Nationwide Diffusion of Technology: Experimental Evidence from Multiple Networks*

Deivy Houeix†

[Click here for the most recent version]

August 5, 2026

Abstract

To what extent and why do firms benefit from locating near one another? I separate the relationships that proximity bundles by randomizing access to a digital payment technology among 2,196 taxi businesses in Dakar, Senegal, and tracing its spread through new data on three networks covering most informal firms in the country: social ties (hashed phone-contact directories), location (agent GPS pings), and shared customers (mobile money transaction ledgers). Each network carries a distinct mechanism. Social ties transmit knowledge and raise adoption across industries and cities. Co-location reduces adoption through competition, as firms withhold information from nearby rivals; the effect halves every 186 meters. Shared customers create strategic complementarities, raising adoption even between firms with no social connection. Because proximity bundles countervailing relationships, closeness can appear beneficial in some settings and harmful in others. A structural decomposition shows that targeting socially connected (high-degree) firms raises adoption by 58% over random assignment. Concentrating treatment at the largest garages, the most visible targeting rule, does no better.

*I am grateful to Isaiah Andrews, David Atkin, Abhijit Banerjee, Luisa Cefalà, Esther Duflo, Dean Eckles, Tishara Garg, Justine Knebelmann, Sahil Loomba, Vishan Nigam, Ben Olken, Kunal Sangani, Frank Schilbach, Tavneet Suri, Nicholas Swanson, Rob Townsend and participants in the various MIT lunches for valuable suggestions and comments. I am also indebted to an amazing research team, particularly Koffi Adjassou, William Amedanou, Justin Chery, Martin Duflo, Cheikh Gaye, Ismael Keita, Loic Laurent, Sara Niang, Yacine Ouahioune, and Francois Soleiman for their outstanding field assistance, as well as Hannah S. Ono, Prabhav Sharma, Alexis Yi, and Lawrence Yong at MIT for their excellent research assistance. I am especially thankful to the entire team at Wave Mobile Money, the implementing partner company, for trusting me to carry out this study independently. I gratefully acknowledge financial support from JPAL-DigiFI Africa, PEDL, The Weiss Fund, and Shultz, and data collection support from SDI. I received IRB approval from MIT, protocol #2012000286. This RCT was registered as AEARCTR-0009155.

† Department of Economics, Harvard University

To what extent, and why, do firms benefit from locating near one another? Since Marshall (1920), the leading paradigm has been that proximity speeds the flow of ideas between firms: the "mysteries of the trade . . . are as it were in the air." A large literature establishes that geography matters: firms in denser environments are more productive and innovate more (Moretti, 2021; Baum-Snow et al., 2024), and diffusion between firms is a first-order determinant of productivity and growth (Arrow, 1962; Romer, 1990; Grossman and Helpman, 1991). Yet separating the mechanisms governing agglomeration spillovers has received less attention.

The difficulty is that proximity bundles several relationships that are rarely observed separately. Firms that are close to each other tend to transmit useful knowledge from one to the other (*learning*). But a competitor next door may have incentives to keep a new technology to itself (*competition*). And competing firms often share customers, so that one firm's adoption raises another's return (*demand*), especially for two-sided technologies like digital payments.

Unpacking these channels requires three ingredients that are rarely available together: a clean technology shock, the adoption data of connected firms, and, importantly, the distinct networks along which information and market incentives operate. This paper brings the three pieces together. I randomize early access to a digital payment technology among 2,196 taxi businesses in Dakar, Senegal, with 58% assigned to treatment. I trace the technology's diffusion well beyond the experimental sample using three economy-wide data sources from the country's leading mobile money company: hashed phone-contact fingerprints covering roughly 7 million people, about 70–80% of the adult population; the complete ledger of mobile money transactions; and agents' GPS pings covering the locations of most informal businesses and mobile money users. The technology spreads on its own to more than 230,000 businesses: other taxis, firms in other industries, and firms in cities hundreds of kilometers from Dakar. This breadth allows me to identify mechanisms often difficult to pin down: standard spillover designs measure indirect effects on a held-out control group inside the experiment that is often local and defined by the study itself (Crépon et al., 2013; Baird et al., 2018). Recent evidence (Chaurey et al., 2025) goes further and traces social learning directly from firm to firm, though still within one sector and a narrow radius. Here, I measure the organic diffusion of the technology across the whole economy from a single experimental seed.

The mechanisms I highlight are general, but the taxi industry is well suited to separating them. For a firm rooted in one location, the three networks largely coincide: its neighbors are potentially its competitors, its customers, and its acquaintances. Drivers move through the city, so the networks further come apart. Social ties stretch across the country,

customers span Dakar, and competition is particularly important in garages, where about 14% of drivers compete for the same passengers.

I measure the three networks separately, each carrying a distinct mechanism: learning through social ties, competition through geographic proximity, and demand complementarities through shared customers. The *social network* is constructed from contact-book fingerprints, which are, to my knowledge, novel in economics. When a user first logs in, the company requests a hashed fingerprint of their phone contact book, with 86% consenting. Concretely, each saved number is hashed, so I observe the set of buckets a user saved and not the individual phone numbers. I measure tie strength between two users using their bias-corrected contact-book overlap—how much of their contact lists they hold in common, net of expected hash collisions—whether or not either saved the other's hashed number. The approach draws on the computer science literature on secure set-similarity measurement, which starts with Broder (1997), and was extended to privacy-preserving settings by De Cristofaro et al. (2013) and Swamidass et al. (2015). Most firms in this setting are owner-operated. The fingerprints therefore provide unusually direct evidence on the social networks of firm owners, covering both mutual contacts and contacts with high overlap. The resulting network links all experimental taxi owners and drivers to other mobile money users across Senegal, at varying degrees of tie strength. I validate the recovered relationships with an independent relationship-listing survey. I draw a random sample of each respondent's past transaction partners and ask them to describe the relationship. This survey confirms that corrected overlap tracks social closeness—wives share 20.8% of contacts, taxi colleagues 13.6%, family members 11.4%, and friends 6–9%, while serendipitous contacts (phone numbers a respondent could not place when asked, typically one-time transaction partners) average only 3.2% and customers average 1.9% corrected overlap, just below the 2% mark. Therefore, I define a *community tie* as a corrected overlap of at least 2%, so that the typical customer relationship falls outside the social network (results are not sensitive to this definition, and thresholds of 2, 5, and 10% are analyzed throughout the paper).

The other two networks come from location and transaction data. First, the *geographic network* captures physical proximity between firms, in particular for the garage-based taxi drivers competing for passengers. Geolocated cash-in and cash-out transactions at mobile money agents not only locate users geographically but also record who moves between cities throughout the experiment. Second, the *customer network* is built from the mobile money transaction ledger: two firms are linked if they share regular customers—people who make repeated payments to both firms, including at least one experimental driver. This ledger is a rare window onto customer-sharing links, which are typically

unobservable in informal economies, where cash transactions and the absence of value-added-tax records leave no trail of who transacts with whom. The network is measured both before the experiment (peer-to-peer, when business payments did not exist) and after (peer-to-business, when treated and control taxis gained access to digital payments and appeared as payees).

My empirical strategy leverages the random assignment: treated drivers gain early access to the technology while control drivers are excluded for the first six months, generating clean exogenous variation in firms' exposure to early adopters. Following Borusyak and Hull (2023), I measure each firm's exposure as the share of its experimental taxi connections assigned to treatment, and identify spillovers from variation in that share conditional on its *expected* exposure under the randomization. Expected exposure depends on how connected a firm is and on which strata its experimental contacts fall into, so conditioning on it holds network position fixed and removes the confound that more central firms, those with many contacts or many nearby rivals, may differ from less central ones for reasons unrelated to treatment.

I have three main findings. First, I identify *social learning* by documenting spillover effects on firms in other industries that are connected to treated taxi drivers. A non-taxi firm neither competes with a taxi nor typically shares its customers, so competition and demand are muted; a positive adoption response to a treated taxi contact therefore primarily reflects learning. In the pooled reduced form the effect is significant on strong ties: non-taxi firms connected to treated taxis through high corrected-overlap contacts ($\geq 5\%$) adopt at a rate $+31.0\%$ above the control mean, while the pooled weak-tie estimate is positive but imprecise. Per-tie transmission is approximately uniform across the strength distribution. Survey evidence corroborates this channel: taxi businesses report sharing advice or information with about 74% of the people they know, and having discussed the technology with, or recommended it to 12% of them. The spillovers appear in the community network but not among mutual contacts. Mutual contacts are more likely to be infra-marginal. Control drivers were told about the technology, so their shared contacts learn as much as they would learn from a treated tie. Community contacts instead learn by observation, either directly or through second-order links (treated drivers anecdotally report posting photos of the QR code on social media).

Spillovers expand to firms in cities beyond Dakar, where effects are roughly twice as large. Such reach contrasts with the spirit of Marshall's local spillovers and with what has been documented in the literature so far. A natural candidate for how diffusion can cross cities is seasonal migration, which is pervasive in this context—roughly a third of Dakar-based business owners are observed moving outside the capital during the experiment.

Yet I find that physical travel is not necessary for transmission: the effect is precisely estimated through treated drivers who never visited the destination city, and it is significant among receiving businesses never observed in Dakar. The effect of exposure through drivers who did travel is larger in magnitude but imprecisely estimated, consistent with migration reinforcing the remote channel. This shows the extent to which information travels across space in this economy, within the community network, without requiring face-to-face interaction.

Second, I identify *competition* in the geographic network, within the taxi industry. A treated driver who finds the technology profitable has a strategic reason not to share that finding with co-located rivals, since their adoption would erode the shared customer base he currently captures—similar to the knowledge-hoarding mechanism documented in contemporaneous work among farmers (Cefalà et al., 2025). This force predicts negative same-industry spillovers. Consistent with this mechanism, adoption among non-experimental taxi businesses falls with exposure to treated drivers at nearby garages, and significantly so within the community network: the effect is stronger within 100 meters and decays with distance (half-life of 186 m). Additional evidence confirms that the competitive force operates through physical co-location. First, taxi businesses outside Dakar—same industry, but too distant to compete for the same passengers—show no negative effect and even a significant positive effect among mutual phone contacts. Second, friendship reverses the sign: a driver with a treated taxi *friend*, as identified in the survey, adopts at a significantly higher rate than one with a control friend, so the positive forces traveling on very close ties outweigh the competition between firm owners.

Third, I identify *demand complementarities* in the customer network, a mechanism that requires no social tie. The channel is especially central for payment technologies: payments are two-sided, so a business's return to adoption depends on whether its customers want to pay digitally, and demand can affect adoption beyond the learning and competition mechanisms. Using shared regular customers measured after the experiment, I find that when such a customer pays a treated driver digitally, the driver's customer-linked peers outside the experiment become 35.7% more likely to adopt. The effect remains large ($+23.6\%$) among businesses with no social connection to the treated driver, ruling out social learning between firms as the only explanation, and is positive but imprecisely estimated across industries. The customer network thus identifies a demand effect, captured through second-order links between firms, regardless of their social ties. The sparse pre-experiment customer network, built from peer-to-peer transfers and therefore immune by construction to treatment-induced link formation, yields a similar conclusion.

I then develop a structural model to quantify the three forces on a common scale. I

estimate a discrete-time adoption model in which each firm's adoption depends on learning through social ties, competition that decays with distance from garages, and demand through shared customers. I identify each channel using networks where the other two channels are muted: cross-industry moments identify learning, since competition and demand across sectors are limited; customer links among non-garage taxis not in the same community network isolate demand effects; and the distance gradient of garage coefficients identifies the level and decay of competition. Survey moments on adoption by relationship type are used as untargeted validation. Learning operates at a uniform rate across the community-tie strength distribution: the strong-tie and weaker community-tie coefficients are statistically indistinguishable ($+0.271$ vs. $+0.267$ on the latent scale), meaning any genuine community tie transmits the technology. Competition is intense but highly local: at 50 meters from a garage the competitive force is roughly 24 times the influence of full community-network exposure, falling to 11 times by 500 meters; the two equalize at roughly 2 km from a garage. Customer demand generates large positive complementarities, including between taxis with no social tie. The estimated parameters help explain some reduced-form patterns. First, the pooled within-industry spillover in the community network is a near-exact null because learning and competition offset each other, as predicted by the model. Second, the negative strong-tie coefficient emerges because same-industry strong ties are spatially correlated with the garages that carry competition. The model also correctly predicts the signs of four untargeted survey moments, including the fact that taxi friends raise adoption even though they are competitors, because demand outweighs competition, and that serendipitous contacts also raise adoption because learning travels through weak ties as well.

This structural estimation allows me to run targeting-policy counterfactuals. Instead of rolling the program out at random, a social planner could have targeted socially connected (high-degree) taxi businesses, which, through social learning and customer demand, would have delivered 58% more adoption than random assignment. Garage-based rules—the most visible targeting strategy—capture almost none of these gains: concentrating treatment at the largest garages, where competitive crowding is maximal, performs marginally worse than random. The gap between the best and worst rules emerges from the interplay of the social, customer, and geographic spillovers.

This paper contributes to three literatures. First, it unbundles key mechanisms governing agglomeration spillovers. A large literature shows that firms are more productive and innovative when they locate near one another (Marshall, 1920; Moretti, 2021; Baum-Snow et al., 2024; Chaurey et al., 2025), in part through face-to-face contact (Cai and Szeidl, 2017; Atkin et al., 2024). Yet estimated spillovers between firms are sometimes positive (Bloom

et al., 2013; Myers and Lanahan, 2022) or negative (Cefalà et al., 2025), depending on the setting. These estimates are hard to reconcile because proximity bundles various relationships: firms that are close tend to know one another, compete, and share customers, so the estimated effect of proximity can take either sign. By observing the social, geographic, and customer networks within a single experiment, I make progress in separating them and find that the forces that spread the technology run through social and customer ties, ties that are largely independent of space, while physical proximity within the taxi industry slows diffusion, as nearby rivals withhold what they learn. Where social and customer ties are themselves spatially concentrated, using distance as a proxy may be reasonable. The competitive cost, long posited even outside economics (Foster, 1965), bites where neighbors compete for the same customers. This bundling may be particularly strong in the local, non-tradable service sector, making it a natural setting to disentangle these agglomeration spillovers—and an increasingly important sector in development (Rodrik, 2016) where evidence remains limited.

Second, the paper makes progress on a long-standing identification problem in the literature on network effects and the adoption of digital payments (Björkegren, 2018; Fafchamps et al., 2023; Alvarez et al., 2023; Crouzet et al., 2023; Higgins, 2024; Crouzet et al., 2024): separating social learning from strategic complementarities. Both predict higher adoption when connected firms adopt. Because payments are two-sided and a firm's incentive to adopt rises with its customers' digital use, the two mechanisms are easily conflated when the data capture only proximity or aggregate exposure. Recent work identifies complementarities from policy-induced geographic variation (Crouzet et al., 2023; Higgins, 2024) and structural models (Alvarez et al., 2023), while noting that learning is hard to quantify in the same setting. I show that within the taxi sector, customer demand is the dominant force, consistent with the recent emphasis on complementarities, while social learning is the primary diffusion channel across industries and cities.

Third, the paper expands the measurement frontier for firm networks in developing economies, where most firms are informal, owner-operated, and absent from administrative records. In these settings, the owner's personal and business networks are deeply intertwined (Bernhardt et al., 2019). Diffusion studies have often been bounded by what a survey can enumerate, typically one industry in a handful of villages or neighborhoods (Cai et al., 2015; Chuang and Schechter, 2015; Suri and Udry, 2024; Duflo et al., 2026; Chaurey et al., 2025), or by administrative records on formal firms in rich countries (Bloom et al., 2013). I leverage hashed contact-book fingerprints, to my knowledge new to economics, which map the social ties of roughly 7 million people (70–80% of Senegal's adult population) without directly revealing phone numbers. I also use the mobile money

ledger to measure the customer links that cash typically leaves untraced. Relationships that earlier work inferred from physical proximity or reconstructed from surveys thus become observable at national scale. The networks I measure reveal the limits of the proxy: ties can be sparse and stretch far beyond the neighborhood, with over half of the members of taxi owners' community networks living outside the capital. Because such data now exist across most lower-income countries, the same measurement is feasible in many other contexts, bringing the questions asked of formal production networks (how firms find partners, how firms upgrade, how shocks and technologies propagate) within reach for informal economies (Bernard and Moxnes, 2018; Breza et al., 2019; Verhoogen, 2023).

The rest of the paper proceeds as follows. Section 1 describes the setting, experimental design, and data. Section 2 documents empirical patterns in firms' social networks. Section 3 presents the empirical framework. Section 4 provides causal estimates of technology adoption spillovers across industry and space. Section 5 examines the mechanisms driving diffusion. Section 6 develops the structural model and estimation. Section 7 presents counterfactual targeting exercises. Section 8 concludes.

1 Setting, Experiment, and Data

1.1 The Technology and Market Environment

The paper studies the diffusion of a new digital payment technology introduced in Senegal in 2022. The product is a peer-to-business mobile payment tool that allows customers to pay a business directly from an existing mobile money account by scanning a QR code displayed by the business, at a 1% fee for drivers (see Figure 1).[1] In the taxi sector, the technology changes payment in two economically meaningful ways. First, business payments are effectively irreversible, unlike ordinary person-to-person transfers that can often be reversed through a simple call to the provider. Second, the QR interface makes payment substantially more convenient than manually typing a driver's phone number into the mobile money application. These features matter in a cash-intensive sector such as taxis, where transactions are frequent, small, and operationally costly.

The technology is especially well suited to studying diffusion because adoption is one-sided. Customers do not need to adopt a new consumer technology; they use existing

[1] I deliberately do not focus on the 1% fee barrier in this paper. The technology is popular, with over two million unique Senegalese users making business payments during the experiment, though cash remains dominant. All drivers who adopted the technology were aware of the fee, and the primary focus of this study is adoption, not subsequent usage. Additionally, for all drivers in this experiment, the fee was waived for the first CFA 50,000 (approximately USD 83) collected digitally.

mobile money accounts that were already widespread in Senegal. The adoption decision therefore lies with the business owner. At the same time, the technology was not specific to taxis. The same business-payment product was rolled out across many sectors—restaurants, pharmacies, boutiques, mechanics, and street vendors, among others—during the same period (see a subset of business adopters in Figure 1). This makes the setting unusually useful for studying whether the same technology diffuses beyond the sector in which access is initially randomized.

The market environment also helps isolate the forces behind adoption. The monetary cost of adoption is low. The main adoption cost is time: the business must locate a registration agent, provide the required paperwork, and install the business application. In practice, the relevant requirements are possession of an official ID, a personal mobile money account, and a smartphone.[2] These features matter because they make the technology similar to many other digital business tools in low- and middle-income settings: adoption is constrained much more by information, trust, and coordination than by large fixed capital costs.

The broader rollout was rapid and followed an S-shaped curve (see Jack and Suri (2014); Suri (2017) for similar digital technologies). At the start of the experiment, essentially no taxis had adopted the technology, and adoption outside the taxi sector was still limited. Over the following months, however, the same product spread quickly across the economy: within a year, nearly 200,000 businesses in Senegal had adopted the digital payment technology (see Figure 2).[3] By early 2024, data show that approximately 17,000 taxis and 220,000 other businesses had adopted the technology, suggesting that by the end of the experiment in October 2022, a substantial share of the potential non-taxi business population had already adopted, while taxi adoption was still growing rapidly ($\sim$7,000 of an eventual $\sim$17,000 by early 2024). The technology adoption was not only located in Dakar, the capital city, where the experiment took place, but also across many cities in Senegal (see Figure A8). This rapid expansion is an asset: it allows us to ask whether a randomized intervention in one urban sector can generate diffusion of the same technology across industries and across cities.

The taxi sector in Dakar provides a particularly informative experimental environment. Taxis are numerous, socially connected, and economically important (taxis are ubiquitous in Dakar, employing about 4-6% of the urban male population, with about

[2]Smartphone adoption in urban Senegal is relatively high, with the vast majority of taxi businesses in this experiment possessing one. This high rate is also documented across Sub-Saharan Africa more broadly—about 51% including rural areas (Figure A1).

[3]The company has one main competitor, less popular but that started digitizing payments around the same time. Some businesses adopted both technologies to satisfy more customers.

21,000 active taxis in 2019 (CETUD)).[4] The sector is predominantly informal and includes two key stakeholders: the taxi business owners and the drivers. In this industry, the owner is the legal entity that owns the car and the majority of the experimental sample consists of single taxi owners who drive their own taxis. In this study, I have grouped all drivers and owners together under the term "taxi businesses," each considered as a business entity. Drivers interact with many customers each day, regularly encounter other drivers at garages and waiting points, and frequently exchange information. On average, a taxi driver meets about 14 customers a day, and 90% of drivers report engaging in conversation with customers at least half of the time, with 50% doing so most of the time. Additionally, a typical taxi driver reports interacting and regularly spending time with about 15 other taxi drivers. These features make all three channels studied in the paper plausible ex ante: social learning through owner networks, competition among nearby same-industry firms, and customer-demand complementarities when customers begin to expect digital payment.

Technological Benefits The economic impact of the technology on businesses is the focus of a separate paper (Houeix, 2025): from the point of view of this paper, the key takeaway is that this technology is beneficial for businesses. To summarize, the adoption of digital payment technology significantly reduced various costs associated with using cash transactions for drivers. Passengers usually negotiate fares at the beginning of a ride and predominantly paid in cash before the technology was introduced. While peer-to-peer (P2P) mobile money payments were available, they were rarely used due to revocability concerns related to transaction reversibility. The technology decreased the time spent searching for small change and the incidence of change-related issues, such as mistakes in giving change. In addition, the technology improved several non-monetary outcomes, including better account management, reduced anxiety about theft, and fewer incidents of digital robbery. Overall, taxi businesses expressed substantial willingness to pay for this technology. See Houeix (2025) for more details.

1.2 Experimental Design

The experiment randomized early access to the technology among taxi businesses in Dakar.

[4]The population under study predominantly consists of urban poor (according to the PPI Index developed by Innovations for Poverty Action), typically male workers aged 30-50. The taxi industry is often a family enterprise, with family members involved in about 50% of the cases. The vast majority of taxi drivers and owners have not completed primary education (68%) and about half of the drivers were unable to save any money over the past three months.

Listing survey. Between March and May 2022, drivers and owners were listed at garages, car washes, meeting points, and taxi stands across Dakar (Figure A4). Non-driving owners were contacted through their drivers, as owner contact details were required for adoption. Employed drivers who refused to share owner contacts were also listed.

Baseline survey. Starting from the listing, I contacted 2,072 taxi owners and asked each to update their current roster of drivers. Baseline data collection was conducted in twelve rolling batches across three Dakar sites, from March to June 2022. In total, 3,011 individuals (owners + drivers) were part of the baseline, with owners and drivers interviewed separately in 45-minute questionnaires. The baseline dataset includes 2,196 unique taxi stakeholders: 1,821 are drivers (including owner-drivers who operate their own taxi) and the remainder are non-driving owners who employ at least one driver. Overall attrition from invitation to participation was 26%, with no significant differences across treatment arms (treatment assignments were disclosed only after baseline surveys to prevent differential attrition). Roughly half of attrition reflects baseline ineligibility (e.g., the driver lacked an Android phone, a license, or an ID), rather than refusal or no-shows.

The unit of randomization is the taxi business. The sample comprises three types of stakeholders: of the 2,196 surveyed, 1,821 are drivers—owner-drivers who operate their own vehicle or employed drivers working for an owner—and the remaining 375 are non-driving owners who employ at least one driver (608 owner–driver pairs are linked at baseline). All types are included as a single unit for the following analysis—the taxi business—regardless of whether the respondent is the owner or the driver. Approximately 58 percent were assigned to treatment and the remainder to a waitlist control group.[5] Randomization was computerized and stratified on baseline characteristics collected during listing and surveying. The realized experimental sample is balanced across treatment arms. Treated businesses were provided with the QR payment card, the business application, and training on how to use the product. For the sake of this study, control businesses were prevented from receiving the technology during the experimental period and were offered access only after the experiment ended. Importantly, the experimental protocol did not encourage treated businesses to market the technology to

[5]Sample accounting: 2,196 stakeholder records randomized at the business level → 2,150 matched to a mobile money account (1,256 treated, 894 control; 46 could not be matched), spanning 1,454 taxi businesses with treatment constant within business → 1,912 with phone fingerprint data → 1,907 in the phone network after minimum-contact filtering. Throughout, "experimental drivers" refers to these matched records—including non-driving owners' accounts—whose phones define the experimental side of the network. All reduced-form analyses use the 2,150 matched records for exposure computation; the phone network analyses condition on the 1,907 appearing in the fingerprint data.

others or to share information about it. Any adoption among non-experimental firms is therefore naturally interpretable as a spillover of the randomized intervention.

Timeline. The key dates are: (i) first treated onboarding: March 28, 2022 (start of outcome window); (ii) last treated onboarding: June 2022 (final baseline wave); (iii) adoption-window cutoff: October 5, 2022 (when control-group rollout began and could freely adopt); (iv) control rollout completed: December 2022. The structural model uses $T = 6$ monthly periods from the first onboarding to the cutoff. Drivers onboarded in later waves had fewer than six months of exposure, but since adoption is measured as a binary indicator by October 5, the six-month window is an upper bound on treated exposure for all drivers.

The design serves two purposes. First, it identifies the private effects of the technology on adopting taxi businesses, which are studied in a separate paper (Houeix, 2025). Second, the focus of this paper, it creates exogenous variation in the treatment status of taxi firms that are embedded in much larger social and commercial networks. The core empirical strategy then compares connected firms that differ in the share of their experimentally linked taxis who happened to be randomized into treatment.

1.3 Measuring Firm Networks at Scale

A central object in the paper is the social proximity between non-experimental firms and treated taxi businesses, which I measure using encrypted phone-contact fingerprints collected by the mobile money company. Unlike the survey links or realized transactions more common in economics, contact-book fingerprints capture the set of contacts a user chose to save in the phone, making them well suited to measuring *durable* social proximity among small, owner-operated firms. The operator collected fingerprints for roughly 7 million users: on first login to the application, users are asked whether they consent to share their contact list (consent rate about 86 percent), and for those who consent the operator stores an encrypted fingerprint that it uses operationally to facilitate transfers to saved contacts. Because most firms in this setting are small and owner-operated, these fingerprints provide unusually direct evidence on the social networks through which business information may travel. Related objects have been studied in computer science for secure set-overlap measurement and private contact discovery, but, to my knowledge, not for economic measurement (De Cristofaro et al., 2013; Swamidass et al., 2015; Demmler et al., 2018).

Constructing the network. Each phone number saved in a user's contact list is hashed into one of $2^{16} = 65{,}536$ buckets using a 16-bit system FNV-1a hash (Noll et al., 2026). Let F_i denote the set of hashed contacts saved by user i, and h_i the hash of user i's own phone number; these fingerprints preserve set operations while preventing recovery of raw phone numbers. The fingerprints yield two complementary objects for any pair of users i and j: directed *save* indicators, $\mathbf{1}\{h_j \in F_i\}$ (i saved j's number) and $\mathbf{1}\{h_i \in F_j\}$, whose conjunction defines a *mutual contact*; and the contact-book *overlap* $|F_i \cap F_j|$, which is defined whether or not either party saved the other. The analysis network is built from the overlap measure—two people who share a meaningful fraction of their saved contacts are socially proximate even if neither stored the other's number—with the save indicators retained as tie attributes and analyzed separately.[6]

Measuring tie strength. Tie intensity is measured by the *bias-corrected overlap share* of the two contact books:

$$\text{OverlapShare}_{ij} = \frac{\max(I_{ij} - \mu_{ij},\, 0)}{\min(|F_i|, |F_j|)}, \qquad \mu_{ij} = \frac{|F_i|\,|F_j|}{65{,}536},$$

where I_{ij} is the observed intersection of the two hashed fingerprints and μ_{ij} is the intersection expected from independent hash collisions, so the numerator retains only overlap in excess of chance. Normalizing by the smaller contact list ensures that users with larger contact books are not mechanically penalized. Economically, this is best interpreted as a measure of *social proximity*: how embedded are the two users in overlapping social circles? The interpretation is consistent with the network-science literature, where common-neighbor and neighborhood-overlap measures are informative about latent social proximity and link formation (Adamic and Adar, 2003; Liben-Nowell and Kleinberg, 2007).

Community ties and tie strength. A *community tie* is a pair with corrected overlap of at least 2 percent—two people who demonstrably share a meaningful fraction of their social environment, whether or not either saved the other's number. The 2 percent floor

[6]The unit of observation in the phone network is a mobile money account, identified by phone number. Using the fingerprint data, 85% of phone numbers can be linked to unique individuals; about 16.5% of individuals hold multiple accounts. Aggregating to the individual level using the operator's legal-entity identifiers changes mutual-contact edge counts by less than 0.5% and strong-tie counts by 1%, because multi-account individuals are typically connected to different drivers through each account. The phone-level analysis is therefore essentially equivalent to a person-level analysis but avoids the loss of outcome linkage that individual-level identifiers would entail.

is set an order of magnitude above the collision-noise scale[7] and is validated directly against survey ground truth: 81–87 percent of surveyed spouses, family members, and taxi colleagues clear the floor, against 23–28 percent of clients and serendipitous contacts (Appendix Table B1), and unsaved pairs above the floor are reported as close relationships 37 percent of the time versus 18 percent below it—half the rate even of pairs without fingerprint data (Appendix Table B2). Because overlap is normalized by the smaller contact book, the community graph is deliberately broader than a contact book: it measures *shared social environment*—before trimming, the median driver is community-tied to roughly 170,000 individuals—and the strict hub trimming described below reduces this to an analysis network of about 12,400 contacts per driver. *Strong ties* are community ties with corrected overlap of at least 5 percent, and the main specifications are reported at the 2, 5, and 10 percent thresholds.[8] Reciprocity is measured separately: a *mutual contact* is a pair in which each party saved the other's number. Mutual ties are a small, distinct layer of the community network (under 1 percent of analysis-network ties) with markedly higher baseline adoption, and the analysis shows they play a specific role—carrying within-industry adoption across markets—rather than amplifying local diffusion (Appendix Table B6).

Links below 5 percent are not discarded as meaningless, however. A threshold calibrated to recover the *core* social network is not automatically the right threshold for every mechanism: competition may operate through weak mutual contacts—nearby rivals, former customers, peripheral industry contacts—and weak or long ties may bridge otherwise distant circles, mattering for cross-industry or cross-space diffusion even if they carry less intense information.

Validating overlap as social proximity. Three pieces of survey evidence support the interpretation of overlap as social closeness. In the endline survey, 1,693 of the 2,196 experimental drivers completed a network module, each reporting on up to 16 randomly selected mobile money contacts—relationship type (wife, family, friend, taxi colleague, client, and so on) and interaction frequency—as described in Section 1.4.1. Of the 23,520 survey pairs, 22,537 (96%) are classified under the corrected hashing scheme, 16,470 with fingerprint data on both sides (Appendix C); location-matched subsets are smaller (Table B3).

[7] For a false match, the expected corrected overlap is centered on zero with noise scaled by $1/\min(|F_i|, |F_j|)$, so a 2 percent corrected floor removes nearly all purely accidental overlap. Within the mutual-contact subgraph, the threshold that minimizes the mean absolute error between inferred and actual contact-book size is 0.4–0.5 percent on the corrected scale, so 2 percent is conservative.

[8] Results are insensitive to the exact strength cutoff: exclusive strength bands from 2 to 17+ percent show approximately uniform per-band transmission for cross-industry adoption (Appendix Figure A10).

First, corrected overlap rises monotonically with reported closeness. Wives share 20.8% of contacts on average (86.5% of pairs above the 2% community floor; 76.8% above the 5% strong-tie threshold), taxi colleagues 13.6% (81.4%; 62.8%), family members 11.4% (80.8%; 64.1%), and friends 6–9%. At the bottom of the distribution, serendipitous contacts—numbers the respondent could not place when prompted, typically incidental or one-time transaction partners—average 3.2% with only 27.7% above the community floor, and taxi clients 1.9% (22.8%): the 2% floor cleanly separates durable relationships from incidental contact. The same gradient appears by interaction frequency: contacts seen weekly or more average 10.6% corrected overlap, versus 5.5% for contacts seen less than monthly, a $1.9\times$ difference (Table B1).

Second, reciprocity marks relationship depth. Mutual contacts are reported as close 61.1% of the time; one-sided contacts (36.8%) and community ties where neither party saved the other (37.1%) are indistinguishable—a one-way save adds nothing beyond community membership—while pairs *below* the community floor are close only 18.2% of the time (Table B2). Reciprocal saving is thus a meaningful marker of tie depth—which is why the mutual layer is analyzed separately (Appendix Table B6)—but overlap without reciprocity still identifies real relationships, and it is these community ties that carry the diffusion effects in Table 1.

Third, overlap carries social-proximity content at any distance. A given overlap level might in principle signal a close relationship when two parties are far apart but a casual acquaintance when they are nearby. Regressing survey-reported closeness on corrected overlap, geographic distance, reciprocal-save status, and their interactions (Table B3) shows that the overlap–closeness slope is unchanged per kilometer of distance ($p = 0.685$ on the continuous interaction); a coarse "contact outside Dakar" indicator reveals some attenuation—the slope for far contacts is about 60% of the near slope—but remains large and highly significant, so overlap identifies genuine relationships everywhere, somewhat less sharply at distance. Conditional on overlap, mutual status and its distance interactions are precise zeros: reciprocity adds no closeness content beyond the overlap measure, at any distance. This licenses reading the cross-city community-tie estimates as (if anything, conservatively measured versions of) the same object as the within-Dakar ones. Overall, corrected overlap separates stronger from weaker ties in a useful—if not perfectly sharp—way.

The network is observed after the experiment. The phone-contact fingerprints are observed in February 2024, after the experiment ended, so they must be interpreted as a post-period proxy for durable *pre-existing* social proximity. Two checks support this. First,

treated and control drivers are balanced on network characteristics: joint F-tests are not significant on either the legacy panel ($p = 0.58$) or the community-tie panel ($p = 0.34$). Three individual comparisons are marginal: control drivers have slightly *more* mutual contacts ($p = 0.056$) and marginally larger contact books (KS $p = 0.047$)—both working *against* finding positive spillovers from treatment—and treated drivers have a trivially higher share of merchant community ties (2.50% vs. 2.41%, $p = 0.028$) (Table B4). Second, a more demanding test for treatment-induced tie strengthening—for example, treated drivers saving more technology-related contacts—finds nothing: treated drivers' overlap distributions do not shift rightward (if anything, they are non-significantly *lower*), and the composition of strong ties ($\geq 5\%$ overlap)—the sample of interest below—is indistinguishable across arms on every dimension tested (number of ties, share taxi, share business, mean overlap; all $p > 0.19$), with a weak-tie placebo ($< 1\%$) and separate contact-type splits (e.g., $p = 0.14$ for taxi contacts) showing the same (Table B5). The takeaway from Tables B4 and B5 together is that the February-2024 network can be treated as pre-determined with respect to treatment: arms are balanced on every community-network dimension (joint F, $p = 0.34$), the few marginal differences run *against* finding positive spillovers, and there is no evidence that treatment created or strengthened the ties on which the spillover estimates ride; any residual bias from differential appearance is bounded directly by the Lee bounds (Appendix Table B8).

Differential appearance in the data. Of the 2,150 experimental drivers with a matched mobile money account, 1,912 (88.9%) have at least one phone number with fingerprint data, and 1,907 (88.7%) appear in the phone network; the 11.1% gap reflects non-collection, as the operator's fingerprint extraction did not reach these users by the February 2024 snapshot. Treated drivers are somewhat more likely to appear (89.9% vs. 87.6%, $p = 0.093$), plausibly because treatment raised app usage, which triggers collection—but the difference is small and, conditional on appearing, the economically relevant variables (overlap composition, share business, share taxi) show no meaningful imbalance (Table B4). To formally bound any resulting bias, tightened Lee bounds (Lee, 2009) computed within contact-book-size quartiles are tight around every main estimate: at the strong-tie threshold the cross-industry Dakar result is bounded at $[+4.08, +4.20]$ (both significant), the cross-industry result outside Dakar at $[+0.680, +0.685]$ (both significant), and the within-taxi competition effect at $[-1.10, -0.98]$ (both $p_{\text{RI}} < 0.01$); on the community network the distant cross-industry effect is bounded at $[+0.554, +0.686]$, both bounds significant (Table B8). The bounds bracket the point estimates and preserve sign and significance throughout, indicating that differential attrition does not drive the estimated

spillovers. After filtering and degree trimming, the community-tie network (corrected overlap $\geq 2\%$) links these drivers to 1.15 million individuals through roughly 24 million edges; about 164,000 of these individuals are additionally connected through the reciprocal-save (mutual) layer.

1.4 Survey, Transaction, and Location Data

To interpret the mechanisms behind diffusion, the paper combines the phone network with survey evidence, transaction records, and location data.

1.4.1 Survey Data

The survey data serve two purposes. First, they provide the baseline characteristics used to describe the experimental sample and assess randomization balance. Second, they characterize the content of economic relationships. Between September and December 2023, respondents were asked about a random subset of 16 links drawn from their transaction network. For each link, the survey recorded the relationship type, communication frequency, whether the respondent exchanges advice or other important information with the contact, whether the respondent would borrow money from the contact, and whether the two discussed the payment technology. The questionnaire follows Banerjee et al. (2013). These data are useful for interpreting the administrative network measures. In Table B1, higher contact-book overlap is associated with ties that respondents themselves describe as closer and more socially meaningful.

1.4.2 Transaction Network Data

The transaction data provide a second network layer. Because the mobile money provider first operated peer-to-peer transfers and only later expanded into business payments, the data contain the universe of person-to-person mobile money transactions on the platform before the experiment. These records are used to construct direct transaction links and measures of interaction intensity. The transaction network is much sparser than the phone network, but it has two advantages: it is observed before treatment, and repeated transfers are a revealed form of economic or social interaction. In the reduced-form analysis, it provides an important cross-check on the phone-based results.

The transaction data are also used to construct the customer-link objects relevant for the customer-demand channel. The paper distinguishes two networks. The first is a *pre-intervention customer network*, built from person-to-person transactions between April

2021 and March 2022. Customer-like transfers are identified using the range typical of taxi payments, and repeated contact is required to exclude incidental transfers. Because this network predates the experiment, it provides the cleanest evidence that shared-customer exposure is not mechanically created by treatment. The second is a *post-period customer network*, built from business-payment data between November 2022 and November 2023. This later network is a cleaner measure of realized customer overlap, but because it is observed after treatment it is used more cautiously: as supporting evidence and, later, as a calibration object rather than the sole basis for identification.

The transaction network is less exhaustive than the phone network. At baseline, the average user transacted with 35 other users, compared with roughly 590 phone contacts for the experimental drivers. It nevertheless contains richer information on financial interaction. In the 12 months before the experiment, the average user sent or received 85 transfers totaling CFA 2,013,490 (USD 3,356). The average transfer size was CFA 23,686 (USD 39). These data make it possible to exploit interaction intensity—for example, the number or frequency of transactions between a taxi business and a given contact.

1.4.3 Location and Migration Data

The location data come from deposits and withdrawals made through the provider's nationwide network of mobile money agents (see Figure A5). Because each cash-in and cash-out transaction is geolocated using the coordinates of the handling agent, these data support a user-month panel of locations. The resulting panel is unusually informative about temporary migration and geographic diffusion. It makes it possible to identify users who remain in a city throughout the experiment, users who temporarily migrate into Dakar, and treated taxi drivers who travel to other cities during the rollout.[9] These measures are essential for distinguishing remote diffusion through social ties from diffusion through physical co-presence.

1.4.4 Technology Adoption Data

Finally, adoption of the business payment technology is observed directly in the provider's administrative records. For each adopting business, the data record the registration date, business category, and registration location. The main outcome is adoption during the experimental window, defined as the period from the start of the baseline survey until the control group became eligible for treatment. These data allow adoption to

[9] As a check on the construction, the data show substantial population movements into Touba during the pilgrimage month; see Figure A6.

be tracked separately for taxis in Dakar, non-taxi businesses in Dakar, and non-taxi businesses elsewhere in Senegal. That distinction is central to the paper because it links the same randomized intervention to diffusion within industry, across industries, and across space (see Figure 2).

2 Empirical Patterns in Firms' Networks

I begin by documenting two key patterns in the networks of small firms, using the comprehensive network data. These patterns motivate the investigation of technology diffusion within firms' social networks.

Stylized Fact 1 - Firms' networks not only are local and intra-industry, but also expand across industries and regions. On average, the taxi businesses in the experiment have approximately 590 phone contacts (unique entries in the encrypted contact-book fingerprint). The community-tie graph is far broader than the contact book by construction—it links people who share social environment, not only saved numbers: before trimming, the median driver is community-tied to roughly 170,000 individuals. After strict hub trimming, the analysis network contains about 12,400 community ties per driver, of which about 1,500 are strong ($\geq 5\%$); the reciprocal-save (mutual) layer is far smaller still ($\sim$100 per driver).[10] The data reveal that the taxi businesses' network extends significantly beyond Dakar, where the experiment took place: over half of their phone contacts reside outside the capital city, and most of these are located in urban areas across the country.

As a descriptive exercise, I identify business links within the phone contact network by matching mutual contacts to all business adopters of digital payments by April 2023. On average, each driver is connected to roughly 250 other business owners through mutual phone contacts, of which 12 are taxi businesses and 237 are in other industries (3.9% of all mutual contacts). Among strong ties specifically, each driver is connected on average to 29 business owners (1 taxi, 28 non-taxi). This extensive cross-industry reach is a critical feature of the data: it allows me to track technology diffusion across both industry and space from a single experimental sample, and it reflects the broader characteristics of firm networks in this low-income, owner-operated context.

[10] These counts are computed on the analysis sample after two filtering steps: (i) keeping only the most informative phone per individual (best-phone-per-person, Appendix C), and (ii) applying top-2% degree trimming (removing hub nodes). Before these steps, the mean mutual-contact count is $\sim$8,200, of which $\sim$700 clear the strong-tie threshold (Table B5, Panel C). Table B4 reports the intermediate pre-trimming values.

Stylized Fact 2 - Substantial information sharing among social networks. Small taxi businesses not only share information and advice among themselves but also with firms beyond the taxi industry. As described earlier, I extracted a random subset of about sixteen mobile money recipients and senders from the experimental participants and surveyed each subject about the nature of their relationships with these individuals. As shown in Table B11, taxi businesses are familiar with 58% of the extracted contacts, the majority being family or friends. Notably, many of these family and friends also own businesses, often outside the taxi industry, providing additional support that many social connections are with other business owners. Among the people they know, businesses communicate frequently, interacting daily or weekly with 68% of them. Taxi businesses share important advice and information with about 74% of the people they know, particularly other businesses and taxi drivers.[11] More specifically, they reported having discussed or recommended the digital payment technology to 12% of the people they know. This underscores the high level of information sharing and potential for social learning within business networks, motivating the exploration of social learning as a key mechanism for technology diffusion.

3 Empirical Framework

The experiment generates exogenous variation in firms' exposure to early adopters. I exploit this variation to estimate spillovers: how the adoption decision of a non-experimental firm responds when a larger share of the taxi businesses in its network were assigned to treatment. The estimand throughout is the causal effect of randomized exposure on connected firms' adoption, estimated separately across network layers, industries, and locations.

3.1 Exposure-Based Reduced-Form Design

The main specification regresses adoption on the share of a focal unit's experimental taxi contacts that were treated. Let $\mathcal{E}$ denote the set of experimental taxi businesses, and let $T_i \in \{0, 1\}$ indicate treatment assignment for experimental taxi $i \in \mathcal{E}$. For focal unit j and network layer n, let $\mathcal{E}_j^{(n)} \subseteq \mathcal{E}$ denote the experimental taxis connected to j in network n,

[11] The exact survey question was "More generally, do you provide important information/advice to [Name]?" in the spirit of Banerjee et al. (2013).

with $N_j^{(n)} = |\mathcal{E}_j^{(n)}| \geq 1$. The share-treated exposure variable is

$$S_j^{(n)} = \frac{1}{N_j^{(n)}} \sum_{i \in \mathcal{E}_j^{(n)}} T_i.$$

The reduced-form specification is

$$Y_j = \alpha_{g,n} + \beta_{g,n} S_j^{(n)} + \lambda_{g,n} \bar{S}_j^{(n)} + u_j, \tag{1}$$

where Y_j indicates whether unit j adopts during the experimental window, g indexes the receiving population, and $\bar{S}_j^{(n)} = \mathbb{E}_{\mathcal{R}}[S_j^{(n)} \mid \mathcal{E}_j^{(n)}]$ is the expected exposure under the randomization protocol $\mathcal{R}$, computed from 1,000 re-randomization draws following Borusyak and Hull (2023). Under complete randomization with a common treatment probability, this control is closely related to the number of experimental contacts; under the stratified design used here, it also absorbs differences in the composition of a focal unit's experimental neighbors across strata.

The coefficient $\beta_{g,n}$ captures the causal effect on adoption of increasing the proportion of treated taxis among a focal unit's experimental contacts in layer n. The share-treated parameterization is preferred because the network layers differ substantially in density: a ten-percentage-point increase in exposure has a comparable interpretation whether measured in the phone network, the transaction network, or the customer network. To make the design concrete, consider a clothing shop in Dakar whose owner has five experimental taxi drivers in her mutual phone network, three of them randomly assigned early access. Her exposure in the phone layer, $S_j^{(\text{phone})}$, equals $3/5$. Equation (1) asks whether her probability of adopting during the experimental window rises with this treated share, holding fixed her expected exposure $\bar{S}_j^{(\text{phone})}$—which is pinned down by how many experimental taxi contacts she has, and so absorbs her network position. A positive $\beta_{g,n}$ means that exposure to early adopters through her phone contacts raises her own adoption. Because she is in a different industry, she neither competes with these drivers nor shares their passengers, so the estimate isolates social learning.

A complementary number-treated specification replaces the share with the count:

$$Y_j = \tilde{\alpha}_{g,n} + \tilde{\beta}_{g,n} T_j^{(n)} + \tilde{\lambda}_{g,n} \bar{T}_j^{(n)} + \tilde{u}_j, \qquad T_j^{(n)} = \sum_{i \in \mathcal{E}_j^{(n)}} T_i, \tag{2}$$

where $\bar{T}_j^{(n)} = \mathbb{E}_{\mathcal{R}}[T_j^{(n)} \mid \mathcal{E}_j^{(n)}]$ is the expected number treated under the randomization protocol. This specification asks how adoption changes when one additional experimen-

tal taxi contact is treated, controlling for expected exposure. The two parameterizations yield similar results.

Equation (1) resembles the linear-in-means formulation in the peer-effects literature (Manski, 1993; Brock and Durlauf, 2001) and is close in spirit to recent exposure-based designs that use randomized peer treatment to study learning and spillovers (Cai et al., 2015; Duflo et al., 2026; Chaurey et al., 2025). The key distinction from a classical peer-outcomes regression is that the regressor is peers' *randomized treatment status*, not their endogenous adoption choices. Conditional on network position, variation in $S_j^{(n)}$ is generated by the experiment itself, so the usual reflection problem does not arise.

The reduced-form coefficients capture *net* spillover effects through the network used to define exposure. A positive coefficient in the cross-industry phone network can be read as evidence of social learning, because same-industry competition is absent in that population. A coefficient in the within-taxi phone network, by contrast, aggregates learning and competition. Estimating the same design across multiple network layers and samples—holding the empirical framework constant while varying the channel content of the exposure measure—is what allows the reduced-form evidence to speak to the mechanism decomposition developed later in the structural model.

3.2 Outcomes, Samples, and Network Definitions

The outcome is an indicator for adopting the business payment technology during the experimental window, from the start of baseline surveying on March 28, 2022, until the control group became eligible for treatment on October 5, 2022. Adoption is observed directly in the administrative records of the mobile money provider.

Three receiving populations define the main spillover samples. The first is non-experimental taxi businesses in Dakar, capturing within-industry diffusion. The second is non-taxi businesses in Dakar, capturing cross-industry diffusion within the treatment city. The third is non-taxi businesses outside Dakar, capturing diffusion across space. In all cases, the experimental taxis themselves are excluded from the spillover regressions, so the estimand is the response of untreated connected firms to randomized exposure in their network (Athey et al., 2018).

The estimating sample is defined by the network layer used to measure exposure: a focal unit enters a regression only if it has at least one experimental taxi contact in that layer. This restriction keeps estimation on the support of feasible exposure and avoids comparing connected firms to structurally isolated ones. Because the relevant connected set differs across network layers, the sample changes across specifications.

Several network layers are used, each loading on different spillover channels. The *community-tie phone network* includes all pairs with bias-corrected contact-book overlap of at least 2 percent—the calibrated boundary of a person's measured social world (Section 1.3)—and captures social exposure whether or not the parties saved each other's numbers; the *strong-tie* specification restricts to corrected overlap of at least 5 percent. The *mutual-contact layer* (reciprocal saves) is analyzed separately and plays a distinct role, carrying within-industry adoption across markets rather than local diffusion. The *pre-experiment transaction network* captures direct economic interaction and provides a cross-check on the phone-network results. The *customer network*, measured both before and after the intervention, captures shared customers between taxi businesses and identifies the customer-demand channel; the pre-intervention version provides the cleanest evidence, while the post-period version is measured more precisely but must be interpreted with greater caution. Finally, the *garage network* links businesses through local spatial proximity and isolates competitive exposure; because few experimental taxis share garages, the main regressor there is an indicator for whether any treated experimental taxi operates within a given radius.

3.3 Inference Under Network Dependence

Overlapping network connections generate dependence across focal units: two businesses may share experimental neighbors, be linked to each other, or belong to the same local market. Rather than imposing a parametric dependence structure, I base inference on the experimental randomization itself.

For each specification, I compute randomization-inference p-values by redrawing treatment assignment according to the original stratified design, recomputing exposure, and re-estimating the coefficient. The focal units, the network, and the specification are held fixed. Because focal units are never directly treated, the relevant sharp null is that the treatment assignments of all experimental taxis have no effect on focal-unit outcomes. Under this null, the randomization distribution provides the appropriate finite-sample benchmark (Athey et al., 2018).

The tables report heteroskedasticity-robust standard errors alongside randomization-inference p-values, but the latter are the primary measure of statistical significance. In the main reduced-form tables, p-values are two-sided. In mechanism specifications where the sign is predicted ex ante—most notably the garage-competition regressions—one-sided p-values are also reported.

4 Technology Diffusion Across Industry and Space

This section documents three main reduced-form facts. First, non-taxi firms in Dakar adopt at higher rates when more of their social contacts among taxi drivers were assigned to treatment. Second, this diffusion extends beyond Dakar through remote interactions along social ties: businesses in other Senegalese cities connected to treated taxis also adopt at higher rates, whether or not they or the drivers physically move between cities. Third, within-industry spillovers are heterogeneous—negative on average in the phone-contact network, but positive through survey-identified friendships—reflecting the opposing forces of social learning and competition. These results establish causal diffusion across industries and space, with the sign and magnitude of spillovers depending on the type of tie through which exposure operates.

4.1 Cross-Industry Diffusion in Dakar

Non-taxi businesses in Dakar adopt at higher rates when more of their taxi contacts were assigned to treatment. Because these firms operate outside the taxi sector and do not compete directly with the experimental taxis, the cross-industry estimate is largely free of the within-industry confounds that complicate taxi-on-taxi results.

Table 1 shows that the effect operates throughout the community network, reporting the specification at the 2, 5, and 10 percent corrected-overlap thresholds (Section 1.3); Appendix Figure A7 maps the resulting adoption across Dakar. At the strong-tie threshold the cross-industry coefficient is positive, significant, and economically large: moving the treated share from zero to one raises adoption by $\hat{\beta} \times 1{,}000 = +4.16$ ($p_{\text{RI}} = 0.017$), a $+31\%$ effect relative to the mean adoption rate, rising to $+5.16$ ($p_{\text{RI}} = 0.015$, $+48\%$) at the 10 percent threshold. Decomposing the effect into exclusive strength bands (Appendix Figure A10), per-band transmission is remarkably uniform—between $+21\%$ and $+43\%$ of baseline from the 2–5% band through 17%+—indicating that any genuine community tie transmits the technology at a roughly constant rate; stronger ties yield more precise estimates, not larger per-tie effects.[12]

This pattern—null on weaker ties, significant on high-overlap ties—is consistent with social learning operating through relationships with enough shared context to transmit credible information about a technology's reliability and profitability. Section 5 confirms

[12]Because the phone network is observed after the experiment, identification relies on balance of network characteristics across treatment arms. Treated and control drivers are statistically indistinguishable on the composition and intensity of their measured phone networks, including the share of businesses, the share of high-overlap contacts, and mean overlap. If anything, control drivers have more detected mutual contacts, working against finding positive spillovers from treatment. See Table B4.

this interpretation: competition and customer-demand complementarities matter primarily within the taxi sector, so the cross-industry estimates isolate a nearly pure social-learning margin. Survey evidence reinforces this conclusion, as discussed in Section 5.1.

4.2 Diffusion Across Cities

Technology adoption also spreads beyond the capital. Table 1 shows that businesses in other Senegalese cities connected to treated Dakar taxis adopt at significantly higher rates at every threshold ($p_{\text{RI}} = 0.011$, 0.009, and 0.030 at the 2, 5, and 10 percent cutoffs), with proportional effects of $+53\%$ to $+76\%$—roughly twice the within-Dakar cross-industry effect sizes, consistent with a lower baseline outside the capital amplifying the same diffusion force. Unlike the within-Dakar effect, the distant effect strengthens with tie intensity across the strength bands (Appendix Figure A10; peaking at $+76\%$ in the 8–11% band and $+72\%$ in the open tail): transmitting adoption across space appears to require more redundant or stronger community connections than local observation does. These results show that knowledge spillovers need not be local in the classic Marshallian sense (Marshall, 1920).

Reciprocity helps distinguish which ties carry each margin (Appendix Table B6). Cross-industry diffusion—local and distant—operates entirely through *non-mutual* community ties: pairs who share a social environment without having saved each other's numbers. Mutual contacts, despite baseline adoption rates several times higher, show no incremental response to treatment; their exposure to the technology is saturated irrespective of the experiment (inframarginality). Conversely, within-industry adoption in markets outside Dakar responds *only* through mutual ties, at every threshold ($+0.90$ to $+4.39$, $p_{\text{RI}} \leq 0.006$)—profession-specific adoption crosses markets through reciprocal relationships, a channel the pre-treatment transaction network reproduces independently (Appendix Table B12). Awareness travels on the broad community network; experiential, trust-intensive content travels on reciprocal ties.

Table 6 (Panel B) shows that this pattern is not driven by a single destination: pooling the eight largest secondary cities with city fixed effects, exposure to treated drivers raises adoption in the destination city even through drivers who never visited it ($p_{\text{RI}} = 0.018$), with the migrant-driver channel positive but less precise ($p_{\text{RI}} = 0.133$), and per-city point estimates are generally positive. Precision varies with the density of network links into each city, but the qualitative pattern is stable. Given the low baseline adoption rate outside Dakar, these treatment-induced increases represent large proportional shifts in adoption behavior.

4.3 Within-Industry Exposure and Tie Heterogeneity

Within the taxi sector, same-industry exposure combines social learning with within-market competition. Learning pushes toward positive spillovers; competition pushes toward negative ones; and the phone-exposure coefficient reflects their net effect. Table 2 identifies this tension directly, and the pattern is correspondingly more nuanced than the cross-industry one. At the community threshold the within-taxi coefficient is a statistical null ($+0.08$, $p_{\text{RI}} = 0.674$): on the broad network, learning and competition offset. At the strong-tie threshold competition dominates—the coefficient is negative and significant ($\hat{\beta} \times 1{,}000 = -0.98$, $p_{\text{RI}} = 0.007$, a -38% effect)—and the band decomposition locates the crowding in the moderate-strength bands (5–8%: -36%, $p_{\text{RI}} = 0.011$), fading among the very strongest ties (Appendix Figure A10). Community exposure to treated same-industry rivals lowers adoption, consistent with the geographic competition documented at garages below.

Disaggregating by tie type reveals that learning remains potent within industry on meaningful relationships. Table 3 reports that exposure through friends generates a substantial positive effect ($+36.1\%$), while exposure through frequent interlocutors produces a similar effect ($+24.3\%$). The contrast with the negative average phone result is suggestive of a force opposing learning on weaker ties. The structural model (Section 6) identifies this force as geographic competition operating through a separate channel (S^{geo}), orthogonal to phone exposure. Under that decomposition, the negative within-taxi phone coefficient reflects learning effects close to zero on the weak-tie majority—in addition to competition running through some social ties (see Section 6 for the interpretation).

This heterogeneity aligns with the broader sociology of diffusion. Embedded ties transmit detailed, credible, and repeated information, while longer ties carry novel information across social distance (Granovetter, 1973; Aral, 2016; Jahani et al., 2023). In the present setting, strong within-industry ties help overcome knowledge-hoarding frictions (Cefalà et al., 2025) and support learning even amid rivalry. Weak same-industry contacts, by contrast, are more exposed to competitive crowding and information withholding, depressing the average spillover.

The reduced-form evidence therefore establishes broad diffusion across industry and space, but also shows that within-industry spillovers are fundamentally heterogeneous. The next section decomposes these effects into three channels: social learning, local competition, and customer demand.

5 Social Learning, Competition, and Customer Demand

The reduced-form results in Section 4 point to multiple channels. Cross-industry and inter-city spillovers are consistent with social learning; heterogeneous within-industry results suggest that competition also matters. This section organizes the evidence around three channels—social learning on meaningful ties, local competitive crowding within industry, and customer-demand complementarities. Each subsection identifies its channel through a distinct empirical variation and ties it to the structural decomposition developed later. Lastly, I document that geographic diffusion operates through remote interactions and does not require seasonal migration, although migration may contribute alongside.

5.1 Social Learning on Meaningful Social Ties

Spillovers are strongest where relationships are most capable of transmitting information, a pattern that generic market complementarities would not generate and that points to social learning.

Survey-elicited relationship types help interpret this result. Table 3 shows that within the taxi sector, spillovers concentrate among friends ($+36.1\%$, $p_{\text{RI}} = 0.007$) and frequent interlocutors ($+24.3\%$, $p_{\text{RI}} = 0.023$), while exposure through non-friends produces a smaller and insignificant effect ($+2.8\%$). Close contacts are the people from whom a firm owner can learn how the technology works, whether customers use it, and whether payments are reliable.

Cross-industry survey evidence shows a complementary pattern. Table 3, Panel C reports that cross-industry spillovers are significant through serendipitous encounters ($+47.9\%$, $p_{\text{RI}} = 0.004$) and positive but imprecise through former or regular taxi clients ($+93.7\%$, $p_{\text{RI}} = 0.134$). The strongest ties (Panel B: friends, frequent contacts) show positive but imprecise cross-industry effects. This fits the idea that novel information often arrives through ties bridging separate circles (Atkin et al., 2024): inside the taxi sector, close ties overcome rivalry-induced information withholding; across industries, weaker ties are valuable precisely because they connect firms to different market environments (note that many of these nominally weaker survey ties still exceed the 5 percent corrected phone-overlap threshold, see Table B1).

The pre-treatment transaction network provides an additional cross-check with a different tie measure (Appendix Table B12). Local effects on money-sending ties are positive but not significant under randomization inference; what the network does carry, robustly, is the cross-market channel: individuals to whom a driver had repeatedly sent money

before the experiment (≥ 2 transfers) respond strongly on the within-taxi margin outside Dakar ($+83\%$ of baseline, $p_{\text{RI}} = 0.001$; $+60\%$ under a ≥ 1-transfer definition), mirroring the mutual-contact result of Appendix Table B6. Financial ties, like reciprocal-save ties, appear to carry the kind of detailed, trusted information that supports adoption across markets.

5.2 Local Competition and Crowding Out Within Industry

Within the taxi sector, competition partially offsets the positive learning force. A treated driver at a garage who discovers that the technology is profitable has a strategic incentive not to share this information with co-located rivals: if nearby untreated drivers also adopt, they compete for the same digital-payment customers the treated driver currently captures. This knowledge-withholding mechanism—analogous to the channel identified by Cefalà et al. (2025) among farmers—generates negative same-industry spillovers that are strongest where market overlap is most direct: at short distances from garages where drivers compete for the same passengers and parking spots. It can explain the negative average same-industry coefficients in Table 1: same-industry exposure activates this competitive channel, which is absent across sectors.[13]

The garage environment, matching, and identifying variation. Garages are Dakar's taxi ranks: fixed stations where drivers queue for fares, park between shifts, and spend idle hours together. They are simultaneously the densest social setting in the industry and the most direct arena of competition—drivers at the same rank draw from the same stream of passengers. Garage affiliation is measured from the baseline survey: 304 of the 2,150 matched experimental drivers report operating from a garage, and 216 driver–garage pairs carry a name that can be cleaned (a hand-built alias map consolidates spelling variants) and located; 36 named garages match the GPS garage list used for the spatial analysis, covering 142 experimental drivers. Rosters are small: the median matched garage hosts 2 experimental drivers (mean 3.9), with a single large outlier (the airport garage, 30); randomization then delivers 30 garages with at least one treated driver and 6 with control-only rosters of one to three drivers. With rosters this size, the treated *share* is an extremely coarse variable—for a two-driver garage it can only be 0, $\frac{1}{2}$, or 1—so the analysis uses the binary AnyTreated contrast, whose 30-versus-6 variation is where the experiment actually operates, with randomization inference clustered

[13]Garage coordinates are computed as the centroid of GPS pings of experimental drivers assigned to that garage, and distance from a non-experimental business to its nearest garage is the Haversine distance to that point.

at the garage level. The ShareTreated results hold: point estimates are uniformly negative with the same monotone distance decay, individually less precise for the reason just given. Non-experimental taxis at a garage play no role in the exposure—treatment status is defined by the experimental roster alone—and enter only as outcome units among the businesses located near the garage's coordinates.

For the roughly 14% of experimental drivers operating in garages (304 of the 2,150 matched), Table 4 estimates spillovers from exposure to treated taxis in the same local garage environment.[14] Panel A, which includes all businesses near garages (identified by the mobile money agents visited throughout the experiment; garage locations are mapped in Appendix Figure A9), shows negative point estimates at every radius from 50 to 500 meters, with the coefficient at 300 meters reaching marginal significance ($\hat{\beta} = -0.0036$, $p_{\text{RI}} = 0.085$). Panel B restricts to businesses inside the garage drivers' social environment—a community tie (corrected overlap $\geq 2\%$) to at least one driver at that garage, excluding mutual contacts. Here the negative effects are large and significant at every radius but one: from $\hat{\beta} = -0.102$ at 50 meters ($p_{\text{RI}} = 0.034$) to -0.015 at 500 meters ($p_{\text{RI}} = 0.023$; the 75-meter estimate is marginal at $p_{\text{RI}} = 0.052$), effect sizes of -260% to -690% of the control mean—twenty to thirty times Panel A. Extending the design outward, the community-tied effect remains detectable at 750 meters (-182%, $p_{\text{RI}} = 0.026$) and one kilometer (-201%, $p_{\text{RI}} = 0.016$), while the all-merchant panel is null under randomization inference throughout. Panel C shows the effect is entirely absent among the drivers' mutual phone contacts. Taken together, the three panels indicate that competitive crowding requires both market overlap (proximity) and community-level exposure to the treated garage, operating through overlapping foot traffic and customer awareness rather than through personal relationships—close contacts of the garage's drivers show no crowding at all. These estimates rest on a small cluster sample—36 garages, with randomization operating across 30 treated and 6 control-only rosters—so their magnitudes should be interpreted with caution; the pattern is nonetheless consistent with competition on every margin examined (distance decay, social-exposure concentration, the null among close contacts, and the other-garage placebo), and it is corroborated by the independent within-taxi network evidence of Table 2. Figure 4 displays the decay of the competition effect with distance for both populations, with 95% randomization-inference confidence intervals obtained by test inversion.

Two qualifications are important. First, the negative effects are local and strongest where market overlap is most direct; they do not imply that the taxi sector as a whole is

[14] AnyTreated $= 1$ if any experimental driver at the nearest garage was treated; see the preceding paragraph for the roster-size rationale and the ShareTreated robustness.

characterized by pure crowding out. Second, as Table 3 shows, strong social ties inside the sector still generate positive spillovers. Local rivalry depresses the average same-industry coefficient, but strong relationships transmit enough information to overcome competitive drag.

A natural question is why customer demand, which is large within the taxi sector, does not offset competition near garages. The two forces act on mostly disjoint populations: fewer than 1% of garage-proximate taxi businesses share regular mobile money customers with experimental drivers (Section 5.3), because garage riders are predominantly one-time passengers rather than the repeat customers who generate demand links. Competition therefore seems to mostly operate within garages from the experimental taxis.

Placebo: adoption at other garages. If the negative coefficients reflect competition in the local garage market rather than some spurious correlate of garage exposure, they should be absent for taxi adoption at *any* garage other than the one the merchant is near. Table B9 tests this directly: every coefficient is small and statistically null under randomization inference, in the full sample and in the community-tied panel alike ($p_{\text{RI}} \geq 0.44$ at every radius for community-tied merchants). The market-level evidence points the same way: within-taxi exposure through the phone network is null or positive outside Dakar (Table 1), so the crowding-out is specific to physically overlapping markets, not a general property of same-industry exposure.

5.3 Customer-Demand Complementarities

When customers become accustomed to paying digitally with treated taxis, other businesses serving the same customers may find adoption more attractive because their customers now expect or prefer digital payments. This channel is conceptually distinct from social learning (it operates even when two businesses do not know each other) and from competition (it raises rather than lowers the private return to adoption). The channel operates through shared customers alone.

The identification strategy exploits the customer network directly. For each non-experimental business j, I observe which experimental drivers share customers with j and compute the treated share among those drivers. The key regressor is therefore the fraction of j's customer-linked experimental contacts that were randomly assigned to treatment, controlling for j's expected exposure—a function of j's network position that absorbs any mechanical correlation between connectivity and adoption propensity. The estimating equation asks: conditional on how connected a business is to the experiment

through shared customers, does it matter whether those connected drivers were actually treated? A positive coefficient implies that customers who experience digital payments with a treated driver subsequently demand them from other businesses they patronize. Importantly, this test does not require the two businesses to interact or even be aware of each other; the customer is the sole conduit. Access to the full universe of mobile money and payment transactions—an unusually rich data structure for studying technology diffusion—is what makes this decomposition possible.[15]

Pre-intervention customer network. Table B14 provides the cleanest test by constructing customer links from person-to-person transactions between April 2021 and March 2022, before the experiment began, ruling out treatment-induced link formation by construction.[16] In this network, exposure to treated taxis through shared customers has a large positive effect among non-social taxi businesses ($\hat{\beta} \times 1{,}000 = 122.16$, $p < 0.001$, $N = 3{,}490$), corresponding to roughly 17.9 percent of the dependent-variable mean.[17] The cross-industry coefficient is a precise zero (-7.44, $p = 0.572$, $N = 6{,}613$). Customer demand thus operates only within the sector in which businesses actually share customers.

Post-experiment customer network. Table 5 provides higher-signal evidence from the business-payment network observed during the rollout, using a deliberately strict link definition: a merchant is linked to a driver only through a shared *regular* customer—at least two transactions with that specific driver *and* at least two with that specific merchant (November 2022–November 2023, when the control group could and did adopt digital payments). The within-taxi coefficient is large and significant ($\hat{\beta} \times 1{,}000 = +255.99$, $p_{\text{RI}} = 0.001$, $N = 8{,}523$), an effect size of $+35.7\%$ (column 1). Restricting to merchants without any community tie to their customer-linked drivers—column 3—the effect remains significant ($+174.54$, $p_{\text{RI}} = 0.001$, $+23.6\%$). That shared-customer links predict adoption even absent any social relationship is the cleanest evidence that customer demand operates independently of learning. When the customer link coincides with a community tie (column 2), the effect nearly triples ($+464.81$, $+67.5\%$). Part of this gap is exposure composition—community-tied merchants share customers with about three times

[15] Most diffusion studies must infer customer overlap from geographic proximity or survey recall. Here, the platform's transaction ledger directly reveals which customers patronize which businesses, permitting a network-based test that does not rely on spatial assumptions.

[16] Because the technology was not rolled out pre-experiment, I proxy customer overlap using P2P transfers of values similar to taxi fares (CFA 500–3,000). This yields a sparse network—as p2p was rarely used for taxi payments at that time due to concerns about revocability—but I present it first precisely because it is free of the endogeneity concern that affects the richer post-experiment measure.

[17] The customer regressions exclude garage-affiliated drivers when isolating the non-social channel, to avoid confounding customer demand with the local-competition effects documented in Section 5.2.

as many drivers—but holding the number of customer-linked drivers fixed, community-tied merchants still respond $2.5\times$ more among the lightly exposed (1–5 linked drivers), consistent with demand and learning being additive at the margin where information is scarce: overlapping customers raise the return to adoption, while social exposure makes it easier to understand and trust the technology. Community ties amplify the demand effect but are not necessary for it: the no-tie coefficient is large and significant by itself. Conversely, the phone-network spillovers of Table 1 are essentially unchanged when every driver–contact pair sharing a regular customer is removed (all coefficients within 4%; Appendix Table B7). Customer demand and social learning therefore operate as separate channels. Varying the shared-customer threshold preserves the pattern in both directions: with any shared customer (≥ 1 transaction with each party) the within-taxi effect is $+55\%$ and cross-industry remains a precise null, and at ≥ 3 transactions the effect and its community-tie gradient persist in the much smaller sample (Appendix Table B13).

The larger post-period magnitude is expected for two reasons. First, the business-payment network is a far more exhaustive measure of actual customer overlap than the pre-period P2P proxy. Second, customer turnover attenuates the predictive power of (pre-experiment) historical links. However, some of this measured overlap may itself be treatment-induced: because not all control drivers had adopted and used the technology between November 2022 and November 2023 (87.6% in treated vs. 75.4% in control), customers of treated drivers are mechanically more visible in the network than customers of controls. The direction of the resulting bias on $\hat{\beta}$ is ambiguous: inflated exposure to treated drivers pushes the regressor up, but the expanded sample may include marginal businesses with weaker true links. The post-period estimate is therefore best viewed as high-powered but potentially biased. The pre-period result remains the primary evidence that the channel is genuine; Table 5 serves as higher-powered supportive evidence and is used in the structural section. [18]

The cross-industry customer coefficients in Table 5 are positive but small and insignificant (column 4: $\hat{\beta} \times 1{,}000 = 20.19$, $p_{\text{RI}} = 0.180$, effect size $+3.3\%$). Customers who transact with treated taxis could in principle also patronize non-taxi businesses, so the specification could have detected a cross-industry effect had one existed. The most likely explanation is that cross-industry customer links in this network are sparse—few non-taxi businesses share a substantial customer base with taxi drivers. The positive sign is consistent with a small force in the right direction, but the data lack power to detect it. In

[18] Table B15 provides a temporal-stability diagnostic, splitting the post-experiment customer network into two windows (November 2022–April 2023 and May–November 2023). The coefficient is positive in both. The W1 magnitude (+28.2) is roughly four times the W2 magnitude (+6.5), an attenuation consistent with customer turnover eroding the effect of shared customers over time.

the structural exercise below, the customer-demand channel is accordingly set to zero for cross-industry spillovers.

5.4 Widespread Diffusion via Remote Interactions

The experiment takes place in Dakar, but Section 4.2 showed that exposure to treated taxis causes cross-industry adoption in other Senegalese cities. One possible explanation is physical movement: treated drivers or their contacts travel to secondary cities and transmit the technology in person. Mobile money geolocation data—which records the city of each transaction during the experiment window (March–October 2022)—allow us to test this directly.

Seasonal migration is widespread. Among the individuals in the community-tie sample (Table 1 Panel A, $\geq 2\%$ corrected overlap), approximately half can be matched to geolocation records. Of these, roughly two-thirds appeared in Dakar at least once during the experiment, while one-third were never observed in the capital despite maintaining strong phone ties with Dakar-based drivers. Seasonal outmigration from Dakar is common: roughly a third of Dakar-based businesses (33.7%) and drivers (34.1%) transacted outside the capital in at least one month during the experiment. The rate is similar across treatment arms (32.6% treated vs. 36.1% control, $p = 0.103$ from a two-proportion z-test), so migration does not seem driven by the technology. These figures establish that physical mobility between Dakar and secondary cities is substantial—precisely the kind of pattern that would sustain a contact-based diffusion channel.

Neither party's movement explains the effect. Table 1 Panel A already establishes cross-city diffusion at every threshold. Table 6 shows the overall effect on the community-tie sample ($+0.068$, $p_{\text{RI}} = 0.009$, $+142\%$ of the control mean) and decomposes it by physical mobility, measured from mobile money geolocation. Panel A splits receiving businesses by their *own* movement: the effect is statistically indistinguishable from zero among those observed in Dakar during the experiment ($p_{\text{RI}} = 0.118$) and is concentrated among businesses *never* observed in the capital, whose coefficient is five times larger ($+0.182$, $p_{\text{RI}} = 0.027$). Physical presence in Dakar is therefore not required for adoption to respond to treated Dakar contacts.

Drivers who migrate to a city do not carry the technology there. Panel B of Table 6 tests the complementary channel: whether diffusion to a secondary city runs through

treated *drivers* who physically travel there. For each destination, the treatment is split into exposure to treated drivers who visited that city at least once during the experiment (per agent geolocation) and exposure to treated drivers who never did, pooling the eight largest secondary cities with city fixed effects. The two channels are not symmetric in what the data can say about them. The migrant-driver coefficient is three times larger than the non-migrant one ($+0.021$ vs. $+0.007$ per 100) but imprecise ($p_{\text{RI}} = 0.133$), so a substantial role for drivers physically carrying the technology cannot be ruled out—seasonal migration may well be part of the story. What the design does establish is that physical travel is not *necessary*: exposure to treated drivers who never visited the destination raises adoption there by $+75\%$ of the control mean ($p_{\text{RI}} = 0.018$), so remote phone interactions alone move the technology across cities. A merchant-level split that appears to concentrate effects among individuals who themselves migrate to the destination city confounds treatment with network geography—individuals connected to city-c visitors disproportionately live near city c; once expected exposure absorbs network position, migration carries no *detectable* additional effect, though the imprecision of the migrant coefficient counsels against reading this as evidence of absence.

Tie strength and cross-industry diffusion across space. The distant cross-industry effects strengthen with tie intensity. Across the exclusive strength bands (Appendix Figure A10), the outside-Dakar effect size rises from $+67\%$ in the 2–5% band to $+76\%$ at 8–11% and $+72\%$ in the open tail (both $p_{\text{RI}} \leq 0.064$), while the within-Dakar effect is flat across bands—consistent with local diffusion operating through simple observation, which any real tie provides, whereas transmission to secondary cities requires the redundancy or trust of more intensely shared social environments. Close ties carry information both locally and to people the driver knows well but who live and work elsewhere.

Taken together, these results show that cross-city diffusion operates through remote social learning, with seasonal migration possibly contributing but not required. Individuals who maintain community phone ties with treated drivers in Dakar learn about the technology through those social connections and act on it where they live and work—without either party needing to travel: businesses that never visited Dakar respond at least as strongly as those that did, and the effect of treated drivers who never visited the destination is significant on its own. The migrant-driver coefficient is larger but imprecisely estimated, so drivers physically carrying the technology may add to the remote channel. Because these are cross-industry businesses (the outcome is non-taxi adoption), competition and customer demand are not active channels; the effect is attributable to social learning transmitted across space through the phone network. Seasonal migra-

tion between Dakar and secondary cities is a large phenomenon in West Africa—a third of businesses and drivers move seasonally—which makes the sufficiency of the remote channel notable: the information moves through the network itself, even if it may also travel with migrants.

6 Structural Model

The reduced-form evidence identifies three channels of technology diffusion, each with a different sign, operating through a different network layer. Social learning transmits through phone contacts and is positive. Geographic competition transmits through physical proximity to garages and is negative. Customer demand transmits through shared transaction partners and is positive. Because these channels have opposite signs, aggregate spillovers can be misleading: the average within-industry coefficient appears negative even though learning is strongly positive on meaningful social ties. This section estimates a structural model that decomposes adoption into these three channels, quantifies each, and uses the estimates for counterfactual targeting policies. Appendix Table B10 consolidates the construction, timing, and use of each network object.

6.1 A Model of Adoption with Learning, Competition, and Demand

Treatment was randomized on March 28, 2022; adoption is measured through October 5, 2022, yielding $T = 6$ monthly periods. Adoption is absorbing: once adopted, the decision is irreversible. At each month, non-adopter j adopts with probability:

$$p_j = \Lambda\Bigg(\alpha_g + \underbrace{\left[\beta^s_{\text{learn}} \cdot \mathbf{1}[\text{strong}_j] + \beta^w_{\text{learn}} \cdot \mathbf{1}[\text{weak}_j]\right] \cdot S^{\text{phone}}_j}_{\text{social learning}} + \underbrace{\beta^{\text{geo}}_{\text{comp}} \cdot S^{\text{geo}}_j \cdot \mathbf{1}[\text{taxi}_j]}_{\text{geographic competition}} + \underbrace{\beta_{\text{demand}} \cdot S^{\text{cust}}_j \cdot \mathbf{1}[\text{taxi}_j]}_{\text{customer demand}}\Bigg), \quad (3)$$

where $\Lambda(\cdot)$ is the logistic CDF and $\alpha_g \in \{\alpha_{\text{cross}}, \alpha_{\text{taxi}}\}$ is an industry-specific intercept. Cumulative adoption over six months is $\Pr(\text{adopt}_j) = 1 - (1 - p_j)^6$.

Each channel operates through a distinct network:

- *Phone network* (S^{phone}): the share of j's community phone contacts (corrected overlap $\geq 2\%$) who are treated. Strong ties ($\geq 5\%$ corrected overlap) and weaker community

ties (2–5%) have separate learning rates. In the model, this is the only channel available to cross-industry businesses.

- *Customer network* (S^{cust}): the share of j's customer-linked drivers who are treated, constructed from 133 million pay-to-business (P2B) mobile money transaction records.[19] When a customer who pays both a treated driver and taxi business j starts using digital payments, the customer creates demand pressure on j to adopt.

- *Geographic network* (S^{geo}): the treatment intensity at j's nearest garage, discounted by distance:

$$S_j^{\text{geo}} = \text{ShareTreated}_{\text{nearest}(j)} \cdot \exp\big(-\delta \cdot d(j, \text{nearest}(j))\big), \tag{4}$$

  where d is distance in km and $\delta > 0$ is an estimated decay rate. The exponential kernel is standard in economic geography (Ahlfeldt et al., 2015) and fits the data: Table 4 reports coefficients at six distance bins (50–500 m), and $\log|\hat{b}(d)|$ is approximately linear in d ($R^2 = 0.96$), precisely what exponential decay predicts.

The model has seven parameters: two intercepts (α_{cross}, α_{taxi}), two learning rates (β^s_{learn}, β^w_{learn}), geographic competition ($\beta^{\text{geo}}_{\text{comp}} \leq 0$), customer demand ($\beta_{\text{demand}}$), and distance decay ($\delta$).

6.2 Identification and Estimation

I match reduced-form moments to the model's predictions via minimum distance estimation (Chamberlain, 1984; Newey and McFadden, 1994). Five moments (m1, m2, m4, m5, m6) each identify one parameter. The within-taxi phone moment (m3) provides a consistency check: the population-weighted combination $w_s\beta^s_{\text{learn}} + w_w\beta^w_{\text{learn}}$ is algebraically implied by m1 and m2 under the indirect inference procedure (which identifies each β from a regression on the same exposure variable), so model-data equality on m3 is guaranteed by construction rather than representing an independent testable restriction.[20] The distance gradient (m7) uses six distance-bin coefficients to identify two parameters (δ from the slope, $\beta^{\text{geo}}_{\text{comp}}$ from the level), leaving four further overidentifying restrictions ($R^2 = 0.96$). In total, the system has seven parameters identified from eleven moment

[19] A customer qualifies if they made ≥ 2 P2B transactions in aggregate across non-garage experimental drivers between November 2022 and November 2023 (not necessarily with the same driver). A business–driver pair is linked if they share ≥ 1 qualifying customer. This yields 95,114 businesses linked to 1,521 experimental drivers. Correlation with S^{phone}: $r = 0.005$.

[20] If the learning coefficients were identified by a different method (e.g., one that matched cumulative rates rather than OLS coefficients), m3 would provide a genuine overidentifying test. Under the current indirect inference approach, the linear structure of OLS ensures exact aggregation.

conditions (five scalar moments plus six distance bins), leaving four overidentifying restrictions (all from the distance gradient). Because all exposure variables are functions of the exogenous treatment assignment, the model has a closed-form adoption probability—no cascade simulation is needed, in contrast to Banerjee et al. (2013).

#	Source	Target	Identifies
m1	Table 1, $\geq 5\%$ overlap, cross-ind.	OLS coef.	β^s_{learn}
m2	2–5% band (Appendix Fig. A10), cross-ind.	OLS coef.	β^w_{learn}
m3	Table 1, pooled (w_s+w_w), within-taxi	OLS coef.	$w_s\beta^s_1 + w_w\beta^w_1$
m4	Table 5, column 1	OLS coef.	β_{demand}
m5	Cross-industry adoption rate	Rate	α_{cross}
m6	Within-taxi adoption rate	Rate	α_{taxi}
m7	Table 4 Panel A gradient (6 bins); Panel B 100 m level	Slope + level (4 overid.)	δ, $\beta^{\text{geo}}_{\text{comp}}$

The identification logic is sequential. Cross-industry moments (m1, m2) identify learning separately for strong and weak ties, because competition and demand are absent across industries. The within-taxi phone moment (m3) loads only on learning ($w_s\beta^s_{\text{learn}} + w_w\beta^w_{\text{learn}}$), since competition operates through the separate geographic exposure S^{geo}. The within-taxi customer moment (m4) isolates demand: among non-garage taxi businesses sharing customers (defined by ≥ 2 P2B transactions, November 2022–November 2023) with experimental drivers, the coefficient is $+255.99^{***}$ ($p_{\text{RI}} = 0.001$, $N = 8{,}523$). Because the post-experiment customer network may itself be partly treatment-induced (Section 5.3), this moment likely represents an upper bound on the demand parameter; it is used here for its statistical power, while the pre-intervention estimate in Table B14 provides the cleanest causal evidence for the channel. Adoption rates (m5, m6) pin down the intercepts. The gradient of garage coefficients across distance bins (m7) identifies δ from the log-linear slope and $\beta^{\text{geo}}_{\text{comp}}$ from the level (Appendix D.1).

6.3 Results

Table 7 presents the estimates.

Learning: uniform across the community network. $\hat{\beta}^s_{\text{learn}} = +0.271^{*}$ versus $\hat{\beta}^w_{\text{learn}} = +0.267$: statistically indistinguishable. Once ties are measured by corrected overlap, information flows at essentially the same per-tie rate through the weaker (2–5%) and stronger ($\geq 5\%$) segments of the community network, matching the uniform per-band transmission in the reduced form (Appendix Figure A10). What predicts diffusion is having a community tie; intensity adds little. A driver's reach—the number of community ties—is correspondingly the policy-relevant statistic.

Geographic competition: intense and local. $\hat{\beta}^{\text{geo}}_{\text{comp}} = -10.004^{***}$, with a half-life of 186 m ($\hat{\delta} = 3.73^{***}$ km^{-1}). At 50 m from a garage, competitive pressure is $\sim$24$\times$ stronger than the influence of full community-network exposure; at 500 m it has fallen to $\sim$11$\times$, and the two forces equalize at roughly 2 km from a garage (Appendix D.3). This reflects the local nature of taxi competition: drivers at the same garage compete for the same customers and parking spots, while drivers a few hundred meters away operate in largely separate markets.

Separating competition from the phone network helps interpret a puzzle in the reduced-form evidence. The within-taxi phone coefficient (m3) is null at the community threshold and negative at strong ties (Table 2; Appendix Figure A10), suggesting that within-industry phone exposure embeds some competitive pressure. The structural model is consistent with this reading but locates the competitive force primarily in geography: same-industry phone ties are spatially correlated—taxi drivers who know each other often operate nearby—so a treated within-industry phone contact also raises S^{geo}. The garage distance gradient, with a half-life of 186 m, pins down competition as a sharply local, geographic channel. Once S^{geo} is netted out, the residual phone coefficient reflects learning, which operates uniformly across tie strength ($\hat{\beta}^{w}_{\text{learn}} = +0.267 \approx \hat{\beta}^{s}_{\text{learn}}$). The near-null pooled m3 coefficient—predicted by the model to within a rounding error without being targeted—thus reflects that within-industry phone exposure is correlated with the geographic proximity that carries competition. Consistent with this, the strong-tie within-industry negative is concentrated within a few hundred meters of garages, though a smaller weak-tie negative away from garages indicates geography does not absorb the competitive force entirely.

Customer demand. $\hat{\beta}_{\text{demand}} = +0.795^{***}$: non-garage taxi businesses sharing P2B customers ($\geq$ 2 transactions, November 2022–November 2023) are 35.7% more likely to adopt (Table 5, column 1; $p_{\text{RI}} = 0.001$). What kind of customers create this demand linkage? Among the 28,609 qualifying customers, 23% are *repeat patrons* of at least one driver (≥ 2 transactions with the same driver), and 86% transact with ≥ 2 experimental drivers. These are regular mobile money taxi riders who rotate across the fleet. Crucially, 99% of them also make P2B payments to non-taxi businesses (median 26 non-taxi businesses per customer, 78 total transactions in the window). When a treated driver onboards a customer into digital payments, that customer's pre-existing spending at dozens of other businesses creates demand pressure for adoption.

The demand channel is not predominant near taxi garages: fewer than 1% of garage-proximate taxi businesses share mobile money customers with experimental drivers, so

the demand channel affects very few businesses in that population even though the conditional effect among the linked few can still be large. Customers who ride with drivers at a garage are predominantly passengers, not the regular transaction customers who generate demand linkages in the mobile money network.

Demand complementarity is identified within the taxi industry, where the customer-link data are richest. The structural model accordingly restricts the channel to within-taxi spillovers; allowing it to operate across industries would, if anything, strengthen the counterfactual results.[21]

Validation. The model correctly predicts the sign of all four untargeted survey moments (Table 3): positive for taxi friends (m9) and non-friends (m10), positive for all-contact friends on the within-taxi outcome (m8), and small in absolute terms for serendipitous contacts (m11 = $+0.018$). Moment m9 is the most demanding of these checks: taxi friends are same-industry competitors, yet the coefficient is positive ($+0.098$), which the model attributes to demand dominating competition. It underpredicts magnitudes by 3–30$\times$ for m8–m10, primarily because survey "friends" are the right tail of tie strength and because baseline adoption is rare; m11 is excluded from this range because the model prediction is essentially zero while the data value, though statistically significant in Table 3 ($p_{\text{RI}} = 0.004$), is small in absolute probability units relative to m8–m10 (Appendix D.2).

7 Counterfactuals and Targeting

For the same onboarding budget of $N_T = 1,256$ drivers, how much additional adoption could alternative targeting rules generate?[22]

7.1 Setup

I fix all three networks at their observed values, recompute each individual's exposure under alternative treatment assignments, and predict six-month adoption. The counterfactuals are evaluated on a *combined reachable network* of 3,103,114 individuals for whom

[21] Among cross-industry businesses who are also community-tied to drivers (Table 5, column 5), the customer coefficient is positive but imprecise ($+40.3$, $p_{\text{RI}} = 0.425$, $N = 2{,}791$), suggesting demand may operate across industries when mediated by a social relationship. This noisier estimate is not used in the structural model.

[22] The experimental sample contains 2,196 drivers ($\sim$58% treated). Of these, 2,150 have a matched mobile money account (1,256 treated, 894 control), and 1,907 appear in the phone network (Section 1.3). The counterfactual treatment vector operates on the 2,150 matched drivers; the network exposure is computed for all individuals connected to these drivers.

at least one structural channel is active (Appendix E).

What changes. Changing which drivers are treated shifts all three exposure variables simultaneously. S^{phone} shifts for every connected individual. S^{cust} shifts for businesses sharing customers with different drivers. S^{geo} shifts because the treatment share at each garage changes—a garage where all drivers are treated has ShareTreated $= 1$, generating intense local competition. The comparison between degree targeting (high $\bar{S}^{\text{phone}}$, diffuse geography) and garage targeting (lower $\bar{S}^{\text{phone}}$, concentrated geography) illustrates the cost of geographic concentration.

7.2 Targeting Rules

All scenarios hold the budget fixed at $N_T = 1,256$:

1. *Random*: the realized experimental assignment ($\bar{S}^{\text{phone}} = 0.586$).
2. *All-garage*: what if all $1,256$ treated drivers had the network characteristics of garage drivers? Phone exposure rises modestly ($\bar{S}^{\text{phone}} = 0.659$ vs. 0.586) because the scenario rescales treated drivers' network intensities to the garage-driver mean; customer demand is similar ($\bar{S}^{\text{demand}} = 0.700$); and geographic competition reaches its maximum (ShareTreated$_g = 1$ at every garage). See Appendix E for the reweighting procedure.
3. *Top mobile money users*: the 1,256 heaviest platform users. Requires no network data.
4. *Highest degree*: the 1,256 drivers with the most phone contacts ($\bar{S}^{\text{phone}} = 0.893$).
5. *Geographic reach*: the 1,256 drivers with the most out-of-Dakar contacts.
6. *Garage-concentrated*: treatment is concentrated at the largest garages until the budget is exhausted.

7.3 Results

Table 8 reports treatment-induced adoption—predicted adoption minus the no-treatment baseline—under each rule, decomposed by margin. Per-margin columns can exceed the total because some individuals belong to multiple populations; the total counts each person once.

Cross-industry adoption scales with phone exposure. This margin accounts for 94% of induced adoption under random onboarding (6,024 of 6,393) and is powered by

community-tie learning. Degree targeting raises it by +59.0 ($\bar{S}^{\text{phone}} = 0.893$ vs. 0.586). All-garage raises it only modestly (+13.7%)—garage drivers reach the community network, but far less than high-degree drivers do.

Within-taxi: a statistical zero at national scale. The reachable network's within-taxi margin is tiny—the model's no-treatment baseline is roughly nine adopters among the 10,442 reachable taxi merchants (the observed community-network taxi rate is 0.058%), and no targeting rule moves this margin by more than a handful: predicted within-taxi adopters range from 8 (all-garage) to 13 (geographic reach) across the rules in Table 8. Competition and demand still shape which taxi businesses adopt (Section 5), but at the national scale nearly all induced adoption occurs on the cross-industry margin.

Total: network targeting dominates. Degree targeting produces 10,102 induced adopters (58% vs. random). Garage-based rules capture at most a quarter of this gain: the all-garage benchmark—which also scales drivers' network intensities to the garage-driver mean—reaches 7,261 (+13.6%), and concentrating treatment at the largest garages, where competitive crowding is maximal, is the only rule below random (-1.2%). The structural decomposition highlights a gap of 59 percentage points between the best and worst rules: garages make treatment visible, but visibility itself does not diffuse the technology, and concentrating treatment at the largest garages also concentrates competitive crowding.

7.4 Policy Implications

Three lessons emerge for a social planner maximizing technology adoption from a fixed onboarding budget. These results contribute to a broader literature on policy design for firm-level upgrading in developing countries (Verhoogen, 2023), where evaluating the cost-effectiveness of specific interventions has been highlighted as a key open question.

Network targeting is powerful. Selecting high-degree drivers produces 58% more adoption than random, implementable from phone fingerprint data already collected for the overlap computation. In budget terms, degree targeting reaches the full random rollout's induced adoption with roughly half the onboarding budget ($\sim$630 of 1,256 drivers), and the naive degree ranking captures essentially all of the gain available to a sophisticated coverage-maximizing (greedy) seeding rule at this scale—the value lies in the network data, not in the optimization. Targeting drivers by their community ties to secondary cities delivers essentially the same portfolio and the same outcomes as degree targeting: drivers' networks are dispersed enough—over half of located reachable contacts live

outside Dakar—that every high-degree driver is also a bridge, so maximizing national diffusion and maximizing secondary-city reach select the same firms. Geographic reach targeting performs comparably (53%). The gains come overwhelmingly from the cross-industry learning channel, which serves the vast majority of non-taxi individuals in the combined network.

Simple heuristics work. Onboarding heavy mobile money users—requiring no phone network data—produces 6,656 induced adopters (4.1% over random). Heavy users tend to occupy more central network positions, so the heuristic partially captures the degree-targeting gains without requiring social-network information.

The cost of geographic concentration. The strategy that might appear most natural from a diffusion standpoint—concentrating treatment at dense, visible garages—captures almost none of the achievable improvement: garage rules cluster within a few percent of random, and the most concentrated variant is the only rule below it. Garage drivers' community reach does not match that of high-degree drivers, and maximal geographic concentration activates the competition channel, whose half-life of 186 m makes the penalty severe within $\sim$200 m of a garage. A planner who targets garages because they are visible and easy to reach maximizes crowding while leaving the network channel, which accounts for 94% of induced adoption, largely unused.

8 Conclusion

This paper combines a randomized experiment with exhaustive network data—phone contact directories covering 70–80% of Senegal's adult population and the complete mobile money transaction ledger—to study when technology diffuses across firms and when it does not.

Three channels shape diffusion, each identified through a distinct empirical variation. Social learning operates throughout the community network—at a uniform per-tie rate across the strength distribution—and is the only channel that reaches across industries and across cities. Within-industry competition generates negative spillovers that are large at short distances and decay steeply (half-life 186 m), partially offsetting learning gains among same-sector firms. Customer-demand complementarities—transmitted through shared customers, not social ties—are the dominant force within the taxi sector. The sign and magnitude of spillovers depend on which channel is activated by which type of tie, which may help explain why studies observing different parts of the diffusion process have reached different conclusions.

The paper makes two contributions that go beyond documenting the existence of

spillovers. First, because the data and experimental design allow me to separately identify and quantify each channel—learning on strong ties, competition at short distances, demand through shared customers—I can address a question the literature has not been able to tackle: how should a social planner target a fixed onboarding budget to maximize diffusion? The answer depends entirely on the channel structure. The counterfactuals show that selecting drivers with the most phone contacts generates 58% more adoption than random assignment, while the intuitively appealing strategy of concentrating treatment at the most visible locations—the largest garages—performs marginally worse than random, because the visibility gain is offset by maximal competitive crowding and weak network reach. The 59-percentage-point gap between the best and worst targeting rules would be challenging to anticipate—or even to sign correctly—without this decomposition.

Second, the paper shows that technology can diffuse over long distances: information travels remotely through phone-contact ties, and physical movement is not required. Adoption in secondary cities does not depend on the adopting business itself visiting Dakar, and exposure to treated drivers who never traveled to the destination raises adoption there on its own (Table 6); the channel through drivers who did travel is larger in magnitude but too imprecise to certify, so migration may add to the remote channel. The result refines the Marshallian view of knowledge spillovers: neither the transmission of information nor its conversion into adoption requires co-presence. Seasonal migration is pervasive in this setting—roughly a third of Dakar-based businesses and drivers move seasonally—which makes the sufficiency of the remote channel all the more notable: social networks carry urban innovations across space faster and more broadly than people do.

More broadly, the combination of random assignment of technology access, exhaustive network data, and nationwide diffusion of a single technology offers a template for studying knowledge spillovers at the scale of an entire economy. Small firms are central to output and employment in low- and middle-income countries, and their adoption decisions are shaped by social networks that existing data rarely capture. How to design interventions that leverage these networks—targeting the right firms, through the right ties, while accounting for the channels that help and those that hurt—remains an important question for both research and policy.

References

Adamic, Lada A. and Eytan Adar, "Friends and Neighbors on the Web," *Social Networks*, 2003, *25* (3), 211–230.

Ahlfeldt, Gabriel M., Stephen J. Redding, Daniel M. Sturm, and Nikolaus Wolf, "The Economics of Density: Evidence from the Berlin Wall," *Econometrica*, 2015, *83* (6), 2127–2189.

Alvarez, Fernando, David Argente, Francesco Lippi, Esteban Mendez, and Diana Van Patten, "Strategic Complementarities in a Dynamic Model of Technology Adoption: P2P Digital Payments," *Working Paper*, 2023.

Aral, Sinan, "The Future of Weak Ties," *American Journal of Sociology*, 2016, *121* (6), 1931–1939.

Arrow, Kenneth J, "The economic implications of learning by doing," 1962.

Athey, Susan, Dean Eckles, and Guido W. Imbens, "Exact p-Values for Network Interference," *Journal of the American Statistical Association*, 2018, *113* (521), 230–240.

Atkin, David, Keith Chen, and Anton Popov, "The Returns to Face-to-Face Interactions: Knowledge Spillovers in Silicon Valley," *Working Paper*, 2024.

Baird, Sarah, J. Aislinn Bohren, Craig McIntosh, and Berk Özler, "Optimal Design of Experiments in the Presence of Interference," *Review of Economics and Statistics*, 2018, *100* (5), 844–860.

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

Baum-Snow, Nathaniel, Nicolas Gendron-Carrier, and Ronni Pavan, "Local Productivity Spillovers," *American Economic Review*, April 2024, *114* (4), 1030–69.

Bernard, Andrew B. and Andreas Moxnes, "Networks and Trade," *Annual Review of Economics*, 2018, *10*, 65–85.

Bernhardt, Arielle, Erica Field, Rohini Pande, and Natalia Rigol, "Household Matters: Revisiting the Returns to Capital among Female Microentrepreneurs," *American Economic Review: Insights*, September 2019, *1* (2), 141–60.

Björkegren, Daniel, "The Adoption of Network Goods: Evidence from the Spread of Mobile Phones in Rwanda," *The Review of Economic Studies*, 07 2018, *86* (3), 1033–1060.

Bloom, Nicholas, Mark Schankerman, and John Van Reenen, "Identifying Technology Spillovers and Product Market Rivalry," *Econometrica*, 2013, *81* (4), 1347–1393.

Borusyak, Kirill and Peter Hull, "Nonrandom Exposure to Exogenous Shocks," *Econometrica*, 2023, *91* (6), 2155–2185.

Breza, Emily, Arun G. Chandrasekhar, Benjamin Golub, and Aneesha Parvathaneni, "Networks in Economic Development," *Oxford Review of Economic Policy*, 2019, *35* (4), 678–721.

Brock, William A. and Steven N. Durlauf, "Discrete Choice with Social Interactions," *Review of Economic Studies*, 2001, *68* (2), 235–260.

Broder, Andrei Z., "On the Resemblance and Containment of Documents," in "Proceedings of the Compression and Complexity of Sequences" IEEE 1997, pp. 21–29.

Cai, Jing, Alain De Janvry, and Elisabeth Sadoulet, "Social Networks and the Decision to Insure," *American Economic Journal: Applied Economics*, 2015, *7* (2), 81–108.

_ **and Adam Szeidl**, "Interfirm Relationships and Business Performance*," *The Quarterly Journal of Economics*, 12 2017, *133* (3), 1229–1282.

Cefalà, Luisa, Franck Irakoze, Pedro Naso, and Nicholas Swanson, "The Economic Consequences of Knowledge Hoarding," *Working Paper*, 2025.

Chamberlain, Gary, "Panel Data," *Handbook of Econometrics*, 1984, *2*, 1247–1318.

Chaurey, Ritam, Gaurav Nayyar, Siddharth Sharma, and Eric Verhoogen, "Social Learning among Urban Manufacturing Firms: Energy-Efficient Motors in Bangladesh.," *Working Paper*, 2025.

Chuang, Yating and Laura Schechter, "Social Networks in Developing Countries," *Annual Review of Resource Economics*, 2015, *7*, 451–472.

Crépon, Bruno, Esther Duflo, Marc Gurgand, Roland Rathelot, and Philippe Zamora, "Do Labor Market Policies Have Displacement Effects? Evidence from a Clustered Randomized Experiment," *Quarterly Journal of Economics*, 2013, *128* (2), 531–580.

Cristofaro, Emiliano De, Mark Manulis, and Bertram Poettering, "Private Discovery of Common Social Contacts," *International Journal of Information Security*, 2013, *12* (1), 49–65.

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

_ **, Pulak Ghosh, Apoorv Gupta, and Filippo Mezzanotti**, "Demographics and Technology Diffusion: Evidence from Mobile Payments," March 2024.

Demmler, Daniel, Peter Rindal, Mike Rosulek, and Ni Trieu, "PIR-PSI: Scaling Private Contact Discovery," *Proceedings on Privacy Enhancing Technologies*, 2018, *2018* (4), 159–178.

Duflo, Esther, Daniel Keniston, Tavneet Suri, and Celine Zipfel, "Chat Over Coffee? Technology Diffusion Through Social and Geographic Networks in Rwanda," Technical Report 31368 2026.

Fafchamps, Marcel, Måns Söderbom, and Monique van den Boogart, "Adoption with Social Learning and Network Externalities," *Oxford Bulletin of Economics and Statistics*, 2023. Open Access.

Foster, George M., "Peasant Society and the Image of Limited Good," *American Anthropologist*, 1965, *67* (2), 293–315.

Gouriéroux, Christian, Alain Monfort, and Eric Renault, "Indirect Inference," *Journal of Applied Econometrics*, 1993, *8*, S85–S118.

Granovetter, Mark, "The Strength of Weak Ties," *American Journal of Sociology*, 1973, *78*, 1360–1380.

Grossman, Gene M. and Elhanan Helpman, "Trade, knowledge spillovers, and growth," *European Economic Review*, 1991, *35* (2-3), 517–526.

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

Houeix, Deivy, "Asymmetric Information and Digital Technology Adoption: Evidence from Senegal," *Working Paper*, 2025.

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

Jahani, Eaman, Samuel P. Fraiberger, Michael Bailey, and Dean Eckles, "Long ties, disruptive life events, and economic prosperity," *Proceedings of the National Academy of Sciences*, 2023, *120* (28), e2211062120.

Lee, David S., "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects," *Review of Economic Studies*, 2009, *76* (3), 1071–1102.

Liben-Nowell, David and Jon Kleinberg, "The Link-Prediction Problem for Social Networks," *Journal of the American Society for Information Science and Technology*, 2007, *58* (7), 1019–1031.

Manski, Charles F., "Identification of Endogenous Social Effects: The Reflection Problem," *The Review of Economic Studies*, 1993, *60* (3), 531–542.

Marshall, Alfred, "Principles of Economics," *MacMillan*, 1920.

Moretti, Enrico, "The Effect of High-Tech Clusters on the Productivity of Top Inventors," *American Economic Review*, October 2021, *111* (10), 3328–75.

Myers, Kyle R. and Lauren Lanahan, "Estimating Spillovers from Publicly Funded R&D: Evidence from the US Department of Energy," *American Economic Review*, July 2022, *112* (7), 2393–2423.

Newey, Whitney K. and Daniel McFadden, "Large Sample Estimation and Hypothesis Testing," *Handbook of Econometrics*, 1994, *4*, 2111–2245.

Noll, Landon Curt, Kiem-Phong Vo, Donald E. Eastlake, and Tony Hansen, "The FNV Non-Cryptographic Hash Algorithm," RFC 9923, RFC Editor February 2026.

Rodrik, Dani, "Premature Deindustrialization," *Journal of Economic Growth*, 2016, *21* (1), 1–33.

Romer, Paul M, "Endogenous technological change," *Journal of Political Economy*, 1990, *98* (5, Part 2), S71–S102.

Suri, Tavneet, "Mobile Money," *Annual Review of Economics*, 2017, *9*, 497–520.

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

Swamidass, S. Joshua, Matthew Matlock, and Leon Rozenblit, "Securely Measuring the Overlap between Private Datasets with Cryptosets," *PLoS ONE*, 2015, *10* (2), e0117898.

Verhoogen, Eric, "Firm-Level Upgrading in Developing Countries," *Journal of Economic Literature*, December 2023, *61* (4), 1410–1464.

Main Figures and Tables



Figure 1: Digital Payments for Businesses — Senegal

Notes: These images illustrate the implementation of the technology across different business settings. The top panels show the deployment within taxi vehicles: the QR code is visible from outside the vehicle through a windshield sticker, allowing potential customers to easily identify it. Drivers were equipped with the corporate mobile application and trained on how to use it by a dedicated mobile money agent. Passengers have the option to pay using their smartphones, or drivers can scan passengers' mobile money cards when they do not have smartphones. The bottom panels show the same QR code technology deployed in other sectors, including open-air market vendors (left) and retail stores (right), where businesses display the QR code as a sticker at their point of sale for customers to scan and pay. Businesses in different sectors were provided with the same QR code technology, with displays adjusted to the sector.

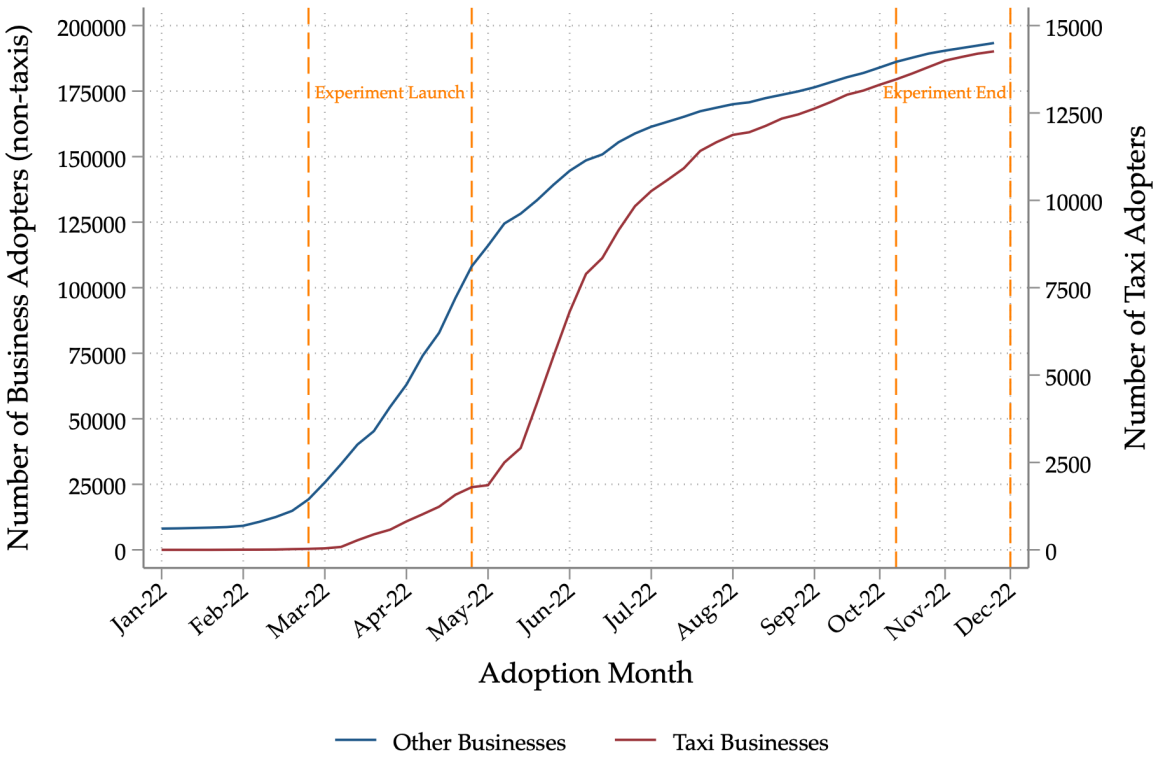


Figure 2: Digital Payments' Adoption in 2022 - *Author's calculations*

Notes: This figure illustrates the cumulative adoption of digital payment technology by businesses in Senegal throughout 2022, covering both taxi businesses and other industries. The adoption rate follows an S-shaped curve for both other businesses and taxi businesses. The experiment commenced on March 28, 2022, when the first taxi businesses were equipped with the technology. A week after the end of the baseline survey, taxis outside the experiment were permitted to adopt the digital payment technology, while the control group remained on a waitlist for the time of the experiment. Throughout the year, all other businesses also had the opportunity to adopt the technology. The first endline survey was conducted on October 5, 2022, and the experiment randomizing access to digital payment technology concluded in late December 2022. By early 2024, approximately 17,000 taxis and 220,000 other businesses had adopted the technology. By October 2022, a substantial share of non-taxi businesses had already adopted, while taxi adoption ($\sim$7,000 of an eventual $\sim$17,000) was still growing.

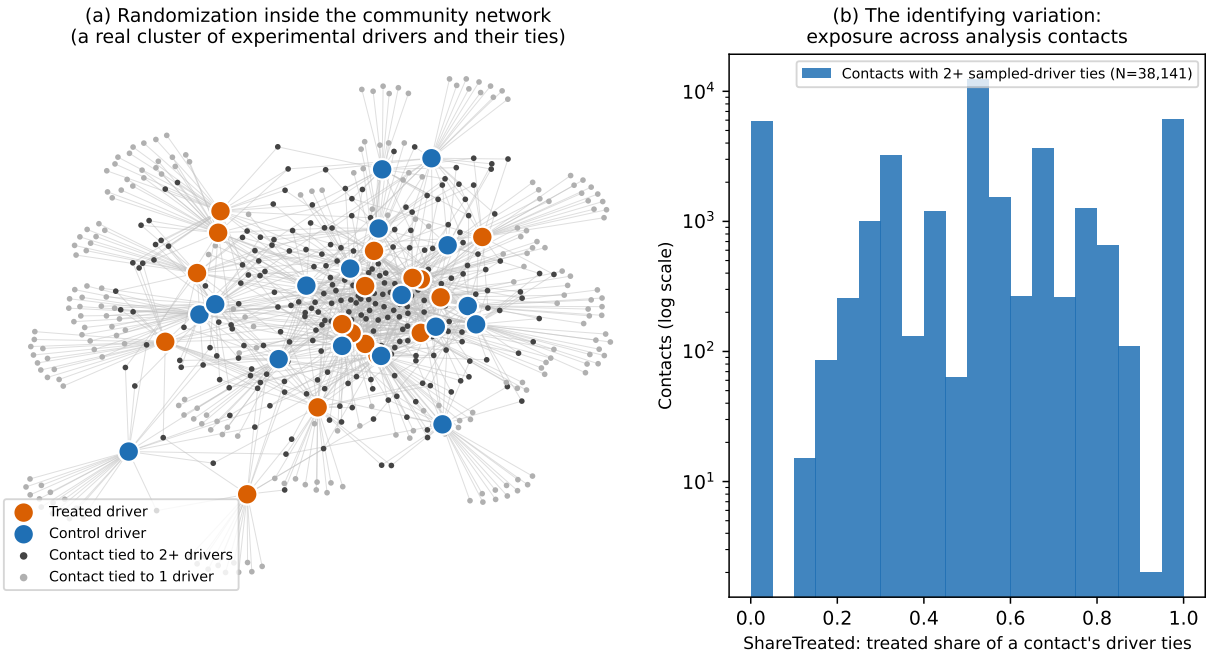


Figure 3: The Experimental Design in the Community Network

Notes: Both panels are drawn from the actual analysis data. *Panel (a)*: a connected cluster of 34 experimental drivers (selected by shared-contact intensity from a random sample of the experimental population, balancing arms) with a sample of their community-tied contacts (bias-corrected contact-book overlap $\geq 2\%$; Section 1.3). Orange and blue circles are drivers as randomized into treatment and control; dark gray dots are contacts tied to two or more experimental drivers—the individuals for whom randomization generates variation in exposure—and light gray dots are single-driver contacts. Treatment was assigned to drivers embedded in one connected social fabric; spillovers are identified from contacts' positions within it. *Panel (b)*: the realized distribution of ShareTreated_j (the treated share of contact j's experimental driver ties) among multi-tie contacts of the sampled drivers—the identifying variation used throughout the analysis (log scale; spikes at coarse fractions reflect contacts with few driver ties). A stylized schematic of the design is in Appendix Figure A2.

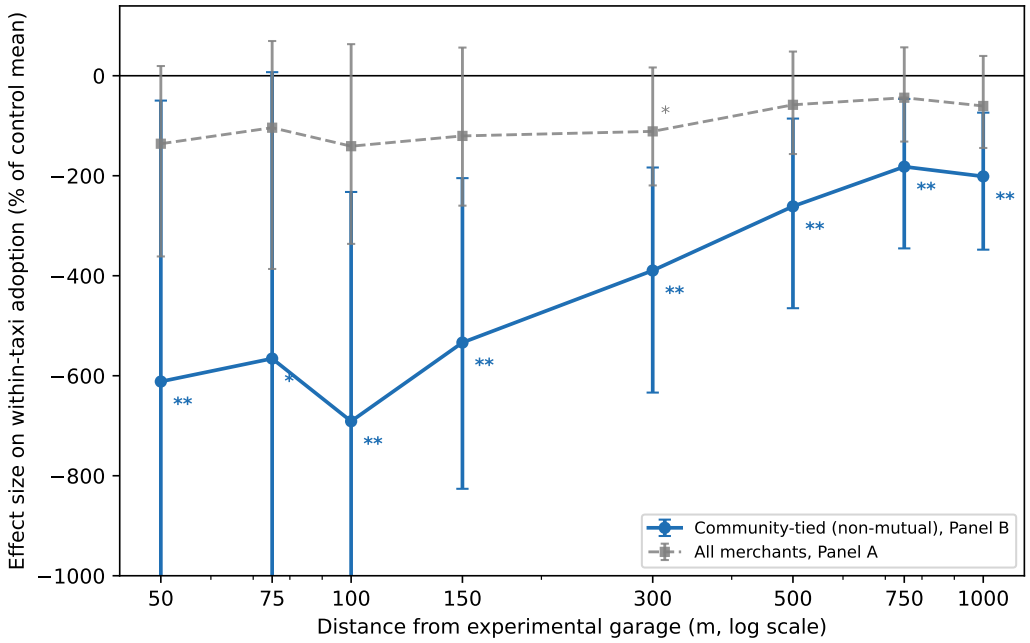


Figure 4: Garage proximity spillovers: effect size by distance

Notes: From Table 4: Panel B (blue, merchants with a community tie to the garage's drivers—corrected contact-book overlap $\geq 2\%$ with at least one of them, excluding mutual contacts) and Panel A (gray, all merchants), extended to 750 m and 1 km. Effect size $= \hat{\beta}_1/\bar{y}$, where $\hat{\beta}_1$ is the coefficient on AnyTreated and $\bar{y}$ is the control mean adoption rate. Error bars are 95% *randomization-inference* confidence intervals obtained by test inversion: the set of constant effects τ_0 whose sharp null—imposed by adjusting outcomes to $y - \tau_0$ AnyTreated—is not rejected at the 5% level across the 1,000 placebo re-randomizations. Bars and stars are thus the same inference object; conventional cluster-robust intervals (20–36 garage clusters) are not relied on. Stars denote RI p-values at the zero null: Panel B has $p_{\text{RI}} \leq 0.05$ at seven of eight radii ($p = 0.052$ at 75 m), Panel A is null throughout. The competition effect concentrates on the socially exposed—peaking at -691% at 100 m—and decays with distance beyond the immediate garage area, to -201% at 1 km, versus -104% to -141% (all RI-null) among all merchants—consistent with competitive crowding out transmitted through market-level community exposure rather than through close relationships (mutual-contact merchants show no effect; Table 4, Panel C). The structural estimation (moment m7) uses Panel B's 100 m coefficient.

Table 1: Cross-Industry Network Spillovers by Tie Strength

	$\geq 2\%$ (community)	$\geq 5\%$ (strong)	$\geq 10\%$
Cross-industry, Dakar			
Share treated	$+2.91$ {+33%}	$+4.16^{**}$ {+31%}	$+5.16^{**}$ {+48%}
	(0.50)	(0.29)	(0.42)
	[0.123]	[0.017]	[0.015]
Expected exposure	-1.78	-22.75	-32.48
	(15.92)	(9.60)	(13.96)
	[0.911]	[0.018]	[0.020]
Cross-industry, outside Dakar			
Share treated	$+0.68^{**}$ {+76%}	$+0.68^{***}$ {+53%}	$+0.73^{**}$ {+73%}
	(0.16)	(0.09)	(0.13)
	[0.011]	[0.009]	[0.030]
Expected exposure	$+8.35$	-2.85	-1.00
	(5.52)	(3.02)	(4.13)
	[0.130]	[0.345]	[0.810]
Dep. var. mean (Dkr / outside)	8.78 / 0.89	13.43 / 1.29	10.77 / 1.00
N	1,145,899	1,020,391	376,871

Notes: All estimates $\times 1{,}000$; effect sizes ($\beta_1/\bar{y}$) in braces; HC1 standard errors in parentheses; p-values in brackets—randomization inference (1,000 permutations) for Share treated, HC1 for Expected exposure (a control, not re-randomized). Each cell estimates $y_j = \beta_1 \,\text{ShareTreated}_j + \beta_2 \,\text{ExpExposure}_j + \alpha$: y_j indicates registration as a non-taxi merchant (in Dakar, or outside Dakar) during the experiment window, ShareTreated_j is the treated share of contact j's experimental driver contacts, and ExpExposure_j is its mean across the 1,000 placebo re-randomizations, absorbing j's network position. Ties are all links with bias-corrected contact-book overlap above the indicated threshold, saved or not (2% floor set above the collision-noise scale and survey-validated; Section 1.3); contacts connected to any top-2% degree driver are excluded entirely. Within-industry effects, where competition operates, are in Table 2; per-band transmission is in Appendix Figure A10; the reciprocity split is in Appendix Table B6; results excluding shared-customer pairs are in Appendix Table B7. $^{*}p_{\text{RI}} < 0.10$, $^{**} < 0.05$, $^{***} < 0.01$.

Table 2: Within-Taxi Network Spillovers by Tie Strength

	$\geq 2\%$ (community)	$\geq 5\%$ (strong)	$\geq 10\%$
Within-taxi, Dakar			
Share treated	$+0.08$ {+13%}	-0.98^{***} {-38%}	-0.56 {-11%}
	(0.09)	(0.12)	(0.28)
	[0.674]	[0.007]	[0.542]
Expected exposure	$+4.09$	-1.09	-0.32
	(2.79)	(3.71)	(9.41)
	[0.143]	[0.770]	[0.973]
Within-taxi, outside Dakar			
Share treated	-0.05 {-67%}	-0.09 {-24%}	$+0.05$ {+5%}
	(0.04)	(0.04)	(0.11)
	[0.162]	[0.190]	[0.787]
Expected exposure	-0.27	-0.37	-3.42
	(0.67)	(1.33)	(3.23)
	[0.685]	[0.780]	[0.290]
Dep. var. mean (Dkr / outside)	0.58 / 0.07	2.62 / 0.37	5.26 / 0.94
N	1,145,899	1,020,391	376,871

Notes: Same specification, samples, community-tie definitions, and trimming as Table 1; outcomes are registration as a *taxi* merchant in Dakar or outside Dakar. The within-industry pattern is more nuanced than the cross-industry one, reflecting learning and competition operating on the same margin with opposite signs. At the community threshold ($\geq 2\%$) the Dakar effect is a statistical null ($+0.08$, $p_{\text{RI}} = 0.674$; the $+13\%$ effect size reflects the very low base rate, 0.58 per 1,000, not a detectable effect); at the strong-tie threshold the effect is *negative* and significant (-0.98, -38%, $p_{\text{RI}} = 0.007$)—drivers socially close enough to share customers and market position compete rather than transmit; at $\geq 10\%$ the point estimate remains negative but loses precision as the sample shrinks and closest ties (family, intimates) dilute the colleague share. Appendix Figure A10 shows the band decomposition: the competitive effect concentrates at 5–8% overlap, and within degree strata it is driven by contacts with one to five strong driver ties—competition bites in small circles. Outside Dakar the community-layer effects are null throughout; the positive cross-market within-taxi channel operates through the reciprocal-save (mutual) layer and the transaction network (Appendix Tables B6 and B12), which this table's community-tie exposure does not isolate. Geographic evidence on the competition mechanism is in Table 4. $^{*}p_{\text{RI}} < 0.10$, $^{**} < 0.05$, $^{***} < 0.01$.

Table 3: Spillover Effects by Relationship Type (Kinship Survey)

	Taxi adoption, Dakar		Cross-ind. adoption		
Subnetwork	Coef	Eff. size	Coef	Eff. size	N
Panel A: Taxi contacts by social proximity					
Taxi × Friends	9.84^{***}	0.361			614
	(3.75)				
	[0.007]				
Taxi × Non-friends	0.88	0.028			549
	(3.94)				
	[0.803]				
Taxi × Freq. daily/weekly	6.95^{**}	0.243			916
	(3.04)				
	[0.023]				
Panel B: All contacts by social proximity					
Friends	3.04^{**}	0.316	0.09	0.013	1,983
	(1.41)		(1.12)		
	[0.025]		[0.920]		
Freq. daily/weekly	1.53^{*}	0.216	0.32	0.065	4,119
	(0.85)		(0.68)		
	[0.055]		[0.630]		
Panel C: Casual ties					
Clients	0.46	0.368	3.51	0.937	395
	(1.23)		(2.27)		
	[0.716]		[0.134]		
Serendipitous encounters	0.29	0.164	1.80^{***}	0.479	5,112
	(0.38)		(0.59)		
	[0.457]		[0.004]		

Notes: Each row estimates $y_j = \alpha + \beta_1 \cdot \text{NumTreated}_{s,j} + \beta_2 \cdot \text{ExpectedExposure}_{s,j} + \varepsilon_j$ on the subsample with ≥ 1 connection in subnetwork s, restricted to baseline survey contacts. Coefficients $\beta_1 \times 100$ reported; HC1 robust standard errors in parentheses. Effect size $= \beta / \bar{y}_{\text{control}}$. "Taxi adoption, Dakar" = within-industry adoption; "Cross-ind. adoption" = across-industry adoption (Dakar + outside). Stars based on randomization inference (RI) p-values from 1,000 permutations of treatment assignment; conventional t-test p-values are similar throughout. Panel A shows within-industry outcomes only, as taxi contacts have no cross-industry adoption variation. "Serendipitous encounters" are contacts classified as casual or unknown relationship. $^{*}p_{\text{RI}} < 0.10$, $^{**}p_{\text{RI}} < 0.05$, $^{***}p_{\text{RI}} < 0.01$.

Table 4: Garage Proximity Spillovers — Within-Taxi Adoption

	50 m	75 m	100 m	150 m	300 m	500 m
Panel A: All merchants near garages						
Any treated	-0.0104	-0.0068	-0.0074	-0.0049	-0.0036^{*}	-0.0016
p_{RI}	[0.107]	[0.368]	[0.194]	[0.179]	[0.085]	[0.277]
Effect size	-136%	-104%	-141%	-120%	-111%	-58%
N	48,691	64,217	119,250	219,530	588,403	1,247,597
Panel B: Community-tied merchants (corrected overlap $\geq 2\%$, excluding mutual contacts)						
Any treated	-0.1025^{**}	-0.0829^{*}	-0.0826^{**}	-0.0524^{**}	-0.0269^{**}	-0.0147^{**}
p_{RI}	[0.034]	[0.052]	[0.031]	[0.021]	[0.017]	[0.023]
Effect size	-612%	-566%	-691%	-534%	-390%	-261%
N	6,027	7,987	15,313	25,254	56,795	122,865
Panel C: Mutual phone contacts only						
Any treated	$+0.0283$	$+0.0141$	-0.0655	-0.0171	-0.0639	-0.0484
p_{RI}	[0.678]	[0.811]	[0.489]	[0.775]	[0.219]	[0.226]
Effect size	$+17\%$	$+9\%$	-44%	-13%	-53%	-52%
N	456	483	613	809	1,101	1,704

Notes: Each column estimates $y_j = \beta_1 \cdot \text{AnyTreated}_j + \beta_2 \cdot \text{ExpAnyTreated}_j + \alpha$ for non-experimental merchants within the specified radius of experimental taxi garages, where ExpAnyTreated_j is the mean of AnyTreated across 1,000 placebo re-randomizations (absorbing each garage's exposure propensity). Outcome: within-industry taxi adoption at the same garage location; adoption at *other* garages is a placebo and is null throughout (Appendix Table B9). Panel A includes all merchants near the garage. Panel B restricts to merchants inside the garage drivers' social environment: a bias-corrected contact-book overlap share $(I - \mu)/\min$ of at least 2% with at least one experimental driver at that garage—the survey-validated community-tie floor (Section 1.3)—excluding mutual contacts, so the panel isolates community exposure from friendship. Panel C restricts to mutual phone contacts (both parties saved each other's number) of the garage's drivers. The competition effect concentrates 20–30$\times$ on community-exposed merchants and is absent among close contacts, indicating transmission through market-level social exposure rather than personal relationships. Effect size $= \beta_1/\bar{y}$. RI p-values (1,000 permutations) in brackets, clustered assignment at the garage level. $^{*}p_{\text{RI}} < 0.10$, $^{**} < 0.05$, $^{***} < 0.01$.

Table 5: Spillovers via Shared Regular Customers (≥ 2 transactions with each party)

	Within-taxi adoption (Dakar)			Cross-industry adoption		
	(1) All	(2) Comm.-tied	(3) No tie	(4) All	(5) Comm.-tied	(6) No tie
Share treated	$\mathbf{255.99}^{***}$	$\mathbf{464.81}^{***}$	$\mathbf{174.54}^{***}$	20.19	40.32	18.60
	(19.29)	(38.72)	(21.48)	(7.74)	(32.18)	(7.97)
	[0.001]	[0.001]	[0.001]	[0.180]	[0.425]	[0.205]
Expected exposure	2388.63	3001.82	2276.40	433.08	921.26	404.99
	(633.69)	(1447.59)	(689.61)	(244.68)	(1019.98)	(252.10)
	[0.958]	[0.951]	[0.934]	[0.864]	[0.875]	[0.845]
Effect size	+35.7%	+67.5%	+23.6%	+3.3%	+6.8%	+3.0%
Dep. var. mean	717.00	688.99	738.08	620.55	594.77	623.34
N	8,523	3,659	4,864	28,512	2,791	25,721

Notes: Coefficients, SEs, and dep. var. means multiplied by 1,000. The sample is non-experimental merchants sharing at least one *regular* customer with a non-garage experimental driver: a customer with ≥ 2 pay-to-business transactions with that specific driver *and* ≥ 2 with that specific merchant, November 2022–November 2023 (both-sides threshold; ≥ 3 in Appendix Table B13). Columns (1)–(3): taxi merchants (TAXI, TAXI_PILOT), dep. var. = within-taxi adoption (Dakar); columns (4)–(6): non-taxi merchants, dep. var. = cross-industry adoption. "Comm.-tied" = merchants with a community tie—bias-corrected contact-book overlap $\max(I-\mu,0)/\min \geq 2\%$ (Section 1.3)—to at least one of the drivers with whom they share qualifying customers; because customer-sharing and book overlap are related by construction, this split measures the intensity of shared social environment rather than an independently assigned relationship. Within-taxi demand spillovers are positive and significant in every column. The larger community-tied coefficient partly reflects exposure composition—community-tied merchants share customers with three times as many drivers (median 20 vs. 7)—and partly genuine amplification: holding the number of customer-linked drivers fixed, community-tied merchants respond $2.5\times$ more among lightly-exposed merchants (1–5 linked drivers: $+255$ vs. $+104$, both significant), while at higher customer overlap the effect is large regardless of social ties ($+337$ to $+680$ across both groups). Cross-industry spillovers are null throughout, and the network results of Table 1 are unchanged when customer-sharing pairs are excluded (Appendix Table B7). Specification: $y_j = \beta_1 \cdot \text{ShareTreated}_j + \beta_2 \cdot \text{ExpectedExposure}_j + \alpha$ over merchant j's customer-linked experimental drivers. Effect size $= \beta_1/\bar{y}$. HC1 robust SEs in parentheses; RI p-values (1,000 draws) in brackets. $^{*}p_{\text{RI}} < 0.10$, $^{**}p_{\text{RI}} < 0.05$, $^{***}p_{\text{RI}} < 0.01$.

Table 6: Diffusion Across Space: Widespread, and Not Only Through Migration

	Share treated	p_{RI}	Effect size	N
Overall effect	$+0.068^{***}$	[0.009]	$+142\%$	1,145,816
Panel A: By the merchant's own mobility				
Merchant **visited** Dakar	$+0.039$	[0.118]	$+115\%$	281,207
Merchant **never** visited Dakar	$+0.182^{**}$	[0.027]	$+87\%$	275,683
Panel B: By the driver's mobility to the destination (pooled across cities)				
Treated driver **visited** the city (migrant)	$+0.021$	[0.133]	$+240\%$	9,167,192
Treated driver **did not** visit the city	$+0.007^{**}$	[0.018]	$+75\%$	9,167,192

Notes: Coefficients are multiplied by 100. The sample is the 1,145,816 individuals in the community-tie network (bias-corrected contact-book overlap $\geq 2\%$, the survey-validated community-tie floor; Section 1.3), under the same trimming as Table 1. Each cell reports β_1 from $y_j = \beta_1 \,\text{ShareTreated}_j + \beta_2 \,\text{ExpExposure}_j + \alpha$, where ShareTreated_j is the share of individual j's experimental driver contacts assigned to treatment and ExpExposure_j is its mean across 1,000 placebo re-randomizations (included to absorb network position); β_2 is insignificant throughout. The overall effect is positive and significant: treatment-induced cross-industry adoption reaches secondary cities across thirteen regions of Senegal—Thiès and Mbour account for only 41% of in-network destination adopters, with Touba, Saint-Louis, Kaolack, Ziguinchor, and Tambacounda among the contributors. *Panel A* (outcome: cross-industry adoption outside Dakar) splits individuals by their *own* mobility—whether they were observed in Dakar at least once during the experiment (mobile-money geolocation): both groups respond, but the never-visited coefficient is five times larger and only it is statistically significant—the effect does not require the merchant to travel to the capital. *Panel B* (outcome: cross-industry adoption in the destination city; the community-tie sample stacked across the eight secondary cities with city fixed effects and SEs clustered by individual; the stacked build yields 1,145,899 individuals per city, differing from the overall row's 1,145,816 by 83 individuals due to independent pipeline construction) splits the treatment by the *drivers'* mobility—whether j's treated driver contacts visited the destination city at least once during the experiment (agent geolocation); see Section 5.4 for interpretation. RI p-values (1,000 permutations) in brackets; effect size $= \beta_1 / \bar{y}_{\text{control}}$. $^{*}p_{\text{RI}} < 0.10$, $^{**} < 0.05$, $^{***} < 0.01$.

Table 7: Structural Parameter Estimates (7-Parameter Model)

	Estimate	SE
Panel A: Structural parameters		
$\hat{\alpha}_{\text{cross}}$ (cross-industry intercept)	-6.261^{***}	(0.177)
$\hat{\alpha}_{\text{taxi}}$ (within-taxi intercept)	-8.821	(8.532)
$\hat{\beta}^{s}_{\text{learn}}$ (strong-tie learning, $\geq 5\%$)	$+0.271^{*}$	(0.153)
$\hat{\beta}^{w}_{\text{learn}}$ (community-tie learning, 2–5%)	$+0.267$	(0.205)
$\hat{\beta}^{\text{geo}}_{\text{comp}}$ (geographic competition)	-10.004^{***}	(1.17)
$\hat{\beta}_{\text{demand}}$ (customer demand)	$+0.795^{***}$	(0.058)
$\hat{\delta}$ (distance decay, km^{-1})	$+3.73^{***}$	
Competition half-life	186 m	

#	Source	Channel	$\hat{b}$ (data)	Model pred.	Sign check
Panel B: Identifying moments					
m1	Table 1, cross-Dkr $\geq 5\%$	β^{s}_{l}	$+0.00416$	$+0.00353$	✓
m2	Fig. A10, cross-Dkr 2–5% band	β^{w}_{l}	$+0.00407$	$+0.00345$	✓
m4	Table 5, within-taxi cust.	β_{d}	$+0.25599$	(matched)	✓
m5	Cross-ind. rate	α_{cross}	1.343%	1.343%	✓
m6	Within-taxi rate	α_{taxi}	0.058%	0.058%	✓
m7	Table 4 Panel B, 100 m level; Panel A gradient	$\beta^{\text{geo}}_{\text{c}}$; δ	-0.08258	-0.08258	✓
Overidentifying restriction (not targeted)					
m3	Table 2, taxi $\geq 2\%$ (pooled)	$w_s\beta^{s}_{\text{l}}+w_w\beta^{w}_{\text{l}}$	$+0.00008$	$+0.00019$	✓
Panel C: Out-of-sample validation (untargeted survey moments, predicted sign)					
m8	Table 3, all friends (within-taxi)	learning	$+0.03044$	$+$	✓
m9	Table 3, taxi friends	learning + demand − comp.	$+0.09845$	$+$	✓
m10	Table 3, taxi non-friends	learning + demand − comp.	$+0.00875$	small	✓
m11	Table 3, serendipitous	community-tie learning	$+0.01800$	$+$	✓
		Sign accuracy (untargeted: m3 and Panel C, 5/5)			11/11

Notes: Minimum distance estimation with indirect inference on the community-tie network (corrected overlap $\geq 2\%$; strong $\geq 5\%$, weaker community band 2–5%). Learning parameters (β^{s}_{learn}, β^{w}_{learn}) are identified from the cross-industry phone-network moments (m1 from Table 1; m2 from the 2–5% exclusive band) via indirect inference (OLS on model-predicted adoption); they are statistically indistinguishable ($+0.271$ vs. $+0.267$)—learning operates at a uniform rate across the community-tie strength distribution. Customer demand (β_{demand}) is identified from the within-taxi customer regression (Table 5, column 1) via average-marginal-effect inversion, so m4 is matched by construction. Distance decay (δ) is identified from the log-linear gradient of the six all-merchant garage coefficients (Table 4, Panel A; $R^2 = 0.96$), and the competition level ($\beta^{\text{geo}}_{\text{comp}}$) from the community-tied 100 m coefficient (Panel B) inverted at the garage-sample adoption rate—the garage population's baseline rate is $\sim$20$\times$ the network sample's, so using the network-sample rate would misstate the level. Moment m3 is not targeted and serves as an overidentification check: the model predicts the near-null pooled within-taxi coefficient ($+0.00019$ vs. $+0.00008$ in the data) as learning and competition offsetting. Standard errors by delta method with numerical Jacobian; the large SE on $\hat{\alpha}_{\text{taxi}}$ reflects the very low within-taxi base rate. The sign-check column marks all 11 moments; targeted moments (m1–m7) match by construction or estimation, so the informative checks are the untargeted ones (m3 and Panel C, 5 of 5). Panel C reports untargeted survey moments (Table 3): with uniform community-tie learning, the model predicts positive responses through friends and—unlike a model in which weak ties transmit nothing—through serendipitous contacts as well (m11), which the data confirm; m9 and m10 load on all three channels, with demand dominating competition. $^{*}p<0.10$, $^{**}p<0.05$, $^{***}p<0.01$.

Table 8: Counterfactual Targeting: Combined Reachable Network ($N = 3{,}103{,}114$)

	$\bar{S}^{\text{phone}}$	$\bar{S}^{\text{demand}}$	Cross-ind., Dakar		Cross-ind., outside		Within-taxi, Dakar		Total	
			Induced	Δ	Induced	Δ	Adopters	Δ	Induced	Δ
No treatment (baseline)	0.000	0.000	*Baseline adopters: 35,196 cross-Dakar + 4,906 cross-other + 9 within-taxi = 40,111*							
Actual random ($N_T = 1{,}256$)	0.586	0.658	6,024	—	369	—	9	—	6,393	—
All-garage	0.659	0.700	6,847	+13.7%	414	+12.2%	8	−10.9%	7,261	+13.6%
Garage-concentrated	0.578	0.650	5,938	−1.4%	377	+2.2%	9	−8.6%	6,315	−1.2%
Top MM users	0.612	0.616	6,317	+4.9%	338	−8.4%	10	+7.0%	6,656	+4.1%
Highest degree	0.893	0.943	9,578	+59.0%	521	+40.9%	12	+32.5%	10,102	+58.0%
Geographic reach	0.865	0.930	9,248	+53.5%	548	+48.4%	13	+35.3%	9,799	+53.3%

Notes: Model-predicted treatment-induced adoption (6-month cumulative − no-treatment baseline) on the combined reachable network built from the community-tie graph (corrected overlap $\geq 2\%$; strong $\geq 5\%$): the union of individuals with community ties to experimental drivers (trimmed), taxi merchants, and businesses within 500 m of a garage. Structural parameters from Table 7. Competition uses the nearest-garage formulation $S_j^{\text{geo}} = \text{ShareTreated}_{\text{nearest}} \times \exp(-\hat{\delta} d)$ with $\hat{\delta} = 3.73^{***}$ km^{-1} (half-life 186 m). Each counterfactual re-assigns the same treatment budget ($N_T = 1{,}256$). "All-garage" scales treated drivers' network and customer intensities to the garage-driver mean and sets $\text{ShareTreated}_g = 1$ at every garage; "Garage-concentrated" concentrates treatment at the largest garages. Δ = % change vs. actual random, from unrounded predictions. Because the within-taxi Dakar margin is tiny (baseline rate 0.058%), treatment-induced counts are sub-integer; the column therefore reports the model's *predicted number of within-taxi adopters among the 10,442 reachable taxi merchants* under each rule (no-treatment baseline: 9) with Δ relative to the random-assignment level, from unrounded predictions—a statistical zero at national scale either way. No within-taxi outside-Dakar column is reported: that margin travels on reciprocal-save and transaction ties (Appendix Tables B6 and B12), channels the structural model does not include, and its community-tie reduced form is null (Table 2). Network-based targeting dominates: treating the highest-degree drivers raises induced adoption by +58% over random assignment, while garage-based strategies range from −1.2% to +13.6%—the visibility of garages does not compensate for their weaker network reach, and concentrating treatment at the largest garages performs (weakly) worst.

Online Appendix

Nationwide Diffusion of Technology: Experimental Evidence from Multiple Networks

Deivy Houeix

A Appendix Figures **60**

B Appendix Tables **68**

C Fingerprint Data Processing **82**

D Structural Estimation: Details, Validation, and Distance Decay **84**
D.1 Estimation Details 84
D.2 Validation Details 86
D.3 Distance Decay of Competition 87

E Counterfactual Details **88**

A Appendix Figures

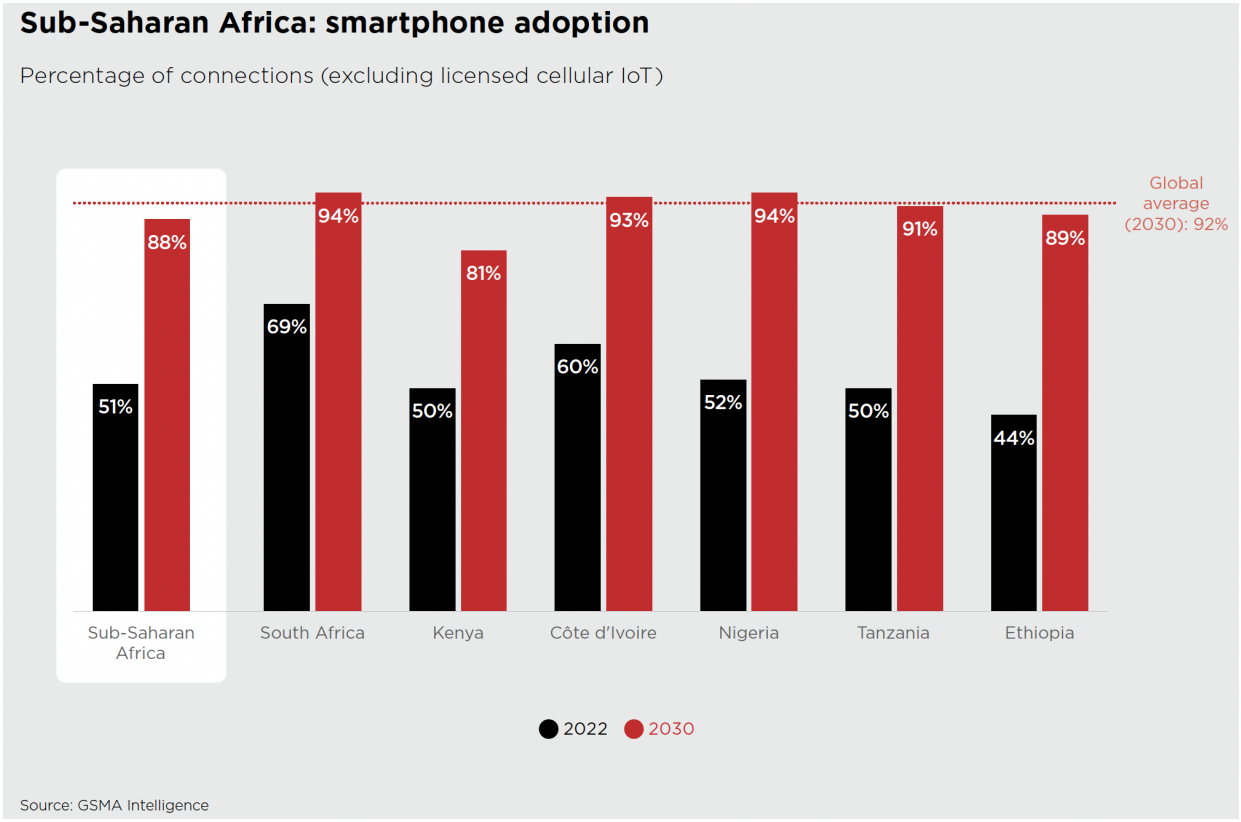


Figure A1: Smartphone Adoption in Sub-Saharan Africa

Notes: This graph is sourced from the GSMA Intelligence 2023 report, "The Mobile Economy Sub-Saharan Africa 2023," on page 15 of 46. The report gathers data and forecasts smartphone adoption across several African countries. Côte d'Ivoire, which is similar to Senegal in size and geography, reports a 60% smartphone adoption rate for the entire country, a statistic that is likely much higher in urban areas.

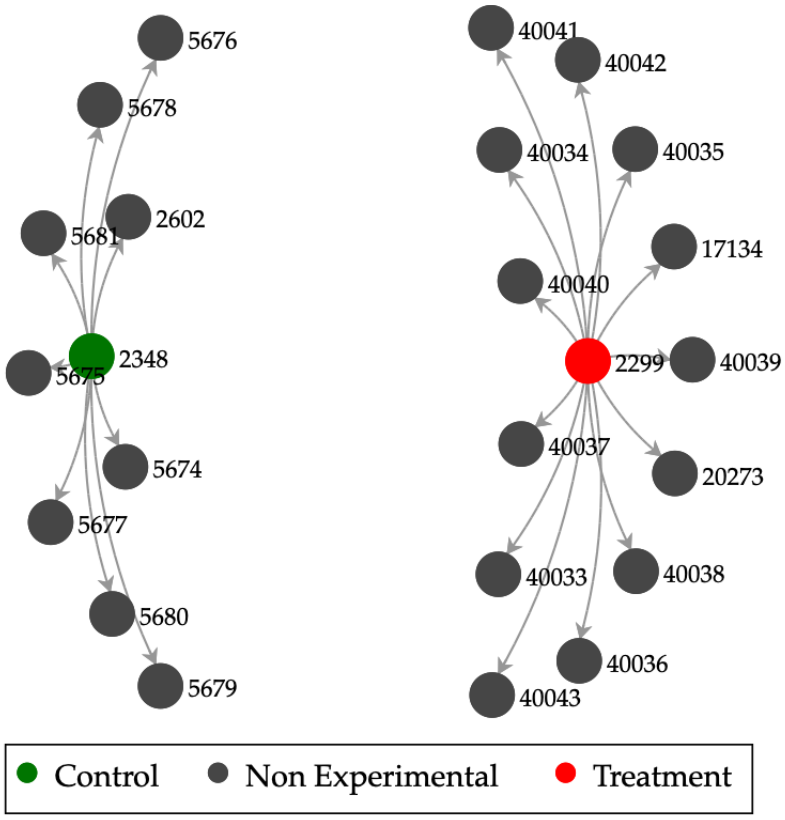


Figure A2: Illustration of the Experimental Design (Schematic)

Notes: This figure illustrates and simplifies the experimental strategy. Colored nodes (red/green) are experimental taxi businesses randomly assigned to treatment or control. Black nodes are non-experimental individuals connected to the experimental sample through the phone network. The study quantifies the likelihood of technology adoption among these non-experimental contacts, comparing adoption rates across the networks of treated versus control taxi businesses. In practice, networks often overlap, leading to network interference, an issue discussed in Section 3.1. Figure 3 shows the design in the actual community network.

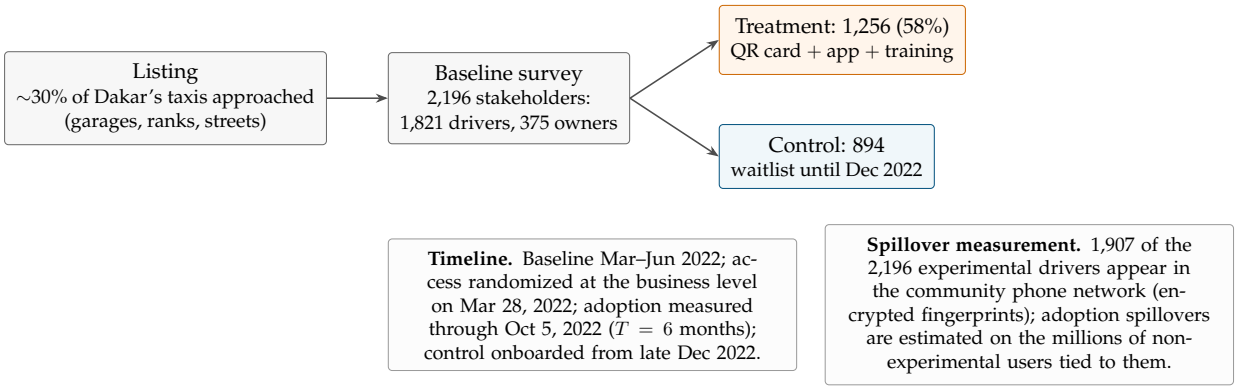


Figure A3: Experimental Design

Notes: This figure outlines the experimental design. Each taxi driver was initially listed at various locations including garages, car wash stations, meeting points, and on the streets of Dakar (e.g., during traffic jams). The listing was comprehensive enough to cover most parts of the city, and a large portion of the taxi drivers working in the city were approached (it is estimated that about 30% of the drivers were approached, with some refusing to be surveyed, others not interested, and some not eligible). Subsequently, taxi owners and drivers were invited to specific locations in Dakar to be surveyed and randomized into the treatment. The primary experimental variation involved randomizing access to the technology. The treatment was then subdivided into three different "observability" treatment arms for owners to monitor their drivers' transactions in various ways. For this paper, these sub-treatments are pooled together as most of the sample considered here consisted of taxi businesses operated solo.

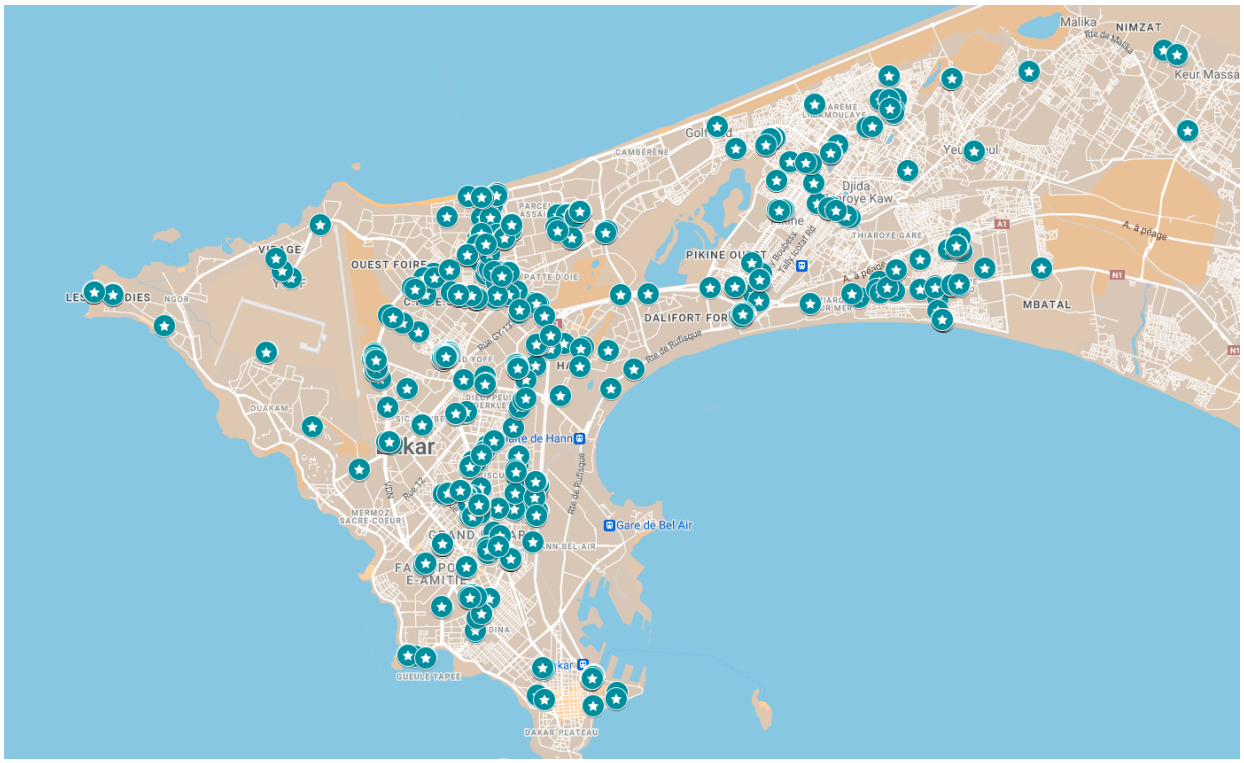


Figure A4: Listing Survey Activity in Dakar - March/April 2022

Notes: The figure displays a map of Dakar, capital city of Senegal. Each blue dot is the GPS location of the listing survey conducted before the experiment. Drivers were recruited in garages, car wash stations, meeting points, and on the streets of Dakar (e.g., during traffic jams). The extent of the listing was so broad that it covers most parts of the city and a large portion of the taxi drivers working in the city were approached (among which some refused to be surveyed, some were not interested, and some were not eligible). Owners not driving were primarily recruited by asking drivers about their owners' contact information during the listing survey.

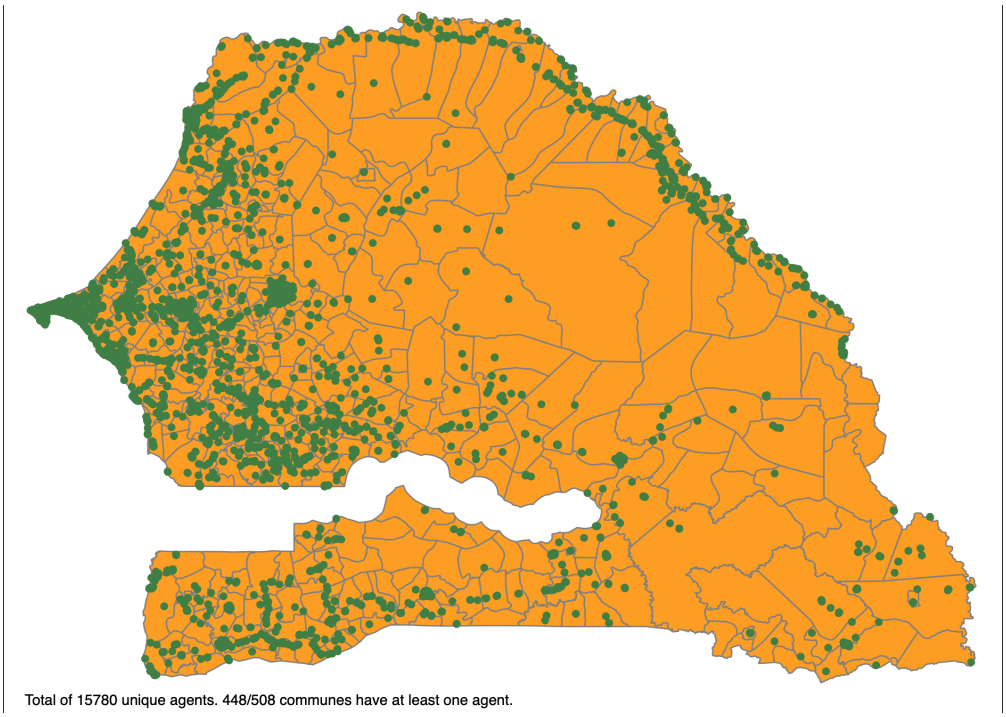


Figure A5: Mobile Money Agents in Senegal in 2022

Notes: This Figure represents a map of Senegal. Each green dot is the GPS location of a mobile money agent for the main mobile money company. There are over 15,000 unique agents in Senegal from this mobile money company, which covers the full set of urban cities and the vast majority of communes/villages.

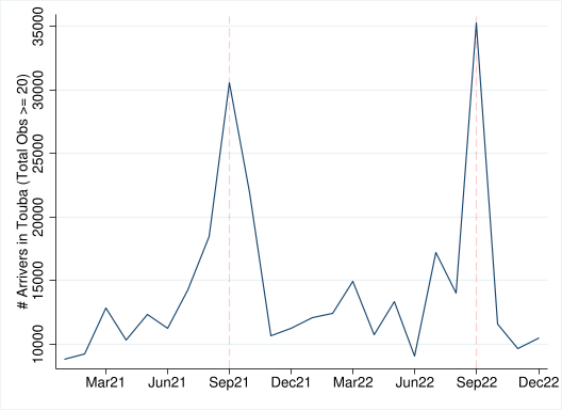


Figure A6: Location Data - Arrivals in Touba

Notes: This Figure represents the number of arrivals in Touba over each month from February 2021 to December 2022, as measured via the withdrawal/deposit of mobile money to agents located in the city of Touba. There are peaks of arrivals in the months of September, which is directly consistent with the pilgrimage dates happening every year in this city.

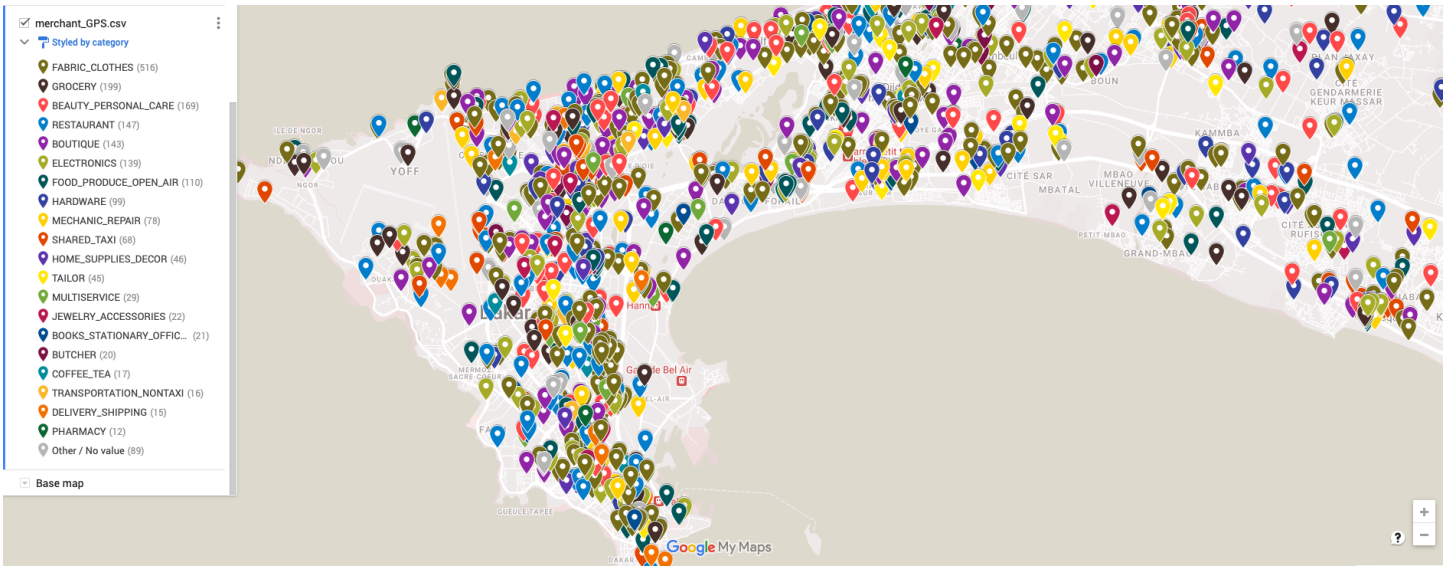


Figure A7: Technology Diffusion in Dakar Throughout the Experiment

Notes: This map shows the technology diffusion, for a randomly selected sample of businesses, across the capital city of Dakar with the firm's sectors. Each colored dot shows the location of a business that adopted the digital payment technology during the time of the experiment.

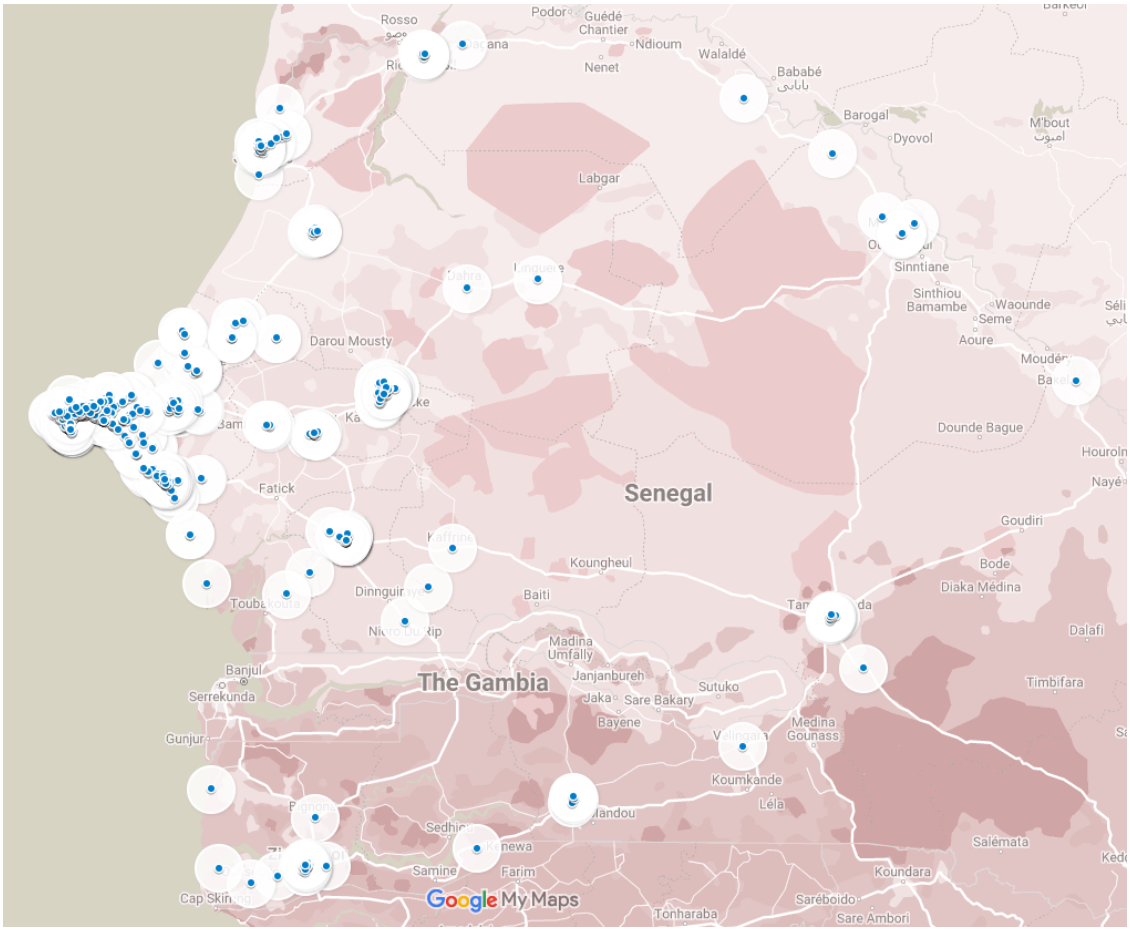


Figure A8: Technology Diffusion Inter-Industry Across Space Throughout the Experiment

Notes: This map shows the technology diffusion, for a randomly selected sample of businesses, across many different urban cities in Senegal, beyond the capital city of Dakar, where the experiment took place. Each blue dot shows the location of a business that adopted the digital payment technology during the time of the experiment.

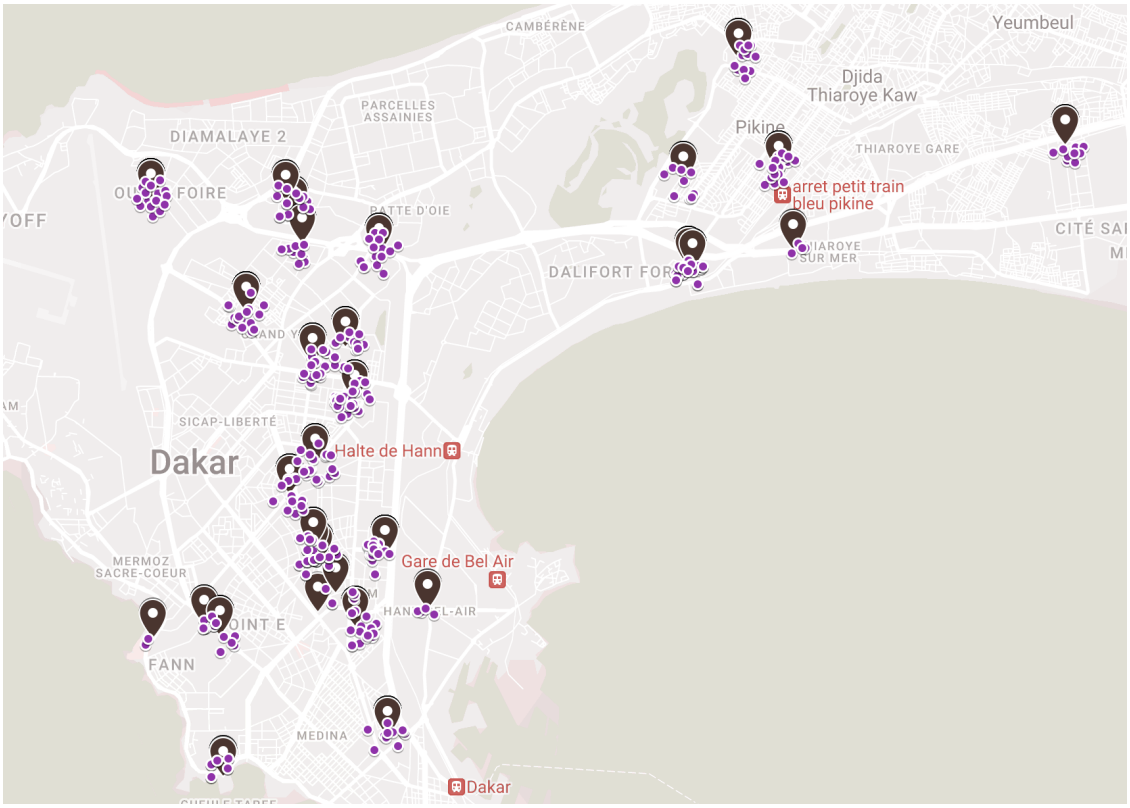


Figure A9: Local Technology Adoption Spillovers Nearby Taxi Garages

Notes: This figure displays a map of Dakar, pinpointing the geo-locations of a subset of taxi garages where at least one taxi driver participated in the experiment. Each garage was manually coded to define a neighborhood. Mobile money agents within a 200-meter radius were identified and tagged, in this illustration in purple. For these tagged agents, I compiled a dataset of mobile money users living in proximity to these agents to delineate the residents of each neighborhood. For these residents, I then analyzed their propensity to adopt the technology depending on the number of treated taxi businesses and the total number of experimental taxi businesses within their neighborhood.

B Appendix Tables

Table B1: Cross-Validating Social Networks via Encrypted Phone Contacts

Relationship category	n	Mean corrected overlap (%)	Share $\geq 2\%$ (community)	Share $\geq 5\%$ (strong)	Share $\geq 10\%$
Wife	495	20.81	86.5%	76.8%	64.4%
Taxi contact	1,882	13.64	81.4%	62.8%	38.6%
Family	3,814	11.44	80.8%	64.1%	41.8%
Friend (family friend)	6,712	9.09	72.5%	53.3%	31.6%
Professional	6,077	8.94	65.7%	46.0%	26.4%
Friend (non-taxi)	2,922	6.00	61.9%	39.1%	18.2%
Serendipitous	7,739	3.20	27.7%	13.7%	6.3%
Client	526	1.94	22.8%	8.7%	3.0%
By interaction frequency					
Weekly or more	6,028	10.56	73.3%	54.4%	33.7%
Less than monthly	2,770	5.54	52.3%	31.4%	14.6%

Notes: This table cross-validates the bias-corrected contact-book overlap measure against survey-reported relationship types. The endline survey asked each experimental driver about up to 16 randomly selected mobile money contacts, eliciting relationship type and interaction frequency. For each surveyed (driver, contact) pair matched to fingerprint data under the corrected E.164 hashing scheme (Appendix C; 22,537 of 23,520 pairs classified, of which 16,470 have fingerprints on both sides), we compute the bias-corrected overlap share $\max(I - \mu, 0)/\min(|F_i|, |F_j|)$, where μ is the expected hash-collision count, and report the category mean and the share of pairs above the 2% community-tie floor, the 5% strong-tie threshold, and 10%. Corrected overlap rises steeply with relationship closeness: wives average 20.8% (86.5% above the community floor; 76.8% strong), taxi colleagues and family 11–14% (about 81% above the floor), friends 6–9%, while serendipitous contacts (numbers the respondent could not place when prompted—typically incidental transaction partners) average 3.2% with only 27.7% above the floor, and taxi clients 1.9% (22.8%)—the calibrated 2% floor cleanly separates durable relationships from incidental contact. Contacts seen weekly or more share 1.9$\times$ the overlap of contacts seen less than monthly.

Table B2: Closeness by Phone Link Type

	Mutual (both saved)	One-sided (either saved)	Comm. tie, not saved ($\geq 2\%$)	No comm. tie ($< 2\%$)	No fingerprint
Share close (wife/family/friend)	61.1%	36.8%	37.1%	18.2%	34.9%
N	6,517	4,173	768	5,012	6,067

Notes: Survey-validated (driver, contact) pairs from the endline kinship module, matched to fingerprint data under the corrected E.164 hashing scheme (Appendix C); 22,537 classified pairs. Contacts were sampled from the pre-experiment mobile money transaction network. "Close" $= 1$ if the driver reports the contact as wife, family, or friend. Categories: *mutual* = both parties' phone hashes appear in each other's contact-book fingerprint; *one-sided* = exactly one does; *community tie, not saved* = neither saved the other, but bias-corrected contact-book overlap is at least 2%—the community-tie floor used throughout the analysis; *no community tie* = both have fingerprints, no save in either direction, and corrected overlap below 2%; *no fingerprint* = the contact lacks fingerprint data on all their numbers. Two facts stand out. First, the community floor is the sharp divide: unsaved pairs above it are reported close 37.1% of the time, versus 18.2% below it—half the rate even of pairs with no fingerprint data at all (34.9%, whose composition mirrors the population). Second, a *one-sided* save adds nothing beyond community membership (36.8% vs. 37.1%): only *reciprocal* saving marks deeper relationships (61.1%). The corrected-overlap floor, not the save indicator, is what separates genuine relationships from incidental transaction partners; reciprocity then marks depth within the community network, which is why the mutual layer is analyzed separately (Appendix Table B6).

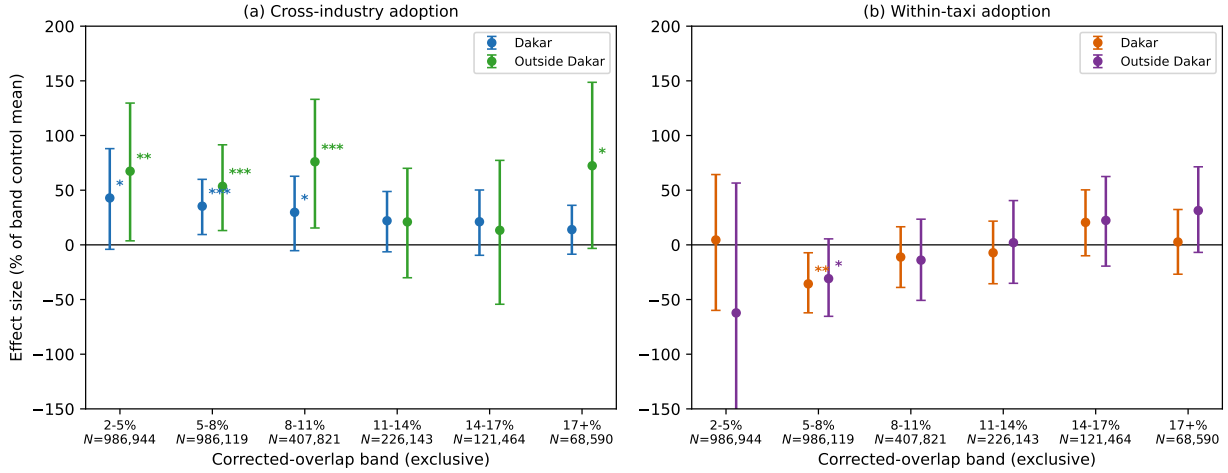


Figure A10: Per-band transmission by tie strength

Notes: Effect sizes ($\hat{\beta}_1/\bar{y}$, in percent of the band control mean) of ShareTreated on adoption, estimated separately within exclusive bands of bias-corrected contact-book overlap (specification, sample construction, and strict trimming as in Table 1; band Ns from 986,944 at 2–5% to 68,590 at 17%+). Error bars are 95% randomization-inference confidence intervals obtained by test inversion: the set of constant effects τ_0 whose sharp null is not rejected at 5% across the 1,000 placebo re-randomizations. Stars denote RI p-values at the zero null ($^* < 0.10$, $^{**} < 0.05$, $^{***} < 0.01$). Panel (a): cross-industry transmission is approximately uniform across the strength distribution—certified in the 5–8% band in Dakar and in the 2–5, 5–8, and 8–11% bands outside Dakar—showing that diffusion operates throughout the community network rather than only at the closest ties. Panel (b): the within-taxi competitive effect concentrates sharply in the 5–8% band (Dakar -36%, CI $[-62, -7]$)—colleagues close enough to compete—and is absent at both the weakest and strongest ties.

Table B3: Overlap Predicts Closeness Across Distance, Controlling for Reciprocity

	(1) Close	(2) Close	(3) Close	(4) Close
Corrected overlap share		$+0.954^{***}$	$+0.980^{***}$	$+1.291^{***}$
		(0.122)	(0.169)	(0.196)
Distance (km)		$+0.0004^{**}$	$+0.0005$	
		(0.0002)	(0.0005)	
Overlap × Distance			-0.0007	
			(0.0017)	
Far (modal city ≠ Dakar)				$+0.147^{**}$
				(0.072)
Overlap × Far				-0.525^{**}
				(0.222)
Mutual (both saved)	-0.089^{***}		-0.017	$+0.005$
	(0.033)		(0.054)	(0.050)
Mutual × Distance / Far			-0.000	-0.018
			(0.0005)	(0.068)
R^2	0.002	0.063	0.063	0.072
N	2,707	1,586	1,586	2,707

Notes: Column (1) shows that reciprocity *alone* does not proxy closeness: unconditionally, mutual pairs are reported close *less* often in this location-matched sample (-0.089, $R^2 = 0.002$)—its non-mutual pairs are heavily community-tied—and adding mutual to the overlap specification leaves R^2 unchanged (columns 2–3), because mutual carries no information beyond corrected overlap. OLS regressions of survey-reported closeness (contact is wife, family, or friend) on the bias-corrected contact-book overlap share, geographic distance, reciprocal-save (mutual) status, and their interactions, for surveyed (driver, contact) pairs matched to fingerprint and geolocation data. Distance is the haversine distance between the pair's transaction-weighted mean agent coordinates (columns 1–2, continuous); column 3 uses an indicator for the contact's modal city being outside Dakar, which retains pairs without precise coordinates. HC1 standard errors in parentheses. Corrected overlap strongly predicts closeness at any distance. The continuous-distance interaction is null (column 2, $p = 0.685$): per kilometer, overlap's social-proximity content does not measurably change. The coarse far indicator (column 3) shows some attenuation—the overlap slope for far contacts is $1.291 - 0.525 = 0.766$, about 60% of the near slope, still large and highly significant—so overlap is informative about social proximity everywhere, somewhat less sharply so at distance. Conditional on overlap, mutual status and its interactions are precise zeros: reciprocal saving adds no closeness content beyond the overlap measure at any distance, consistent with the community-tie network (not the save indicators) being the primary measure of social proximity. $^{*}p < 0.10$, $^{**} < 0.05$, $^{***} < 0.01$.

Table B4: Phone Contact Network Balance: Treatment vs. Control Drivers

	Treated	Control	Difference	SE	p-value
Panel A: Contact book size					
Contact book size (fingerprint)	585	614	-28.8	24.5	0.240
Mutual contacts detected	8,309	8,982	-673	352	0.056*
Panel B: Contact composition					
Share merchant (among mutual)	0.041	0.041	0.0000	0.0002	0.868
Share taxi (among mutual)	0.0034	0.0034	-0.0001	0.0000	0.244
Panel C: Overlap distribution					
Share overlap $\geq$5% (among mutual)	0.0031	0.0024	0.0006	0.0008	0.438
Mean overlap (among mutual)	0.0039	0.0038	0.0000	0.0001	0.727
Median overlap (among mutual)	0.0007	0.0007	-0.0000	0.0000	0.550
Panel D: Community-tie network (corrected overlap $\geq 2\%$)					
Community ties per driver ($\times$1,000)	295.3	305.0	-9.8	14.0	0.486
Strong ties per driver ($\times$1,000)	46.6	46.9	-0.4	4.7	0.935
Mean corrected overlap (driver-level)	0.0337	0.0335	0.0002	0.0005	0.715
Share merchant (community ties)	0.0250	0.0241	0.0010	0.0004	0.028**
Share taxi (community ties)	0.0026	0.0026	-0.0000	0.0001	0.993
Joint F-test (community panel)	$F(5, 1901) = 1.126$			$p = 0.344$	
Observations	1,907				
Kolmogorov–Smirnov tests					
Contact book size (fingerprint)	KS $= 0.0632$			$p = 0.0473$	
Mutual contacts detected	KS $= 0.0587$			$p = 0.0791$	
Mean overlap (among mutual)	KS $= 0.0443$			$p = 0.3135$	
Mutual contacts $\geq$5% overlap	KS $= 0.0256$			$p = 0.9125$	
Share taxi (among mutual)	KS $= 0.0544$			$p = 0.1248$	
Joint F-test	$F = 0.763$			$p = 0.5765$	

Notes: Each row reports treated and control means, the difference, robust SE, and p-value from a regression on treatment status. Fingerprints observed February 2024. Panels A–C use raw (uncorrected) quantities: contact-book size counts 16-bit fingerprint buckets, "Mutual contacts detected" counts reciprocal-save edges, and both are inflated by hash collisions, which is why the raw share of mutual pairs above 5% overlap is ∼0.3% (Section 1.3; same comparison on the full matched sample in Table B5, Panel C). Panel D uses the corrected, trimmed analysis network (∼12,400 community ties per driver, ∼1,500 strong). Merchant classification uses only pre-fingerprint registrations, so treatment-induced registrations cannot contaminate it. The phone network is observed post-experiment (February 2024) and is used as a proxy for pre-existing social proximity; see Section 1.3 for discussion. Panel D reports balance on the community-tie network used in the analysis (bias-corrected overlap $\geq 2\%$, pre-trimming counts; strong $=\geq 5\%$; merchant/taxi classification restricted to registrations before the fingerprint date): treated and control drivers are jointly indistinguishable (F-test $p = 0.344$); the only individually significant difference in this panel (share of community ties to merchants, 0.0250 vs. 0.0241, $p = 0.028$) is economically negligible. A Kolmogorov–Smirnov test on the pooled tie-strength distribution rejects exact equality at the massive tie-level sample size ($n >$ 11M sampled ties) but the difference is trivial and *downward* for treated drivers (pooled tie-level mean corrected overlap 0.0389 vs. 0.0397; the driver-level means above weight drivers equally and differ accordingly)—working against treatment-induced tie strengthening. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table B5: Test for Treatment-Induced Tie Strengthening

	Treated	Control	Diff	t-stat	p-value
Panel A: Overlap distribution (edge-level, 12.4M mutual edges)					
Mean overlap	0.0242	0.0244	-0.0002		
Median overlap	0.0193	0.0194	-0.0001		
p75 overlap	0.0307	0.0305	$+0.0002$		
p99 overlap	0.0959	0.0997	-0.0038		
KS statistic	0.008		$p < 0.001$		
Panel B: Composition of strong ties ($\geq 5\%$ overlap, driver-level)					
Number of strong ties	668	749	-81	-0.51	0.608
Taxi contacts (strong)	0.79	0.85	-0.06	-0.22	0.825
Merchant contacts (strong)	32.1	35.3	-3.2	-0.57	0.571
Share taxi (strong)	0.070%	0.060%	$+0.010\%$	$+1.29$	0.196
Share merchant (strong)	5.65%	5.61%	$+0.05\%$	$+0.64$	0.521
Mean overlap (strong)	0.0738	0.0738	-0.0001	-0.28	0.781
Panel C: Total ties (driver-level)					
Mutual contacts	7,887	8,557	-671	-1.89	0.059
Strong ties ($\geq 5\%$)	668	749	-81	-0.51	0.608
Weak ties ($< 5\%$)	7,219	7,808	-589	-1.98	0.048
Panel D: Overlap by contact type (KS tests)					
Taxi contacts	KS $= 0.015$		$p = 0.14$		
Merchant contacts	KS $= 0.007$		$p < 0.001$		
Non-merchant contacts	KS $= 0.008$		$p < 0.001$		

Notes: Overlap in Panels A–C is the *raw* (uncorrected) share on the legacy mutual-edge build; the analysis measure throughout the paper is the bias-corrected share (Section 1.3), whose driver-level balance is reported in Table B4, Panel D. Tests for whether treatment caused selective tie strengthening in the phone contact network, observed February 2024. Panel A compares the full overlap distribution across 12.4 million mutual edges; the KS statistic is mechanically significant due to sample size but the quantile differences are economically negligible at the center of the distribution (< 0.04 percentage points at the median and p75; 0.38 points at p99) and go in the wrong direction for a strengthening story (treated overlap is weakly *lower*). Panel B compares the composition of strong ties ($\geq 5\%$ overlap) at the driver level; no dimension is significant at 10%. Panel C compares total tie counts; treated drivers have marginally *fewer* contacts, inconsistent with tie formation. Panel D tests overlap distributions separately by contact type; the taxi-contact test—the most relevant for treatment-induced strengthening—has $p = 0.14$. Welch t-tests with unequal variances; $N_T = 1{,}256$, $N_C = 894$ (the full matched sample; the corresponding counts in Table B4 condition on the 1,907 fingerprinted drivers, hence the small level differences in mutual-contact means—the marginal imbalance is the same fact in both tables, not two findings).
Community-network check: the same tests on the community-tie network (bias-corrected overlap $\geq 2\%$) confirm no treatment-induced strengthening: per-driver community-tie counts, strong-tie counts, mean corrected overlap, and tie composition are all balanced (joint F $p = 0.344$; Table B4, Panel D), and the pooled tie-strength distribution of treated drivers is marginally *weaker*, not stronger (mean corrected overlap 0.0389 vs. 0.0397).

Table B6: Reciprocity and the Cross-Market Channel: Mutual vs. Non-Mutual Ties

	$\geq 2\%$	$\geq 5\%$	$\geq 10\%$
Panel A: Non-mutual community ties			
Cross-industry, Dakar	$+2.97$ [.117] {+34%}	$+4.57^{**}$ [.010] {+35%}	$+5.54^{**}$ [.015] {+55%}
Cross-industry, outside	$+0.68^{**}$ [.010] {+76%}	$+0.70^{***}$ [.008] {+54%}	$+0.80^{**}$ [.019] {+82%}
Within-taxi, Dakar	$+0.11$ [.529] {+18%}	-0.81^{**} [.017] {-33%}	-0.22 [.771] {-6%}
Within-taxi, outside Dakar	-0.05 [.138] {-73%}	-0.15^{**} [.015] {-43%}	-0.13 [.438] {-18%}
N	1,146,765	1,006,046	363,672
Panel B: Mutual phone contacts			
Cross-industry, Dakar	$+0.14$ [.925] {+0.4%}	-0.09 [.970] {-0.2%}	$+0.81$ [.791] {+2.5%}
Cross-industry, outside	-0.15 [.617] {-7%}	-0.36 [.398] {-21%}	-0.27 [.592] {-18%}
Within-taxi, Dakar	$+0.05$ [.974] {+0.2%}	$+0.39$ [.885] {+0.8%}	-3.89 [.483] {-5.8%}
Within-taxi, outside Dakar	$+\mathbf{0.90}^{***}$ [.004] {+23%}	$+\mathbf{2.60}^{***}$ [.001] {+30%}	$+\mathbf{4.39}^{***}$ [.006] {+32%}
N	163,543	64,010	27,030

Notes: Coefficients $\times 1{,}000$; RI p-values (1,000 permutations) in brackets; effect sizes ($\beta_1/\bar{y}$) in braces. Each panel re-estimates the specification of Table 1 with exposure computed over the indicated relationship class only: *mutual* pairs are those in which each party saved the other's number; *non-mutual* community ties are all remaining pairs with the indicated corrected overlap (one-sided saves and shared-environment pairs). Exposure is recomputed and the top-2% trimming reapplied within each relationship class, so sample sizes differ slightly from Table 1 (e.g., 1,146,765 vs. 1,145,899 at $\geq 2\%$). The two channels separate completely. Cross-industry learning operates entirely through non-mutual community ties; mutual contacts—despite baseline adoption rates 2.5–7$\times$ higher—show no incremental response, consistent with their exposure to the technology being saturated irrespective of the experiment (inframarginality). Conversely, within-industry adoption in the Thiès market (the only market outside Dakar where the taxi product was rolled out; 98% of outside-Dakar taxi adopters) responds *only* through mutual ties, at every threshold: profession-specific adoption crosses markets through reciprocal relationships, consistent with experiential, trust-intensive content requiring high-bandwidth ties, while awareness—abundant in a launched market—does not bind. The same channel appears in the pre-treatment transaction network (Table B12). Both patterns are unchanged when driver–contact pairs sharing regular customers are excluded (Appendix Table B7). $^{*}p_{\text{RI}} < 0.10$, $^{**} < 0.05$, $^{***} < 0.01$.

Table B7: Network Spillovers Excluding Pairs with Shared Regular Customers

	$\geq 2\%$ (community)	$\geq 5\%$ (strong)	$\geq 10\%$
Cross-industry, Dakar	$+2.88$ [.129] {+33%}	$+4.15^{**}$ [.018] {+31%}	$+5.15^{**}$ [.015] {+48%}
Cross-industry, outside	$+0.68^{***}$ [.010] {+76%}	$+0.68^{***}$ [.010] {+53%}	$+0.73^{**}$ [.030] {+73%}
Within-taxi, Dakar	$+0.04$ [.829] {+6%}	-1.01^{***} [.004] {-39%}	-0.65 [.467] {-13%}
Within-taxi, outside Dakar	-0.04 [.258] {-54%}	-0.10 [.131] {-29%}	$+0.08$ [.660] {+9%}
N	1,145,978	1,020,354	376,759

Notes: Coefficients $\times 1{,}000$; RI p-values (1,000 permutations) in brackets; effect sizes ($\beta_1/\bar{y}$) in braces. This table re-estimates Panel A of Table 1 after removing every driver–contact pair that shares a regular customer—a customer with at least two transactions with the driver and at least two with the contact during the rollout (the link definition of Table 5). Specification, community-tie definition, and trimming as in Table 1. All coefficients are within 4% of the full-network estimates: the cross-industry learning channel, the distant diffusion channel, and the within-taxi competition effect are all intact, indicating that phone-network spillovers are not proxying for shared customer demand. Likewise, the mutual-tie Thiès channel of Appendix Table B6 survives the purge ($+2.28$, $p_{\text{RI}} = 0.003$ at $\geq 5\%$; $+4.12$, $p_{\text{RI}} = 0.009$ at $\geq 10\%$). The two mechanisms—social learning through community ties and customer-demand spillovers through the payment network—operate on separate margins. $^{*}p_{\text{RI}} < 0.10$, $^{**} < 0.05$, $^{***} < 0.01$.

Table B8: Lee Bounds for Differential Fingerprint Attrition

	Taxi, Dakar (within)	Cross-ind., Dakar (across, local)	Cross-ind., outside (across, distant)
Community ties ($\geq 2\%$ corrected overlap)			
Share treated	$+0.075$	$+2.907$	$\mathbf{+0.680}^{**}$
Lee lower bound	$+0.027$	$+1.994$	$+0.554^{**}$
Lee upper bound	$+0.079$	$+2.996$	$+0.686^{***}$
N		1,145,899	
Strong ties ($\geq 5\%$ corrected overlap)			
Share treated	$\mathbf{-0.984}^{***}$	$\mathbf{+4.162}^{**}$	$\mathbf{+0.681}^{***}$
Lee lower bound	-1.103^{***}	$+4.079^{**}$	$+0.680^{***}$
Lee upper bound	-0.976^{***}	$+4.198^{**}$	$+0.685^{***}$
N		1,020,391	

Notes: Tightened Lee bounds (Lee, 2009) for differential attrition from the experimental sample to the fingerprint network sample, computed on the community-tie network (bias-corrected overlap; same universe, first-row dedupe, and strict top-2% trimming as Table 1—the Share treated rows replicate Table 1 exactly). Bounds are computed within contact-book-size quartiles: within each quartile, excess treated drivers are trimmed from the top (lower bound) or bottom (upper bound) of the connected-outcome distribution, and the network and estimates are rebuilt. The bounds are tight throughout—at the strong-tie threshold every bounded estimate retains its sign and significance, including the within-taxi competition effect ($[-1.103, -0.976]$, both $p_{\text{RI}} < 0.01$)—indicating that differential fingerprint attrition does not drive the results. Coefficients $\times 1{,}000$. RI p-values from 1,000 permutations. $^{*}p_{\text{RI}} < 0.10$, $^{**}p_{\text{RI}} < 0.05$, $^{***}p_{\text{RI}} < 0.01$.

Table B9: Placebo: Garage Proximity and Taxi Adoption at *Any Other* Garage

	50 m	75 m	100 m	150 m	300 m	500 m
Panel A: All merchants near garages						
Any treated	−0.0006	−0.0005	+0.0001	+0.0001	−0.0003	+0.0001
p_{RI}	[0.656]	[0.676]	[0.976]	[0.925]	[0.633]	[0.797]
N	48,691	64,217	119,250	219,530	588,403	1,247,597
Panel B: Community-tied merchants (corrected overlap $\geq 2\%$, excluding mutual contacts)						
Any treated	−0.0021	−0.0035	−0.0027	−0.0016	−0.0006	−0.0006
p_{RI}	[0.850]	[0.448]	[0.490]	[0.487]	[0.760]	[0.804]
N	6,027	7,987	15,313	25,254	56,795	122,865
Panel C: Mutual phone contacts only						
Any treated	+0.0309	+0.0250	+0.0162	+0.0176	+0.0330	+0.0225
p_{RI}	[0.375]	[0.650]	[0.601]	[0.583]	[0.206]	[0.192]
N	456	483	613	809	1,101	1,704

Notes: The outcome equals one if the merchant adopted as a Dakar taxi whose home garage is *any* garage other than the one whose treatment exposure is being tested; exposure is measured at the nearby (regression) garage. Same design, samples, and community-tie definitions as Table 4, but the outcome is within-industry taxi adoption at a garage *other* than the one the merchant is near. If the negative effects in Table 4 reflected competition in the local market rather than a spurious correlate of garage exposure, they should not appear for adoption at distant garages—and they do not: every coefficient in Panels A and B is small and statistically null under randomization inference ($p_{\text{RI}} \geq 0.44$ in the community-tied panel at every radius). Mutual contacts (Panel C) show positive but insignificant point estimates, consistent with the cross-market mutual-tie channel of Appendix Table B6 rather than local competition. RI p-values (1,000 permutations) in brackets, clustered assignment at the garage level. $^{*}p_{\text{RI}} < 0.10$, $^{**} < 0.05$, $^{***} < 0.01$.

Table B10: Network Objects: Construction, Timing, and Use

Network	Construction	Observation period	Scale	Used in
Panel A: Phone network (social ties)				
Community ties ($\geq 2\%$)	Bias-corrected contact-book overlap $\max(I - \mu, 0)/\min(\lvert F_i\rvert, \lvert F_j\rvert) \geq 2\%$ (encrypted fingerprints); saves not required. Floor set above the collision-noise scale; survey-validated (Tables B1, B2)	Feb 2024	23.7M driver–contact edges, 1.15M individuals (regression sample)	Table 1 (main); Tables 4 (Panel B), 6; structural model; counterfactuals
Strong ties ($\geq 5\%$)	Subset: corrected overlap $\geq 5\%$	Feb 2024	2.8M edges, 1.02M individuals	Table 1 ($\geq 5\%$ column and bands); structural β_{learn}
Mutual contacts (reciprocal saves)	User i saved j's number AND j saved i's number; a small, distinct layer (under 1% of community ties) analyzed separately	Feb 2024	186K edges, 164K individuals (at $\geq 2\%$)	Appendix Table B6 (cross-market channel); Table 4 (Panel C)
Panel B: Customer network (demand channel)				
Pre-intervention customers (P2P)	Person-to-person transfers in taxi-fare range (CFA 500–3,000), $\geq$ 2 transactions with same driver	Apr 2021–Mar 2022	—	Reduced-form Table B14 (pre-intervention specification)
Post-experiment customers (P2B)	Pay-to-business transactions; a shared *regular* customer has $\geq$ 2 transactions with the driver AND $\geq$ 2 with the merchant	Nov 2022–Nov 2023	28,609 customers, 95,114 merchants, 3.4M edges	Reduced-form Table 5 (main); structural β_{demand}; counterfactual S^{cust}
Panel C: Geographic network (competition channel)				
Nearest-garage proximity	Haversine distance from individual (imputed via driver-contact GPS average) to nearest garage	Cross-section	42 garages (36 with experimental drivers), 40,710 non-taxi within 500 m	Table 4; structural $\beta^{\text{geo}}_{\text{comp}}$, δ; counterfactual S^{geo}
Panel D: Transaction network (financial ties)				
P2P transaction ties	Person-to-person mobile money transfers *sent by the driver* before the experiment	2020–2022Q1	66K driver–recipient edges, 60K individuals (trimmed)	Appendix Table B12 (cross-market money-tie channel)

Notes: This table summarizes the network objects used in the analysis. The phone network (Panel A) is the primary network for measuring social proximity: the *community-tie* graph (corrected overlap $\geq 2\%$) is the analysis network throughout, with the $\geq 5\%$ strong-tie restriction and the mutual (reciprocal-save) layer as nested objects; scales are reported for the analysis sample after strict top-2% degree trimming (Section 1.3). The network is observed post-experiment (February 2024) and interpreted as a proxy for pre-existing social ties; Table B4 and Table B5 support this interpretation. The customer network (Panel B) links merchants and drivers through shared regular customers; the post-experiment P2B version is used in the structural model because it directly observes business payments (which did not exist pre-experiment for taxis); the structural exposure S^{cust} uses the aggregate qualifying rule ($\geq$ 2 transactions across non-garage experimental drivers, Section 6), while Table 5 uses the stricter per-pair rule shown in the row. The geographic network (Panel C) captures competitive proximity to taxi garages. The transaction network (Panel D) is observed pre-experiment and carries the within-industry cross-market channel, mirroring the mutual-contact layer.

Table B11: Kinship Survey in the Transaction Network - Summary Statistics

	Share
Panel A. Nature of the Relationships	
Family	0.23
	(0.16)
Friends	0.16
	(0.15)
Any Business	0.34
	(0.20)
Taxi Drivers	0.10
	(0.11)
Taxi Drivers Same Garage	0.01
	(0.04)
Business Owners	0.19
	(0.15)
Both Personal and Professional	0.23
	(0.17)
Casual Interactions - Do Not Remember	0.42
	(0.23)
Panel B. Frequency of Conversations in Past 3 Months	
Every Day	0.37
	(0.26)
Every Week	0.31
	(0.23)
At Least Once a Month	0.21
	(0.21)
Less Than Once a Month	0.06
	(0.12)
Never	0.02
	(0.07)
Would Borrow Money From If Needed	0.51
	(0.29)
Panel C. Information Sharing	
Generally Share Important Advice/Information	0.74
	(0.28)
Discussed or Recommended the Digital Payment Technology	0.12
	(0.18)
Observations	1,693

Notes: This table displays summary statistics about the kinship survey data collected from September to December 2023 among RCT-taxi businesses. The survey involved businesses successfully followed up with and who consented to participate in this kinship survey. Sixteen links were identified for each respondent. Eight links were randomly selected from the pre-experiment transaction network, while the other eight were chosen from transactions occurring during or after the experiment. Respondents were asked to define their relationship and answer several survey questions about each of these links. $N = 1,693$ is the number of driver respondents; each observation is the driver-level share across their 16 links. Panel B statistics are computed over the 1,605 respondents with frequency answers; the advice/information question in Panel C was added after the survey began and is available for 1,371 respondents (shares among known contacts in both panels). The 27,720 survey pairs referenced in Section 1.3 come from the drivers who completed the network validation module.

Table B12: Spillovers in the Pre-Experiment Transaction Network ($\geq$ 2 P2P transfers)

	Taxi, Dakar	Cross-ind., Dakar	Cross-ind., other	Taxi, outside
Share treated	$+1.53$	$+0.33$	$+0.14$	$+\mathbf{11.19}^{***}$
p_{RI}	$[0.605]$	$[0.936]$	$[0.845]$	$[0.001]$
Effect size	$+4.8\%$	$+0.6\%$	$+3.4\%$	$+83\%$
Dep. var. mean	0.0319	0.0542	0.0041	0.0134
Adopters	739	1,254	95	311
N		23,146		

Notes: Coefficients $\times 1{,}000$. The network is constructed from pre-treatment mobile-money transfer records (2020–2022Q1): an individual j is linked to an experimental driver if the driver *sent* at least two P2P transfers to j before the experiment (repeated-transfer ties). Specification and trimming as in Table 1. Money-sending ties carry within-industry adoption into markets outside Dakar ($+83\%$ of baseline, $p_{\text{RI}} = 0.001$)—mirroring the mutual-contact channel of Appendix Table B6—and carry no significant local effects. The pattern is identical with a ≥ 1-transfer definition ($N = 60{,}381$; outside-Dakar $+5.23$, $+60\%$, $p_{\text{RI}} = 0.001$; all other margins null under randomization inference). RI p-values (1,000 permutations) in brackets. $^{*}p_{\text{RI}} < 0.10$, $^{**} < 0.05$, $^{***} < 0.01$.

Table B13: Shared-Customer Spillovers: Varying the Regular-Customer Threshold

	Within-taxi adoption (Dakar)			Cross-industry adoption		
	(1) All	(2) Comm.-tied	(3) No tie	(4) All	(5) Comm.-tied	(6) No tie
Panel A: ≥ 1 transaction with each party (any shared customer)						
Share treated	$\mathbf{405.61}^{***}$	660.40^{***}	242.12^{***}	0.60	-10.10	1.36
	(38.07)	(61.56)	(47.19)	(6.52)	(24.43)	(6.76)
	[0.001]	[0.001]	[0.002]	[0.973]	[0.773]	[0.952]
Effect size	+55.0%	+88.1%	+34.3%	+0.1%	−1.6%	+0.2%
Dep. var. mean	737.50	749.73	705.68	627.97	626.23	628.37
N	10,339	7,468	2,871	66,419	12,382	54,037
Panel B: ≥ 3 transactions with each party						
Share treated	161.48^{*}	$\mathbf{321.22}^{**}$	118.99	46.51	71.76	44.04
	(38.53)	(81.01)	(43.30)	(17.27)	(74.00)	(17.79)
	[0.051]	[0.017]	[0.142]	[0.145]	[0.441]	[0.170]
Effect size	+32.1%	+58.1%	+24.9%	+8.3%	+12.9%	+7.9%
Dep. var. mean	502.47	552.78	478.43	559.31	557.25	559.45
N	2,227	720	1,507	4,375	262	4,113

Notes: Identical construction to Table 5, varying the shared-customer threshold around the main table's ≥ 2 transactions with each party. Panel A requires only one transaction with the driver and one with the merchant (any shared customer); Panel B requires three with each, under which only 1,797 driver–customer pairs meet the driver-side requirement and samples shrink sharply. The within-taxi effect and its community-tie gradient persist at every threshold—all, community-tied, and no-tie coefficients are positive throughout, with the community-tied effect roughly twice the no-tie effect—while cross-industry effects are null at every threshold. Coefficients $\times 1{,}000$; HC1 SEs in parentheses; RI p-values in brackets. Regressions control for expected exposure. $^{*}p_{\text{RI}} < 0.10$, $^{**} < 0.05$, $^{***} < 0.01$.

Table B14: Pre-Intervention P2P Customer Network (≥ 2 transactions, Apr 2021–Mar 2022)

	Within-taxi, Dakar	Cross-industry
Share treated	$+\mathbf{122.16}^{***}$	-7.44
	(23.57)	(13.17)
Effect size	$+17.9\%$	-1.0%
Dep. var. mean	680.80	760.62
N	3,490	6,613

Notes: Coefficients, SEs, and dep. var. means multiplied by 1,000. Sample: non-experimental merchants linked to non-garage experimental drivers through shared customers identified from pre-intervention P2P transaction data (April 2021–March 2022, CFA 500–3,000 per transaction, ≥ 2 transactions per customer). Merchants who adopted before March 2022 are excluded ($N = 1{,}785$), as are matched (experimental) merchants. Column 1: taxi merchants, dep. var. = within-taxi adoption (Dakar); column 2: non-taxi merchants, dep. var. = cross-industry adoption (Dakar or outside). Both columns exclude merchants with mutual phone contacts to experimental drivers (non-social specification); including them changes the coefficients by less than 1% (only 20 taxi and 22 non-taxi merchants have mutual ties in this sample). Regressions control for expected exposure (leave-one-out mean; coefficients not shown). The main post-experiment customer table (Table 5) applies the stricter community-tie split (corrected overlap $\geq 2\%$) and reaches the same conclusion: customer-demand effects do not require a social relationship. HC1 robust SEs in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table B15: Customer Network Temporal Stability Diagnostic

	W1 Nov 2022–Apr 2023	W2 May–Nov 2023
Panel A: Non-Social subsample (no mutual phone ties)		
ShareTreated ($\times 100$)	$+27.28^{***}$	$+6.59^{***}$
	(2.33)	(2.51)
N	4,450	7,757
$\bar{y}$	0.784	0.744
Panel B: All customer-linked merchants		
ShareTreated ($\times 100$)	$+28.18^{***}$	$+6.47^{***}$
	(2.33)	(2.50)
N	4,495	7,906

Notes: Each column constructs the shared-customer network G^{cust} from transactions within that window only (≥ 2 transactions per customer). Garage-affiliated drivers excluded. Outcome: within-taxi DKR mobile payment adoption by endline. ShareTreated $= \sum T_i / N_j^{\text{cust}}$ is the share of customer-linked experimental drivers assigned to treatment. Robust SEs in parentheses. The Non-Social subsample restricts to merchants with no mutual phone ties ($I_{\text{obs}} < 2$) with any experimental driver; the main customer table (Table 5) applies the stricter community-tie split (corrected overlap $\geq 2\%$) with the same conclusion. $^{*}p<0.10$, $^{**}p<0.05$, $^{***}p<0.01$.

C Fingerprint Data Processing

The phone contact network used throughout the paper is constructed from encrypted contact-book fingerprints collected by the mobile money operator. This appendix documents the processing steps from raw fingerprint data to the analysis-ready edge list.

Step 1: Fingerprint extraction. When a user first logs into the mobile money application, the app requests access to the phone's contact list (consent rate $\approx 86\%$). For consenting users, each saved phone number is hashed into one of $2^{16} = 65{,}536$ buckets using a 16-bit FNV-1a hash (Noll et al., 2026). The resulting fingerprint is a set of unique bucket values representing the user's contact book. The February 2024 snapshot covers approximately 7 million unique users.

Step 2: Matching experimental drivers. Each of the 2,196 experimental drivers has up to three registered phone numbers. I search the fingerprint file for any number matching any driver, keeping the fingerprint with the most buckets when a driver has multiple records (best-phone-per-person). Of the 2,150 drivers with a mobile money account match (`has_match`), 1,912 (88.9%) have at least one phone number with fingerprint data (Section 1.3). The remaining 238 drivers were never fingerprinted by the application.

Step 3: Minimum contact threshold. Drivers with fewer than 10 unique hash buckets in their fingerprint are excluded, as the overlap computation is dominated by hash collision noise at such small set sizes. This removes 5 additional drivers, leaving 1,907 in the network analysis sample.

Step 4: Candidate pairs and overlap measurement. For each fingerprinted user ("individual j") and each driver i, the raw bucket intersection $I_{ij} = |F_i \cap F_j|$ is computed. At 16-bit resolution, part of any observed intersection is mechanical: two independent contact books of sizes $|F_i|$ and $|F_j|$ collide in

$$\mu_{ij} = \frac{|F_i|\,|F_j|}{65{,}536}$$

buckets in expectation. The analysis therefore measures ties by the *bias-corrected overlap share*

$$\text{OverlapShare}_{ij}^{\text{corr}} = \frac{\max(I_{ij} - \mu_{ij},\, 0)}{\min(|F_i|, |F_j|)},$$

which subtracts the expected collision count before normalizing by the smaller contact book. For a truly unrelated pair, the corrected share is centered on zero with noise of order $1/\min(|F_i|, |F_j|)$; a 2 percent floor therefore removes nearly all purely accidental overlap while retaining genuine shared-environment pairs. The 2 percent *community-tie* floor is set an order of magnitude above this noise scale and is validated against survey-reported relationships (Section 1.3; Appendix Tables B1 and B2). Because the share is normalized by the smaller contact book, the resulting graph is deliberately broader than a contact book—it measures shared social environment—and is subsequently reduced by strict hub trimming (Step 7).

Step 5: Save indicators as tie attributes. Directed *save* indicators are computed separately: driver i's self-hash bucket appearing in individual j's fingerprint ("j saved i"), and vice versa. The conjunction defines a *mutual contact*. Unlike earlier versions of this pipeline, these indicators do *not* define the network: the analysis network is the community-tie graph (corrected overlap $\geq 2\%$), and the save indicators are retained only as tie *attributes*—to identify the small reciprocal-save (mutual) layer, which plays a distinct role (carrying within-industry adoption across markets; Appendix Table B6), and to know the direction from which a link originates. Because a single self-hash bucket match can arise from collision, save indicators for any individual pair are noisy; they are informative in aggregate (mutual pairs are reported as substantially closer in survey data; Table B2) but are never used as a stand-alone existence criterion for a tie.

Step 6: Pair deduplication. Individuals can appear under multiple phone numbers. Driver–contact pairs are deduplicated keeping the first record per pair (best-phone convention) *before* any threshold is applied, so that thresholding operates on a unique-pair edge list.

Step 7: Degree trimming. Contacts in the top 2% of the degree distribution are dropped, and every contact connected to any top-2% degree driver is then excluded entirely (strict trimming). This removes hub nodes whose high connectivity may reflect institutional roles (e.g., agents, dispatchers) rather than personal social ties, and is applied in all main specifications.

Final analysis sample. After all processing steps, the community-tie network (corrected overlap $\geq 2\%$) links the experimental drivers to 1,145,899 individuals through 23.7 million

edges; 1,020,391 individuals have at least one strong tie ($\geq 5\%$), and about 164,000 are additionally connected through the reciprocal-save (mutual) layer.

D Structural Estimation: Details, Validation, and Distance Decay

D.1 Estimation Details

Overview. The seven parameters are solved sequentially, exploiting the block structure of the moment conditions. All exposure variables are functions of the exogenous treatment assignment, so the model has a closed-form adoption probability—no cascade simulation is needed.

Learning (β^s_{learn}, β^w_{learn}). For each candidate β, I compute model-implied adoption probabilities $P_j(\theta) = 1 - (1 - \Lambda(\alpha + \beta \cdot S_j^{\text{phone}}))^6$ for all cross-industry individuals, then run the same OLS regression on P_j that produced the reduced-form Table 1 coefficients—an indirect inference step (Gouriéroux et al., 1993). The value of β that equates the model-predicted OLS coefficient to the data coefficient is found separately for the strong-tie subsample ($\geq 5\%$ corrected overlap; m1 $\rightarrow \beta^s_{\text{learn}}$) and the weaker community band (2–5%; m2 $\rightarrow \beta^w_{\text{learn}}$). Each is a one-dimensional root-finding problem (Brent's method).

Customer demand (β_{demand}). Identified via the average marginal effect of the absorbing model: $\hat{\beta}_{\text{demand}} = \hat{b}_{\text{m4}}/[T(1-q)^{T-1} \cdot q(1-q)]$, where $q = 1-(1-\bar{P})^{1/T}$ is the per-period adoption probability implied by the six-month cumulative rate $\bar{P} = 0.717$ among customer-linked non-garage taxi businesses (Table 5, column 1, reports the mean of adoption$\times$1,000 as 717.00), and $T = 6$. The AME of the absorbing binomial model is $T(1-q)^{T-1} \cdot q(1-q)$, which accounts for the fact that a latent-index shift in any single period compounds into a smaller cumulative probability change than the naive logistic $\bar{P}(1-\bar{P})$ would imply.

Intercepts (α_{cross}, α_{taxi}). Each solved by bisection from the rate moments (m5, m6): find α such that the sample average of $P_j(\alpha, \hat{\beta})$ matches the observed six-month adoption rate. The two intercepts are calibrated against different subsamples: m5 uses the cross-industry adoption rate from the strong-tie sample ($\geq 5\%$; Table 1), and m6 uses the within-taxi rate from the community sample ($\geq 2\%$; Table 2). Note that $\hat{\alpha}_{\text{taxi}}$ is not the log-odds of the target rate at zero exposure: the bisection averages P_j over observed (nonzero) exposures, and because $\hat{\beta}^{\text{geo}}_{\text{comp}}$ is large and negative, geographic competition pulls most taxi

businesses' adoption probabilities well below $\Lambda(\hat{\alpha}_{\text{taxi}})$, requiring a less negative intercept to match the observed rate. This also explains the large SE on $\hat{\alpha}_{\text{taxi}}$: the intercept and competition level partially trade off. For $\hat{\alpha}_{\text{cross}}$, the calibration is nearly exact because cross-industry businesses have no competition or demand channel.

Distance decay (δ) and geographic competition ($\beta^{\text{geo}}_{\text{comp}}$). The reduced-form coefficient at distance d is approximately:

$$\hat{b}(d) \;\approx\; \beta^{\text{geo}}_{\text{comp}} \cdot \exp(-\delta \cdot d) \;\times\; T \cdot \bar{p}(1-\bar{p}), \tag{A1}$$

where the second factor is the average marginal effect of the logistic-binomial model.[23] Taking logs gives $\log|\hat{b}(d)| \approx \text{const} - \delta \cdot d$. Regressing across the six all-merchant (Panel A) bins yields $\hat{\delta} = 3.73^{***}$ km^{-1} (half-life $= 186$ m); Panel A is used for the gradient because its samples are largest and the *ratio* of coefficients identifies δ regardless of the population's level. The level then pins down $\hat{\beta}^{\text{geo}}_{\text{comp}} = \hat{b}(d_0)/[\exp(-\hat{\delta} \cdot d_0) \cdot T\bar{p}_g(1-\bar{p}_g)]$ from the *community-tied* (Panel B) coefficient at 100 m ($\hat{b}(0.1) = -0.0826$), where $\bar{p}_g$ is the per-period adoption probability implied by the garage-sample base rate (1.20% cumulative). Using the garage-sample rate is essential: garage-area adoption rates are roughly twenty times the network-sample rate, so inverting the level at the network-sample AME would misstate the competition parameter by an order of magnitude. This gives $\hat{\beta}^{\text{geo}}_{\text{comp}} = -10.004^{***}$.

Fit of the distance gradient. The log-linear specification fits well ($R^2 = 0.96$). Predicted-to-actual ratios range from 0.83 (150 m) to 1.22 (50 m), with 4 of six bins within 15%.

Standard errors. Standard errors for $\hat{\alpha}_{\text{cross}}$, $\hat{\alpha}_{\text{taxi}}$, $\hat{\beta}^{s}_{\text{learn}}$, and $\hat{\beta}^{w}_{\text{learn}}$ use the delta method with a numerical Jacobian of the moment conditions.

The remaining three parameters ($\hat{\delta}$, $\hat{\beta}^{\text{geo}}_{\text{comp}}$, $\hat{\beta}_{\text{demand}}$) require different approaches because the AME inversion that identifies them amplifies noise on the logit-index scale. A naive delta-method SE for $\hat{\beta}^{\text{geo}}_{\text{comp}}$ would divide by the tiny network-sample AME, inflating the SE by orders of magnitude even though the underlying reduced-form garage coefficients are significant under randomization inference (Table 4, Panel B).

$\hat{\delta}$ and $\hat{\beta}^{geo}_{comp}$: Both are identified from the log-linear regression $\log|\hat{b}(d)| = a - \delta \cdot d$ across six distance bins, which has four overidentifying restrictions. The OLS standard errors of this regression directly give SE($\hat{\delta}$) from the slope; SE($\hat{\beta}^{\text{geo}}_{\text{comp}}$) propagates the sampling error of the Panel B 100 m coefficient through the garage-sample AME inversion.

[23] Using $\partial\Lambda/\partial\eta = \bar{p}(1-\bar{p})$ and $\partial P/\partial p = T(1-p)^{T-1} \approx T$ when $p \ll 1$. Exact in the limit $\bar{p} \to 0$.

$\hat{\beta}_{demand}$: Identified via the absorbing-model AME: $\hat{\beta}_{\text{demand}} = \hat{b}_{\text{m4}}/[T(1-q)^{T-1} \cdot q(1-q)]$, where $q = 1-(1-\bar{P})^{1/T}$ and $\bar{P}$ is the six-month adoption rate among customer-linked taxi businesses. The SE is computed by moment bootstrap: drawing $\hat{b}_{\text{m4}}$ from $\mathcal{N}(\hat{b}, \hat{\sigma}_b^2)$ and perturbing $\bar{P}$, re-solving for $\hat{\beta}_{\text{demand}}$ in each of 500 draws, and taking the empirical standard deviation. This correctly propagates uncertainty through the nonlinear AME inversion.

ShareTreated vs. AnyTreated. The reduced-form Table 4 uses binary AnyTreated for RI power. The structural model requires continuous ShareTreated because counterfactuals change treatment intensity. Of the 36 GPS-matched garages with experimental drivers, rosters are small—a median of 2 and mean of 3.9 experimental drivers per garage (142 in total; the largest, the airport garage, has 30)—so ShareTreated is an extremely coarse variable (for a two-driver garage it can only be 0, $\frac{1}{2}$, or 1), and the realized variation is essentially binary: 30 garages have AnyTreated $= 1$ and 6 have AnyTreated $= 0$ (control-only rosters of one to three drivers). AnyTreated captures the empirically relevant contrast; RI provides the power.

The structural identification of δ from the AnyTreated gradient is valid because δ is identified from the *ratio* of coefficients across distance bins, not from their level. Taking logs: $\log|\hat{b}(d_1)| - \log|\hat{b}(d_2)| = -\delta(d_1 - d_2)$, regardless of whether the regressor is binary or continuous. The level pins down $\beta^{\text{geo}}_{\text{comp}}$; since AnyTreated ≈ 1 at most garages, this is interpretable as the effect of full treatment intensity. ShareTreated regressions show the same sign and monotone decay, though with less power than the binary contrast.

D.2 Validation Details

Sign predictions. The model correctly predicts the sign of all four untargeted survey moments and the correct ordering (m9 $>$ m10; m8 $>$ m11):

- m8 (all friends, within-taxi outcome): data $+0.030$; the model predicts a positive response, driven by community-tie learning (most friends are cross-industry contacts, so learning is the dominant active channel).
- m9 (taxi friends): data $+0.098$; the model predicts a positive response despite same-industry competition, because at survey-sample exposure levels the demand index dominates the competition index.
- m10 (taxi non-friends): data $+0.009$; the model predicts a small positive response, with demand and competition both attenuated.

- m11 (serendipitous): data $+0.018$, statistically significant in Table 3 ($p_{\text{RI}} = 0.004$) and large relative to the control mean ($+47.9\%$). Under the community-network model this moment is no longer a puzzle: learning operates at a uniform rate across the community-tie strength distribution ($\hat{\beta}^{w}_{\text{learn}} \approx \hat{\beta}^{s}_{\text{learn}}$), and 28% of survey "serendipitous" contacts are community ties by the corrected-overlap measure (Table B1), so the model predicts a positive response in this category, consistent with the data. A specification in which weak ties transmit nothing would predict no response here.

Magnitude attenuation. Model-implied responses evaluated at representative population-mean exposures are an order of magnitude smaller than the survey coefficients, so Panel C is read as a sign test rather than a magnitude test. Three factors drive the attenuation. First, *logistic compression*: baseline adoption is rare (0.058% for taxi, 1.343% for cross-industry), placing the model in the left tail where large index effects produce small probability differences. Second, *subsample selection*: survey "friends" are the right tail of tie strength, not the population average that identifies the parameters. Third, *unmodeled channels*: taxi friends likely share customers at above-average rates; serendipitous contacts may transmit through in-person interactions.

Because Panel C predictions use representative population-mean exposures (not individual-level data from the survey sample), they are indicative of qualitative implications rather than exact forecasts. The sign predictions are the appropriate validation criterion.

D.3 Distance Decay of Competition

Amplification factors. I recover $w(r)$ by converting each Panel B garage coefficient to the latent scale ($\hat{b}(r)/[\bar{p}(r)(1-\bar{p}(r))]$), subtracting learning ($\hat{\beta}^{s}_{\text{learn}}$), and dividing by the phone-network baseline ($\approx +0.271$):

Radius (m)	$\hat{b}(r)$	$\bar{p}(r)$	$w(r)$
50	-0.102	0.017	24×
75	-0.083	0.015	22×
100	-0.083	0.012	27×
150	-0.052	0.010	21×
300	-0.027	0.007	15×
500	-0.015	0.006	11×

Parametric fits. An exponential fit to the amplification factors ($w(d) = 26 \cdot e^{-d/577\text{m}}$) implies a half-life of 400 m; setting $w(d^*) = 1$ gives $d^* \approx 2$ km—the distance from a garage

at which the competitive force has decayed to the influence of full community-network exposure. The model-side calculation, $|\hat{\beta}^{\text{geo}}_{\text{comp}}|e^{-\hat{\delta}d^*} = \hat{\beta}^{s}_{\text{learn}}$, gives $d^* \approx 1\,\text{km}$; both place the crossover at walking-to-driving scale rather than city scale. The data-side amplification half-life exceeds the 186 m structural half-life. The difference arises because the amplification factors $w(r)$ incorporate both the exponential kernel *and* variation in baseline adoption $\bar{p}(r)$ across distance bins: businesses closer to garages have higher baseline adoption, so the AME conversion $[\bar{p}(1-\bar{p})]^{-1}$ amplifies the latent coefficient more at short distances, making $w(r)$ decline faster than the pure structural decay $\exp(-\hat{\delta} \cdot d)$.

E Counterfactual Details

Relevant population. The full analysis sample ($N \approx 3.9$M) is not the right basis: the vast majority have only weak ties ($\hat{\beta}^{w}_{\text{learn}} \approx 0$), making them structurally unresponsive to targeting. The combined reachable network includes three populations: (A) 3,082,043 individuals with community phone ties (corrected overlap $\geq 2\%$; hub-contact trimming only—the estimation samples additionally exclude every contact of a top-2% driver, which is why the regression N of 1,145,899 is smaller), reached through learning—99.3% have at least one strong ($\geq 5\%$) tie and receive β^{s}_{learn}, the remainder β^{w}_{learn}; because the two learning rates are statistically indistinguishable (Table 7), this classification is immaterial to the counterfactuals; (B) 10,442 taxi businesses in the phone network, for whom all three channels are active—learning through phone ties, competition through geographic proximity to garages, and customer demand through shared P2B customers; and (C) 40,710 non-taxi businesses within 500 m of a garage, reached through phone learning only (competition and demand require $\mathbf{1}[\text{taxi}_j]$ per equation (3)). Customer demand is an attribute of taxi businesses in population B who share customers with experimental drivers, not a separate population. Total: 3,103,114.

Customer exposure from transactions. For taxi businesses in population B, S^{cust}_j is constructed from P2B mobile money transactions between November 2022 and November 2023. A customer qualifies if they made ≥ 2 P2B transactions in aggregate across non-garage experimental drivers during this window (not necessarily with the same driver); a merchant–driver pair is linked if they share at least one qualifying customer. This yields 95,114 businesses linked to 1,521 drivers. Correlation with S^{phone}: $r = 0.005$, ensuring demand and learning respond independently to targeting. For non-taxi individuals and taxi businesses without customer links, $S^{\text{cust}}_j = 0$.

All-garage reweighting. The counterfactual asks: what if all $1,256$ treated drivers had the network characteristics of garage drivers? The same $1,256$ drivers remain treated (preserving the budget and network positions), but each driver's contribution to exposure is scaled by the ratio of average garage-driver to average treated-driver characteristics.

Formally, define the driver-level treatment vector $\mathbf{t} \in \mathbb{R}^{n_d}$. Under random assignment, $t_i = 1$ for treated drivers. Under all-garage, $t_i^{\text{phone}} = \bar{d}_{\text{garage}}/\bar{d}_{\text{treated}}$ for all $1,256$ treated drivers, where $\bar{d}$ denotes mean phone degree. The phone exposure for individual j is:

$$S_j^{\text{phone}} = \min\left(1,\ \frac{\sum_i B_{ji}\, t_i^{\text{phone}}}{\sum_i B_{ji}}\right), \tag{A2}$$

where $B_{ji} = 1$ if driver i is a mutual phone contact of j. Customer demand exposure S_j^{cust} is computed analogously, scaling by the ratio of average customer links (from the incidence matrix B^{cust}). Geographic exposure is set to its maximum: $\text{ShareTreated}_g = 1$ at every garage.

The ratios are computed from the incidence matrices directly. Under the community-tie network, scaling treated drivers to the garage-driver mean raises both phone and customer exposure ($\bar{S}^{\text{phone}}$: $0.586 \to 0.659$; $\bar{S}^{\text{demand}}$: $0.658 \to 0.700$), which is why the all-garage benchmark modestly exceeds random (+13.6%) despite maximal geographic competition (ShareTreated $= 1$ at all garages); the garage-concentrated rule, which does not rescale intensities, falls below random.