimated effects of the various information sets on the uptake of money services and savings. For uptake of services, the effects are positively much larger for the joint program (elasticity of 50.6%, p-value=0.035), compared to other individual information sets. The results are similar for savings likelihood on M-Money at vendor points. Customers are significantly more likely to save on M-Money with much larger impacts for the joint program (12.3pp = +20.2% of control mean, p-value=0.024), compared to 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 reporting-only information set (p-value=0.066).

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

IV.5.2 Effects on Mitigation of Shocks and Poverty

The estimated impacts of the various information sets on consumer welfare (shock mitigation and poverty) are reported in Table D.6. For shock mitigation, the joint information program shows significantly negative larger impacts, compared to individual information counterparts. As in the pooled estimate, this effect is mainly driven by mitigation of unexpected shocks related to household expenses. Effects from the MR-only program are relatively small and insignificant. For poverty, we do not find evidence across the various programs for an impact on overall poverty levels. These results agree with our earlier findings that the two individual information sets are informationally complementary and that the impact on poverty as a structural or composite outcome may be distributional.

Treatment Effects on Businesses

IV.6 Treatment Effects on Business Transactions

Did market vendors experience an increase in sales revenue? If consumer records, and hence the estimated treatment effects, are accurate, then one might expect direct increases in business transactions (all else being equal). Table 6 (alternatively, Table D.7) reports the estimated broad impacts on businesses. Meaningful positive treatment effects are reported separately for M-Money services (+40%), other non-M-Money business services (+57%), and total business sales (+36%), which combines the previous two sales revenues. The large positive impacts on non-M-Money transactions (for bundled stores) suggests positive spillover effects of the information programs on overall local market activities.[10] In Table D.8, we find limited significant difference in treatment effects across the different information programs.

For context, the typical transaction is about GHS100 (based on the audit transactions of GHS50, GHS160 and GHS1100 which were chosen to be typical of the market setting, Table

[10]We do not find evidence for an impact on the number of customers (similar to Brown, Hossain, and Morgan 2010) (results omitted but available). Did the interventions cause businesses to exit the local market? Defining business exits (or deaths) as vendors that were unreachable and/or inactive registered phone numbers during our endline phone surveys, we do not find evidence for an impact on exits (see column (4) of Table 6). This is consistent with the very low attrition rate (0.8%) of the audit transactional exercises that require physical visits to the vendor. Recall that we make repeated endline calls (see Figure C.2), varying the days and time of day. In our market environment, defining business exists as unreachable vendors seems relevant because active vendor phone numbers are required for the business of M-Money to be in operation. However, it is possible that businesses could simply be switching their registered numbers, which seems unlikely: one can replace the vendor phone number at no cost if lost; obtaining a completely new vendor number is costly and entails more paperwork.

C.2). The regular and official fee will be 1% of the transaction value, which implies a fee value of GHS1.0. The experiment leads the total fee (regular fee + misconduct) to fall from about 1.75% to about 1.10% (Table 2), about a 40% reduction of the transaction fee. The 40% increase in consumer demand (36% increase in total vendor sales) in response to a 40% fee reduction is reasonable; it is an elasticity of about 1.0 (0.9).

Decreased vendor misconduct and increased demand for financial services are beneficial to consumers; increased business sales revenue is beneficial to service providers; but is the average vendor better or worse off? From the audit transaction data, we estimate an average effective price of about GHS1.75 per GHS100 transaction volume for control vendors, versus GHS1.10 per GHS100 transaction volume for treated vendors. With a treatment effect of +GHS450 increase in M-Money services, the treatment increases M-Money sales revenue from about GHS800 to GHS1250. Vendors earn sales commission as profits, so the treatment changed their average profits from $\frac{1.75}{100}\times 800$=GHS14.0 to $\frac{1.10}{100}\times 1250$=GHS13.75. This implies that vendors are unaffected, which is consistent with the estimated elasticity of 0.9. If we account for the additional improvements in vendors' non-M-Money services (the positive externalities from bundling), then the average vendor is better off.

The improvement in business and vendor outcomes is interesting and calls for further discussions.[11] Why are vendors engaging in misconduct in the first place, if it is not profitable to do so? What model of market structure can explain these results? This market operates on a vertical structure: service providers at the upstream set up vendors at the downstream who work for them and earn commissions on sales, which leads to a version of the well-known single versus double marginalization problem (Tirole 1988, Chapter 4; Janssen and Shelegia 2015). Providers have limited visibility into the behavior of their vendors.

Because service providers fix prices of transactions at vendor points (price forcing to charge

[11]To further explore the market-level effects, we solicit administrative data from the largest provider about market transactions that originate from localities in our study area during the endline period. A limitation of this provider's data is that it does not cover all our experimental localities and hence does not provide much variation across the separate market-level treatment arms. However, pooling all the treatments, we find an overall increase in transaction activities in the treated markets relative to the control markets, which is qualitatively consistent with our baseline treatment effects on household and vendor outcomes (results omitted and available upon request).

marginal cost), vendors cannot reduce their sales in an attempt to marginalize. Through misconduct, vendors impose illegal mark-ups on transactions, which results in fewer sales than is optimal from the viewpoint of the provider. The treatment pushed the double marginalization high price (high misconduct) to a single marginalization lower price (low to no misconduct). In this case, lowered misconduct results in benefits not only to providers and vendors, but also to consumers. In our setting, there are vendors who earn profits not only from the M-Money business, but also from selling other products. When the treatment leads to less misconduct, customers conduct larger money transactions and also purchase other non-M-Money items at the vendors premises. Thus, we show that the vendor can also be better off under the single marginalization result, which is a novel and interesting result.[12]

V Framework: Interpreting the Results

We present a framework to guide the information programs and interpretation of our results. We seek to understand what happens when we give relevant seller misconduct information to both (potentially dishonest and informed) vendors 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 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. Following Macchiavello and Morjaria (2015), we think of a vendor's reputation as consumer perceptions about the vendor's tendency to commit misconduct. Negative perceptions trigger direct punishments and affect vendor reputation (via a reduction of vendor sales in other joint lines of business and of customer referrals, including 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).

[12]We thank Matt Shum for pointing this out.

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.

V.1 Baseline Setting

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

V.1.1 Feature 1: Asymmetric (Imperfect) Information

`Are customers 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 of true transactional prices compared to customers, and thus creating incentives for vendor misconduct. In a series of tests, both vendors and customers are asked to indicate the true charges for two randomly chosen transactions of sizes GHS200 (small to medium) and GHS1200 (large). We are careful to inform vendors at the beginning that we are not there to perform any actual transactions, but to rather assess their overall knowledge about the market. Knowledge tests are taken towards the end of the surveys for both sets of subjects. With reference to the official charges, this provides us an estimate of their knowledge about the true charges, specifically the percentage of subjects whose answers are correct across markets.

Results are displayed in Figure C.4, showing strong evidence of asymmetric information: vendors have superior knowledge of true transactional charges relative to customers. Overall, consumers are correct 48% (median=42%) of the time, while vendors are correct 73% (median=79%) of the time. These results are expected because unlike customers, vendors receive formal training about the market for M-Money before they start their businesses.

V.1.2 Feature 2: Reputation

`I. Vendors: how important is good market reputation?`

We ask a random sample of vendors in the control group of our experiment post-endline about the importance of demonstrating good market reputation (or image and responsibility) to potential customers through their market transactions. As shown in Figure C.5, the vast majority of vendors (81% [SD=0.391]) consider good market reputation or image as important, suggesting that there is likely a positive return or reward to vendors for good market reputation.

`II. Customers: how trustworthy are vendors carrying out M-Money transactions (at baseline)?`

Our baseline census solicits information about customers' level of trust in carrying out their market transactions. Figure C.6 reports the results, suggesting limited level of trust. About 62% [SD=0.48, n=1275] of customers indicate distrust in transacting at vendor points, while the rest (38% [SD=0.48, n=779]) indicate trust. This suggests that vendors have low reputation in the market, perhaps because customers are unable to infer the vendors behavior. This is consistent with the evidence of imperfect information about transaction tariffs above, and of misperceived consumer beliefs about vendor misconduct below.

V.1.3 Feature 3: Perceptions about Misconduct

`Do customers hold misperceived beliefs about misconduct (at baseline)?`

Figure C.7 compares 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 (Table C.2 and Figure C.3). We denote this by $(1-\pi)$, implying that π is the percentage of honest transactions (i.e., transactions not overcharged). Next, we also ask 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, our market environment reflects an empirical setting where (i) consumers are objectively less sophisticated (uninformed); (ii) market vendors value their reputation in the market, but good reputation is difficult to establish because consumers, not knowing official prices, cannot determine whether vendors are being honest, which suggests that vendors receive a positive return for good market reputation (extended sales revenue) if they are viewed by customers as responsible; and (iii) at baseline, consumers underperceive the level of honest transactions (upwardly-biased beliefs about vendor misconduct). Our setup and information programs work to reduce vendors misconduct and enhance consumers' subjective perception of the level of honest vendor behavior. Moreover, consumers might transact more if misconduct (equivalently, the marginal cost of transactions) is low.

V.2 Model: Misconduct, Punishment, and Reputation

V.2.1 Environment

We assume a continuum of local markets, defined by the pair (i, j), where i denotes a randomly selected vendor and j denotes potential customer(s). This is akin to our experiment's design, whereby we construct a local market using a randomly selected vendor and nearby households as customers per locality to maximize statistical power. In each locality, 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). 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 percentage of honest transactions (that is, the 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.

The vendor receives revenue in two ways: reputation (from honest behavior) and "uncertain" direct benefits (from dishonest behavior). First, given public information σ, consumers' willingness to pay is $\mathbb{E}[\hat{s} = 1|\sigma]$; this equals the vendor's reputational payoff given the signal. We call this reputational payoff as the vendor cares about $\mathbb{E}[\hat{s} = 1|\sigma]$ that customers compute (i.e., posterior that the vendor is honest) and assigns 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. The vendor's reputational revenue is proportional to the market size (denoted by $\eta > 0$) and his/her belief that customers perceive his/her actions as honest. Second, if the vendor chooses $s = 0$ (a dishonest action), 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 $Y^r \mathbf{I}_{s=0} < Y \mathbf{I}_{s=0}$ 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[13]

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

V.2.2 Mapping Model to Experiment

Before analyzing the framework, it is useful to discuss how our model and analysis map to our experimental setup. Market vendor(s) decide whether to commit misconduct ($s = 0$) or not ($s = 1$). Consumers (uniformed vs informed) learn about the transactional action through public signals σ. Based on their inference about a vendor's action given the available signal, a customer either assigns a reputational payoff ($\mathbb{E}[\hat{s} = 1|\sigma]$) to the vendor (via either PT or MR information programs) or reports the vendor's dishonest behavior as a direct punishment (via MR information program). If customers perceive (via $\hat{\pi}$) that the vendor is honest, then the vendor receives higher revenue (i.e., through repeated or large transactions and not being reported) 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 technology to detect and reward misconduct behavior (reputation effects, where $\sigma = s$), another with technology to directly report and punish misconduct behavior (reputation and punishment effects), and one with both. We model assignment of the anti-misconduct market information as either a shift in the distribution of $\hat{\pi}$ or $\mathbb{E}[\hat{s} = 1|\sigma]$. As we show (and as implied by the model), the information assignment (i) increases customers beliefs about the percentage of honest transactions $\hat{\pi}$; (ii) cause customers to update their beliefs about honest vendor behavior (thus to assign $\mathbb{E}[\hat{s} = 1|\sigma]$); and (iii) cause vendors themselves to update their beliefs about how informed consumers are and the likelihood of direct punishment. Together, these increase honest market vendor actions ($s = 1$) and improve market outcomes by increasing consumer demand

[13] $[qY^r + (1-q)Y]$ is the vendor's opportunity cost of being honest. Our simple sanctioning and reputation formulation provides a moral hazard analog of the labor supply and stigma (adverse selection) model of Bursztyn, González, and Yanagizawa-Drott (2020).

for services and vendor sales revenue.

V.2.3 Analysis

In the game, we are interested in Perfect Bayesian Equilibria. Denote $\hat{\pi}^* = \frac{qY^r+(1-q)Y}{2\eta} + 1/2$ (assume $\hat{\pi}^* < 1$) (In Appendix A, we provide detail foundations for η).

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:

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

Proof. See Appendix A. ■

In our experiment, when we provide symmetric two-sided information about the official prices of transactions, consumers' signal σ is the same as the s action chosen by the vendor ($s = \sigma$). There is revelation of the imperfectly observed vendor's actions and beliefs are updated to the posterior $\Pr[\hat{s} = 1|\sigma = s = 1] = 1$ and $\Pr[\hat{s} = 1|\sigma = s = 0] = 0$. The maximal extent of reputation gain is given by the difference: $\triangle\mathbb{E}[\hat{s} = 1|\sigma] = \mathbb{E}[\hat{s} = 1|\sigma = s = 1] - \mathbb{E}[\hat{s} = 0|\sigma = s = 0]$ 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. Informed consumers thus exert a positive externality on uninformed consumers by driving up honest vendor behavior.* (iii) *Increasing either $\triangle\mathbb{E}[\hat{s} = 1|\sigma]$ (PT or MR information programs) or q (MR information program) increases the number of $s = 1$.*

Proof. See Appendix A. ■

V.3 Reputation Effects – Subjective Beliefs and Belief Updates

We empirically evaluate predictions from the model to measure consumer beliefs about misconduct and reputation concerns for vendors to explain our treatment effects.

V.3.1 Subjective Beliefs

From the assumed lower $\hat{\pi}$, Proposition 2 indicates that an upward shift in perceptions of honest vendor behavior $\hat{\pi}$ (as well as the number of informed customers θ) should increase honest vendor actions $\Pr(s = 1)$. Thus, a necessary requirement for our information program to reduce misconduct (with impacts on the allied market outcomes) is to check whether consumer perceptions about honest vendor behavior $\hat{\pi}$ increase. Did our information intervention actually increase consumer perceptions of honest vendor behavior? Following Bursztyn, González, and Yanagizawa-Drott (2020), we elicit perceptions about seller misconduct (or honest behavior, otherwise) by asking customers to indicate their belief (`Agree or Disagree`) to the statement: `In my view, M-Money vendors generally overcharge customer transactions at vendor points`. Next, to incentivize our beliefs elicitation about seller misconduct, we jointly ask consumers: `What's your estimate of the % of others (all vendors and customers in this locality) that will Agree (Disagree, otherwise) with this statement?` In each local market, the respondent with the closest guess receives 10GHS (see Appendix I for details). These questions provide two alternative measures of customers' perception of vendor misconduct or honest behavior (non-incentivized versus incentivized). For a third measure, we also ask whether customers believe they have experienced transactional overcharges at vendor points, as in the baseline. The three subjective belief measures, which reflect consumer belief in vendors' trustworthiness, are significantly positively correlated (p-value = 0.000).

First, we ask if customer views about honest vendor behavior at endline shifted in the direction of the information treatments. In Figure E.1, we plot the distribution of $\hat{\pi}$ at endline, reflecting consumers' subjective beliefs about honest vendor transactions (`Disagree` to the belief statement) by treatment status. These are displayed for the various treatments together (pooled) and separately. Second, we estimate

$$\hat{\pi}_{jvd} = \beta \mathbf{I}_{jvd} + \eta_d + \beta_0 \hat{\pi}_{base,jvd} + \mathbf{X}'_{jvd}\xi + \epsilon_{jvd}$$

Table E.1 reports the estimated effects of the information program on consumer perceptions of honest vendor behavior. There is strong evidence (both visual and formal) that the intervention meaningfully increases perceptions of honest vendor behavior. The results robustly replicate across all three alternative measures of consumer beliefs about vendor honesty. We estimate a pooled effect of +7.0 pp (+30% of control mean) increase in subjective customer belief about honest vendor behavior at endline. The perceived effects appear to be much larger for the joint program. The change in perceptions reflect the reality that consumers now have the technologies to enforce vendors trustworthy behavior using the channels activated – social sanctions and/or punishment. Consistent with the reputational revenue definition, consumers perceive vendors as honest in treated markets (Macchiavello and Morjaria 2015).

V.3.2 Belief Updates and Calibration

We next measure reputation based on two tests: (i) how customers are able to infer vendor misconduct, or (ii) how vendors themselves are able to detect informed customers who might reward honest vendor actions and/or report dishonest behavior. Providing market information about misconduct makes consumers more likely to detect or report vendor misconduct, which raises the importance of vendor reputation.

Empirical Test I: Do consumers update their beliefs about vendors misconduct? We define consumer beliefs update as the probability of a customer correctly inferring vendor misconduct (or honest behavior, otherwise) given the information treatment. To test this, we estimate

$$(1-\hat{\pi})_{jvd} = \gamma_1 Misconduct_{ivd} + \gamma_2 \mathbf{I}_{jvd} + \gamma_3 Misconduct_{ivd} \times \mathbf{I}_{jvd} + \gamma_0 (1-\hat{\pi})_{base,jvd} + \mathbf{X}'_{jvd}\xi + \epsilon_{jvd}$$

This specification has perception of misconduct (`Agree` to the belief statement) as the dependent variable and examines how the intervention cause consumer perceptions to more closely correlate ($\gamma_3 > 0$) with the audit measure of misconduct, $Misconduct$. Our results robustly replicate across all three alternative measures of beliefs about vendor misconduct behavior. Tables 7 and E.2 show the results. We estimate a pooled effect of $\hat{\gamma}_3$=27 pp (+51% of control mean, p-value=0.034) increase in customers' ability to correctly guess misconduct behavior. For the separate treatment effects, we find evidence that the joint information program had economically larger effects. These results (i) provide evidence of updated consumer belief i.e., increased ability of customers to detect or report vendor behavior, and (ii) show increased consumer sophistication. Treated customers are significantly better calibrated about prevailing vendor behavior relative to the control group. If vendors recognize this shift, then they might update their beliefs about the sophistication of consumers or the likelihood they will report dishonest behavior. This leads to the second empirical test:

Empirical Test II: Do vendors update their beliefs about customer sophistication (i.e., θ) and likelihood of reporting dishonest behavior? We define vendor beliefs update 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. Descriptive results from follow-up surveys (not reported) also provide corroborative evidence that vendors update their beliefs about consumer sophistication. Overall, these results are strongly consistent with our proposed reputation-based interpretation: providing symmetric information about misconduct to both parties (uninformed customers and informed vendors) attenuates consumer misbelief about misconduct, and raises vendor concern for market reputation. In response, vendors reduce misconduct with market-wide impacts.

VI Further Results and Discussion

Before concluding, we discuss a number of corroborative results and potential alternative explanations. Lastly, we compute the value of the anti-misconduct information programs.

VI.1 Additional Heterogeneity and Alternative Explanations

We present further heterogeneous results that lend corroborative support for the reputational and/or direct punishment effects. First, sellers who operate bundled stores are likely to be more concerned for reputation following our information programs 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 at endline to be larger for vendors who bundle M-Money with other services, relative to market vendors who operate only M-Money services. This will be consistent with our earlier evidence indicating large positive spillover impacts of the information program on vendors' non-M-Money sales revenue (Tables 6-D.8). Tables E.3 and E.4 show robust evidence that the information effects on misconduct are concentrated on vendors who bundled services.

Second, under much asymmetric information about the true transactional tariffs (Figure C.4), consumers might find it difficult to detect, report, and thus reward good vendor behavior, which would be especially true for customers who were vulnerable (illiterate, marginalized) at baseline. Tables E.3 and E.4 show consistent evidence. The negative impact of the intervention on vendor misconduct is larger in markets with high fraction of customers having no formal education at baseline, who also performed poorly in our baseline knowledge tests about the official tariffs.

Finally, we have shown that a combination of more accurate consumer beliefs about misconduct and reputation concerns for vendors drives the estimated impacts of our anti-misconduct information programs. However, since the market (particularly consumers) becomes more informed about the true tariffs, the information interventions might also turn on two interesting alternative mechanisms: price effect or bargaining effect for the real consumer and business outcomes. Price effects can considered as a by-product of reputation: vendors take 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-affirm reputation. Regarding bargaining effects, M-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 the market transaction. We believe that bargaining is not driving our findings.

VI.2 The Value of Anti-Misconduct Information

The value of information arises from empowering consumers with technologies to enforce market vendors' trustworthy behavior by relying on social sanctions and/or punishment. How cost-effective is our anti-misconduct information intervention? Does this compare well to financial information interventions? When computing cost-effectiveness, we focus on usage of money services-only measure for customers and sales revenue-only measure for vendors. This is a very conservative approach that does not consider the significant treatment effects on savings, risk mitigation outcomes, and other positive externalities, such as increased non-M-Money sales revenue for bundled stores.

We first compute the total cost of interventions to be GHS15,165.[14] With about 730 subjects, we then estimate $\frac{\text{GHS15,165}}{730}$ = GHS20.8 cost 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 takes roughly ten minutes per visit to deliver the information intervention. Compared to Ghana's minimum wage over the period (GHS10.65 per day), the time-use and 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$19.3 increase in the usage of financial services for customers (Table D.4), with sizable implications for consumer welfare (including risk mitigation; see Table 4). This alone suggests a large return of 383%. For vendors, the treatment effect (+GHS437 = +$US85.4; see Table 6) implies a ratio of 1:21 improvement in vendor outcomes.

These rough cost-effectiveness estimates compare favorably with other financial information programs. Frisancho (2018) reports a cost per pupil of US$4.80 and a US$1 increase in financial education program expenses for a 3.3 point improvement in financial literacy. For comparison, we estimate a pooled treatment effect of +27 pp (=+51%) increase in customers' ability to correctly guess seller misconduct behavior. In a recent meta-analysis of financial education interventions, Kaiser et al. (2020) report 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 information that reduces seller misconduct could be a cost-effective way to improve local markets.

VII Conclusion

Misconduct in markets matters for efficiency. The provision of information sets that deter and reduce vendor misconduct has broader market impacts. Customers meaningfully increase

[14] This is based on the number of trained field officers utilized (3 officers), the number of visits to the treated subjects to deliver interventions (3 rounds), transportation costs (GHS385 per officer × 3 officers × 3 rounds = GHS3,465), remuneration and allowance for officers (GHS1,200 per officer × 3 officers × 3 rounds = GHS10,800), and occasional accommodation for officers during field visits (GHS100 per officer × 3 officers × 3 rounds = GHS900). The total cost equals GHS15,165. We reach about 632 panel of treated customers ($\frac{1}{3}\sum_{r=1}^{3}$ number of subjects reachable per round$_r$ = $\frac{629+617+642}{3}$ = 632) and about 97 panel of treated vendors ($\frac{98+96+98}{3}$ = 97), bringing the total panel number of subjects to 730. Almost all subjects are reached once or twice.

their take-up of transactional services and savings behavior at vendor points, which enables them to better mitigate unexpected shocks. Businesses experience meaningful increases in their sales revenue with limited impact on vendor profits/commissions, suggesting improved market efficiency.

Reputation does matter for misconduct. In rural financial environments, where markets are subject to a high degree of imperfect information, the importance of reputation as a discipline device against market misconduct is limited. When customers do not know official and mandated prices, they cannot observe whether they are being cheated, making it difficult for vendors to establish good reputation–which may increase vendor misconduct. However, reputation becomes effective and disciplinary if there is a high probability that customers will infer misconduct (Shapiro 1982, Burkhardt 2018) and if 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 vendor misconduct through cost-effective information disclosure programs; (ii) quantify the programs' impact on important economic outcomes on both sides of the market; and (iii) show that these effects are driven by a combination of more accurate consumer beliefs about misconduct and increased vendor concern for reputation. 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 are underprovided (Jensen and Miller 2018; Bai 2019) due to seller misconduct, which diminishes overall market efficiency.

Commerce requires reputation and/or consumer trust, but reputation in markets might be difficult to build and thus low due to imperfect information. In developing countries, where consumers are either uninformed about FinTech or lack experience with it and many market digitization initiatives are underway, consumers suffer significant market misconduct which can lead to consumer mistrust in payment markets; act as a barrier to market activities; and reduce households' welfare. Our study is the first, to our knowledge, to provide quantitative estimates on vendor misconduct and the value of anti-misconduct information programs in payment markets, emphasizing the effect of social sanctions and punishment.

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

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

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

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

[8] Banerjee, Abhijit, Rema Hanna and Sendhil Mullainathan S. 2012. "Corruption." Ch. 27, pages 1109-1147 of: Handbook of Organizational Economics.

[9] Behaghel, Luc, Bruno Crepon, Marc Gurgand, and Thomas Le Barbanchon. 2015. "Please Call Again: Correcting Non-Response Bias in Treatment Effect Models." Review of Economics and Statistics, 97 (5): 1070-1080.

[10] 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.

[11] Blackmon, William, Rafe Mazer and Shana Warren. 2021, "IPA Consumer Protection Research Initiative: RFP Overview." https://www.poverty-action.org/sites/default/files/presentation/IPA-Consumer-Protection-RFP-Info-Session-Feb-10-11-2021-Presentation-Final.pdf

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

[13] Breza, Emily L. and Arun Chandrasekhar. 2019. "Social Networks, Reputation, and Commitment: Evidence from a Savings Monitors Experiment." Econometrica, 87 (1): 175-216.

[14] Brown, Jennifer, Tanjim Hossain, and John Morgan. 2010. "Shrouded Attributes and Information Suppression: Evidence from the Field." Quarterly Journal of Economics, 125 (2): 859-876.

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

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

[17] Bursztyn, Leonardo, Alessandra L. González, and David Yanagizawa-Drott. 2020. "Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia." American Economic Review, 110 (10): 2997-3029.

[18] Cai, Jing and Adam Szeidl. 2017. "Interfirm Relationships and Business Performance." Quarterly Journal of Economics, 133 (3): 1229-1282.

[19] Coppejans, Mark, Donna Gilleskie, Holger Sieg, and Koleman Strumpf. 2007. "Consumer Demand under Price Uncertainty: Empirical Evidence from the Market for Cigarettes." Review of Economics and Statistics, 89 (3): 510-521.

[20] De Mel, Suresh, David McKenzie, and Christopher Woodruff. 2008. "Returns to Capital in Microenterprises: Evidence from a Field Experiment." Quarterly Journal of Economics, 123 (4): 1329-1372.

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

[22] Dupas, Pascaline and Jonathan Robinson. 2013. "Why Don't the Poor Save More? Evidence from Health Savings Experiments." American Economic Review, 103 (4): 1138-1171.

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

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

[25] Gabaix, Xavier and David Laibson. 2006. "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets." Quarterly Journal of Economics, 121 (2): 505-540.

[26] Garz, Seth, Xavier Giné, Dean Karlan, Rafe Mazer, Caitlin Sanford, and Jonathan Zinman. 2021. "Consumer Protection for Financial Inclusion in Low and Middle Income Countries: Bridging Regulator and Academic Perspectives." NBER Working Paper No. 28262.

[27] 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.

[28] Gibbons, Robert S. and John Roberts. 2012. "The Handbook of Organizational Economics." Princeton University Press.

[29] Higgins, Sean. 2020. "Financial Technology Adoption." Revise and Resubmit, American Economic Review.

[30] Imbens, Guido and Charles Manski. 2004. "Confidence Intervals for Partially Identified Parameters." Econometrica, 72 (6): 1845-1857.

[31] 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.

[32] Janssen, Maarten, and Sandro Shelegia. 2015. "Consumer Search and Double Marginalization." American Economic Review, 105 (6): 1683-1710.

[33] 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.

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

[35] Kaiser, Tim, Annamaria Lusardi, Lukas Menkhoff, and Carly J. Urban. 2020. "Financial Education Affects Financial Knowledge and Downstream Behaviors." NBER Working Paper No. 27057.

[36] Karpoff, Jonathan M. 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

[37] 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.

[38] Klein, Tobias J., Christian Lambertz, and Konrad O. Stahl. 2016. "Market Transparency, Adverse Selection, and Moral Hazard." Journal of Political Economy, 124 (6): 1677-1713.

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

[40] Macchiavello, Rocco and Ameet Morjaria. 2015. "The Value of Relationships: Evidence from a Supply Shock to Kenyan Rose Exports." American Economic Review, 105 (9): 2911-2945.

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

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

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

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

[45] Rosenthal, Robert W. 1980. "A Model in Which an Increase in the Number of Sellers Leads to a Higher Price." Econometrica, 48 (6): 1575-1579.

[46] Satterthwaite, Mark A. 1979. "Consumer Information, Equilibrium Industry Price, and the Number of Sellers." Bell Journal of Economics, 10 (2): 483-502.

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

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

[49] Shapiro, Carl. 1983. “Premiums for High Quality Products as Returns to Reputations.” Quarterly Journal of Economics, XCVIII: 658-679.

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

[51] Tirole, Jean. 1988. *The Theory of Industrial Organization*. Cambridge, MA: MIT Press.

[52] Zitzewitz, Eric. 2012. “Forensic Economics.” Journal of Economic Literature, 50 (3): 731-769.

Main Results for Text

Table 1: **Study timeline, sample, and randomization**

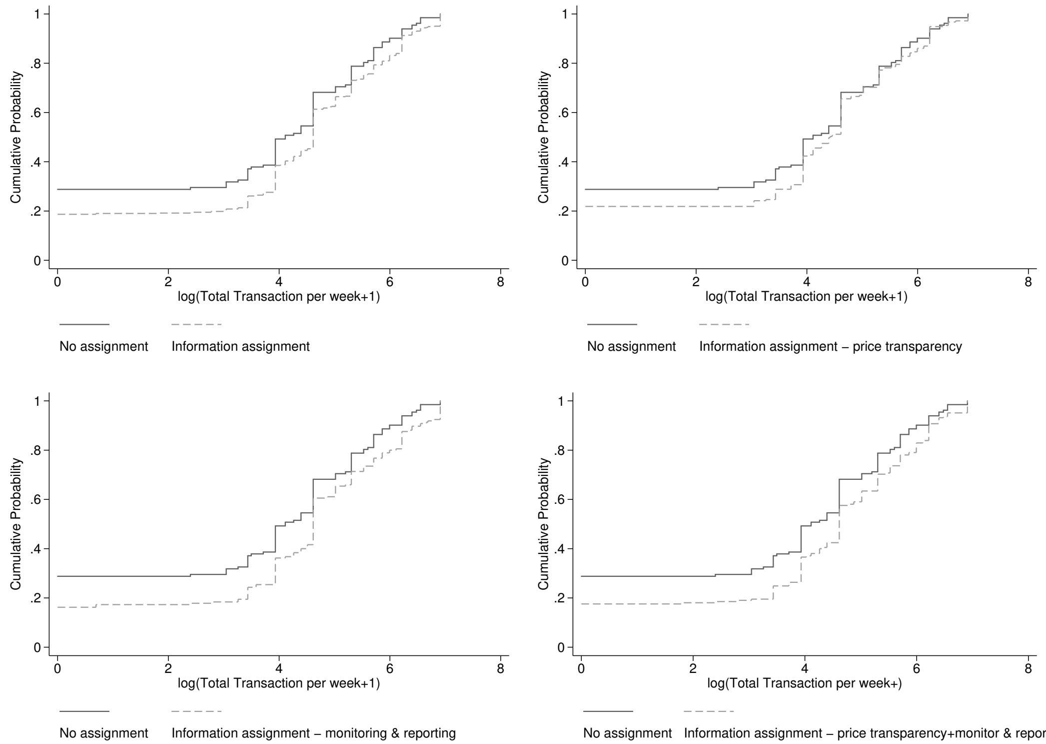
	DATE	ACTIVITY
Part 1	February 2017	**Pilots:** Misconduct – Incidence and severity, correlates (Annan 2017)
Part 2	Feb 15- Mar 20, 2019	**Baseline:** Market census – Listing all markets, detailed market records, demographics, market outcomes, consumer beliefs about misconduct + trust Total Markets: n=130 localities = 333 vendors×1,921 nearby customers
	Aug 1- Aug 15, 2019	**Select sample (Experiment):** n=130 localities =130 random vendors×990 customers **Intervention:** Information assignment Control: *No Information* (n=32 markets: 32 random vendors, 185 nearby customers) Treatment I: *Price Transparency* (n=31 markets: 31 random vendors, 272 customers) Treatment II: *Monitoring and Reporting* (n=32 markets: 32 vendors, 257 customers) Treatment III: *Joint program* (n=35 markets: 35 random vendors, 276 customers) [Randomization stratified based on 9 large administrative districts]
	Sep 01- Oct 15, 2019	**Audit study I:** Estimate vendor misconduct, $\geq$ 1 in 5 transactions (22%)
Part 3	Oct 15- Dec 15, 2019	**Intervention:** Information deployment [repeated 3 times] **Transaction networks data:** family vs friends vs strangers
Part 4	May 15- May 30, 2020	**Endline:** Phone survey + manual tracing supplement main market outcomes, consumer beliefs about misconduct elicitation
	Aug 15- Sep 01, 2020	**Audit study II:** Re-estimate vendor misconduct Spans select sample (Experiment) + non-select sample vendors
	> Sep 15, 2020	**Administrative data:** market transaction records from service provider

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

Model: Linear-OLS	**1**(Misconduct=Yes) (1)	(2)	Amount-Misconduct, GHS (3)	(4)	`asinh` (Amount-Misconduct) (5)
PANEL A					
Treatment: Information Assignment (β)	-0.231	-0.211	-0.675	-0.551	-0.323
	(0.055)	(0.086)	(0.185)	(0.255)	(0.138)
	[-0.324, -0.138]	[-0.354, -0.067]	[-0.984, -0.367]	[-0.975, -0.125]	[-0.553, -0.093]
Baseline misconduct	X	X	X	X	X
Market District F.E.	X		X		
Market District × Transaction × Date F.E.		X		X	X
Observations	335	335	335	335	335
Mean of dependent variable (control)	0.294	0.294	0.778	0.778	0.453
Lee (2009) Attrition Bounds	<-0.174, -0.164>		<-0.484, -0.1435>		<-0.280, -0.276>
Imbens and Manski (2004) CS	[-0.225, -0.094]		[-0.642, -0.085]		[-0.369, -0.109]
PANEL B					
Price Transparency (β_1)	-0.177	-0.184	-0.550	-0.439	-0.248
	(0.065)	(0.094)	(0.199)	(0.276)	(0.148)
	[-0.285, -0.069]	[-0.342, -0.027]	[-1.881, -0.219]	[-0.898, 0.020]	[-0.496, -0.001]
Monitor and Report (β_2)	-0.257	-0.217	-0.687	-0.574	-0.341
	(0.063)	(0.093)	(0.222)	(0.275)	(0.148)
	[-0.363 0.151]	[-0.373, -0.061]	[-1.057, -0.317]	[-1.032, -0.117]	[-0.588, -0.093]
Joint program: PT + MR (δ)	-0.233	-0.212	-0.718	-0.555	-0.325
	(0.064)	(0.089)	(0.198)	(0.279)	(0.148)
	[-0.340, -0.127]	[-0.360, -0.062]	[-1.048, -0.388]	[-1.019, -0.089]	[-0.572, -0.078]
Baseline misconduct	X	X	X	X	X
Market District F.E.	X		X		
Market District × Transaction × Date F.E.		X		X	X
Observations	335	335	335	335	335
Mean of dependent variable (control)	0.294	0.294	0.778	0.778	0.453
p-value (test: $\beta_1 = \delta$)	0.327	0.670	0.280	0.553	0.451
p-value (test: $\beta_2 = \delta$)	0.660	0.921	0.860	0.923	0.880
p-value (test: $\beta_1 = \beta_2$)	0.104	0.563	0.347	0.411	0.311
p-value (test: $\beta_1 + \beta_2 = \delta$)	0.027	0.108	0.074	0.204	0.164

Note: **1**(.) is a logical indicator that equals 1 if argument in parenthesis is true, 0 otherwise. Includes randomization strata (market district) dummies, baseline outcomes, and additional controls (age, marital status, ethnic group status, employment status, business experience, and bundled store status). Observations are at the vendor x transaction type x transaction 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 pooled 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 total consumer 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. Observations are at the customer level. From a Kolmogorov–Smirnov test for the equality of distributions, p-values equal 0.091, 0.481, 0.068 and 0.065, respectively (transaction data trimmed at the 5% level). Results are similar using an inverse hyperbolic sine transformation `asinh` of the total transactions. We report our main results for both transformations (`log`, `asinh`) including the levels of total consumer transactions.

Table 3: **Effect of treatment on usage of transactional services and savings**

Model: Linear-OLS	`log` (Total Transactions per week+1)			`asinh` (Total Transactions per week)	Saved (last month)	Saved (last month)
	(1)	(1)	(2)	(3)	(4)	(5)
Treatment: Information Assignment (β)	0.561	0.416	0.402	0.449	0.080	0.076
	(0.225)	(0.220)	(0.213)	(0.238)	(0.046)	(0.045)
	[0.189, 0.932]	[0.052, 0.779]	[0.050, 0.755]	[0.057, 0.841]	[0.004, 0.157]	[0.001, 0.151]
Inference Robustness (β)						
CI: Clustered S.E.	[0.059,1.062]	[0.096, 0.843]	[0.082, 0.723]	[0.093, 0.805]	[0.007, 0.153]	[0.004, 0.148]
CI: Wild Bootstrap	[0.191, 0.922]	[0.113, 0.821]	[0.024, 0.789]	[0.043, 0.825]	[0.004, 0.156]	[0.004, 0.149]
p-value: Permutation Test	0.015	0.041	0.048	0.046	0.080	0.099
p-value: L-S-X MHT Corr (2019)	0.012				0.048	
Market District F.E.		X	X	X	X	X
Baseline outcomes		X	X	X	X	X
Controls			X	X		X
Observations		763	723	763	689	689
R-squared		0.064	0.108	0.108	0.075	0.105
Mean of dependent variable (control)		3.583	3.583	4.589	0.605	0.605
Lee (2009) Attrition Bounds	[0.432, 0.805]			[0.561, 0.814]	[0.070, 0.125]	
Imbens and Manski (2004) CS	[0.076, 1.197]			[0.124, 1.293]	[0.001, 0.201]	
Behaghel et al. (2015) Attrition Bounds	[0.430, 0.738]			[0.547, 0.751]	[0.078, 0.120]	

Note: Market district is the randomization strata. Observations are at the customer level. Total Transactions per week is the value of M-Money transactions customer conducted in the local market per week at endline. Saved (last month) is a 0-1 indicator for whether the customer saved money on M-Money at endline. Controls in columns (4)-(5) include: gender, age, marital status, ethnic group status, employment status, education, and income. 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 shock mitigation and poverty**

Model: Linear-OLS

	u-Shocks Experience (any) (1)	u-Shocks HH Revenue (2)	u-Shocks Health (3)	u-Shocks HH Expenditure (4)	Poverty Likelihood (5)
Treatment: Information Assignment (β)	-0.068	-0.072	-0.056	-0.107	1.033
	(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 and is a 0-1 indicator for whether consumer experienced unexpected shock(s) that s/he cannot financially remedy or pay for. Includes randomization strata (market district) dummies, baseline outcomes, and additional controls (gender, age, marital status, ethnic group status, employment status, education, and income). 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 5: **Effect of information sets on usage and savings**

	Log Total Transaction per week+1	Total Transaction per week, GHS	Used M-Money (last month)	Saved (last month)	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 (gender, age, marital status, ethnic group status, employment status, education, and income). 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 6: **Effect of treatment on business sales revenue**

Model: Linear-OLS	`log` Sales (M-Money) per day (GHS)	`log` Sales (Non-M-Money) per day (GHS)	`log` Sales (Total) per day (GHS)	**1**(Business Exit =Yes)
	(1)	(2)	(3)	(4)
Treatment: Information Assignment (β)	0.401	0.565	0.359	-0.069
	(0.238)	(0.267)	(0.215)	(0.0585)
Robust S.E.	[0.005, 0.797]	[0.118, 1.012]	[0.001, 0.718]	[-0.166, 0.028]
Clustered S.E.	[0.005 0.797]	[0.119, 1.011]	[0.002, 0.717]	[-0.165, 0.027]
Wild Bootstrap	[-0.017, 0.807]	[0.127, 1.001]	[-0.010, 0.725]	[-0.168, 0.032]
p-value: Permutation Test	0.074	0.048	0.100	0.195
Lee (2009) Attrition Bounds	<0.284, 0.618>	<-0.028, 0.3502>	<0.106, 0.422>	All obs. selected
Behaghel et al. (2015) Attrition Bounds	[0.447, 0.499]	[0.172, 0.243]	[0.273, 0.321]	All obs. selected
Market District F.E.	X	X	X	X
Baseline sales revenue	X	X	X	
Controls	X	X	X	X
Observations	107	81	84	129
Mean of dependent variable (control)	6.289	4.723	6.703	0.218

Note: **1**(.) is a logical indicator that equals 1 if argument in parenthesis is true, 0 otherwise. Market district is the randomization strata. Observations are at the vendor level. Controls in columns (1)-(4) include: age, marital status, ethnic group status, employment status, business experience, and bundled store status. 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 (≡vendor) level. Differential attrition bounds are reported. Results similar to post-double-selection LASSO estimates clustered at local market level. At baseline, 75% of vendors bundled M-Money with other services, hence the variation in sample sizes.

Table 7: **Updates: information effects on correct customer inference of misconduct**

DV: 0-1 Indicator for belief about vendor misconduct $(1-\hat{\pi})$			
	(1)	(2)	(3)
Treatment: Information Assignment (β)	-0.196	-.276	-0.257
	(0.065)	(0.072)	(0.083)
	[-0.304, -0.088]	[-0.396, -0.155]	[-0.395, -0.119]
x Objective Misconduct measure	0.252	0.295	0.266
	(0.111)	(0.113)	(0.124)
	[0.067, 0.436]	[0.106, 0.484]	[0.060, 0.472]
Objective Misconduct measure	-0.132	-.217	-0.194
	(0.076)	(0.084)	(0.094)
	[-0.258, -0.006]	[-0.357, -0.077]	[-0.351, -0.036]
Market District F.E.		X	X
Baseline belief about vendor misconduct			X
Controls			X
Observations	678	678	678
Mean of dependent variable (control)	0.529	0.529	0.529

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. Controls in column (3) include: gender, age, marital status, ethnic group status, employment status, education, and income. Observations are at the customer 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, to alternative inference procedures (Wild cluster bootstrap and permutation test clustered at the market level), and to the alternative beliefs measures (non-incentivized vs incentivized outcomes).

Supplementary Appendix
(For Online Publication)

A Proofs

A.1 Proof of Proposition 1

Proof. $s = 1$ IFF

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

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

$\mathbb{E}[\hat{s}=1|\sigma, s] = \Pr[\hat{s}=1|\sigma, s]$, so we write:

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

We get

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

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

where $\triangle\mu = \mu(1,1) - \mu(1,0) = \triangle\mathbb{E}[\hat{s}=1|\sigma]$. In this PBE:
If $\hat{\pi} > \hat{\pi}^*$, then $\mu(1,1) = \Pr(\hat{s}=1|\sigma, s=1) = 1$ and $\mu(1,0) = \Pr(\hat{s}=1|\sigma, s=0) = 0$. 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 $\hat{\pi} > \frac{qY^s+(1-q)Y}{2\eta(1-0)} + 1/2 \geq \hat{\pi}^*$. If $\hat{\pi} < \hat{\pi}^*$, then $\mu(1,1) = \Pr(\hat{s}=1|\sigma, s=1) = x \in (0,1)$ (it can be anything), $\mu(1,0) = \Pr(\hat{s}=1|\sigma, s=0) = \hat{\pi}$, $\Delta\mu < 1$ and

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

The vendor does not find it worthwhile to choose an honest action $s = 1$ 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 opportunity cost of being honest $[qY^s + (1-q)Y]$ is too high. In our experiment, by providing symmetric two-sided information about official and mandated prices of transactions, consumers' signal σ is the same as the s action chosen by the vendor ($s = \sigma$). There is revelation of the imperfectly observed vendor's actions and beliefs are updated to the posterior $\Pr[\hat{s} = 1|\sigma = s = 1] = 1$ and $\Pr[\hat{s} = 1|\sigma = s = 0] = 0$. ∎

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 (market size).

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: $v(\Pr[\hat{s} = 1|\sigma, s]) = \Pr[\hat{s} = 1|\sigma, s]v_G + (1 - \Pr[\hat{s} = 1|\sigma, s])v_B$, with a reduced form demand function: $D_i(\Pr[\hat{s} = 1|\sigma, s] = 1) = v(\Pr[\hat{s} = 1|\sigma, s] = 1) = \theta v_G$ for informed customers versus $D_u(\Pr[\hat{s} = 1|\sigma, s]) = v(\Pr[\hat{s} = 1|\sigma, s]) = (1-\theta)v(\Pr[\hat{s} = 1|\sigma, s])$ for uninformed customers. Thus, the aggregate market demand for honest transactions is

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

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

Effects: Letting η equal the aggregate demand D_s, and observing that $\frac{\partial D_{s=1}}{\partial \theta} = v_G - v(\Pr[\hat{s} = 1|\sigma, s]) = v_G - \Pr[\hat{s} = 1|\sigma, s]v_G - (1 - \Pr[\hat{s} = 1|\sigma, s])v_B \geq 0|_{\Pr[\hat{s}=1|\sigma,s]=1}$ in equilibrium. For dishonest transactions, $\frac{\partial D_{s=0}}{\partial \theta} = v_B - v(\Pr[\hat{s} = 1|\sigma, s]) \leq 0|_{\Pr[\hat{s}=1|\sigma,s]=1}$. 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 $\hat{\pi}^{**} = \frac{qY^s+(1-q)Y+0}{2\eta\triangle\mu} + 1/2$ 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, $\frac{\partial \hat{\pi}}{\partial \triangle\mu}|_{\hat{\pi}=\hat{\pi}^{**}} = -\frac{2\eta(qY^s+(1-q)Y+0)}{(2\eta\triangle\mu)^2} < 0$. ∎

B Balance and Attrition

Table B.1: **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***	8.96e-04
	(0.054)	(0.076)
Age (years)	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		
Head of household reads English	4.104***	0.102
	(0.163)	(0.223)
Outer wall uses cement	3.909***	-0.306
	(0.222)	(0.342)
Toilet facility	4.617***	-0.349
	(0.140)	(0.268)
Number of 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 B.2: **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 (years)	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		
Head of household reads English	3.428***	-0.124
	(0.114)	(0.152)
Outer wall uses cement	3.664***	-0.272
	(0.196)	(0.195)
Toilet facility	4.372***	-0.584
	(0.137)	(0.182)
Number of 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 overcharged/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 B.3: **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 (years)	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				
Head of household reads English	4.248***	-0.213	-0.093	0.139
	(0.295)	(0.506)	(0.480)	(0.4178)
Outer wall uses 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 of 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 (Schreiner 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 pooled 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. Mean baseline characteristics are also balanced across treatment arms.

Table B.4: **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 (years)	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				
Head of household reads English	3.353***	-0.081	-0.345	0.226
	(0.212)	(0.321)	(0.347)	(0.305)
Outer wall uses 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 of 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 (Schreiner 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 overcharged/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 pooled 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. Mean baseline characteristics are also balanced across treatment arms.

Table B.5: **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 summary statistics for the subsample that was successfully reached for follow-up and for the subsample that was not successfully reached in endline phone surveys or manual contact tracing. Shown for both sides of the market (vendors versus customers). We fail to reject the null that attrition is non-differential (i) between the separate treatment arms and control arm and (ii) between pooled treatment arms and control arm at the 5% significance level. Attrition for endline audit exercises is 0.8%: 129 out of the 130 randomly selected vendors were reached. There was only one unreachable vendor in the joint PT + MR program.

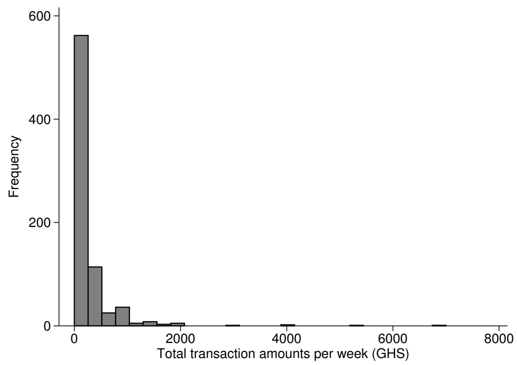
C Descriptive Statistics

Table C.1: **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
Head of household reads English	0.769	0.421	0.606	0.488
Outer wall uses 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 overcharged			0.191	0.403
Ever overcharged + unauthorized account use			0.293	0.455
Number of observations	333		1,921	

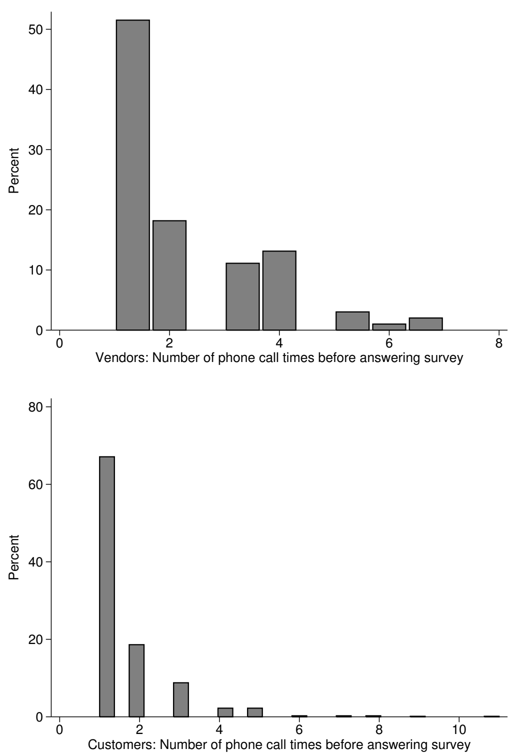
Note: Table reports summary statistics of relevant variables from our market census separately for both sides of the market: vendors *versus* customers. This includes information about demographics, poverty indicators, and market outcomes, respectively. Customers' borrowing and savings behavior and their subjective assessment of market misconduct on M-Money are also shown. The census covers 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 C.1: **Distribution (histogram) of total transactions at endline**



Observations are at the customer level.

Figure C.2: **Phone calls and reachability of subjects**



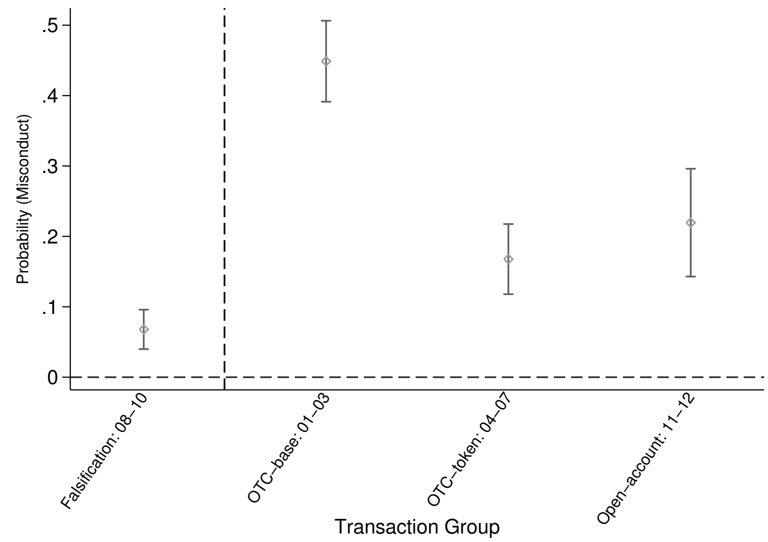
Observations are at the subject (vendor, customer, respectively) level.

Table C.2: **Misconduct at baseline: descriptive statistics based on transactional audit exercise, details**

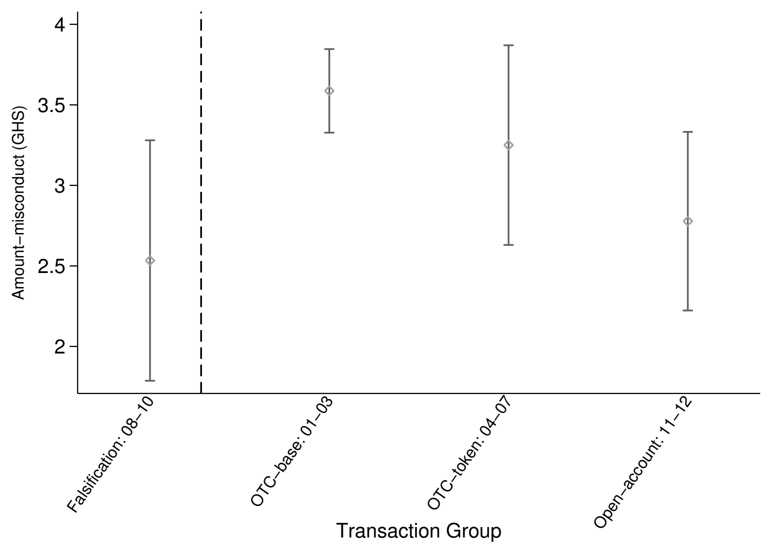
# Transaction type (description)	Outcome variable	Mean	SD	Transaction group	Mean	SD
01 Cash-in GHS50 - to others' wallet	**1**[Misconduct=Yes]	0.35	0.480		0.44	0.498
	Overcharged [GHS]	4.65	1.093		3.58	1.498
02 Cash-in GHS160 - to others' wallet	**1**[Misconduct=Yes]	0.52	0.502	$\{= \quad OTC - base$		
	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 successful transactions		663			663	

Note: Table reports the specific transactions used for the actual transactional exercises and shows the descriptive statistics of misconduct (n=663). These misconduct outcomes are based on the transactional exercises. Transactions are categorized into four groups: 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., leave 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 amount overcharged 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 three varying transaction sizes, is approximately GHS4.03.

Figure C.3: **Misconduct at baseline: descriptive statistics based on transactional audit exercise, graphical**



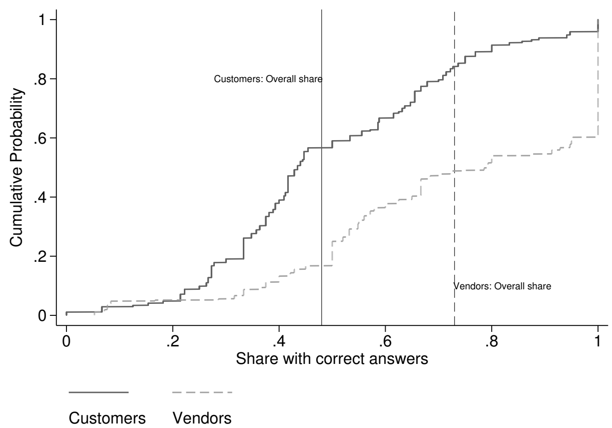
(a) **MISCONDUCT INCIDENCE × TRANSACTION GROUP**



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

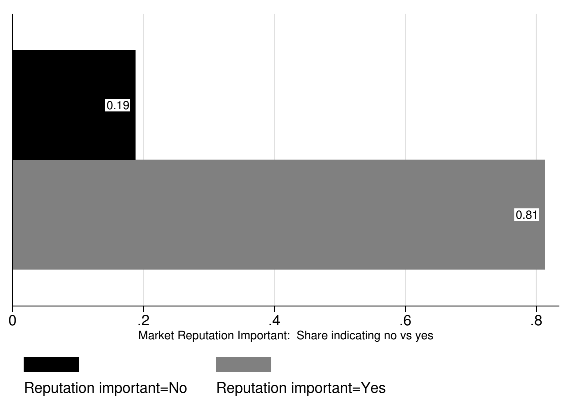
Note: Figures display the distribution of misconduct (n=663), measured as either the probability of the vendor committing a misconduct "incidence" (Figure (a)) or the amount overcharged as result of misconduct "severity" (Figure (b)) using actual transactional exercises at baseline. Transactions are categorized into four groups: 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 Table C.2. 90% confidence intervals (CI) are displayed around the estimates. As expected, misconduct is much higher in 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 C.4: **Asymmetric information about transactional prices**



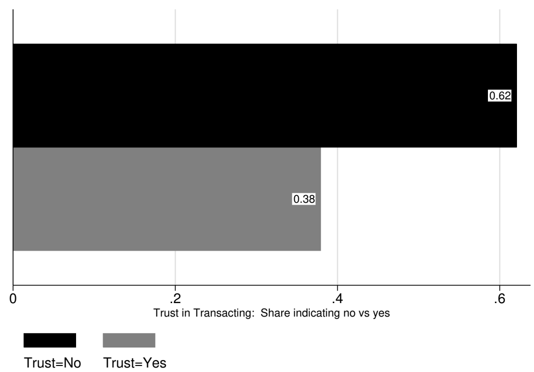
Note: Figure plots the distributions (CDFs) of the share of subjects with accurate answers for charges on randomly selected popular transactions (GHS200; GHS1200) derived with reference to their official or mandated rates (2GHS; 10GHS, respectively). A subject (customer, vendor) is correct if his/her answer matches the mandated rate. Observations are at the subject level. In each local market, we compute the share of subjects who answered correctly. Shown separately for customers and vendors. Trimmed to exclude unrealistic zero vendor knowledge/ correctness at the local market level. From a Kolmogorov–Smirnov test for the equality of distributions, p-value < 0.01.

Figure C.5: **Importance of reputation to vendors**



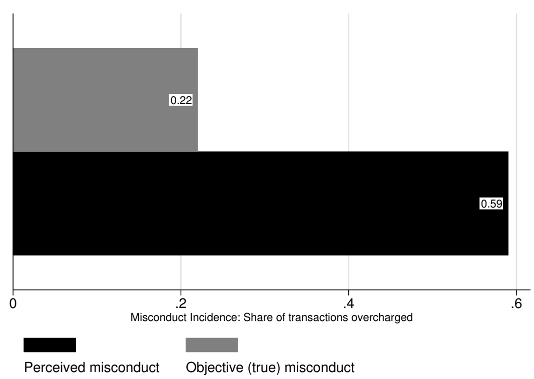
Note: Figure plots the share of vendors who value good market reputation through their money market transactions. Subjects (vendors) are asked to indicate how important it is to show a high degree of good market image and responsibility to potential customers when carrying out M-Money transactions on 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 C.6: **Consumer trust in performing money transactions at vendor points**



Note: Figure plots the share of customers, at baseline, who trust or do not trust the money transactions they make at vendor banking points. Subjects (customers) are asked to indicate their level of trust for carrying out M-Money transactions at 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 C.7: **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. π = the share of transactions not overcharged.

D Further Results: Treatment Effects

Table D.1: **Spillover effect of information sets on vendor misconduct**

Model: Linear-OLS	$\mathbf{1}$(Misconduct=Yes)		Amount-Misconduct, GHS		`asinh` (Amount-Misconduct) (5)
	(1)	(2)	(3)	(4)	
PANEL A					
Treatment: Information Assignment (β)	-0.153	-0.218	-0.473	-0.648	-0.363
	(0.055)	(0.065)	(0.173)	(0.206)	(0.112)
	[-0.245, -0.060]	[-0.326, -0.109]	[-0.763, -0.184]	[-0.992, -0.303]	[-0.552, -0.174]
Baseline misconduct					
Market District F.E.	X		X		
Market District × Transaction × Date F.E.		X		X	X
Observations	411	411	411	411	411
Mean of dependent variable (control)	0.278	0.278	0.783	0.783	0.445
Lee (2009) Attrition Bounds	<-0.170, -0.155>		<-0.569, -0.479>		<-0.280, -0.276>
Imbens and Manski (2004) CS	[-0.305, -0.076]		[-1.211, -0.220]		[-0.369, -0.109]
PANEL B					
Price Transparency (β_1)	-0.163	-0.232	-0.567	-0.720	-0.402
	(0.058)	(0.070)	(0.172)	(0.196)	(0.112)
	[-0.260, -0.065]	[-0.351 -0.114]	[-0.856, -0.279]	[-1.048, 0.391]	[-0.590, -0.215]
Monitor and Report (β_2)	-0.182	-0.239	-0.470	-0.693	-0.390
	(0.056)	(0.075)	(0.191)	(0.242)	(0.131)
	[-0.277 0.087]	[-0.364, -0.113]	[-0.789, -0.151]	[-1.098, -0.287]	[-0.609, -0.171]
Joint program: PT + MR (δ)	-0.122	-0.178	-0.409	-0.524	-0.292
	(0.069)	(0.070)	(0.211)	(0.224)	(0.122)
	[-0.238, -0.006]	[-0.296, -0.060]	[-0.762, -0.055]	[-0.900 -0.149]	[-0.496, -0.089]
Baseline misconduct					
Market District F.E.	X		X		
Market District × Transaction × Date F.E.		X		X	X
Observations	405	405	405	405	405
Mean of dependent variable (control)	0.278	0.278	0.783	0.783	0.445
p-value (test: $\beta_1 = \delta$)	0.512	0.315	0.353	0.179	0.186
p-value (test: $\beta_2 = \delta$)	0.235	0.235	0.712	0.323	0.276
p-value (test: $\beta_1 = \beta_2$)	0.640	0.915	0.482	0.859	0.890
p-value (test: $\beta_1 + \beta_2 = \delta$)	0.007	0.001	0.011	0.001	0.001

Note: For spillover effects, estimations compare non-treated vendors located in treated localities (or markets) to the pure control localities. $\mathbf{1}(.)$ is a logical indicator that equals 1 if the argument in the parenthesis is true, 0 otherwise. Includes randomization strata (market district) dummies. Observations are at the vendor x transaction type x transaction 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 pooled 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 D.2: **Heterogeneity in effect of information sets on vendor misconduct**

Model: Linear-OLS		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 × ...		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

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 (age, marital status, ethnic group status, employment status, business experience, and bundled store status). Observations are at the vendor x transaction type x transaction 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, hence the variation in sample sizes.

Table D.3: **Heterogeneity in effect of information sets on vendor misconduct**

Model: Linear-OLS		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 (b_1)	-0.728	-2.802	**x** Female (b_1)	0.155	0.358
		(0.731)	(3.202)		(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 (b_2)	-0.742	-2.410	**x** Female (b_2)	0.086	0.324
		(0.786)	(3.059)		(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 (d)	-1.502	-5.028	**x** Female (d)	0.197	0.350
		(0.702)	(2.953)		(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 × ... Transaction × 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
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, baseline outcomes, and additional controls (age, marital status, ethnic group status, employment status, business experience, and bundled store status). 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, hence the variation in sample sizes.

Table D.4: **Effect of treatment on usage of transactional services**

Model: Tobit
DV: Total Transactions 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: Tobit regressions censored at 0. Market district is the randomization strata. Observations are at the customer level. Controls in column (4) include: gender, age, marital status, ethnic group status, employment status, education, and income.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 D.5: **Effect of treatment on usage of transactional services**

Model: Linear-OLS
DV: Used M-Money (last month)

	(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. Used M-Money (last month) is a 0-1 indicator for whether the customer used M-Money at endline. Controls in column (4) include: gender, age, marital status, ethnic group status, employment status, education, and income. 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.

Table D.6: **Effect of information sets on shock mitigation and poverty**

Model: Linear

Model: Linear-OLS	u-Shocks Experience (any) (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 and is a 0-1 indicator for whether consumer experienced unexpected shock(s) that s/he cannot financially remedy or pay for. Includes randomization strata (market district) dummies, baseline outcomes, and additional controls (gender, age, marital status, ethnic group status, employment status, education, and income). 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 D.7: **Effect of treatment on business sales revenue**

Model: Linear-OLS	(1)	(2)	(3)	(4)
	Sales (M-Money) per day (GHS)	Sales (Non-M-Money) per day (GHS)	Sales (Total) per day (GHS)	**1**(Business Exit=Yes)
Treatment: Information	436.6	116.1	577.2	-0.069
Assignment (β)	(178.4)	(66.11)	(233.1)	(0.0585)
Robust S.E.	[140.1, 733.2]	[5.961, 226.3]	[188.7, 965.7]	[-0.166, 0.028]
Clustered S.E.	[140.4, 732.7]	[6.207, 226.1]	[189.5, 964.9]	[-0.165, 0.027]
Wild Bootstrap	[142.2, 722.9]	[4.891, 225.7]	[190.6, 965.9]	[-0.168, 0.032]
p-value: Permutation Test	0.025	0.135	0.035	0.195
Lee (2009) Attrition Bounds	<242.9, 486.9>	[-68.07, 21.81]	[223.9, 499.1]	All obs. selected
Behaghel et al. (2015) Attrition Bounds	[403.4, 452.2]	[0.546, 18.08]	[404.0, 468.6]	All obs. selected
Market District F.E.	X	X	X	X
Baseline sales revenue	X	X	X	
Controls	X	X	X	X
Observations	107	85	85	129
Mean of dependent variable (control)	792.8	239.5	1032	0.218

Note: Market district is the randomization strata. Observations are at the vendor level. Controls in columns (1)-(4) include: age, marital status, ethnic group status, employment status, business experience, and bundled store status. 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 (≡vendor) level. Differential attrition bounds are reported. Results similar to post-double-selection LASSO estimates clustered at local market level. At baseline, 75% of vendors bundle M-Money with other services, hence the variation in sample sizes across columns.

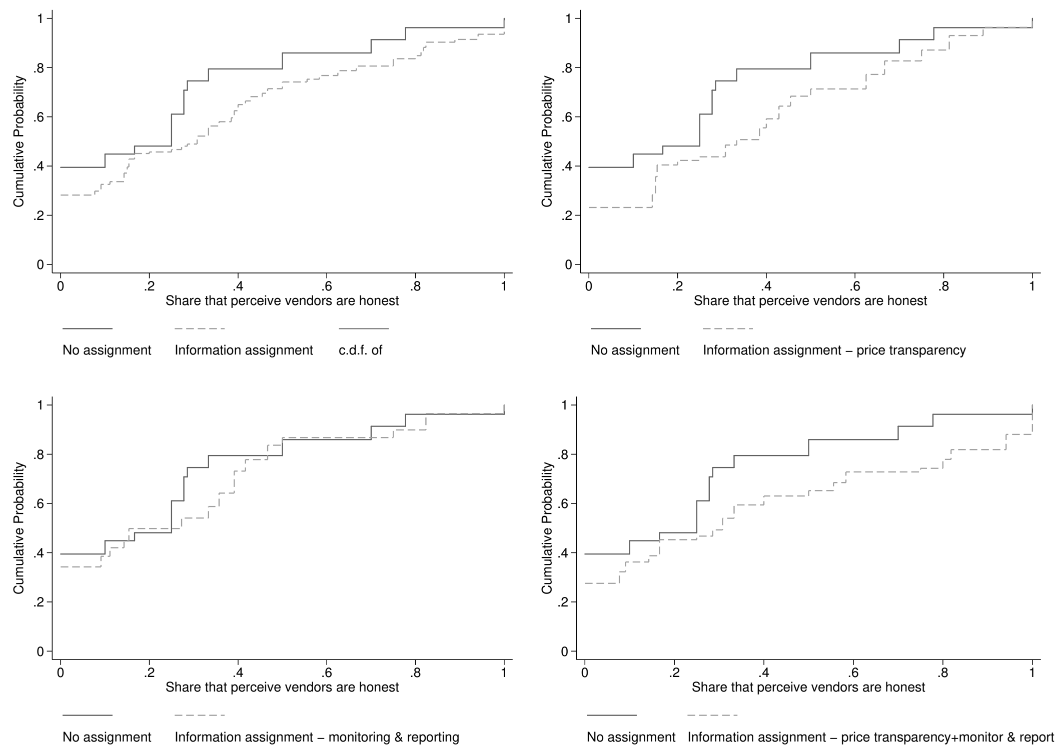
Table D.8: **Effect of treatment on business sales revenue**

Model: Linear-OLS	Sales (M-Money) per day (GHS)	Sales (Non-M-Money) per day (GHS)	Sales (Total) per day (GHS)
	(1)	(2)	(3)
Price Transparency (β_1)	523.6	154.1	788.8
	(222.0)	(86.59)	(295.7)
Robust S.E.	[154.5, 892.7]	[9.720, 298.5]	[295.5, 1282]
Clustered S.E.	[155.1, 892.0]	[10.09, 298.1]	[296.8, 1280]
Wild Bootstrap	[139.3, 879.0]	[8.048, 301.7]	[283.3, 1287]
Lee (2009) Attrition Bounds	<373.94, 580.32>	<-38.74, 58.52>	<335.1, 636.4>
Behaghel et al. (2015) Attrition Bounds	[419.0, 665.5]	[-38.79, 81.38]	[380.2, 744.0]
Monitor and Report (β_2)	418.4	45.69	438.3
	(259.8)	(81.69)	(360.3)
Robust S.E.	[-13.51, 850.4]	[-90.52, 181.9]	[-162.5, 1039]
Clustered S.E.	[-12.78, 849.7]	[-90.16, 181.5]	[-160.9, 1037]
Wild Bootstrap	[-29.54, 863.3]	[-90.9, 175.8]	[-134.7, 1029]
Lee (2009) Attrition Bounds	<75.36, 497.6>	[<-100.5, -31.44>	<-1.400, 449.32>
Behaghel et al. (2015) Attrition Bounds	[134.6, 419.1]	[-88.64, -39.12]	[60.84, 367.8]
Joint program: PT + MR (δ)	358.1	130.9	434.5
	(198.1)	(89.42)	(269.0)
Robust S.E.	[28.80, 687.4]	[-18.13, 280.1]	[-14.03, 883.2]
Clustered S.E.	[29.35, 686.9]	[-17.74, 279.7]	[-12.85, 882.0]
Wild Bootstrap	[39.79, 678.8]	[-14.97, 279.7]	[-4.134, 882.9]
Lee (2009) Attrition Bounds	<301.4, 372.5>	<5.117, 37.10]>	<334.9, 406.6>
Behaghel et al. (2015) Attrition Bounds	[270.3, 514.1]	[-39.60, 70.85]	[332.5, 571.8]
Market District F.E.	X	X	X
Baseline sales revenue	X	X	X
Controls	X	X	X
Observations	107	85	85
Mean of dependent variable (control)	792.8	239.5	1032
p-value (test: $\beta_1 = \delta$)	0.436	0.817	0.258
p-value (test: $\beta_2 = \delta$)	0.810	0.317	0.991
p-value (test: $\beta_1 = \beta_2$)	0.680	0.240	0.332
p-value (test: $\beta_1 + \beta_2 = \delta$)	0.096	0.575	0.109

Note: Market district is the randomization strata. Observations are at the vendor level. Controls in columns (1)-(3) include: age, marital status, ethnic group status, employment status, business experience, and bundled store status. 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 (≡vendor) level. Differential attrition bounds are reported. Results similar to post-double-selection LASSO estimates clustered at the local market level. At baseline, 75% of vendors bundle M-Money with other business services, hence the variation in sample sizes across columns.

E Further Results: Belief Updates and Heterogeneity

Figure E.1: **Distributions of subjective customer belief about vendor honesty at endline by treatment status**



Note: Figure plots distributions (CDFs) of customer belief about "honest vendor behavior" at endline for different experimental subsamples. **1**(.) is a logical indicator that equals 1 if argument in the parenthesis is true, 0 otherwise. Observations are at the customer level. Beliefs denote customers' perception that they are not being overcharged at vendor points (or perception that they have not experienced seller misconduct). In each local market, we compute the share of experimental customers who indicate they believe they are not experiencing misconduct (indicating belief in honest vendor behavior) at endline. From a Kolmogorov–Smirnov test for the equality of distributions, p-value $= 0.000$ for all cases.

Table E.1: **Subjective customer belief about vendor honesty increase at endline**

DV: 0-1 Indicator for belief about vendor honesty $\hat{\pi}$

	(1)	(2)	(3)
PANEL A			
Treatment: Information Assignment (β)	0.099	0.074	0.073
	(0.055)	(0.036)	(0.038)
	[0.008, 0.190]	[0.012, 0.135]	[0.010, 0.137]
Market District F.E.		X	X
Baseline belief about vendor honesty			X
Controls			X
Observations	943	941	941
Mean of dependent variable (control)	0.243	0.243	0.243
PANEL B			
Price Transparency (β_1)	0.120	0.118	0.120
	(0.072)	(0.052)	(0.053)
	[0.001, 0.240]	[0.031, 0.206]	[0.032, 0.208]
Monitor and Report (β_2)	0.033	-0.022	-0.026
	(0.066)	(0.047)	(0.0417)
	[-0.077, 0.143]	[-.0101, 0.055]	[-0.105, 0.052]
Joint program: PT + MR (δ)	0.140	0.119	0.121
	(0.080)	(0.046)	(0.048)
	[0.007, 0.274]	[0.0425, 0.196]	[0.041, 0.201]
Market District F.E.		X	X
Baseline belief about vendors' honesty			X
Controls			X
Observations	943	941	941
Mean of dependent variable (control)	0.243	0.243	0.243
p-value (test: $\beta_1 = \delta$)	0.823	0.989	0.988
p-value (test: $\beta_2 = \delta$)	0.208	0.006	0.004
p-value (test: $\beta_1 = \beta_2$)	0.261	0.016	0.012
p-value (test: $\beta_1 + \beta_2 = \delta$)	0.908	0.745	0.706

Note: Includes randomization strata (market district) dummies. Observations are at the customer level. In each market, we compute the baseline outcome as the share of experimental customers who think vendors do not overcharge transactions (belief about honest vendor behavior). Controls in column (3) include: gender, age, marital status, ethnic group status, employment status, education, and income. Clustered standard errors at local market level reported in parenthesis. 90% confidence intervals (CI) are reported in brackets. Panel A reports the pooled estimate of treatment effects, while panel B shows the effects separately for each information program. Results similar to post-double-selection LASSO estimates clustered at the local market level, to alternative inference procedures (Wild cluster bootstrap and permutation test clustered at the market level), and to the alternative beliefs measures (non-incentivized vs incentivized outcomes).

Table E.2: **BELIEF UPDATE: EFFECT OF INFORMATION SETS ON CORRECT CUSTOMER INFERENCE OF VENDOR MISCONDUCT**

DV: 0-1 Indicator for belief about vendor misconduct $(1-\hat{\pi})$

	(1)	(2)	(3)
Price Transparency (β_1)	-.178	-0.361	-0.328
	(0.086)	(0.079)	(0.093)
	[-0.321, -0.035]	[-0.493, -0.229]	[-0.484, -0.172]
x Objective Misconduct measure	0.132	0.318	0.266
	(0.163)	(0.119)	(0.134)
	[-0.139, 0.404]	[0.120, 0.517]	[0.043, 0.489]
Monitor and Report (β_2)	-0.119	-0.133	-0.110
	(0.118)	(0.082)	(0.093)
	[-0.316, 0.077]	[-0.270, 0.003]	[-0.264, 0.044]
x Objective Misconduct measure	0.271	0.280	0.247
	(0.147)	(0.123)	(0.135)
	[0.026, 0.516]	[0.075, 0.484]	[0.021, 0.472]
Joint program: PT + MR (δ)	-0.294	-0.338	-0.326
	(0.107)	(0.065)	(0.081)
	[-0.473, -0.116]	[-0.447, -0.230]	[-0.460, -0.191]
x Audit Objective Misconduct measure	0.360	0.326	0.305
	(0.160)	(0.111)	(0.119)
	[0.094, 0.627]	[0.141, 0.511]	[0.107, 0.503]
Objective Misconduct measure	-0.132	-0.228	-0.200
	(0.076)	(0.076)	(0.093)
	[-0.259, -0.005]	[-0.355, -0.101]	[-0.355, -0.046]
Market District F.E.		X	X
Baseline belief about vendor misconduct			X
Controls			X
Observations	678	678	678
Mean of dependent variable (control)	0.529	0.529	0.529
p-value (test: $\beta_1 = \delta$)	0.381	0.752	0.968
p-value (test: $\beta_2 = \delta$)	0.263	0.019	0.011
p-value (test: $\beta_1 = \beta_2$)	0.678	0.023	0.024
p-value (test: $\beta_1 + \beta_2 = \delta$)	0.983	0.149	0.333

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. Observations are at the customer level. Controls in column (3) include: gender, age, marital status, ethnic group status, employment status, education, and income. 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, to alternative inference procedures (Wild cluster bootstrap and permutation test clustered at the market level), and to the alternative beliefs measures (non-incentivized vs incentivized outcomes).

Table E.3: **Corroborative evidence for reputation effects**

Model: Linear-OLS		BUNDLED STORES			CUSTOMER ILLITERACY [EDUCATION]	
		1(Misconduct=Yes)	Misconduct, GHS		**1**(Misconduct=Yes)	Misconduct, GHS
Treatment: Information Assignment (β)		-0.138	-0.361		-0.044	-0.056
		(0.092)	(0.282)		(0.090)	(0.269)
		[-0.292, 0.015]	[-831, -0.108]		[-0.194, 0.105]	[-0.503, 0.390]
	x Bundled	-0.350	-0.927	**x** Illiteracy	-1.139	-3.443
		(0.174)	(0.427)		(0.539)	(1.635)
		[-0.639, -0.060]	[-1.63, -0.217]		[-2.035, -0.242]	[-6.16, -0.724]
	Bundled	0.234	0.654	Illiteracy	1.063	3.546
		(0.162)	(0.378)		(0.524)	(1.730)
		[-0.036, 0.505]	[0.26, 1.283]		[0.191, 1.935]	[0.670, 6.422]
Baseline misconduct		X	X		X	X
Market District × ... Transaction × Date F.E.		X	X		X	X
Observations		332	332		332	332
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, baseline outcomes, and additional controls (age, marital status, ethnic group status, employment status, business experience, and bundled store status). Observations are at the vendor x transaction type x transaction 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). Bundled is a 0-1 indicator for whether vendor operates bundled stores (M-Money and non-M-Money services). Illiteracy is defined as the market-level fraction of consumers around the vendor that have no formal education at baseline.

Table E.4: **Corroborative evidence for reputation effects**

Model: Linear-OLS		BUNDLED STORES			CUSTOMER ILLITERACY [EDUCATION]	
		1(Misconduct=Yes)	Misconduct, GHS		**1**(Misconduct=Yes)	Misconduct, GHS
Price		-0.065	-0.134		-0.014	-0.084
Transparency (β_1)		(0.091)	(0.296)		(0.113)	(0.374)
		[-0.217, 0.086]	[-0.627, 0.358]		[-0.202, 0.173]	[-0.707, 0.538]
	x Bundled	-0.573	-1.442	**x** Illiteracy	-1.227	-2.228
	(b_1)	(0.233)	(0.642)	(b_1)	(0.704)	(2.112)
		[-.962, -0.185]	[-2.510, -0.373]		[-2.397, -0.056]	[-5.740, 1.282]
Monitor and		-0.136	-0.375		-0.033	-0.016
Report (β_2)		(0.098)	(0.311)		(0.101)	(0.307)
		[-0.299, 0.026]	[-0.892, 0.141]		[-0.201, 0.135]	[-0.527, 0.494]
	x Bundled	-0.293	-0.710	**x** Illiteracy	-1.344	-4.215
	(b_2)	(0.188)	(0.469)	(b_2)	(0.610)	(2.159)
		[-0.677, 0.019]	[-1.490, 0.070]		[-2.359, -0.329]	[-7.803, -0.626]
Joint program:		-0.135	-0.338		-0.076	-0.112
PT+MR (δ)		(0.098)	(0.309)		(0.101)	(0.324)
		[-0.299, 0.028]	[-0.853, 0.176]		[-0.245, 0.092]	[-0.652, 0.427]
	x Bundled	-0.381	-1.078	**x** Illiteracy	-0.878	-2.875
	(d)	(0.178)	(0.454)	(d)	(0.635)	(1.625)
		[-0.677, -0.085]	[-1.834, -0.322]		[-1.933, 0.177]	[-5.577, -0.173]
	Bundled	0.250	0.675	Illiteracy	1.124	3.720
		(0.168)	(0.393)		(0.531)	(1.801)
		[-0.029, 0.530]	[0.021, 1.328]		[0.241, 2.007]	[0.725, 6.714]
Baseline misconduct		X	X		X	X
Market District × ...		X	X		X	X
Transaction × Date F.E.						
Observations		332	332		332	332
Mean of dep var (control)		0.294	0.778		0.294	0.778
p-value (test: $b_1 = d$)		0.221	0.461		0.555	0.757
p-value (test: $b_2 = d$)		0.392	0.315		0.430	0.523
p-value (test: $b_1 = b_2$)		0.068	0.153		0.838	0.334
p-value (test:$b_1 + b_2 = d$)		0.069	0.148		0.066	0.285

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 (age, marital status, ethnic group status, employment status, business experience, and bundled store status). 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). Bundled is a 0-1 indicator for whether vendor operates bundled stores (M-Money and non-M-Money services). Illiteracy is defined as the market-level fraction of consumers around the vendor that have no formal education at baseline.

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:

- 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: (1) Ask customer to repeat information provided. (2) Ask customer to rate the usefulness of the provided information for their financial decision-making 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 - MONITORING AND REPORTING, MR

Our message is simple. We want to remind you:

- If you suspect any discrepancy or glitches in tariffs 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: (1) Ask customer to repeat information provided. (2) Ask customer to rate the usefulness of the provided information for their financial decision-making 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 in tariffs 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: (1) Ask customer to repeat information provided. (2) Ask customer to rate the usefulness of the provided information for their financial decision-making 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 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 they should not pay fees – deposit should be credited to their account, they should 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: (1) Ask vendor to rate the usefulness of the provided information for their 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 - MONITORING AND REPORTING, MR

- If they suspect any discrepancy or glitches in tariffs 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: (1) Ask vendor to rate the usefulness of the provided information for their 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, they should 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 in tariffs 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: (1) Ask vendor to rate the usefulness of the provided information for their 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 G.1: **Vendor banking points**

Note: Providers – MTN Mobile Money, AirtelTigo Money, Voda Cash, GCB Ltd.'s G-Money (new provider)

H Auditors' Training - Measuring Seller Misconduct

INSTRUCTIONS:
VENDOR-BASED APPROVED TRANSACTION TARIFFS

- Welcome: You have been "assigned" to vendor shops, where you will make specific Mobile Money transactions.
- You will be required to use the same language while transacting at vendor shops (details below).
- Our focus will be vendor- or merchant-based Mobile Money transactions.
- Throughout, we pay fees whenever we are sending money at the vendor to guarantee the receiver receives XGHS-amount.
- Most times picking up money from the vendor should be free (details below).
- Here are the approved rates that we will be working or transacting with at vendors' premises (Let's memorize them. You will be given copies, so you can refer these rates any time you are in doubt):

KEY: TRANSACTIONAL CODES
OVER-THE-COUNTER, OTC

- T1: Put GHS50 on someone's (XX/Yourselves) M-Money wallet {GHS50 => PAY GHS0.5}
- T2: Put GHS160 on someone's (XX/Yourselves) M-Money wallet {GHS160 => PAY GHS1.6}
- T3: Put GHS1100 on someone's (XX/Yourselves) M-Money wallet {GHS1100 => PAY GHS10}

TOKEN

- T4: Send a Token of GHS50 to someone (XX/Yourselves) {GHS50 => PAY GHS2.5}
- T5: Send a Token of GHS1100 to someone (XX/Yourselves) {GHS1100 => PAY GHS55}
- T6: Receive a Token of GHS50 from someone (XX/Yourselves) **{GHS50 => FREE}
- T7: Receive a Token of GHS1100 from someone (XX/Yourselves) **{GHS1100 => FREE}

FALSIFY [INSTANT VERIFIABILITY PROVIDED BY PROVIDER]

- T8: Put or Cash-in GHS50 on your own M-Money wallet {GHC50 => FREE}
- T9: Put or Cash-in GHS110 on your own M-Money wallet {GHS110 => FREE}
- T10: Take or Cash-out GHS50 from your own M-Money wallet {GHS50 => FREE}

ACCOUNT OPENING

- T11: Buy a new SIM card {SIM (or ATTEMPT it) => PAY GHS2}
- T12: Then use T11 to register for Mobile Money Account {REGISTER (or ATTEMPT it) => FREE; initial deposit of GHS5 minimum required but this GHS5 must be on your account, merchant should not take it, verify}.

TRANSACTION APPROACH

**DURING VISIT (Very simple language, no deviations allowed): `Good morning/afternoon/evening.`

`I want to make a M-Money transaction [USE CODES: T1...T12].`

- `Present necessary details: phone number, and sender or recipient details`
- `Thank you for your service`

**AFTER VISIT: Immediately complete the questionnaire (see Table H.1) right after the transaction using your Tablets.

ADDITIONAL NOTES

- [1] The order of transactions to make at vendor points will always be determined (randomly) by the CAPI data entry software on your Tablets (you don't choose it). CAPI will also display the various tariffs in case you are in doubt.
- [2] Please leave spaces blank if a specific transaction-type is not feasible (the software will randomly switch to another transaction-type).
- [3] Practicing: let's take turns to practice repeatedly the transaction approach, using yourselves as vendors and other nearby M-Money vendors. Your supervisors will be monitoring... Any questions or clarifications? Let's discuss.

Table H.1: **QUESTIONNAIRE: AUDITOR'S UNIQUE ID...**

Q0	Q1a	Q1b	Q1c	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13
No.	VISIT DATE			Locality code?	"Rep" Vendor code?	TRANSACTION TYPE? USE CODES: T1...T12	Transaction OVERCHARGED? 1=YES; 2=NO=>Q7	How much DIFFERENCE? GHS	Transaction successful? 1=YES; 2=NO; 3=NO CASH	Appx wait time transaction took? MINS	Related to Vendor just visited? 1=YES; 2=NO => Q11	How are you related to Vendor? 1=RELATIVE; 2=FRIEND; 3=OTHER	Vendor's Gender? 1=MALE 2=FEMALE	Vendor involved in non-Mobile Money businesses? 1=YES 2=NO	Tariffs posted? 1=YES; 2=NO
	MM	DD	TIME												
1															
2															
3															
4															
5															
6															
7															
8															
9															
10															
11															
12															
.															
.															
.															
.															
.															
.															

. . . . . .

I Definition of Relevant Select Variables – Questions

Consumer outcomes:

1. Uptake of transactional services:

   (a) `Consider the last month - What is the typical value of Mobile Money transactions (cash-in and cash-out) you conducted in this locality per week? (NOTE 1: Please only include transactional estimates in seven (7) days. NOTE 2: Ask the customer to refer to his or her records/diaries for past days in case forgotten)` **`GHS/week...`**

   (b) 0-1 Indicator for whether consumer used M-Money (last month): **`If 1(a) > 0GHS`**

2. Savings likelihood:

   (a) `Consider the last month - From a scale of 1 (low) to 5 (high), how likely are you to save generally on M-Money now?` **`1=Very low, 2=Low, 3=Medium, 4=High, 5=Very high`**

   (b) 0-1 Indicator for whether consumer saved on M-Money (last month): **`If uptake if 2(a) > 2 (median)`**

3. Unexpected shocks mitigation: `Have you experienced any of the following common shocks within the past three (3) months where you or your household did not have enough cash or M-Money resources on hand to cover costs?` **`0=No, 1=Yes`**

   (a) `Death of a close person, relative, or friend? (death)`

   (b) `Unexpected loss of revenue or wages, e.g., via unemployment, bad business? (revenue)`

   (c) `Unexpected illness, accident, or health condition? (health)`

   (d) `Any general floods or droughts? (weather)`

   (e) `Unexpected (high/low) food prices? (prices)`

   (f) `Other unexpected shocks (i.e., weather, input prices, diseases) that affect your farm production or house expenses? (house expenses)`

   (g) Shocks experience = 1 `if any of` **`3(a) - 3 (f) = Yes`**

4. Subjective beliefs (perception about seller misconduct) - Highly correlated responses:

(a) Non-incentivized beliefs statement: `Consider the past four (4) months (interventions in force) - In my [research enumerator's] view, M-Money vendors generally overcharge customers' transactions at vendor points (seller misconduct).` **`1=Agree, 2=Disagree`**

(b) Incentivized beliefs: `What's your [customer's] estimate of the % of others (all vendors and customers in this locality) that will Agree with statement 4(a)?` **`%...`**

- NOTE: Customers are jointly asked to guess the percentage of others (all vendors and customers in their locality) who would `Agree` with statement 4(a) (beliefs about others' beliefs). To incentivize their reports, among all respondents in a locality, the respondent with the closest guess (to the locality-level estimate) immediately receives 10GHS after all respondents have answered, either in cash to their M-Money accounts or in-kind through a phone calling-credit. All respondents are informed of this payoff before answering.

(c) `Consider the past four (4) months (interventions in force) - Any experiences of overcharged M-Money fees at Mobile Money centers?` **`0=No, 1=Yes`**

Business outcomes:

1. Sales revenue (Mobile Money): `Consider the last month - What was the total sales the Mobile Money business made daily? (NOTE 1: think about all cash-in and cash-out transaction volume records. NOTE 2: Ask the vendor to refer to his or her records/diaries for past days in case forgotten)` **`GHS/day...`**

2. Sales revenue (non-Mobile Money): `What was the total sales the non-M-Money business made daily considering the last month (NOTE 2: Ask the vendor to refer to his or her records/diaries for past days in case forgotten)?` **`GHS/day...`**

3. Total sales revenue = `1+2,` **`GHS/day...`**

Control set:

1. Bundling (bundled stores): `Currently do you [vendor] offer other services at your business center, other than M-Money? Example - sell provisions, airtime, phones, accessories, appliances, etc.` **`0=No, 1=Yes`** `(Alternative measure: see Q12 in Table H.1)`

2. Tariff posting: `Consider the last thirty (30) days or last month: How often do you [vendor] post your tariff sheets at your banking point in a typical week?` **`1=Never (less than 1 time in 7 days), 2=Sometimes (1-2 times in 7 days), 3=Often (3-4 times in 7 days), 4=Very often (5-7 times in 7 days)`** `(Alternative measure: see Q13 in Table H.1)`
3. Age: `What is your [vendor/customer] age?` **`Years`**
4. Married: `Are you [vendor/customer] married?` **`0=No, 1=Yes`**
5. Akan: `What is ethnicity do you [vendor/customer] identify with?` **`1=Akan, 2=Ewe, 3=Ga-Dangme, 4=Others`**
6. Self-employed: `Are you [vendor/customer] self-employed?` **`0=No, 1=Yes`**
7. Business experience: `How long have you been in the Mobile Money service business?` **`Years`**

Poverty Scorecard (Schreiner 2005):

1. `How many members does the household have?` **`Use codes: 0=Eight+ 4=Seven 9=Six 13=Five 14=Four 21=Three 24=Two 29=One`**
2. `Are all household members ages 5 to 17 currently in school?` **`0=No 2=Yes 3=No one ages 5 to 17`**
3. `Can the male head/spouse read a phrase/sentence in English?` **`0=No 2=No male head/spouse 5=Yes`**
4. `What is the main construction material used for the outer wall?` **`0=Mud bricks/earth, wood, bamboo, metal sheet/slate/asbestos, palm leaves/thatch (grass/raffia), or other 5=Cement/concrete blocks, landcrete, stone, or burnt bricks`**
5. `What type of toilet facility is usually used by the household?` **`0=No toilet facility (bush, beach), or other 4=Pit latrine, bucket/pan 4=Public toilet (e.g., WC, KVIP, pit pan) 6=KVIP or WC`**
6. `What is the main fuel used by the household for cooking?` **`0=None, no cooking 6=Wood, crop residue, sawdust, animal waste, or other 13=Charcoal, or kerosene 22=Gas or electricity`**

7. Does any household member own a working box iron or electric iron? **0=No 4=Yes**

8. Does any household member own a working television, video player, VCD/DVD/MP3/MP4 player/iPod, or satellite dish? **0=No 2=Only television 3=Video player, VCD/DVD/MP3/MP4 player/iPod, or satellite dish (regardless of T.V.)**

9. How many working mobile phones are owned by members of the household? **0=None 4=One 8=Two 10=Three+**

10. Does any household member own a working bicycle, motor cycle, or car? **0=No 3=Only bicycle 8=Motor cycle or car (regardless of bicycle)**

View publication stats