Nationwide Diffusion of Technology: Experimental Evidence from Multiple Networks*

Deivy Houeix†

[Click here for the most recent version]

August 21, 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 at the taxi-business level among 2,196 surveyed taxi owners and drivers in Dakar, Senegal. I trace 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, including weak ones, transmit knowledge and raise adoption across industries and cities. Co-location reduces adoption through a sharply local competitive force: the effect halves roughly every 448 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 counterfactual exercise disciplined by the reduced-form estimates predicts that targeting the businesses with the largest phone-network reach raises induced adoption by about 57% relative to random. Concentrating treatment at garages, a natural targeting rule, does worse than random.

*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 the 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 proven challenging.

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: an exogenous 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 new digital payment technology among 2,196 taxi owners and drivers, spanning 1,454 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). 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. Instead, taxi drivers move through the city, hence the networks further come apart. Social ties stretch across the country, customers span Dakar, and competition is potentially impor-

tant 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. On first login, users are asked to share with the app a hashed fingerprint of their phone contact book, which the operator uses to facilitate transfers to saved contacts; 86% consent. Each saved number is hashed into one of 65,536 buckets, so I observe the set of buckets a user saved without recovering any number. I link two users when each appears in the other's contact book and measure tie strength by their overlap share: how much of their contact lists they hold in common, net of expected hash collisions. 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). Since most firms in this setting are owner-operated, the fingerprints provide an unusually direct measure of the social networks of firm owners, covering both mutual contacts and contacts with high overlap. The resulting network links all experimental taxi business owners to other mobile money users across Senegal, at varying degrees of tie strength. To validate this novel measure of social network, I conduct an independent relationship-listing survey. I draw a random sample of each respondent's past transaction partners and ask them to describe their relationship. This survey confirms that corrected overlap tracks social closeness—spouses 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. 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, especially 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 both treated and control taxis gain access to digital payments and appear 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. 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 *expected* exposure under the randomization (Borusyak and Hull, 2023); a count parameterization yields similar results. 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 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 $+39.0\%$ above the control mean ($p = 0.017$), while the 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. Interestingly, the spillovers appear in the community network, but not among direct mutual phone contacts. Because control drivers were also told about the technology, mutual contacts are more likely to be infra-marginal. Their mutual 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 digital technology on social media).

Spillovers expand to firms in cities beyond Dakar, where effect sizes 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 experi-

ment. 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 ($p = 0.018$), and it is significant among receiving businesses never observed in Dakar ($p = 0.027$). 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 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 ($p = 0.031$) and decays with distance. 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 ($p = 0.003$). 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 ($p = 0.007$). Hence, learning outweighs competition on very close ties.

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 44.8% more likely to adopt ($p = 0.001$). The effect remains large ($+27.4\%$, $p = 0.001$) 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 qualitatively.

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 isolate demand effects; the profile of garage coefficients across eight radii, from 50 meters to one kilometer, identifies the level and decay of competition; and within-taxi responses on mutual ties, inside and outside Dakar, separate within-industry learning from the competition carried by the ties themselves. The eight parameters are estimated jointly by minimum distance. I use survey moments on adoption by relationship type as untargeted validation. Competition is intense but highly local: the competitive force halves every 448 meters and falls below the strong-tie learning response only beyond roughly 1.3 kilometers from a treated garage. Customer demand generates large positive complementarities, including between taxis with no social tie. The estimated parameters help rationalize some reduced-form patterns. First, the within-industry spillover in the community network is null because learning and competition offset each other. 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 and that serendipitous contacts also raise adoption, because learning also travels through weak community ties.

This structural estimation allows me to run targeting-policy counterfactuals. Relative to random assignment, targeting the taxi businesses with the greatest measured phone-network reach would have raised adoption by about 57 percent, almost entirely through the social-learning channel. 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 clarifies what distance measures in the diffusion of technology between firms. Development economics has typically focused on local networks: neighbors, villages, clusters, narrow radii (Cai et al., 2015; Suri and Udry, 2024; Duflo et al., 2026; Chaurey et al., 2025), and the broader literature treats proximity and face-to-face contact as the leading conduits (Marshall, 1920; Moretti, 2021; Baum-Snow et al., 2024; Cai and Szeidl, 2017; Atkin et al., 2024). That focus may well have

been a good approximation: when exchanging information requires co-presence, the local network is the diffusion network. However, there is good reason to think diffusion is not only local in today's economy. Mobile phones hold ties together across the country, and mobile money moves value along them. Consistent with this, I find that the forces spreading the technology do not require closeness: learning runs through social ties at any distance, friendship transmits it between direct rivals, and shared customers carry it between firms with no social tie at all. The force that is distinctly spatial is the one that slows diffusion: co-located rivals withhold what they learn (Cefalà et al., 2025), with a competitive effect that halves every 448 meters. This also helps reconcile spillover estimates of opposite sign depending on the relationships proximity bundles across settings (Bloom et al., 2013; Myers and Lanahan, 2022; Cefalà et al., 2025). The competition effect may be largest in non-tradable services—a growing share of employment in developing economies (Rodrik, 2016)—where firms compete on local customers (Foster, 1965).

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., 2025; 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., 2025), while noting that learning is hard to quantify in the same setting. Observing the two conduits separately shows that the forces differ less in existence than in shape. Strategic complementarities are the stronger force where they operate—customer demand is the dominant channel within the taxi sector—but their reach is bounded by the customer network itself. Demand effects mostly operate between firms that share customers and the cross-industry customer effect is thus limited. Social learning, however, may be weaker per tie, but spread across each firm's thousands of community ties. It is the primary channel that crosses industries and cities in this setting, and the one that carries the aggregate diffusion—nearly all of the gain from network-based targeting in the counterfactual exercise runs through learning. This breadth also helps explain why the reach of learning has been hard to quantify: a design confined to one sector or one neighborhood retains only the small fraction of ties that stay inside its boundary.

Third, the paper expands the measurement frontier for firm networks in developing economies, where most firms are informal, owner-operated, and absent from administra-

tive 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 richer 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 show the limits of the distance proxy: ties can be sparse and stretch far beyond the neighborhood, and over half of the members of taxi owners' community networks live outside the capital. The ties that carry the diffusion are community ties, rarely appearing in survey rosters. Such data now exist across lower-income countries, so the same measurement is feasible in many other settings. The questions asked of formal production networks (how firms find partners, how they upgrade, how shocks and technologies propagate) can then be further addressed in 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, the experimental design, and the network data. Section 2 documents empirical patterns in firms' networks. Section 3 presents the empirical framework. Section 4 estimates social learning across industries and cities. Section 5 identifies competition and customer-demand complementarities within the taxi industry. Section 6 develops and estimates the structural model and reports the counterfactual targeting exercises. Section 7 concludes.

1 Setting, Experiment, and Data

1.1 The Technology and Market Environment

The Digital Payment Technology. 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 (p2p) 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 (see Houeix (2026)).

The technology is especially well suited to studying diffusion. Here, customers use existing mobile money accounts that were already widespread in Senegal. The adoption decision therefore lies with the taxi 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 particularly useful for studying whether the same technology diffuses beyond the sector in which access is initially randomized.

The market environment also helps isolate the mechanisms behind adoption beyond monetary cost. The latter is low in this setting as the main adoption cost is time: the business must locate a registration agent, provide the required paperwork, and install the business application for free. 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, where adoption is constrained by information, trust, and coordination, and less so 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 payment 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, over 230,000 businesses in Senegal had adopted the digital payment technology (see Figure 2).[3] In particular, by early 2024, data show that

[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.

[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.

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 (taxi adoption was still growing rapidly: ∼7,000 of an eventual ∼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 represents an opportunity to study whether a randomized intervention in one urban service sector can generate diffusion of the same technology across industries and across cities.

The Taxi Sector. 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 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 individually 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. 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 social networks, competition among nearby same-industry firms, and customer-demand complementarities when customers begin to expect digital payment.

Overall, the taxi industry is well suited to isolate different mechanisms. For a firm rooted in one location, the three networks largely coincide: its neighbors are at once its potential competitors, its customers, and its acquaintances, so exposure measured on any one network proxies for the others. Taxi drivers move through the city, which pulls the networks apart—social ties possibly stretch across the country, customers span 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.

and competition concentrates at the garages where about 14% of drivers queue for the same passengers—so each mechanism can be measured on a network where the other two are muted.

Technological Benefits. The economic impact of the technology on businesses is the focus of a separate paper (Houeix, 2026): from the point of view of this paper, the key takeaway is that this technology is beneficial for business adopters. 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 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. See Houeix (2026) for more details. This paper abstracts from the owner–driver relationship: the observability sub-treatments are pooled, and diffusion is studied at the level of the business. Many individuals in the experiment are owners driving their own cab, and most adopters outside the experiment are owner-operated firms with no employees, so the internal monitoring margin plays limited role in the spillovers studied here.

1.2 Experimental Design

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

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.

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 invited to the baseline survey, with owners and drivers

interviewed separately in 45-minute questionnaires. The baseline dataset eventually 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 others or to share information about it. The control businesses were still made aware of the technology and could possibly share information about it. Any adoption among non-experimental firms therefore reflects transmission from the experimental sample rather than direct contact with the study.

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) 74.1% of control vehicles (79.4% by November 2023). The estimation exercise

[5]Sample accounting: 2,196 stakeholder records randomized at the business level $\rightarrow$ 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 $\rightarrow$ 1,912 with phone fingerprint data $\rightarrow$ 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.

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 experimental 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, 2026). Second, the focus of this paper, it creates exogenous variation in the treatment status of taxi firms that are embedded in much larger social, local, and customer 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

1.3.1 Social Networks from Hashed Phone Contact Fingerprints

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 at first login, with about 86% of users consenting. Unlike the survey links more common in economics, contact-book fingerprints capture the set of contacts a user chose to save in the phone and the contact overlap, making them well suited to measuring social proximity among informal 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. This is a routine request among consumer apps. 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 to recover set-overlap measurement and private contact securely, 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. The correction matters because hashing maps every saved number into one of 65,536 buckets: two users not belonging to the same community network will still share some buckets by accident. For two typical contact books of about 600 numbers each, $\mu_{ij} = 600 \times 600/65{,}536 \approx 5.5$, so roughly five to six of their buckets coincide by chance alone—about 0.9 percent of either book.[7] An observed intersection of that size carries no information about a relationship, and the corrected share nets it out. Normalizing by the smaller contact list ensures that users with larger contact books are not mechanically penalized.

The correction follows the chemoinformatics literature, which subtracts the overlap expected under random hashing when comparing molecular fingerprints folded into a small bit space (Swamidass and Baldi, 2007). In economics, this is best interpreted as a measure of *social proximity*: how embedded are the two users in overlapping social circles? The interpretation is also consistent with the network-science literature, where neighborhood-overlap measures are informative about latent social proximity and link formation (Adamic and Adar, 2003; Liben-Nowell and Kleinberg, 2007).

[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 can be linked to multiple mobile money 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 few individuals appear in the network through more than one account. The phone-level analysis is therefore essentially equivalent to a person-level analysis.

[7]With $B = 65{,}536$ buckets, if the two fingerprints were independent uniform subsets of their observed sizes, each bucket occupied by i would be occupied by j with probability $|F_j|/B$, so $\mathbb{E}[I_{ij}] = \sum_{b \in F_i} \Pr(b \in F_j) = |F_i||F_j|/B = \mu_{ij}$. The corrected share thus subtracts the intersection that pure chance would generate between two unrelated users.

Tie Strength and the Community Network. I define a *community tie* as 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 sits well above the collision-noise scale[8] and is validated directly against survey ground truth (clients and serendipitous contacts average only 1.9 and 3.2 percent corrected overlap and clear it just 22.8 to 27.7 percent of the time, see Appendix Table B1 and below).

The community graph is deliberately broader than a contact book: it measures *shared social environment*—after hub trimming[9], the median driver is connected to 7,500 individuals (mean 12,700). *Strong ties* are defined as community ties with corrected overlap of at least 5 percent: a median of 600 per driver (mean 1,500). The main specifications are reported at the 2, 5, and 10 percent thresholds.[10] Reciprocity is measured separately: a *mutual contact* is a pair in which each party saved the other's hashed phone number. Mutual ties are a smaller subset of the community network with higher baseline adoption, and the analysis later shows they carry within-industry adoption across cities, but do not amplify diffusion across industry (Appendix Table B5). Non-mutual contact pairs above the floor are reported as close relationships 37 percent of the time versus 18 percent below it (see further discussion below and details in Appendix Table B2).

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. Of the 23,520 survey pairs, 16,470 had fingerprint data on both sides (Appendix C).

First, corrected overlap rises monotonically with reported closeness. Spouses 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%;

[8] After the μ correction, the overlap of two unrelated users is centered on zero, with standard deviation $\sqrt{\mu_{ij}}/\min(|F_i|, |F_j|) \approx 1/\sqrt{65{,}536} \approx 0.4$ percent for comparable contact books; the 2 percent floor is roughly five standard deviations above this null, so almost no purely accidental overlap clears it.

[9] The community graph includes ties with widely-held numbers (agents, popular businesses, commercial phone numbers, etc.) that likely reflect no personal relationship. I remove these by dropping the 2% most-connected contacts and, in a second step, every contact linked to a top-2-percent-connectivity driver, which leaves each driver's idiosyncratic ties (median 172,700 $\rightarrow$ 7,500). The point estimates are essentially unchanged by this procedure; trimming primarily sharpens inference by removing contacts whose treatment exposure is near-collinear across the sample.

[10] 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).

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 aims to separate 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. In contrast, 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" (spouse, family, friends) 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 B5)—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. Regressing survey-reported closeness on corrected overlap, geographic distance, reciprocal-save (mutual) status, and their interactions (Table B3) shows that overlap predicts closeness at every distance—and is, if anything, *more* informative for distant pairs, with the overlap slope roughly five times steeper for contacts outside Dakar: nearby, shared environment generates overlap even among casual acquaintances, while at distance only genuine relationships sustain it. Conditional on overlap, reciprocal saving carries substantial *additional* closeness content (a 31-percentage-point higher probability of a close relationship, stable across distance): overlap measures the shared social environment, while reciprocity marks the depth of the relationship within it.

Network Observed After the Experiment. The phone-contact fingerprints are observed in February 2024, after the experiment ended, and are interpreted as a post-period proxy for durable *pre-existing* social proximity. Table B4 supports this interpretation in two steps. First, treated and control drivers are balanced on network levels: no count or composition measure differs across arms, either on the raw fingerprint network (Panel A) or on the community-tie network used in the analysis (Panel B; joint F-test, $p = 0.69$), and the two marginal raw-network comparisons—control drivers have slightly *more* mutual contacts ($p = 0.056$) and marginally larger contact books—work *against* finding positive spillovers from treatment. Second, a more demanding test asks whether treatment *strengthened* ties, e.g., through treated drivers saving more technology-related contacts

(Panel C): treated drivers' tie-strength distributions do not shift rightward—pooled overlap is, if anything, weakly *lower* for treated, on both the raw edge-level and the corrected community measures—the composition of strong ties ($\geq 5\%$), the sample of interest below, is indistinguishable across arms on every dimension tested (all $p > 0.19$), and contact-type splits show the same (taxi contacts, the most relevant for the mechanism, $p = 0.14$). The February-2024 network can therefore be treated as pre-determined with respect to treatment; any residual bias from differential appearance is bounded below by the Lee bounds, see Table B9 and further discussion below.

Test for Differential Appearance Across Treatment. 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 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, I also compute tightened Lee bounds (Lee, 2009) within contact-book-size quartiles (see Table B9 and results below).

1.3.2 Customer Network from Transaction Data

The transaction data are primarily used to construct the customer-link objects relevant for the customer-demand channel. I distinguish 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 broader measure of realized customer overlap, but because it is observed after treatment it is interpreted more cautiously.

In addition, the transaction data provide a second social 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 network is less exhaustive than the phone network. It nevertheless contains richer information on financial interaction. In the 12 months before the experiment, the median experimental driver sent or received 119 transfers totaling CFA 1,492,655 (USD 2,488)—predominantly as a sender, consistent with fares not yet being payable digitally—with a median transfer size of CFA 12,397 (USD 21).

1.3.3 Geographic Network from 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 particularly 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.[11] These measures are essential for distinguishing remote diffusion through social ties from diffusion through physical co-presence.

1.4 Survey and Technology Adoption Data

Survey Data. To interpret the mechanisms behind diffusion, the paper combines the phone network with survey evidence. 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

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

with ties that respondents themselves describe as closer and more socially meaningful. Because this survey was conducted after the experiment completion, I use it mostly as suggestive and reinforcing evidence of spillovers.

Technology Adoption Data. Adoption of the business payment technology is observed directly in the provider's administrative records. For each adopting business, the data record the exact 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 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 three key patterns in the networks of small firms, using the comprehensive network data. These patterns motivate the investigation of technology diffusion within firms' 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. After hub trimming, the median driver has 7,500 community ties (mean 12,700), of which 600 are strong ($\geq 5\%$); the reciprocal-save (mutual) layer is far smaller still ($\sim$100 per driver).[12] 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 contacts to all business adopters of digital payments as of April 2023. Through

[12]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 B4, Panels A and C, which also reports the intermediate pre-trimming values).

mutual phone contacts alone (before hub trimming), the average driver is connected to 59 other business owners—15 taxi businesses and 44 in other industries (5.6% of all mutual contacts). Among strong mutual ties ($\geq 5\%$ corrected overlap), the composition reverses: of 10 business contacts, 8 are taxi businesses—the closest professional relationships concentrate *within* the industry, foreshadowing the within-industry, cross-market role the mutual layer plays below. The community-tie network extends this reach further (Table B4). Together, these layers allow me to track technology diffusion across both industry and space from a single experimental sample, and they reflect how deeply firm networks are interwoven with social networks in this low-income, owner-operated context.

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 B13, 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.[13] More specifically, they reported having discussed or recommended this digital payment technology to 12% of the people they know. This underscores the significant 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.

Stylized Fact 3 - Social relationships and the customer network are largely distinct. The social network and the customer network connect different pairs of people. In the validation survey, taxi clients share on average 1.9% of contacts with the driver—below the community-tie floor—and only 22.8% of client relationships clear it (Appendix Table B1). Matching each surveyed taxi contact to the customer network, taxi *friends* share customers with the respondent less than half as often as non-friend taxi colleagues (9% versus 24% of pairs, with 4.46 versus 23.55 shared customers on average). A driver's

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

friends typically serve other markets. Physical co-location does not determine friendship either: even among same-garage colleagues close enough to appear in a driver's transaction circle, 29% are not reported as friends, so social closeness varies within the set of co-located rivals. Because social closeness, customer overlap, and physical proximity sort into largely different relationships, each layer can be measured separately and can carry a different economic force, which is the premise of the empirical design that follows.

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, 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 distinction 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 experimental 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 learn-

ing 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 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. 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. 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 re-

gressor 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. 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 randomization-inference p-values as 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 Social Learning Across Industry and Space

This section documents two 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. In both populations, competition and customer demand are muted, so the estimates identify social learning; survey evidence on relationship types corroborates this reading. Section 5 turns to the two market forces that operate within the taxi sector.

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 for a selected sample. 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 $+39\%$ effect relative to predicted adoption at zero exposure, rising to $+5.16$ ($p_{\text{RI}} = 0.015$, $+69\%$) at the 10 percent threshold. Decomposing the effect into exclusive strength bands (Appendix Figure A10), per-band transmission operates throughout the strength distribution—effect sizes of $+15\%$ to $+58\%$ of the zero-exposure baseline from the 2–5% band through 17%+—consistent with transmission throughout the community-tie strength distribution; stronger ties yield more precise estimates, without evidence of larger per-tie effects.[14]

Reciprocity helps distinguish which ties carry each margin (Appendix Table B5). 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). Two features of the setting explain why the community layer, and not the mutual layer, carries the marginal information. First, control drivers were offered and told about the technology, so the close contacts of a control driver could have nearly the same direct information as those of a treated driver (no direct change in *awareness*). The treatment contrast on mutual ties is potentially small by design. Second, community contacts, who share the driver's environment, could learn instead through observation and second-order links, which may respond to whether the driver actually carries the technology. E.g., treated drivers anecdotally report posting photos of their QR code on social media, a channel that reaches the broad community further than close contacts alone.

Count Specification. Appendix Table B8 replaces the treated *share* with the *number* of treated ties. Cross-industry learning survives: an additional treated strong tie raises

[14] 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.

adoption by $+1.67$ per thousand at the $\geq 5\%$ threshold ($p_{\text{RI}} = 0.015$; $+2.01$, $p_{\text{RI}} = 0.008$ at $\geq 10\%$), in Dakar and beyond ($+0.30$, $p_{\text{RI}} = 0.007$).

Lee Bounds for Differential Appearance. Because drivers enter the network sample only if the February-2024 snapshot contains their fingerprint, and treated drivers are slightly more likely to appear (89.9% vs. 87.6%, $p = 0.093$, see Section 1.3), the estimation sample is potentially selected by treatment. Following Lee (2009), I bound the resulting bias by trimming the outcome distribution of the over-represented arm by the appearance-rate differential—comparing equally selected samples and bracketing the true effect under worst-case selection, under the assumption that treatment only (weakly) raises appearance—with trimming within contact-book-size quartiles to tighten the bounds. The bounds are tight around every main estimate (Table B9): $[+4.079^{**}, +4.198^{**}]$ for cross-industry Dakar and $[+0.680^{***}, +0.685^{***}]$ outside Dakar; on the community network, the distant cross-industry effect is bounded at $[+0.554^{**}, +0.686^{***}]$. Even worst-case trimming preserves sign, magnitude, and significance throughout, so differential appearance does not drive the estimated spillovers.

Removing the Demand and Competition Channels Entirely. As an additional check, Appendix Table B7 re-estimates the cross-industry specification of Table 1 after removing both confounding channels, detailed in the next Section 5: every driver–contact pair sharing a regular customer (the demand channel of Table 5) and every pair of a garage-exposed driver (the population carrying local competition). The estimates are essentially unchanged—$+3.63^{**}$ [0.037] in Dakar and $+0.61^{**}$ [0.027] outside Dakar at the strong-tie threshold, against $+4.162**$ and $+0.681***$ in the full sample—so the cross-industry diffusion documented here is carried by social learning alone, with neither shared demand nor competitive exposure contributing to the estimates.

4.2 Widespread Diffusion via Remote Interactions

Technology adoption also spreads beyond the capital. Table 1 shows that businesses in other cities connected to treated Dakar taxis adopt at significantly higher rates at every threshold, with proportional effects of $+81\%$ to $+139\%$ ($p_{\text{RI}} = 0.011$, 0.009, and 0.030 at the 2, 5, and 10 percent cutoffs), roughly twice the within-Dakar cross-industry effect sizes, consistent with a lower baseline outside the capital amplifying the same diffusion force. These results show that knowledge spillovers need not be local in the classic Marshallian sense (Marshall, 1920).

One possible mechanism is physical movement: treated drivers or their contacts temporarily 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 me 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%) cashed in/out 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.

Business Owner's Physical Travel is not Required for Diffusion. Table 6 shows the overall effect on the community-tie sample ($+0.68$ per 1,000, $p_{\text{RI}} = 0.009$, $+138\%$ of the zero-exposure baseline) 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 ($+1.82$ per 1,000, $p_{\text{RI}} = 0.027$), although a similar effect size. Physical presence in Dakar therefore does not appear required for adoption to respond to treated Dakar contacts.

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 of the eight largest secondary cities c, I code each experimental driver as having *visited* c if any of their mobile-money wallets transacted at an agent located in c at least once during the experiment, and I split individual j's treated share into two city-specific exposures: the share of j's driver contacts that are treated *and* visited c, and the share treated but never observed there. The outcome is adoption *in city* c, and the two exposures enter a single regression stacked across the eight cities with city fixed effects, an expected-exposure control for each share (its mean across the 1,000 placebo assignments, which absorbs the fact that individuals connected to c-visiting drivers disproportionately live near c), standard er-

rors clustered by individual, and randomization inference over the driver assignment. I find that the visiting-driver coefficient is three times larger ($+0.021$ vs. $+0.007$ per 100) but imprecise ($p_{\text{RI}} = 0.133$), because few treated drivers visit any given city—so I cannot rule out a substantial role for drivers physically carrying the technology, and 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 $+125\%$ of the control mean ($p_{\text{RI}} = 0.018$). It indicates that remote phone interactions move the technology across cities.

Tie Strength and Cross-Industry Diffusion Across Space. Across the exclusive strength bands (Appendix Figure A10), the distant cross-industry effect is large wherever the data can measure it: effect sizes of +83% to +144% in the 2–5, 5–8, and 8–11% bands and +126% in the open tail (all $p_{\text{RI}} \leq 0.064$), with two statistically null cells in the thin 11–17% range (252 and 153 adopters). As within Dakar, transmission does not rise systematically with tie strength—but at every band the distant effect size exceeds the local one: any community tie carries the technology across cities, to people the driver knows who live and work elsewhere.

Because these are cross-industry businesses (the outcome is non-taxi adoption), competition and customer demand are relatively muted channels; the effect is attributable to social learning transmitted across space through the phone network.

4.3 Survey Evidence on the Learning Channel

Survey-elicited relationship types corroborate the learning interpretation. Table 3, Panel C reports that cross-industry spillovers are significant through serendipitous encounters ($+47.5\%$, $p_{\text{RI}} = 0.004$) and positive but imprecise through former or regular taxi clients ($+94.4\%$, $p_{\text{RI}} = 0.134$). The strongest ties (Panel B: friends, frequent contacts) show positive but highly imprecise cross-industry effects. This is consistent with the idea that novel information often arrives through ties bridging separate circles (Atkin et al., 2024): 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 but need not be mutual contacts, see Table B1). The marginal ties are community ties, weak ones included, rather than mutual contacts: the survey pattern matches the administrative evidence that cross-industry diffusion operates entirely through non-mutual community ties, with approximately uniform per-band transmission (Section 4.1; Appendix Table B5 and Figure A10).

5 Competition and Customer Demand Within Industry

Two market forces operate alongside learning within the taxi sector: competition among nearby rivals and demand complementarities through shared customers. This section identifies each force through a distinct empirical variation and ties it to the structural decomposition developed later.

5.1 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 highly imprecise and appears null ($+0.08$, $p_{\text{RI}} = 0.674$). At the stronger-tie threshold competition dominates—the coefficient is negative and significant ($\hat{\beta} \times 1{,}000 = -0.98$, $p_{\text{RI}} = 0.007$, a -30% effect)—and the band decomposition locates the crowding in the moderate-strength bands (5–8%: -29%, $p_{\text{RI}} = 0.011$), 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 important within industry on friendship and regular relationships. Table 3 shows that within the taxi sector, spillovers concentrate among friends ($+36.2\%$, $p_{\text{RI}} = 0.007$) and frequent interlocutors ($+24.0\%$, $p_{\text{RI}} = 0.023$), while exposure through non-friends produces a null effect. Close contacts are the people from whom another taxi businesses can share information on how the technology works, whether customers use it, and whether payments are reliable. 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}). Under that decomposition, the negative within-taxi phone coefficient emerges because same-industry ties are spatially correlated with the garages that carry competition, with some competition also running through the social ties themselves (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 within industry. Weak same-industry

contacts, by contrast, are more exposed to competitive crowding — plausibly including information withholding — depressing the average spillover.

Count Specification. In the count specification of Appendix Table B8, the within-taxi coefficients, negative and significant in shares, are null at every threshold. One possible explanation is that competition responds further to the fraction of one's circle that is treated than the count. At the count specification, learning may exactly offset competition.

Lee Bounds. The tightened Lee bounds of Section 4.1 bracket the within-taxi estimate at $[-1.103^{***}, -0.976^{***}]$ (Table B9): even worst-case trimming for differential appearance preserves sign, magnitude, and significance.

5.2 Local Competition and Crowding Out Within Industry

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. The experiment identifies the resulting footprint — a negative local same-industry exposure response, concentrated around garages and decaying with distance — and strategic knowledge withholding, analogous to the channel identified by Cefalà et al. (2025) among farmers, is one mechanism consistent with it. The competitive force is 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 2: same-industry exposure activates this competitive channel, which is absent across sectors.[15]

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 area of competition since drivers at the same rank compete 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 36 named garages—covering 142 experimental drivers—can be matched to the GPS garage

[15]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.

list used for the spatial analysis. 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 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 above.

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.[16] 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, i.e., a community tie as defined earlier (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 -76% to -91% of the zero-exposure baseline—coefficients an order of magnitude larger than in Panel A. Extending the design outward, the community-tied effect remains detectable at 750 meters (-68%, $p_{\text{RI}} = 0.026$). Panel C shows the effect is entirely absent among the drivers' mutual phone contacts. Taken together, this indicates that competitive crowding requires both market overlap (proximity) and community-level exposure to the treated garage, and mutual 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 of distance decay is nonetheless consistent with competition, 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

[16] 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.

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 B10 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).

Robustness: Locating the Entrant's Market at Onboarding Survey. The same-/other-garage split of Tables 4 and B10 assigns each entrant's home garage from their modal agent location in the mobility data. Appendix Table B11 re-estimates the split using an independent measure: the GPS coordinates recorded at merchant *onboarding*, during a separate survey the company conducted in partnership with the research team. The two assignments disagree in about 30% of cases on the nearest garage, so this is a demanding check on measurement error in the market assignment. I find that the competition pattern remains unchanged, and, if anything, reinforced: among community-tied merchants, same-garage entry is negative and significant at every radius (-0.0809 at 50 m, $p_{\text{RI}} = 0.024$; -0.0091 at 500 m, $p_{\text{RI}} = 0.036$; $p_{\text{RI}} \leq 0.036$ throughout), decays monotonically with distance, and carries effect sizes 1.4 to 1.7 times the other-garage margin. The other-garage estimates are small and mostly insignificant; the marginal negatives at intermediate radii are consistent with onboarding coordinates occasionally misassigning entrants between adjacent central-Dakar garages, which leaks same-market competition into the other-garage category. Under either assignment, the crowding-out concentrates on the entrant's own market.

Adoption Within-Industry in Thiès. Within-industry adoption in the Thiès market—the only market outside Dakar where the taxi product was rolled out—responds positively through mutual ties at every threshold ($+1.10$ to $+5.16$, $p_{\text{RI}} \leq 0.003$)—and on the

mutual layer with no overlap floor at all: exposure through a treated mutual tie raises adoption by 18% of the base rate ($p_{\text{RI}} = 0.033$), see Table B5. This channel is also confirmed in the pre-treatment transaction network (Appendix Table B14). This corroborates the idea that competition leads to negative spillovers, and this force is not present among mutual contacts across cities.

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 conduit. This decomposition is made possible thanks to the access to the full universe of mobile money and payment transactions—an unusually rich data structure for studying payment technology diffusion.[17]

Customer Network. Table 5 provides 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

[17]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.

coefficient is large and significant ($\hat{\beta} \times 1{,}000 = +255.99$, $p_{\text{RI}} = 0.001$, $N = 8{,}523$), an effect size of $+44.8\%$ (column 1). Restricting to merchants without any community tie to their customer-linked drivers—column 3—the effect remains significant ($+174.54$, $p_{\text{RI}} = 0.001$, $+27.4\%$). 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 roughly quadruples ($+464.81$, $+108.2\%$). Part of this gap is exposure composition—community-tied merchants share customers with about three times as many drivers—but holding the number of customer-linked drivers fixed, community-tied merchants still respond $2.5\times$ more among the less exposed drivers (1–5 links), see Table B16. This is 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. As discussed earlier, customer demand and social learning therefore mostly operate on separate networks (see Table B7). Varying the shared-customer threshold preserves the pattern in both directions (Table B15): with any shared customer (≥ 1 transaction with each party) the within-taxi effect is $+82\%$ and cross-industry remains a precise null, and at ≥ 3 transactions the effect and its community-tie gradient persist in the much smaller sample, with positive but insignificant effect cross-industry.

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

Pre-Intervention Customer Network. Some of the measured overlap may itself be treatment-induced. Because not all control drivers had adopted and used the technology between November 2022 and November 2023 (by end of 2022, it had reached 88% of the control group), customers of treated drivers may mechanically be 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 estimate is therefore best viewed as high-powered but potentially biased. It is the estimate used in the structural section, because it is the only measure that directly observes business payments.

Table B17 constructs customer links from person-to-person transactions between January 2021 and March 2022, before the experiment began, ruling out treatment-induced link formation by construction.[18] In this network, exposure to treated taxis through shared customers has a positive within-taxi effect ($\hat{\beta} \times 1{,}000 = 36.35$, $p_{\text{RI}} = 0.136$, $N = 3{,}037$), corresponding to 4.6 percent of the dependent-variable mean; the estimate is stable across qualification thresholds of one to two fare transactions per customer.[19] The cross-industry coefficient is small and insignificant ($+14.79$, $p_{\text{RI}} = 0.273$, $N = 13{,}000$).

Two things explain the gap in magnitude. First, the business-payment network is a far more exhaustive measure of actual customer overlap than this P2P proxy, almost by definition of the technology designed to solve the p2p revocability issue. Second, customer turnover attenuates the predictive power of (pre-experiment) historical links.[20] The pre-period network is too sparse to certify the within-taxi effect on its own, but its sign pattern—positive within taxi, null across industries—matches the well-powered post-experiment network reported above: customer demand operates within the sector in which businesses actually share customers.

6 A Structural Model of Diffusion and Counterfactual Targeting

The reduced-form results above provide experimentally generated exposure-response moments for four network channels: social learning through phone contacts, customer demand, local competition, and the within-industry tie responses that separate learning from competition inside the taxi market itself. To study alternative allocations of the same onboarding effort, I embed these moments in a structural adoption model and evaluate counterfactual targeting rules at the realized matched business-level budget. The exer-

[18] 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—which limits statistical power, but the measure is free of the endogeneity concern that affects the richer post-experiment network.

[19] The customer regressions exclude garage-affiliated drivers, to avoid confounding customer demand with the local-competition effects documented in Section 5.2.

[20] Table B18 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 ($+282$ per 1,000) is roughly four times the W2 magnitude ($+65$), an attenuation consistent with customer turnover eroding the effect of shared customers over time.

cise is deliberately partial-equilibrium: it transports estimated exposure responses while holding the observed networks fixed.

6.1 Model

The logic of the exercise is simple. Each experimental table estimates how adoption responds to one kind of exposure, in one population, at the exposure levels the randomization happened to generate: a coefficient on the treated share of phone ties, a coefficient on treated customer links, a profile of coefficients on garage treatment at increasing distances. These are causal but local facts, each in its own units. A counterfactual assignment changes all exposures at once — a business selected under one rule stops transmitting under another — so evaluating it requires a single mapping from any exposure configuration to an adoption probability. The model supplies the minimal such mapping: the channels enter one latent index, and the experimental coefficients discipline how strongly each channel moves it.

In each month, a non-adopter i in auxiliary sample s adopts with per-period probability

$$p_{is}(A) = \Lambda\Big(a_s + \underbrace{\beta^{s}_{\text{learn}}\, S_i^{\text{strong}}(A) + \beta^{w}_{\text{learn}}\, S_i^{\text{weak}}(A)}_{\text{social learning}} + \underbrace{\big(\beta_{\text{demand}} + \beta_{DI}\,\mathbf{1}\{\text{embedded}_i\}\big)\, S_i^{\text{cust}}(A)}_{\text{customer demand}} + \underbrace{\beta_{\text{comp}}\,\mathbf{1}\{\text{gate}_i\}\,\mathbf{1}\{\text{AnyTreated}_{g(i)}\}\, e^{-\delta d_i}}_{\text{spatial competition}} + \underbrace{\big(\beta^{\text{taxi}}_{s} - c\,\mathbf{1}\{\text{Dakar}_i\}\big) S_i^{\mu\text{-taxi}}(A) \;-\; c\,\mathbf{1}\{\text{Dakar}_i\}\, S_i^{\nu\text{-taxi}}(A)}_{\text{within-taxi ties: learning and tie-borne competition}}\Big), \quad (3)$$

where Λ is the logistic function and A is a business-level treatment assignment. Adoption is absorbing. Cumulative adoption over the experimental window is

$$P_{is}(A) \;=\; 1 - \big(1 - p_{is}(A)\big)^{T}, \qquad T = 6. \quad (4)$$

The horizon is $T = 6$ monthly periods, from the first onboarding on March 28, 2022 to the adoption cutoff on October 5, 2022 (Section 1.2).

Each exposure is defined on its own auxiliary sample and enters the index as zero elsewhere:

- S_i^{strong}, S_i^{weak}: treated shares of i's phone ties in the strong ($\geq 5\%$ corrected overlap) and weaker community (2–5%) bands, with separate learning coefficients.

- S_i^{cust}: the treated share of i's pair-specific customer-linked experimental taxi businesses; a shared regular customer has ≥ 2 transactions with the specific driver and ≥ 2 with the specific merchant. Customer links that are also community phone ties are *socially embedded*; receivers with at least one embedded link ($\mathbf{1}\{\text{embedded}_i\}$) respond with the amplified coefficient $\beta_{\text{demand}} + \beta_{DI}$: demand transmitted through a relationship carries more than demand alone.
- spatial competition: the gated receiver-level garage exposure $\mathbf{1}\{\text{gate}_i\}\,\mathbf{1}\{\text{AnyTreated}_{g(i)}\}\,e^{-\delta d_i}$, where d_i is the receiver's exact distance to the garage and δ an estimated decay rate.
- within-taxi ties: $S_i^{\mu\text{-taxi}}$ and $S_i^{\nu\text{-taxi}}$ are the treated shares of i's *mutual* (reciprocally saved) and *non-mutual* community phone ties in the within-taxi reciprocity samples (Table B5; Appendix D.4). Strong (mutual) ties carry within-industry learning β_s^{taxi}; ties to a market-proximate treated driver also carry a single tie-borne competition cost c, common to both tie layers, which applies in the saturated Dakar market and not for receivers outside it, where the treated drivers do not operate. Non-mutual ties carry no learning ($\beta_w^{\text{taxi}} = 0$; the held-out outside-Dakar non-mutual cell is consistent with this once location-mislabeled adopters are removed).

Compactly, $P_{is}(A) = 1-(1-\Lambda(a_s+x_{is}(A)'\psi))^T$, where $x_{is}(A)$ stacks the exposures defined for sample s. The common channel parameters are $\psi = (\beta^s_{\text{learn}}, \beta^w_{\text{learn}}, \beta_{\text{demand}}, \beta_{DI}, \beta_{\text{comp}}, \delta, \beta_s^{\text{taxi}}, c)$. I interpret these as *structural exposure-response parameters* disciplined by experimentally induced exposure variation. The model is structural in the sense that it imposes a common nonlinear mapping from network exposure to adoption and uses that mapping to evaluate counterfactual assignments. The coefficients summarize causal exposure responses; reading them as deep primitives of preferences, technology, or network formation would require a fully specified forward-looking adoption problem, which I leave outside the scope of this paper.

Each auxiliary sample s has its own intercept a_s, which sets that sample's baseline adoption level: for each candidate ψ, a_s is chosen so that the model reproduces the sample's observed mean adoption rate. A diagnostic restriction that imposed one common taxi-side intercept across the demand and garage samples produced severe lack of fit; the preferred specification therefore asks the model to explain exposure *responses* while treating cross-sample differences in unconditional adoption levels as nuisance heterogeneity. In the policy exercise below, these estimation baselines are discarded and replaced by policy-risk-set baselines calibrated under the realized assignment (Appendix D).

6.2 Estimation and inference

Let $\hat{m}$ stack the fourteen experimental exposure-response coefficients: the adoption response to the treated share of strong ties ($\geq 5\%$ corrected overlap), the response to the treated share of community-band ties (2–5%), the two customer-demand responses (the no-community sample and the full customer-linked sample), the eight garage coefficients at radii from 50 m to 1 km, and the two within-taxi mutual-tie reciprocity cells described below. For a candidate ψ, let $m(\psi)$ be the same fourteen coefficients computed on the model's adoption probabilities instead of observed adoption — the same receiver samples, the same regressors, the same trimming. The estimator is minimum distance:

$$\hat{\psi} \;=\; \arg\min_{\psi} \big(m(\psi) - \hat{m}\big)' \, \widehat{W} \, \big(m(\psi) - \hat{m}\big), \tag{5}$$

The moments and parameters pair into four blocks — the phone bands, the customer coefficients, the within-taxi cells, and the garage profile — and no parameter enters more than one block. With each sample's baseline profiled within its own block and a diagonal weight, the criterion separates: minimizing it jointly and block by block give numerically identical estimates. In the first three blocks the parameters are exactly identified, and the estimator is a change of units: an auxiliary coefficient equals its structural parameter scaled by the sample's average marginal effect, which the profiled baseline pins to the data. Thus, estimation converts coefficients measured in incompatible probability units onto a common latent-index scale. Only the garage block is asked to fit more moments than it has parameters. Estimation is joint in the full-covariance weighting, which couples the blocks through the cross-moment design covariance, and in the randomization draws, whose common placebo index delivers the joint distribution of the parameters used by the policy counterfactuals.

The estimator targets fourteen auxiliary exposure-response moments, selected to discipline the eight common channel parameters: the strong-band ($\geq 5\%$) phone coefficient, the weaker-community-band (2–5%) phone coefficient, the pair-specific no-community customer-demand coefficient, the full pair-specific customer coefficient (whose excess over the no-community response identifies the embedded-link amplification β_{DI}), the eight Panel-B garage coefficients at 50, 75, 100, 150, 300, 500, 750, and 1,000 m, and two within-taxi reciprocity cells: the mutual within-taxi responses in Dakar and outside it, estimated under the strict trimming of Table 1 (Table B5 reports the mutual layer; Appendix D.4 gives the estimation cells). Each is a causal coefficient validated by randomization inference, and each corresponds to an exposure that a counterfactual assignment would move. I keep the two community-tie bands separate because pooling them would

impose a single per-tie effect on stronger and weaker community ties — precisely the tie-strength question at stake. The 2% community threshold is pre-specified and survey-validated, and lies comfortably above the scale of expected hash-collision noise. The eight radii are what identify the decay: a single garage coefficient would pin down only the product of the competition level and its distance profile, and the two outermost bins — individually significant at 750 m and 1 km — force the profile to fit the measured tail rather than extrapolate the steep near-field decline. The two mutual reciprocity cells identify the within-taxi parameters; the non-mutual cells are not targeted, and the outside-Dakar non-mutual cell serves as a held-out test of the $\beta_w^{\text{taxi}} = 0$ restriction. The survey moments are excluded from estimation and held out as validation (Appendix D.10), as is the embedded-link customer subsample coefficient, which the model must reproduce from $\beta_{\text{demand}} + \beta_{DI}$ without targeting it. Because the non-garage blocks are exactly identified, these held-out moments, together with the outside-Dakar non-mutual cell, are the model's only testable content outside the garage profile. Other reduced-form moments serve as validation, robustness, or held-out evidence rather than entering the objective.

The identifying trick is that the model is never asked to fit the raw data; it is asked to reproduce the experiment's own regression coefficients (Gouriéroux et al., 1993). Concretely, estimation is a loop over candidate parameter vectors. Given a candidate ψ, I first solve for each intercept a_s so that the model matches sample s's mean adoption rate. I then compute every receiver's exposures under the actual treatment assignment, evaluate the implied adoption probabilities P_{is}, and run on these probabilities exactly the regressions that produced the experimental coefficients — same receivers, same regressors, same trimming. This delivers $m(\psi)$, and the outer minimization of (5) does the rest. Because the intercepts absorb levels, the channel parameters are identified only by how adoption responds to experimentally assigned exposure: the two phone moments identify the learning responses, the two demand moments identify β_{demand} and β_{DI}, the eight garage coefficients jointly identify the level and decay of spatial competition, and the two mutual reciprocity cells identify the within-taxi parameters — the outside-Dakar mutual cell pins down β_s^{taxi} where no treated driver competes, and the Dakar mutual cell then reveals c as the cost that offsets it. With fourteen moments and eight parameters, the model is overidentified by six restrictions, all of them in the garage profile, where one level and one decay rate must fit eight coefficients. The model is exactly identified where it translates and overidentified where it tests; the minimized criterion measures the fit of the garage profile, not of the model globally. What the model adds to the reduced form, and where the model risk sits, is the nonlinear index: it converts coefficients measured at the exposure levels the experiment generated into predictions at the exposure levels

a policy would generate. The primary criterion uses diagonal inverse design variances computed from the stored bank of 1,000 business-level placebo assignments; the full covariance, including the strong cross-radius correlation of the garage moments, is used in weighting robustness exercises. Estimation details, moment accounting, and diagnostics are in Appendix D.

The point estimates are $\hat{\beta}^{s}_{\text{learn}} = 0.320$, $\hat{\beta}^{w}_{\text{learn}} = 0.433$, $\hat{\beta}_{\text{demand}} = 0.555$ with embedded-link amplification $\hat{\beta}_{DI} = 1.44$ $[0.20,\ 2.78]$, $\hat{\beta}_{\text{comp}} = -2.483$ with decay $\hat{\delta}$ implying a spatial half-life of competitive exposure of 448 m, and within-taxi parameters $\hat{\beta}^{\text{taxi}}_{s} = 0.311$ and $\hat{c} = 0.30$. Table 7 reports the estimates with their RI intervals.

The distance decay merits discussion, because it is the parameter that turns the garage experiment into a statement about the spatial reach of competition. The garage design measures the competition response at eight nested radii, from 50 m to 1 km, with the two outermost coefficients individually significant. The model requires these eight coefficients to lie on a single exponential profile $\beta_{\text{comp}}\, e^{-\delta d}$, evaluated at each receiver's exact distance to the treated garage: the *ratios* of the coefficients across radii identify δ irrespective of the population's adoption level, while the level identifies β_{comp}. The estimated half-life of 448 m means the competitive force halves roughly every four city blocks. Targeting the measured 750 m and 1 km bins matters: a profile fit to the six inner radii alone extrapolates a steeper decay (half-life 339 m) and understates the reach of competition that the outer bins directly measure. At the point estimates the competitive force falls below the strong-tie learning response beyond roughly 1.3 km from a treated garage; competition is a neighborhood force while learning travels city-wide. The scale matches the most local estimates of urban interaction effects: networking benefits among Manhattan advertising agencies dissipate within roughly 500 m (Arzaghi and Henderson, 2008), production externalities in Berlin halve at a similar distance (Ahlfeldt et al., 2015), and the closest comparable design — randomized diffusion of energy-efficient motors among Bangladeshi manufacturers — finds within-industry spillovers operating through exposure within a 500 m walk (Chaurey et al., 2025), at the same spatial grain but with the opposite sign: energy savings are non-rivalrous in the local product market, while a payment-acceptance advantage is contested passenger by passenger. Appendix D.3 reports the profile fit, a transparent two-step diagnostic of the decay, and its conversion into data units. For inference, I use the stored bank of 1,000 business-level placebo assignments. For draw b I form $m^{*}_{b} = m_{\text{obs}} + (m^{RI}_{b} - \bar{m}^{RI})$ and re-estimate the minimum-distance problem, re-profiling the nuisance baselines while holding the estimator and weight matrix fixed; the re-estimation uses a first-order update around the point estimate, validated against full nonlinear re-estimation on 150 draws. The resulting intervals are RI empirical intervals:

the stored bank re-executes the experiment's stratified assignment mechanism, so the moment perturbations are draws from the randomization distribution itself, pushed through the estimator. They propagate parameter uncertainty conditional on the maintained networks, samples, and model; policy-assignment and specification uncertainty are reported separately. Across all re-estimates the demand response is positive and the competition response negative in every draw; the competition half-life lies in $[266, 820]$ m. The ranking between the strong-band and weak-band learning responses is *not* identified: it flips across draws and specifications, and I leave it uninterpreted.

6.3 Recovering learning within the taxi industry

The reduced form leaves a puzzle: within-taxi adoption responds *negatively* to treated community ties in Dakar (Table 2), yet survey-reported taxi *friends* of treated drivers adopt strongly (effect size $+36.1\%$, Table 3). The within-taxi block of the model resolves it with one economic distinction: learning requires tie strength, competition requires market proximity. Four within-taxi reciprocity cells discipline this claim — the response on mutual and non-mutual ties, in Dakar and outside it (Table B5 reports the mutual cells; Appendix D.4 all four). Outside Dakar no treated driver competes, so the outside-Dakar mutual cell measures within-industry learning through strong ties directly: $\hat{\beta}_s^{\text{taxi}} = 0.311$ $[0.108,\ 0.502]$, statistically indistinguishable from the cross-industry strong-tie learning response $\hat{\beta}^s_{\text{learn}} = 0.320$. Learning transmits through a strong tie at the same rate whether or not the receiver competes with the sender. The Dakar mutual cell is null in the reduced form because the proximity cost almost exactly offsets it, and the Dakar non-mutual cell is negative (-0.81 per 1,000 [.017]) because weak ties carry the cost without the learning ($\beta_w^{\text{taxi}} = 0$). The cost is one parameter, $\hat{c} = 0.30$, estimated from the mutual cells and applied to both tie layers: estimated layer by layer on the full cell set the two costs are indistinguishable (0.303 vs 0.321); the large difference between the two Dakar reduced-form coefficients reflects baseline rates and degree composition, not different competition.

Three held-out checks discipline the block. First, the model predicts exactly zero for the fourth cell (non-mutual ties outside Dakar: no learning, no competitor); the untrimmed estimate is a null (-0.05 per 1,000, $p_{RI} = .45$), and the small negative that appears under the estimation trimming (-0.147 [.015]) is one I do not overinterpret.[21]

[21] The trimmed negative is fragile: significant only at the $\geq 5\%$ threshold (-0.05 [.14] at $\geq 2\%$, -0.13 [.44] at $\geq 10\%$), absent at every threshold on the untrimmed layer, and it vanishes once adopters are located by their observed transactions rather than their registration region (-0.003 [$p_{RI} = .83$]; 91 percent of "outside-Dakar" taxi adopters in fact operate in the Dakar market, Appendix D.4). A small residual negative would in any case not contradict the model: Thiès is some 60 km from Dakar and taxis on intercity routes compete with Dakar drivers, so the Dakar-only competition dummy is an approximation.

Second, the survey friends moment: reported taxi friends are strong ties (they sit at the top of the validated overlap distribution) that rarely share customers or garages, so the model prices them at β_s^{taxi} with no proximity cost — a predicted effect size of +30.0% $[+12.3,\ +47.4]$ against the observed $+36.1\%$ $[+9.1,\ +63.1]$, and a predicted null for non-friends (observed $+2.8\%$, insignificant). A moment estimated entirely from administrative cells reproduces the survey coefficient out of sample. Third, misspecified alternatives fail: imposing the cost equality directly on the reduced-form coefficients (rather than in index units, where it holds) predicts the fourth cell at $+1.41$, and a cost proportional to tie strength implies negative learning on weak ties.

The economics that survives is sharp. Within the taxi industry, ordinary community ties transmit net competition; learning flows only through relationships that are either socially close (mutual contacts, friends) or economically embedded (shared customers, Appendix D.10). The within-taxi sign pattern in the reduced form is the fingerprint of learning and competition traveling on the same ties and canceling in a saturated market.

6.4 Counterfactual targeting policies

I hold the program budget at its realized size — exactly 856 of the 1,454 matched taxi businesses — and compare targeting rules. A business is the indivisible policy unit: selecting it activates all of its matched stakeholder records. Each policy is evaluated on the estimation-aligned receiver populations only (3,024,017 unique receivers after deduplication), with group baselines re-calibrated so the model reproduces observed adoption rates under the realized assignment. For aggregate policy value each receiver is counted once: cross-industry adoption is defined as adoption in either geographic margin, and the Dakar and outside-Dakar outcomes are reported separately as descriptive margins. The benchmark is the *exact-budget random business targeting benchmark*: 2,000 uniform draws of 856 businesses, each evaluated exactly. The budget is held fixed at exactly 856 businesses so that rules are compared at identical scale; the experimental mechanism leaves the realized budget random, so this is a random-targeting benchmark rather than a replay of the experimental randomization. I report induced adoption — adopters under the policy minus adopters under a no-treatment counterfactual — for six targeting rules:

- *Phone reach (degree).* Treat the 856 businesses with the most community ties: the number of distinct individuals connected to the business in the phone network (corrected overlap $\geq 2\%$), each counted once. This rule requires knowing each business's ties, and nothing else.
- *Greedy coverage.* Select businesses one at a time, each time adding the one whose ties

add the most individuals not already reached by the selected set. Two well-connected businesses may share many of their ties: the degree rule counts shared ties twice, this rule counts them once.

- *Top mobile-money users.* Treat the 856 businesses with the highest mobile-money activity before the experiment (November 2017–February 2022). The provider can implement this rule from its own records, with no network data.
- *Garage-concentrated.* Treat all garage-member businesses and fill the remaining slots at random, averaging over 200 random fills. This is the most natural field strategy: recruit where taxis congregate.
- *Customer reach.* Treat the 856 businesses with the most linked customer merchants in the measured transaction network. This rule targets the demand channel the way phone reach targets learning.
- *Phone + customer blend.* Rank businesses by the average of their phone-reach and customer-reach ranks and treat the top 856.

A geographic-reach rule cannot be scored without verified receiver geography and is excluded rather than proxied.

Table 8 reports the results. Four findings stand out. First, targeting by phone reach induces 8,839 adopters — $+56.9\%$ relative to the mean of exact-budget random business assignments, above every draw in the random bank, with a RI parameter interval of $[+32.9\%, +78.2\%]$ and a positive gain in 0.99 of parameter draws. The greedy coverage rule reaches only 0.24% more unique receivers than the simple reach ranking and does not improve predicted adoption (8,692 versus 8,839 induced adopters). Thus almost all of the attainable coverage advantage is captured by the simple business-level reach score; additional overlap management adds little at this budget. Second, a simple observable heuristic — targeting the top pre-treatment mobile-money users — captures roughly two-fifths of the reach-targeting gain ($+23.3\%$). Third, concentrating treatment within garages *underperforms* random assignment (-3.7%, with an estimated gain below zero in 86 percent of parameter draws). The decomposition attributes most of this shortfall to stronger local competitive displacement, with the remainder coming from weaker phone-network reach; garage concentration therefore performs poorly for two distinct reasons rather than competition alone. Fourth, the customer-demand channel is a *level* channel rather than a *targeting* channel: it induces roughly 1,500 adopters among the customer-linked merchants under essentially any placement of the budget — the linked merchants connect to many businesses, so their exposure barely depends on which businesses are

treated. Its adoption gains therefore accrue under random assignment as much as under reach targeting, and targeting customer reach directly earns only $+2.1\%$ over random. One consequence is that the within-taxi margin is *positive* under every rule (+497 induced under the realized assignment): customer demand and mutual-tie learning together outweigh garage displacement and tie-borne competition on the taxi side. Figure 5 shows the full distribution of each policy's gain across the 1,000 placebo draws.

Learning contributes essentially all of the gain: decomposing the reach-targeting gain across channels (with policy baselines recalibrated under the realized assignment for each channel configuration; Appendix D.8), the learning channel alone reproduces the full gain, turning it off eliminates the gain, and the demand, within-taxi tie, and spatial-competition contributions are each within a few percent of zero — the first two because their exposures barely differ across placements, the last slightly positive because reach targeting disperses treatment away from garages. The gain is essentially unchanged when β_{demand} is set to zero. The conclusions are unchanged under the alternative (business-collapsed) phone specification and under a coarser policy-baseline stratification.

The measured customer network is a lower bound on the true one: links are observed only when a customer pays both the driver and the merchant digitally within the transaction window, so cash and infrequent customers are invisible. To assess what undermeasurement means for policy, I re-evaluate the rules in a world where the true network is *double* the measured one — unobserved linked merchants responding like observed ones, with the random benchmark rescaled identically. Three things happen. First, total program impact rises by about a quarter under every placement (the random benchmark moves from 5,633 to 7,017 induced adopters; reach targeting from 8,839 to 10,129): most of the missing network's contribution arrives as a level, whoever is treated. Second, customer-reach targeting becomes modestly useful ($+13\%$ over random, against $+2.1\%$ at measured density) — doubling the network begins to de-saturate the demand channel, so placement starts to move exposure — but it remains far behind phone reach ($+44\%$). Third, the blended phone-plus-customer rule is the best rule in the doubled world ($+56\%$), as it is at measured density ($+58.8\%$): a planner uncertain about how much of the customer network the transaction data capture loses essentially nothing by ranking on the blend. Learning's share of total induced adoption falls from 87 to 70 percent under the random benchmark, but its contribution to the *targeting gain* is unchanged: reach targeting buys the same $\approx 3{,}280$ learning-channel adopters at either density.

Three caveats bound these results. First, the static model holds networks and channel parameters fixed under counterfactuals, excluding adoption feedback and general-equilibrium responses. Second, the learning response is identified from the Dakar cross-

industry margin. Receivers outside Dakar enter the counterfactuals with the same coefficients and their own calibrated baseline: I assume a treated tie moves adoption outside Dakar as it does in Dakar. The outside-Dakar reduced form is consistent with this assumption, but the data are too thin to estimate a separate outside-Dakar learning parameter, so predictions for that margin rest on Dakar variation. Third, reach-targeted policies expose receivers more intensely than the experiment did: under the realized assignment, 27 percent of receivers had every strong-band tie treated; under reach targeting, 48 percent do. Fully exposed receivers exist in the estimation sample, so the model is never evaluated at exposure levels it did not see, but the policies place twice as much weight there, and the predictions rely on the fitted response at high exposure. The demand channel is additionally conditional on the measured customer network: it operates only through the 8,523 merchants whose customers link them to experimental businesses in the transaction data. Merchants without such links have zero demand exposure under every assignment, so their exclusion cannot affect the comparison of policies; it does bound the *level* of the demand contribution, which is why the density counterfactuals above treat the measured network as a floor. The measured links are also held fixed across assignments, although some may exist because a driver was treated. Both concerns are confined to the taxi-side margin: the reach-targeting headline is essentially unchanged when the demand channel is set to zero. These qualifications imply that the counterfactuals are model-based projections with RI parameter intervals; experimental policy effects would require fielding the alternative rules.

6.5 Policy Implications

Three lessons emerge for a planner maximizing adoption from a fixed onboarding budget (Verhoogen, 2023).

Network targeting delivers large gains. Relative to the mean of exact-budget random business assignments, targeting the 856 businesses with the greatest measured phone-network reach raises predicted induced adoption by $+56.9$ percent, and the rule is implementable from the phone fingerprint data already collected for the overlap computation. The simple reach ranking captures essentially all of the gain available to the coverage-maximizing greedy rule. The gains come almost entirely from the cross-industry learning channel: learning contributes essentially the entire targeting gain, while the demand and within-taxi tie channels raise adoption under every placement and therefore contribute levels rather than gains. The gains are also geographically dispersed: over half of taxi businesses' community contacts live outside Dakar (Stylized Fact 1, Section 2), so the

same high-reach businesses carry both geographic margins. Reach targeting raises induced adoption by about two-thirds relative to the benchmark in the Dakar and outside-Dakar cross-industry margins alike.

Simple heuristics capture part of the gain. Targeting the heaviest pre-treatment mobile-money users — a rule requiring only the provider's own administrative records — raises predicted induced adoption by $+23.3$ percent relative to the same benchmark, roughly two-fifths of the reach-targeting gain. Because the model excludes the save and transaction ties on which heavy users are central, this gain is likely a lower bound. And if the measured customer network understates the true one, the blended phone-plus-customer ranking is the robust rule: it matches reach targeting at measured density and dominates it at higher densities.

Geographic concentration is costly. Concentrating treatment at garages performs about 4 percent worse than random assignment: garage businesses' network reach is below average, and maximal spatial concentration activates the competition channel, whose 448 m half-life makes the penalty severe in the immediate garage area. A planner who targets garages because they are visible maximizes crowding while leaving the learning channel largely unused.

7 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—with responses that the data do not distinguish across the strength distribution, and not incrementally through mutual contacts—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 448 m), partially offsetting learning gains among same-sector firms. Customer-demand complementarities—transmitted through shared customers, and strongest when the customer link is also a social tie—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 through community 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, relative to the mean of exact-budget random business assignments, targeting the businesses with the greatest measured phone-network reach raises predicted induced adoption by about 57 percent, 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 62-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. 2003. "Friends and Neighbors on the Web." *Social Networks*, 25(3): 211–230.

Ahlfeldt, Gabriel M., et al. 2015. "The Economics of Density: Evidence from the Berlin Wall." *Econometrica*, 83(6): 2127–2189.

Alvarez, Fernando, et al. 2025. "Strategic Complementarities in a Dynamic Model of Technology Adoption: P2P Digital Payments." *Working Paper*.

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

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

Arzaghi, Mohammad, and J. Vernon Henderson. 2008. "Networking off Madison Avenue." *Review of Economic Studies*, 75(4): 1011–1038.

Athey, Susan, et al. 2018. "Exact p-Values for Network Interference." *Journal of the American Statistical Association*, 113(521): 230–240.

Atkin, David, et al. 2024. "The Returns to Face-to-Face Interactions: Knowledge Spillovers in Silicon Valley." *Working Paper*.

Baird, Sarah, et al. 2018. "Optimal Design of Experiments in the Presence of Interference." *Review of Economics and Statistics*, 100(5): 844–860.

Banerjee, Abhijit, et al. 2013. "The Diffusion of Microfinance." *Science*, 341(6144): 1236498.

Baum-Snow, Nathaniel, et al. 2024. "Local Productivity Spillovers." *American Economic Review*, 114(4): 1030–69.

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

Bernhardt, Arielle, et al. 2019. "Household Matters: Revisiting the Returns to Capital among Female Microentrepreneurs." *American Economic Review: Insights*, 1(2): 141–60.

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

Bloom, Nicholas, et al. 2013. "Identifying Technology Spillovers and Product Market Rivalry." *Econometrica*, 81(4): 1347–1393.

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

Breza, Emily, et al. 2019. "Networks in Economic Development." *Oxford Review of Economic Policy*, 35(4): 678–721.

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

Broder, Andrei Z. 1997. "On the Resemblance and Containment of Documents." 21–29. IEEE.

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

Cai, Jing, et al. 2015. "Social Networks and the Decision to Insure." *American Economic Journal: Applied Economics*, 7(2): 81–108.

Cefalà, Luisa, et al. 2025. "The Economic Consequences of Knowledge Hoarding." *Working Paper*.

Chaurey, Ritam, et al. 2025. "Social Learning among Urban Manufacturing Firms: Energy-Efficient Motors in Bangladesh." National Bureau of Economic Research Working Paper 34296.

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

Crépon, Bruno, et al. 2013. "Do Labor Market Policies Have Displacement Effects? Evidence from a Clustered Randomized Experiment." *Quarterly Journal of Economics*, 128(2): 531–580.

Crouzet, Nicolas, et al. 2023. "Shocks and Technology Adoption: Evidence from Electronic Payment Systems." *Journal of Political Economy*.

Crouzet, Nicolas, et al. 2024. "Demographics and Technology Diffusion: Evidence from Mobile Payments."

De Cristofaro, Emiliano, et al. 2013. "Private Discovery of Common Social Contacts." *International Journal of Information Security*, 12(1): 49–65.

Demmler, Daniel, et al. 2018. "PIR-PSI: Scaling Private Contact Discovery." *Proceedings on Privacy Enhancing Technologies*, 2018(4): 159–178.

Duflo, Esther, et al. 2026. "Chat Over Coffee? Technology Diffusion Through Social and Geographic Networks in Rwanda." 31368.

Fafchamps, Marcel, et al. 2023. "Adoption with Social Learning and Network Externalities." *Oxford Bulletin of Economics and Statistics*. Open Access.

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

Gouriéroux, Christian, et al. 1993. "Indirect Inference." *Journal of Applied Econometrics*, 8: S85–S118.

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

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

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

Houeix, Deivy. 2026. "Asymmetric Information and Digital Technology Adoption: Evidence from Senegal." *Working Paper*.

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

Jahani, Eaman, et al. 2023. "Long ties, disruptive life events, and economic prosperity." *Proceedings of the National Academy of Sciences*, 120(28): e2211062120.

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

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

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

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

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

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

Noll, Landon Curt, et al. 2026. "The FNV Non-Cryptographic Hash Algorithm." RFC Editor RFC 9923.

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

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

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

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

Swamidass, S. Joshua, and Pierre Baldi. 2007. "Mathematical Correction for Fingerprint Similarity Measures to Improve Chemical Retrieval." *Journal of Chemical Information and Modeling*, 47(3): 952–964.

Swamidass, S. Joshua, et al. 2015. "Securely Measuring the Overlap between Private Datasets with Cryptosets." *PLoS ONE*, 10(2): e0117898.

Verhoogen, Eric. 2023. "Firm-Level Upgrading in Developing Countries." *Journal of Economic Literature*, 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.

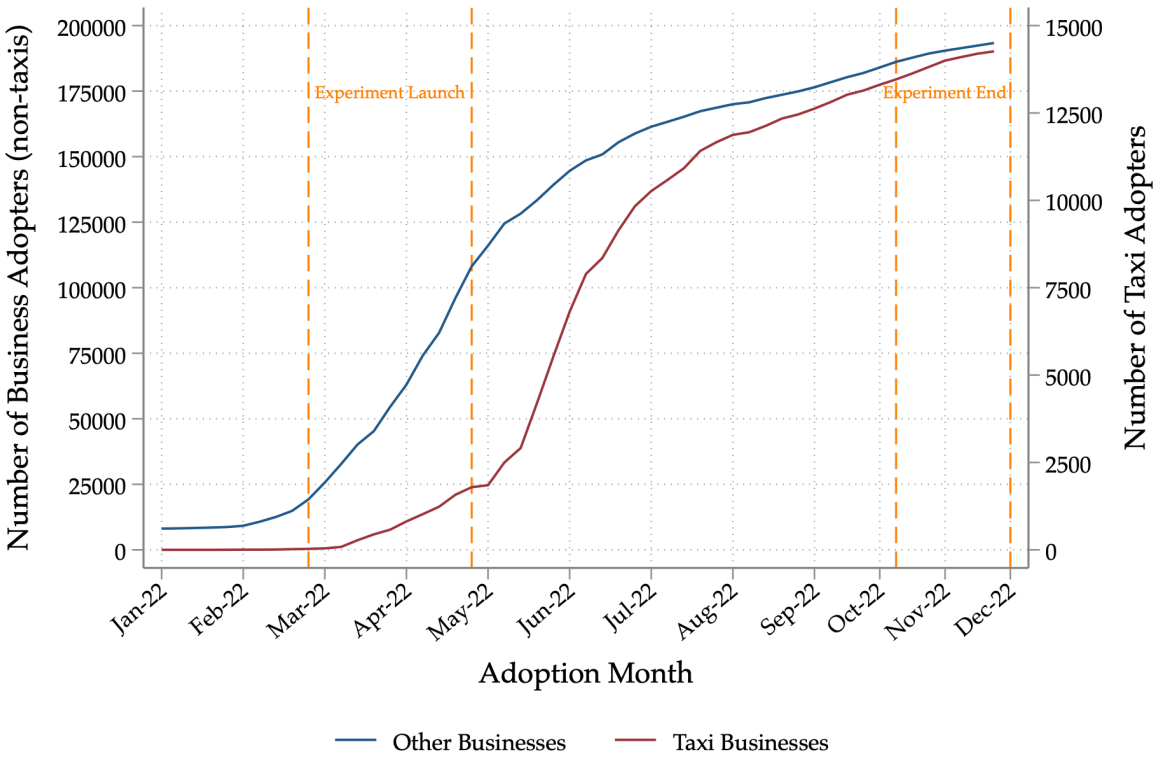


Figure 2: Digital Payments' Adoption in 2022 - *Author's calculations*

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.

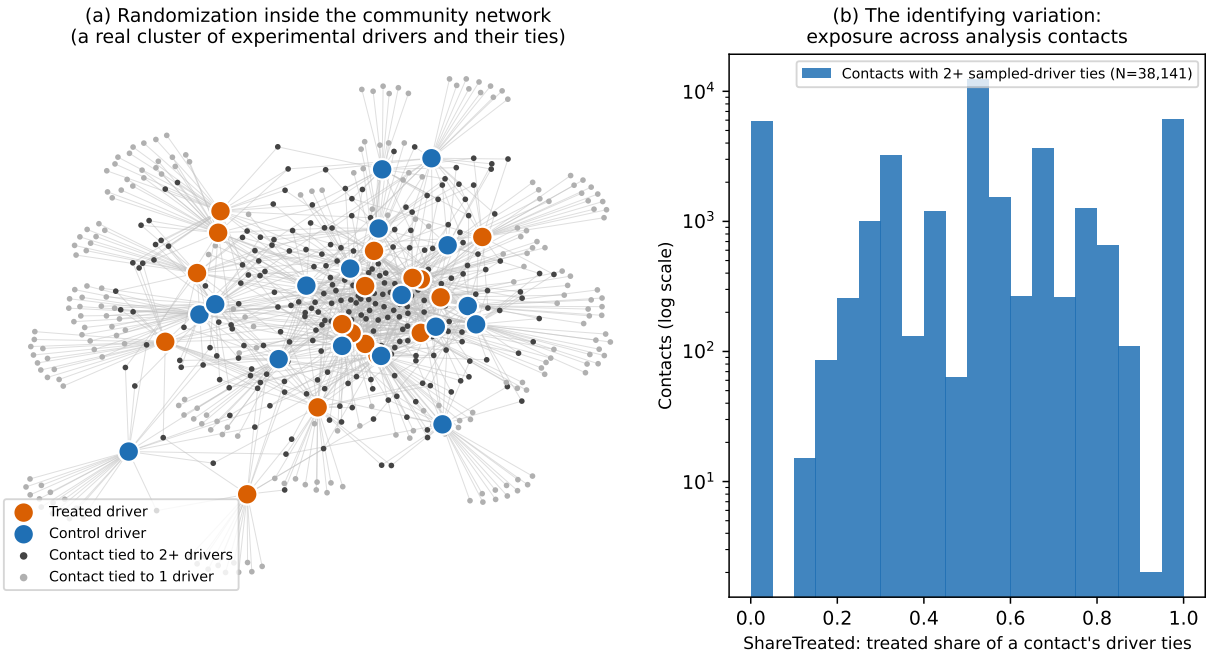


Figure 3: The Experimental Design in the Community Network

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.

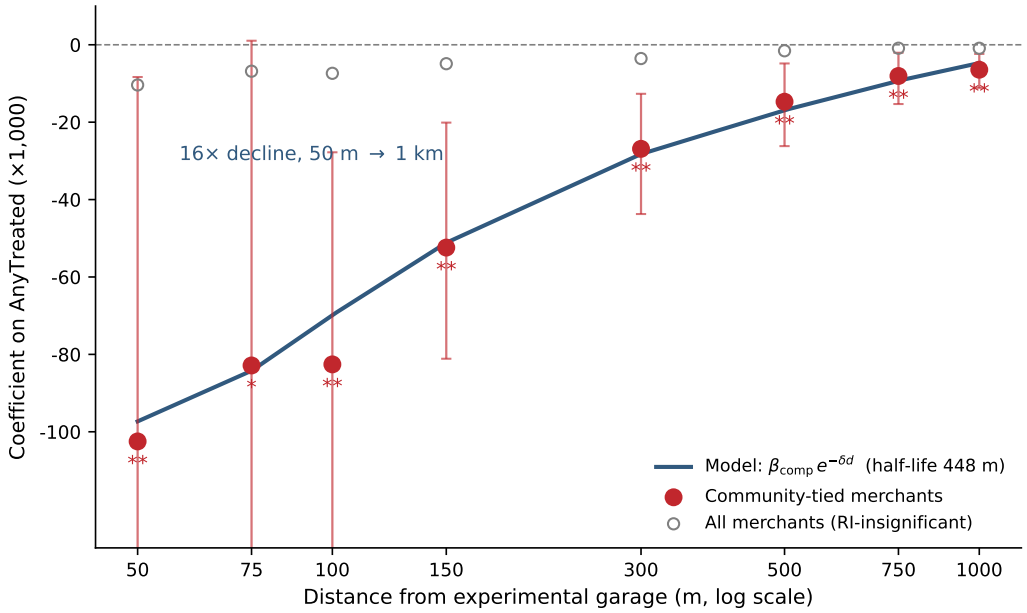


Figure 4: Garage proximity spillovers: coefficient by distance

Notes: From Table 4: Panel B (red, 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), at all eight radii from 50 m to 1 km. The vertical axis is the coefficient on AnyTreated ($\times 1{,}000$); coefficient units avoid conflating the decay of the force with the spatial gradient of baseline adoption, which itself falls fivefold across the radii (effect sizes in the definition of Table 4 range from -91% at 100 m to -70% at 1 km in Panel B). 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; lower bounds at the innermost radii are truncated for display. Bars and stars are 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 solid line is the model-implied coefficient profile at the structural estimates, $\beta_{\text{comp}}\, e^{-\delta d}$ mapped through the auxiliary regression (half-life 448 m); all eight Panel B coefficients are targeted moments of the estimator, and the two outermost bins discipline the tail of the decay (Appendix D.3). The competition effect concentrates on the socially exposed and is an order of magnitude smaller and RI-null among all merchants—consistent with competitive crowding transmitted through market-level community exposure rather than through close relationships (mutual-contact merchants show no robust effect; Table 4, Panel C).

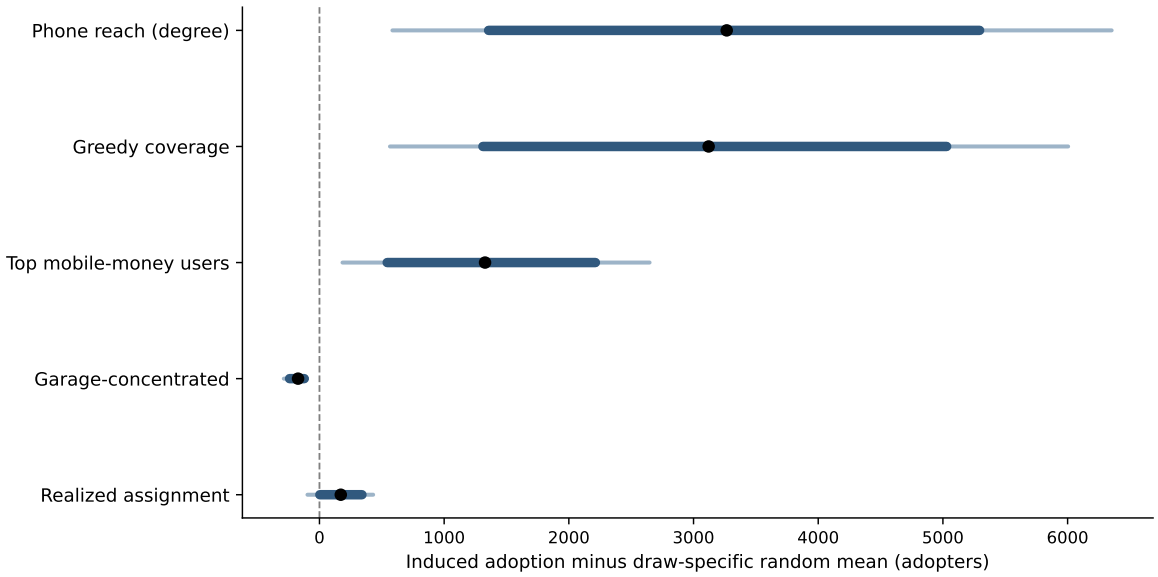


Figure 5: Policy gains under parameter uncertainty. Distribution, across the 991 design-respecting parameter re-estimates, of each policy's induced adoption minus the draw-specific mean of exact-budget random business assignments; dots mark medians, thick and thin bars the 80% and 95% intervals. Assignment variation is not included in these distributions and is reported separately in Table 8.

Table 1: Cross-Industry Network Spillovers by Tie Strength

	$\geq 2\%$ (community)	$\geq 5\%$ (strong)	$\geq 10\%$
Cross-industry, Dakar			
Share treated	$+2.91$ {+41%}	$+4.16^{**}$ {+39%}	$+5.16^{**}$ {+69%}
	(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	$+0.68^{**}$ {+138%}	$+0.68^{***}$ {+81%}	$+0.73^{**}$ {+137%}
	(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 in braces, computed as $\beta_1/\hat{y}(S{=}0)$, where $\hat{y}(S{=}0) = \bar{y} - \hat{\beta}_1 \bar{S}$ is predicted adoption at zero treated exposure with controls at their means ($\bar{S} = 0.590, 0.666, 0.628$ at the three thresholds); 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 B5; results excluding shared-customer pairs are in Appendix Table B6, and additionally excluding garage-exposed drivers 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	$+0.08$ {+14%}	-0.98^{***} {-30%}	-0.56 {-10%}
	(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]
Dep. var. mean	0.58	2.62	5.26
N	1,145,899	1,020,391	376,871

Notes: Same specification, samples, community-tie definitions, and trimming as Table 1; the outcome is registration as a *taxi* merchant in 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 $+14\%$ 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, -30%, $p_{\text{RI}} = 0.007$)—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. In Thiès (the only market outside Dakar where the taxi product was rolled out) the community-layer effects are null throughout (not shown); the positive cross-market within-taxi channel operates through the reciprocal-save (mutual) layer and the transaction network (Appendix Tables B5 and B14), 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	98.45***	0.362			614
	(37.45)				
	[0.007]				
Taxi × Non-friends	8.75	0.028			549
	(39.39)				
	[0.803]				
Taxi × Freq. daily/weekly	69.54**	0.240			916
	(30.42)				
	[0.023]				
Panel B: All contacts by social proximity					
Friends	30.44**	0.316	0.86	0.013	1,983
	(14.12)		(11.18)		
	[0.025]		[0.920]		
Freq. daily/weekly	15.27*	0.213	3.23	0.065	4,119
	(8.47)		(6.80)		
	[0.055]		[0.630]		
Panel C: Casual ties					
Clients	4.59	0.369	35.15	0.944	395
	(12.25)		(22.75)		
	[0.716]		[0.134]		
Serendipitous encounters	2.87	0.165	18.00***	0.475	5,112
	(3.84)		(5.90)		
	[0.457]		[0.004]		

Notes: Each row is a *separate* regression: $y_j = \alpha + \beta_1 \cdot \text{NumTreated}_{s,j} + \beta_2 \cdot \text{ExpectedExposure}_{s,j} + \varepsilon_j$, where exposure is computed over the indicated relationship-defined subnetwork s only, on the subsample with ≥ 1 connection in s, restricted to baseline survey contacts. "Friends" and "Non-friends" partition the surveyed taxi ties, so the two rows compare transmission through the two halves of an exclusive split; an individual connected through both tie types (to different drivers) may contribute to both rows, and the frequency rows re-slice the same ties along a different dimension. Coefficients $\beta_1 \times 1{,}000$ reported; HC1 robust standard errors in parentheses. Effect size $= \beta_1/\hat{y}(0)$ with $\hat{y}(0)$ predicted adoption at zero treated contacts (controls at their means). "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	750 m	1,000 m
Panel A: All merchants near garages								
Any treated	-0.0104	-0.0068	-0.0074	-0.0049	-0.0036^{*}	-0.0016	-0.0009	-0.0009
p_{RI}	[0.107]	[0.368]	[0.194]	[0.179]	[0.085]	[0.277]	[0.349]	[0.211]
Effect size	-62%	-55%	-63%	-60%	-58%	-39%	-33%	-41%
N	48,691	64,217	119,250	219,530	588,403	1,247,597	2,347,566	3,773,560
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^{**}	-0.0081^{**}	-0.0064^{**}
p_{RI}	[0.034]	[0.052]	[0.031]	[0.021]	[0.017]	[0.023]	[0.026]	[0.016]
Effect size	-90%	-89%	-91%	-89%	-84%	-76%	-68%	-70%
N	6,027	7,987	15,313	25,254	56,795	122,865	219,857	352,044
Panel C: Mutual phone contacts only								
Any treated	$+0.0283$	$+0.0141$	-0.0655	-0.0171	-0.0639	-0.0484	-0.0421	-0.0482^{*}
p_{RI}	[0.678]	[0.811]	[0.489]	[0.775]	[0.219]	[0.226]	[0.181]	[0.064]
Effect size	$+20\%$	$+10\%$	-31%	-12%	-36%	-35%	-39%	-49%
N	456	483	613	809	1,101	1,704	2,443	3,222

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 B10). 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 coefficients are an order of magnitude larger on community-exposed merchants and is absent among close contacts, indicating transmission through market-level social exposure rather than personal relationships. Effect size $= \beta_1/\hat{y}(\text{AnyTreated}{=}0)$, predicted adoption with no treated garage (controls at their means; mean AnyTreated 0.83–0.95 across samples). RI p-values (1,000 permutations) in brackets, clustered assignment at the garage level. All eight Panel-B coefficients are targeted moments of the structural estimator (Section 6); the 750 m and 1 km bins discipline the tail of the distance decay. $^{*}p_{\text{RI}} < 0.10$, $^{**} < 0.05$, $^{***} < 0.01$.

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	+44.8%	+108.2%	+27.4%	+3.3%	+7.1%	+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 B15). 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 B6). 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 $= \beta_1/\hat{y}(S{=}0)$ with $\hat{y}(S{=}0) = \bar{y} - \hat{\beta}_1\bar{S}$, predicted adoption at zero treated exposure. 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	$+0.68^{***}$	[0.009]	$+138\%$	1,145,816
Panel A: By the user's own mobility				
User **visited** Dakar	$+0.39$	[0.118]	$+182\%$	281,207
User **never** visited Dakar	$+1.82^{**}$	[0.027]	$+88\%$	275,683
Panel B: By the driver's mobility to the destination (pooled across cities)				
Treated driver **visited** the city (migrant)	$+0.21$	[0.133]	$+400\%$	1,145,899×8
Treated driver **did not** visit the city	$+0.07^{**}$	[0.018]	$+125\%$	1,145,899×8

Notes: Coefficients are multiplied by 1,000. 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 user to travel to the capital. *Panel B* (outcome: cross-industry adoption in the destination city c) splits the treatment by the *drivers'* mobility. Both rows come from the single stacked regression $y_{jc} = \beta_1 \,\text{ShareTreated}_{jc}^{\text{visit}} + \beta_2 \,\text{ShareTreated}_{jc}^{\text{no-visit}} + \gamma_1 \,\overline{S}_{jc}^{\text{visit}} + \gamma_2 \,\overline{S}_{jc}^{\text{no-visit}} + \alpha_c + \varepsilon_{jc}$, where the two shares partition j's treated driver contacts by whether the driver transacted at an agent in city c at least once during the experiment (agent geolocation), so they sum to ShareTreated_j for every city. The specification controls for *both* an expected-exposure term for each share ($\overline{S}$, its mean across the 1,000 placebo re-randomizations, absorbing the fact that individuals connected to c-visiting drivers disproportionately live near c) *and* city fixed effects α_c; the sample stacks the eight largest secondary cities ($1{,}145{,}899 \times 8 = 9{,}167{,}192$ individual–city observations, differing from the overall row's 1,145,816 by 83 individuals due to independent pipeline construction), with SEs clustered by individual and RI re-randomizing the driver assignment and rebuilding both shares in each draw; see Section 4.2 for interpretation. RI p-values (1,000 permutations) in brackets; effect size $= \beta_1 / \hat{y}(S{=}0)$, predicted adoption at zero treated exposure (controls at their means). $^{*}p_{\text{RI}} < 0.10$, $^{**} < 0.05$, $^{***} < 0.01$.

Table 7: Structural Estimates and Targeted-Moment Fit

Panel A: Parameter estimates

	Estimate	95% design-respecting interval	Pr(sign)
$\hat{\beta}^{s}_{\text{learn}}$ (strong-tie learning, $\geq 5\%$)	0.320	$[0.041,\ 0.573]$	0.99
$\hat{\beta}^{w}_{\text{learn}}$ (community-tie learning, 2–5%)	0.433	$[-0.053,\ 0.905]$	0.95
$\hat{\beta}_{\text{demand}}$ (customer demand)	0.555	$[0.257,\ 0.863]$	1.00
$\hat{\beta}_{DI}$ (embedded-link amplification)	1.44	$[0.20,\ 2.78]$	0.99
$\hat{\beta}_{\text{comp}}$ (garage competition)	-2.483	$[-3.733,\ -1.609]$	1.00
$\hat{\delta}$ (distance decay, km^{-1})	1.546	$[0.845,\ 2.605]$	—
$\hat{\beta}^{\text{taxi}}_{s}$ (within-taxi learning, strong/mutual ties)	0.311	$[0.108,\ 0.502]$	1.00
$\hat{c}$ (tie-borne competition, both tie layers)	0.30	$[0.10,\ 0.48]$	1.00
Competition half-life (m)	448	$[266,\ 820]$	—

Panel B: Targeted moments — data vs. model-implied

Moment	Source	Data	Model
Phone, strong band ($\geq 5\%$)	Table 1	0.00416	0.00416
Phone, weaker band (2–5%)	Table 1	0.00407	0.00407
Demand, pair-specific no-community	Table 5	0.17454	0.17454
Customer, pair-specific full sample	Table 5	0.25599	0.25599
Garage, 50 m	Table 4, Panel B	-0.10249	-0.09732
Garage, 75 m	Table 4, Panel B	-0.08286	-0.08424
Garage, 100 m	Table 4, Panel B	-0.08258	-0.06983
Garage, 150 m	Table 4, Panel B	-0.05241	-0.05116
Garage, 300 m	Table 4, Panel B	-0.02689	-0.02824
Garage, 500 m	Table 4, Panel B	-0.01471	-0.01689
Garage, 750 m	Table 4, Panel B	-0.00807	-0.00929
Garage, 1,000 m	Table 4, Panel B	-0.00644	-0.00472
Within-taxi, mutual, Dakar	App. D.4	0.00039	0.00039
Within-taxi, mutual, outside	App. D.4	0.00260	0.00265
Held out (never targeted):			
Within-taxi, non-mutual, outside	App. D.4	-0.00015	0.00000
Customer, embedded-link subsample	Table 5	0.46481	0.49380
Survey, taxi friends (effect size)	Table 3	$+36.1\%$	$+30.0\%$

Notes: Panel A: joint minimum-distance estimates of the structural exposure-response parameters ψ in equation (3), primary (person-level) phone specification, targeting the fourteen auxiliary moments of Panel B. Each auxiliary sample's mean adoption level is matched exactly by its own nuisance baseline; weak-tie within-taxi learning is zero ($\beta^{\text{taxi}}_{w} = 0$; confirmed exactly on location-corrected outcomes, Appendix D.4), and the held-out outside-Dakar non-mutual cell, which the model then predicts at exactly zero, tests this restriction. The criterion uses diagonal inverse-design-variance weights from the 991 jointly complete stored placebo assignments. Intervals are design-respecting empirical 95% intervals: placebo moment deviations — from re-executions of the experiment's stratified assignment mechanism — are recentered at the observed moments and the model re-estimated on each draw (first-order updates, validated by full nonlinear re-estimation on 150 draws; the kernel and half-life intervals are approximate). They condition on the observed networks, samples, and specification. Pr(sign) is the share of re-estimates with the reported sign; $\delta > 0$ is imposed by the parameterization. The ranking between $\hat{\beta}^{s}_{\text{learn}}$ and $\hat{\beta}^{w}_{\text{learn}}$ is not identified and left uninterpreted. Panel B reports the targeted coefficients and their model-implied counterparts in natural probability units (the main tables report the same coefficients $\times 1{,}000$), followed by the held-out moments, which are never targeted; the friends prediction prices a strong market-distant tie at β^{taxi}_{s} with no proximity cost. Single-coefficient moments are fit exactly; the six overidentifying restrictions bind within the garage profile. Moment construction, weighting, identification, and diagnostics: Appendix D.

Table 8: Counterfactual Targeting at the Realized Budget (856 of 1,454 Businesses)

	Cross-industry (learning)		Within-taxi, Dakar		Within-taxi, outside		Total		$\Pr(\text{gain} > 0)$
	Induced	Δ	Induced	Δ	Induced	Δ	Induced	Δ	across draws
Random mean (benchmark)	4,912	—	+638	—	+73	—	5,633	—	—
Phone reach (degree)	8,190	+3,279	+552	−86	+97	+24	8,839	+56.9%	0.99
Greedy coverage	8,066	+3,154	+527	−111	+98	+26	8,692	+54.3%	0.99
Phone + customer blend	7,586	+2,675	+1,268	+630	+89	+16	8,944	+58.8%	0.99
Top mobile-money users	6,253	+1,341	+600	−38	+91	+19	6,944	+23.3%	0.99
Customer reach	4,210	−702	+1,479	+841	+64	−8	5,753	+2.1%	0.56
Garage-concentrated	4,855	−57	+499	−139	+72	−1	5,426	−3.7%	0.14

Notes: Model-predicted induced adoption $V(A) - V(A_\emptyset)$ on the estimation-aligned receiver populations (3,024,017 unique receivers), primary (person-level) phone specification, structural exposure-response parameters from Table 7. Each receiver is counted once: the cross-industry column is adoption in either geographic margin on the unique phone-receiver population, and it coincides with the learning channel's contribution to each rule — cross-industry learning is the only channel operating on phone receivers, and cross-channel interactions are negligible (Appendix D.8) — so learning accounts for 87% of induced adoption under the random benchmark and 93% under reach targeting; the within-taxi Dakar column sums the customer-linked (all 8,523 linked merchants, with the embedded-link amplification active), garage-area, and within-taxi tie risk sets, with cross-layer receivers assigned to a single risk set before evaluation; the within-taxi outside column is taxi adoption outside Dakar through mutual ties. As descriptive margins, Dakar and outside-Dakar cross-industry adoption are mutually exclusive outcomes on the same receivers; their induced counts (benchmark: 4,459 and 458; reach targeting: 7,436 and 764) are reported for reference and are never added into totals. Every policy treats exactly 856 of the 1,454 matched taxi businesses; a business is indivisible, and selecting it activates all of its matched stakeholder records. Group baselines are calibrated so the model reproduces observed adoption rates under the realized assignment. The benchmark row's total is the mean of 2,000 exact-budget random business assignments evaluated exactly (assignment SD 184); its group columns are means over the first 300 of those assignments. The budget is held fixed at 856 so rules are compared at identical scale; the experimental mechanism leaves the realized budget random, so this is a random-targeting benchmark rather than a replay of the experimental randomization. The within-taxi margin is positive under every assignment: customer demand and mutual-tie learning outweigh garage displacement and tie-borne competition on this margin, and its level depends on placement far less than the cross-industry margin because linked merchants connect to many businesses. Δ is the difference from the benchmark row in induced adopters (in percent for the total column). The last column is the share of the 1,000 design-respecting parameter draws in which the policy's gain over the draw-specific random mean is positive (reach-targeting gain interval: $[+33\%, +78\%]$); parameter and assignment uncertainty are never combined. "Customer reach" ranks businesses by linked customer merchants (draw gain interval $[-9\%, +38\%]$); the blend averages the phone-reach and customer-reach ranks (draw interval $[+47\%, +82\%]$). "Garage-concentrated" treats all garage-member businesses and fills the remaining slots randomly (mean over 200 documented fills; group columns are means over the same fills); a geographic-reach rule is excluded because no verified receiver geography exists. Negative competitive effects are never truncated. Appendix D details the policy machinery, channel decompositions, the doubled-customer-network counterfactual, and robustness.

Online Appendix

Nationwide Diffusion of Technology: Experimental Evidence from Multiple Networks

Deivy Houeix

A Appendix Figures **65**

B Appendix Tables **75**

C Fingerprint Data Processing **93**

D Structural Estimation and Policy Evaluation: Procedures **95**
D.1 Model, samples, and moment conditions 95
D.2 Joint minimum-distance estimation 96
D.3 Identification of the distance decay 97
D.4 Within-taxi identification from the reciprocity cells 99
D.5 Baseline coherence: what the data impose on the model's structure 100
D.6 Design-respecting parameter uncertainty 101
D.7 Policy evaluation 102
D.8 Channel decomposition 104
D.9 Robustness 105
D.10 Validation against untargeted survey moments 106
D.11 What the reported intervals do and do not contain 108

A Appendix Figures

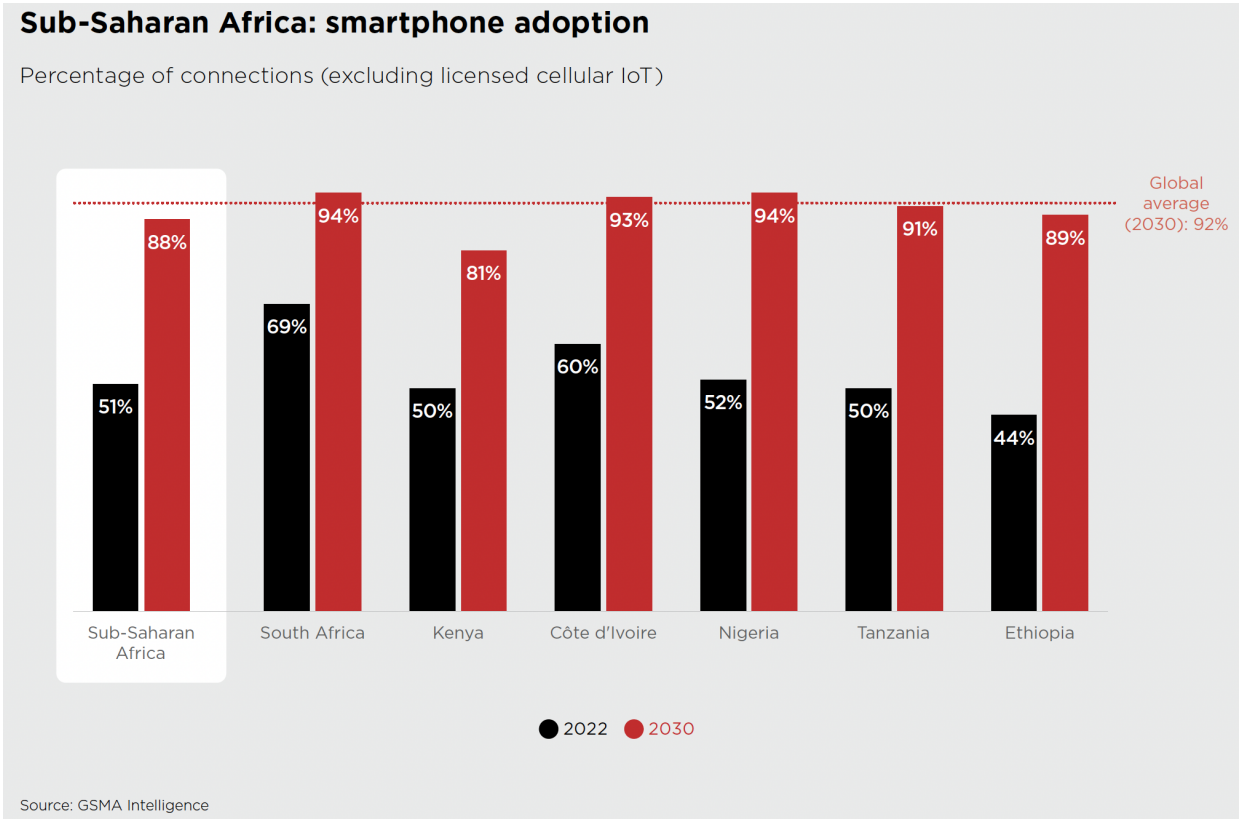


Figure A1: Smartphone Adoption in Sub-Saharan Africa

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.

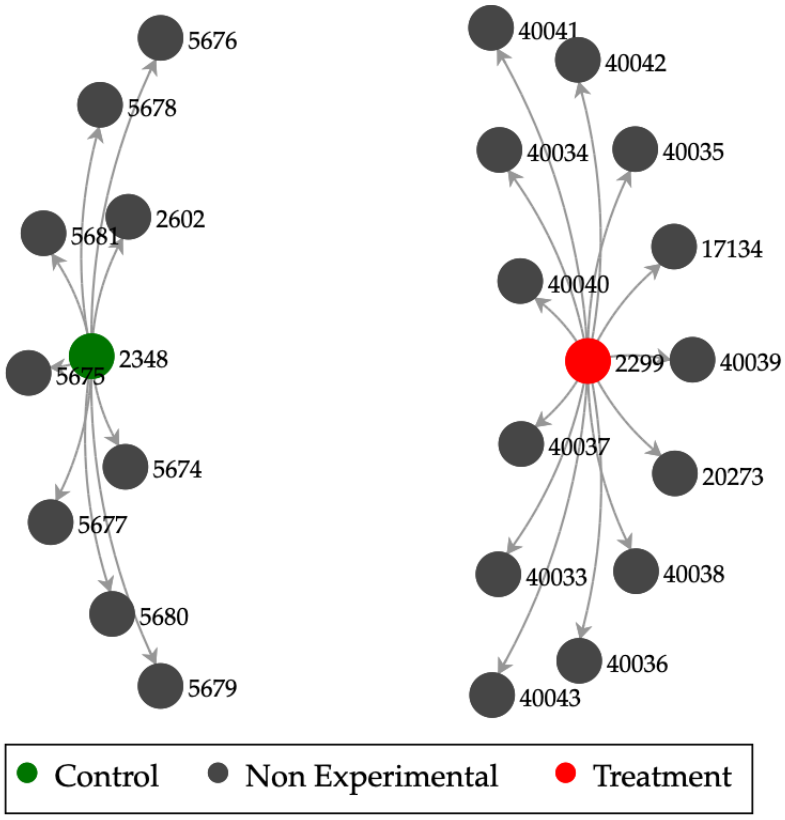


Figure A2: Illustration of the Experimental Design (Schematic)

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 3.3. Figure 3 shows the design in the actual community network.

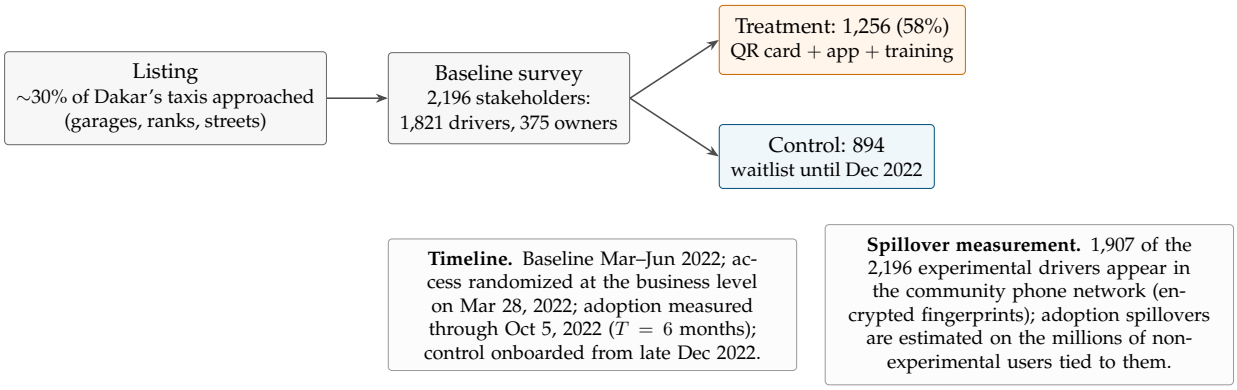


Figure A3: Experimental Design

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.

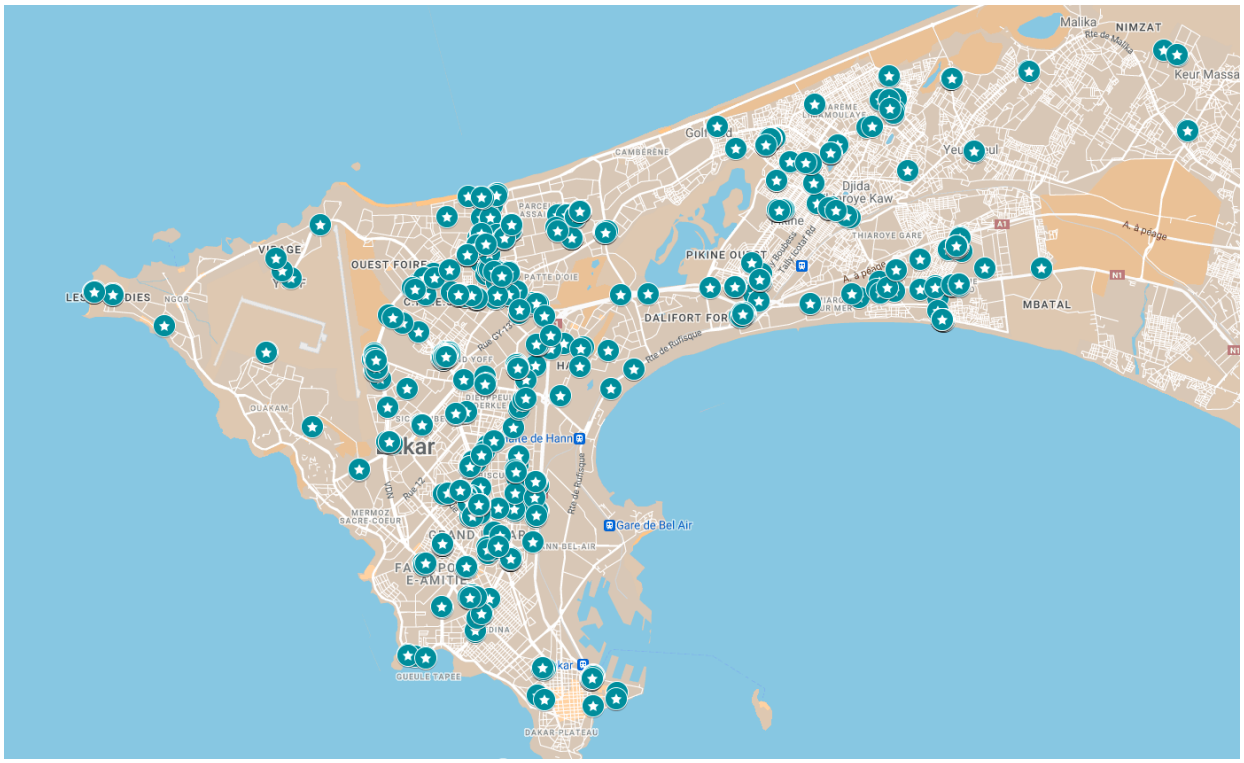


Figure A4: Listing Survey Activity in Dakar - March/April 2022

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.

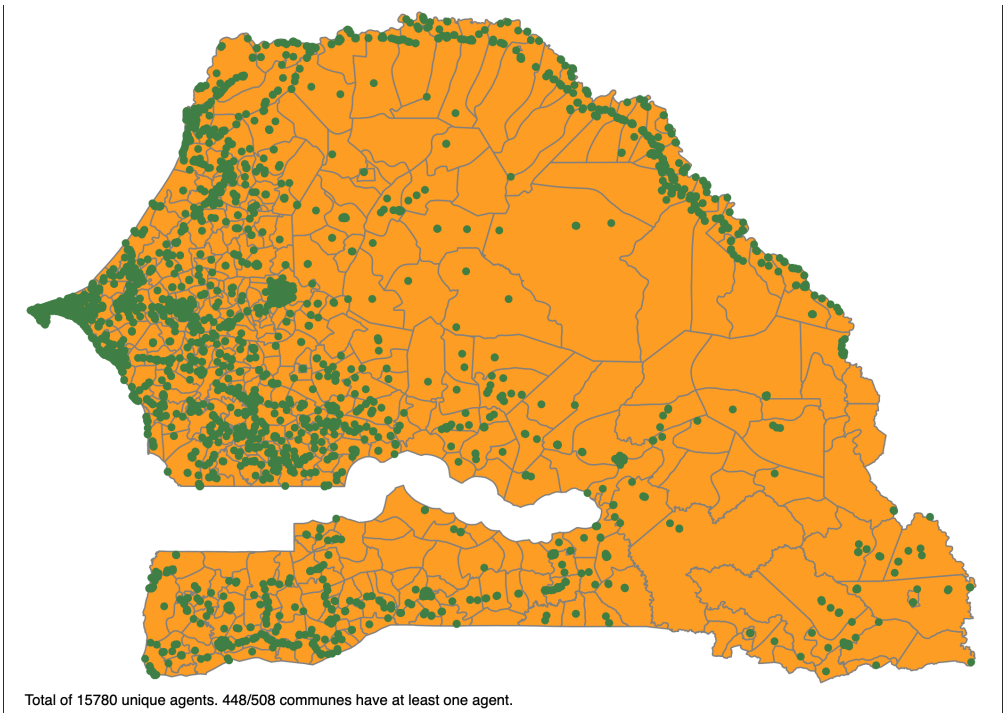


Figure A5: Mobile Money Agents in Senegal in 2022

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.

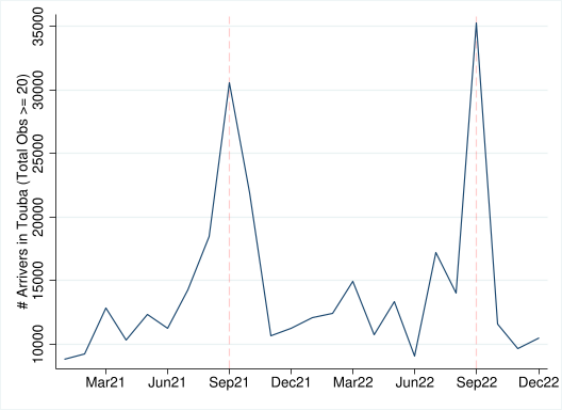


Figure A6: Location Data - Arrivals in Touba

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.

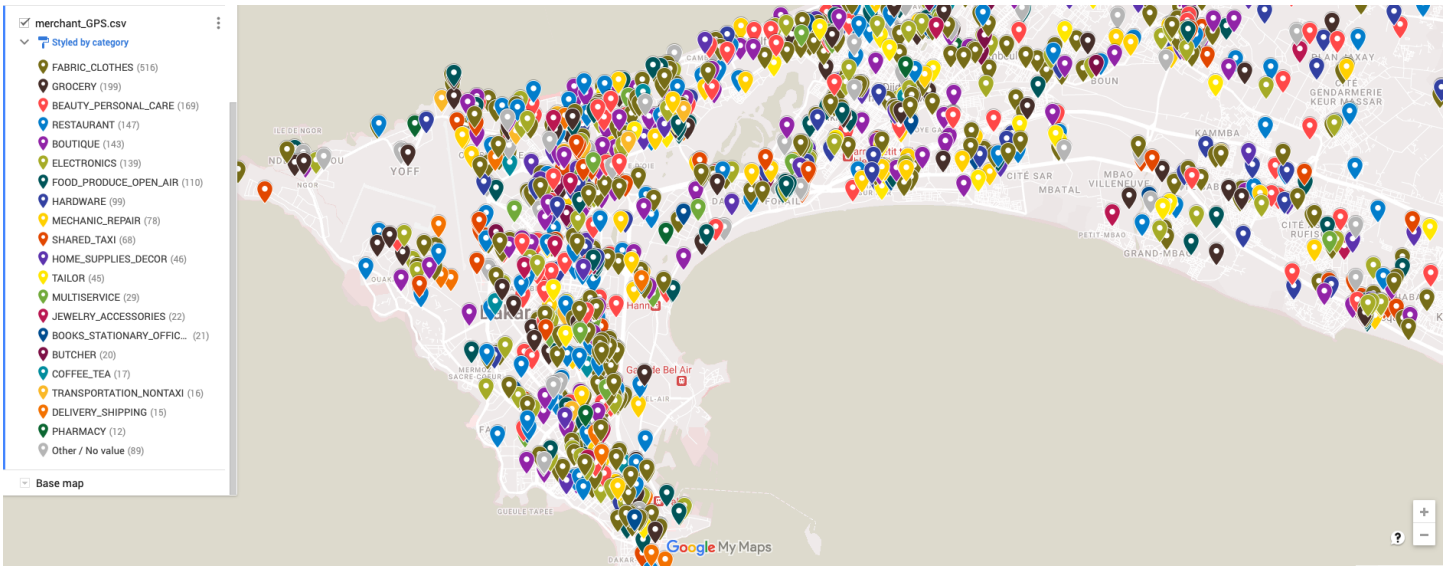


Figure A7: Technology Diffusion in Dakar Throughout the Experiment

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.

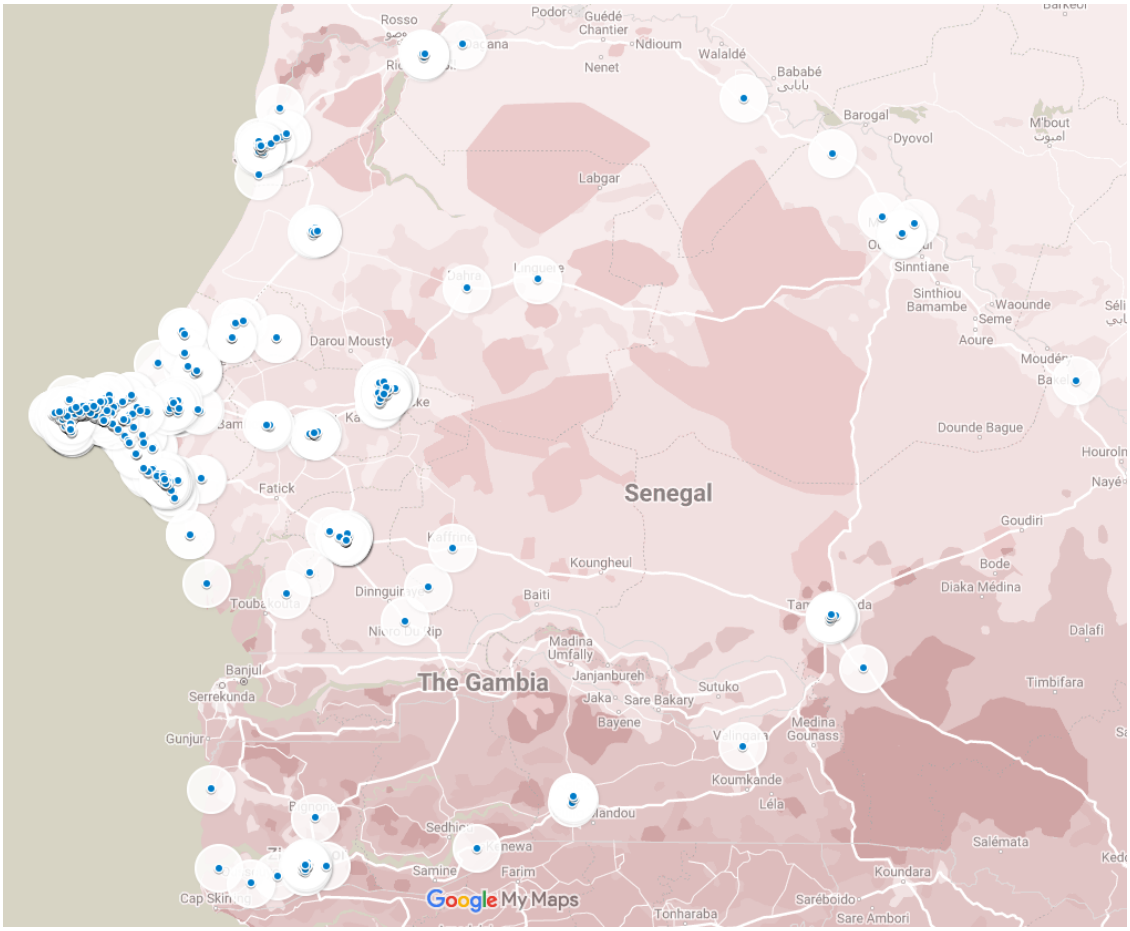


Figure A8: Technology Diffusion Inter-Industry Across Space Throughout the Experiment

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.

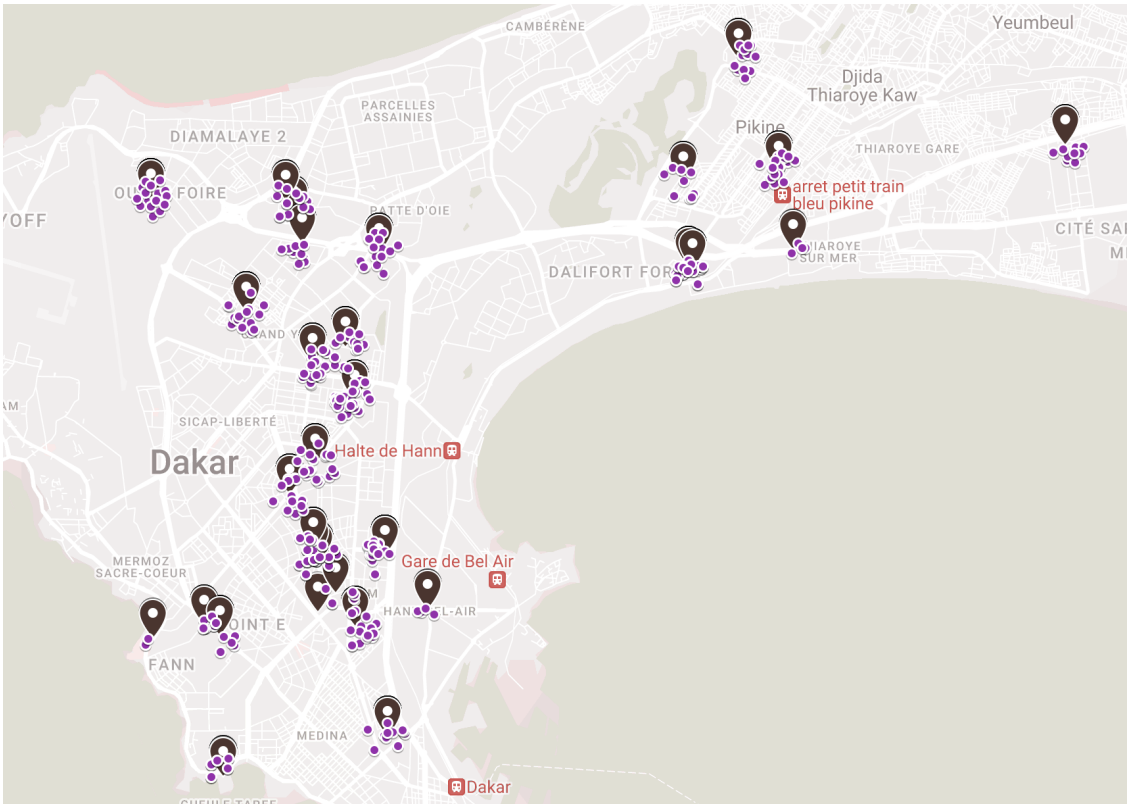


Figure A9: Local Technology Adoption Spillovers Nearby Taxi Garages

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.

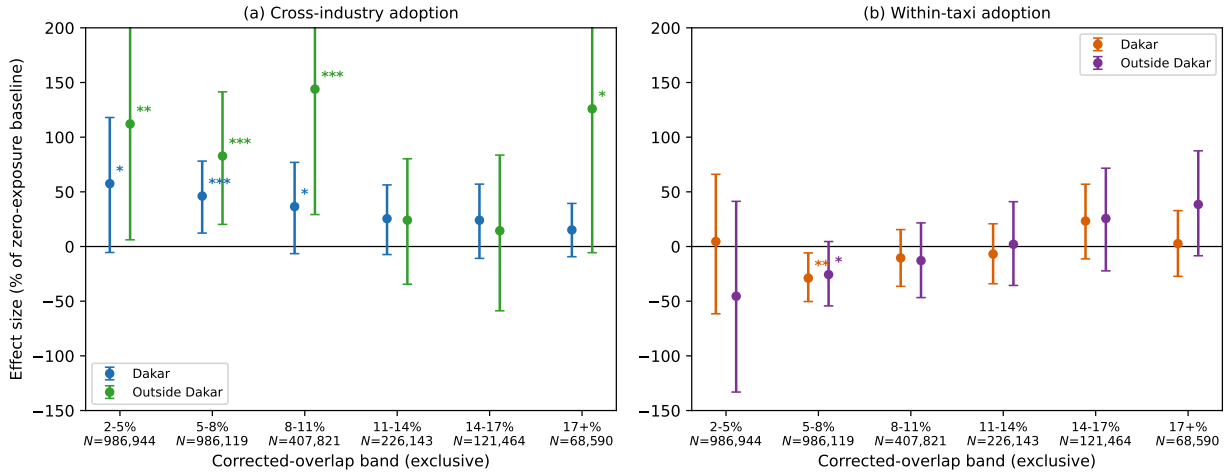


Figure A10: Per-band transmission by tie strength

Notes: Effect sizes ($\hat{\beta}_1/\hat{y}(S{=}0)$, in percent of predicted adoption at zero exposure, with $\hat{y}(S{=}0) = \bar{y} - \hat{\beta}_1\bar{S}$) 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 operates throughout 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 -29%, CI $[-50, -6]$)—colleagues close enough to compete—and is absent at both the weakest and strongest ties.

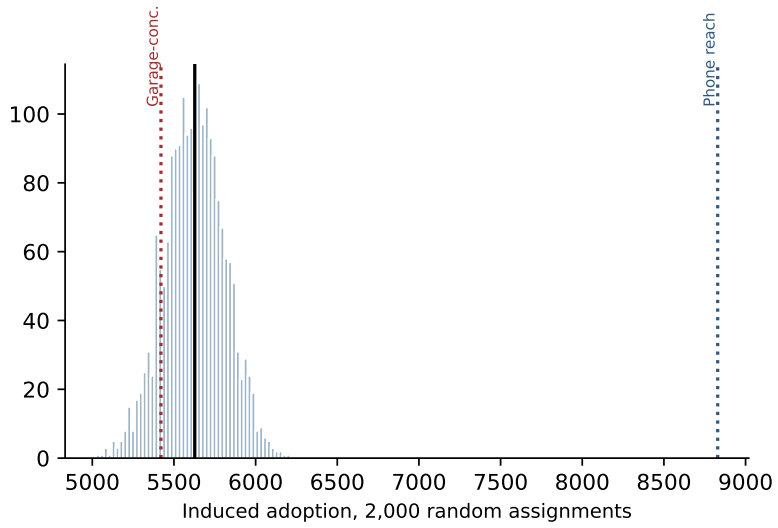


Figure A11: Random-benchmark level under parameter uncertainty: induced adoption of the mean exact-budget random assignment, across the 991 parameter re-estimates. This is parameter uncertainty in the benchmark's level, distinct from assignment-to-assignment variation within the random bank (SD $\approx$ 195).

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 $\geq 10\%$	Share mutual
Wife	495	20.81	86.5%	76.8%	64.4%	72.7%
Taxi contact	1,882	13.64	81.4%	62.8%	38.6%	62.3%
Family	3,814	11.44	80.8%	64.1%	41.8%	63.8%
Friend (family friend)	6,712	9.09	72.5%	53.3%	31.6%	59.3%
Professional	6,077	8.94	65.7%	46.0%	26.4%	55.6%
Friend (non-family)	2,922	6.00	61.9%	39.1%	18.2%	53.5%
Serendipitous	7,739	3.20	27.7%	13.7%	6.3%	20.1%
Client	526	1.94	22.8%	8.7%	3.0%	36.7%
By interaction frequency						
Weekly or more	6,028	10.56	73.3%	54.4%	33.7%	61.5%
Less than monthly	2,770	5.54	52.3%	31.4%	14.6%	46.2%

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× the overlap of contacts seen less than monthly. The last column reports the share of each category in which both parties saved each other's number (mutual): reciprocity follows the same gradient—72.7% of wives versus 36.7% of clients and 20.1% of serendipitous contacts—consistent with reciprocal saving marking relationship depth (Table B3).

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 B5).

Table B3: Overlap, Reciprocity, and Closeness Across Distance

	(1) Close	(2) Close	(3) Close	(4) Close
Corrected overlap share		$+0.555^{***}$	$+0.221^{***}$	$+0.083^{**}$
		(0.044)	(0.040)	(0.042)
Distance (100 km)		$+0.0006$	$+0.0009$	
		(0.0008)	(0.0010)	
Overlap × Distance			-0.0052	
			(0.0049)	
Far (modal city ≠ Dakar)				$+0.112^{***}$
				(0.011)
Overlap × Far				$+0.391^{***}$
				(0.082)
Mutual (both saved)	$+0.331^{***}$		$+0.310^{***}$	$+0.332^{***}$
	(0.008)		(0.009)	(0.011)
Mutual × Distance			$+0.0004$	
			(0.0016)	
Mutual × Far				-0.073^{***}
				(0.019)
R^2	0.109	0.025	0.112	0.126
N	13,355	12,991	12,991	13,355

Notes: OLS regressions of survey-reported closeness (contact is wife, family, or friend) on the bias-corrected contact-book overlap share, reciprocal-save (mutual) status, geographic distance, and their interactions, for all surveyed (driver, contact) pairs with fingerprint data on both sides—overlap is computed directly from the fingerprints, so the sample is not restricted to pairs appearing in the network edge lists. Locations are transaction-weighted mean agent coordinates from the 2020–2024 individual–agent panels, supplemented by merchant-registry coordinates; distance is the haversine distance between the pair's locations (100 km units); column (4) uses an indicator for the contact's modal city lying outside Dakar. HC1 standard errors in parentheses. Three patterns. First, reciprocity marks relationship depth: mutual pairs are reported close far more often both unconditionally (column 1; 62.8% vs. 29.7%) and conditional on overlap (columns 3–4), and the premium persists at distance ($0.332 - 0.073 = 0.259$, column 4). Second, overlap predicts closeness at all distances, and is *more* informative for distant pairs (far slope $0.083 + 0.391 = 0.474$): nearby, overlap partly reflects incidental shared environment, while at distance only genuine relationships sustain it—licensing the interpretation of cross-city community ties as real relationships. Third, the continuous distance level effect is a precise zero; the positive far indicator reflects the kin composition of long-distance transaction partners (wives and family are over-represented among distant contacts). $^{*}p < 0.10$, $^{**} < 0.05$, $^{***}p < 0.01$.

Table B4: Phone Network: Balance and Tests for Treatment-Induced Rewiring

	Treated	Control	Difference	SE / t	p-value
Panel A: Raw network levels (fingerprinted drivers, $N = 1{,}907$)					
Contact book size (fingerprint)	585	614	-28.8	24.5	0.240
Mutual contacts detected	8,309	8,982	-673	352	0.056*
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
Mean overlap (among mutual)	0.0039	0.0038	0.0000	0.0001	0.727
Panel B: Community-tie network (corrected overlap $\geq 2\%$, pre-trim; $N = 1{,}907$)					
Community ties per driver ($\times$1,000)	295.3	305.0	-9.8	14.0	0.486
Strong ties per driver ($\times$1,000)	46.6	46.9	-0.4	4.7	0.935
Mean corrected overlap (driver-level)	0.0337	0.0335	0.0002	0.0005	0.715
Merchant community ties ($\times$1,000)	6.50	6.57	-0.07	0.35	0.839
Taxi community ties	717	742	-25	35	0.469
Joint F-test (Panel B)	$F(5, 1901) = 0.615$			$p = 0.688$	
Panel C: Tie-strengthening tests (full matched sample, $N_T = 1{,}256$, $N_C = 894$)					
Strong ties per driver ($\geq 5\%$ raw)	668	749	-81	-0.51	0.608
Mean overlap among strong ties	0.0738	0.0738	-0.0001	-0.28	0.781
Share taxi among strong ties	0.070%	0.060%	$+0.010\%$	$+1.29$	0.196
Taxi-contact overlap distribution	KS $= 0.015$			$p = 0.14$	
Pooled corrected overlap (community net)	0.0389 vs. 0.0397, treated *lower*				

Notes: Fingerprints observed February 2024; the network is used as a proxy for pre-existing social proximity (Section 1.3). Panels A–B report treated and control means, the difference, robust SE, and p-value from a regression on treatment status. Panel A uses raw (uncorrected) quantities, inflated by hash collisions. Panel B reports the corrected community-tie network used in the analysis (strong $=\geq 5\%$), before the symmetric top-2% hub trim—hence the large counts; arms are jointly indistinguishable ($p = 0.688$). Panel C tests for treatment-induced tie strengthening on the full matched sample (Welch t-tests; level differences from Panel A reflect the sample, not new findings). Because selective saving of technology-related contacts would fatten the upper tail of the tie-strength distribution without necessarily moving its mean, the Kolmogorov–Smirnov test compares the *full* overlap distribution of treated and control drivers' taxi contacts—the margin where treatment-induced saving would appear—and detects no shift ($p = 0.14$); pooled tie strength is, if anything, *lower* for treated drivers, the wrong direction for a strengthening story. Merchant and taxi classifications use only pre-fingerprint registrations. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table B5: Reciprocity and the Cross-Market Channel: Mutual Phone Contacts

	$\geq 2\%$	$\geq 5\%$	$\geq 10\%$
Cross-industry, Dakar	$+0.96$ [.558] {+2.9%}	$+0.87$ [.701] {+2.6%}	$+0.88$ [.790] {+2.9%}
Cross-industry, outside	-0.22 [.432] {-10%}	-0.45 [.247] {-23%}	-0.57 [.214] {-33%}
Within-taxi, Dakar	$+0.74$ [.656] {+2.7%}	$+0.07$ [.990] {+0.1%}	-0.32 [.948] {-0.4%}
Within-taxi, Thiès	$+\mathbf{1.10}^{***}$ [.003] {+26%}	$+\mathbf{2.84}^{***}$ [.001] {+35%}	$+\mathbf{5.16}^{***}$ [.001] {+43%}
N	188,948	73,393	31,880

Notes: Coefficients $\times 1{,}000$; RI p-values (1,000 permutations) in brackets; effect sizes in braces, computed as $\beta_1/\hat{y}(S{=}0)$ with $\hat{y}(S{=}0) = \bar{y} - \hat{\beta}_1\bar{S}$, predicted adoption at zero treated exposure (controls at their means). Each column re-estimates the specification of Table 1 with exposure computed over *mutual* phone contacts only—pairs in which each party saved the other's number—at the indicated corrected-overlap floor. The mutual layer is estimated *untrimmed*: because mutual status requires explicit reciprocal saves, the hub and collision contamination that the top-2% trimming targets is largely absent. Results are insensitive to this choice: with the strict trim reapplied, the within-taxi Thiès estimates are $+0.90$ to $+4.39$ ($p_{\text{RI}} \leq .006$), and on the full mutual layer with no overlap floor (1.30 million contacts, strict trim) the effect remains $+18\%$ of the base rate ($p_{\text{RI}} = .033$). Two patterns stand out. First, cross-industry: mutual contacts—despite baseline adoption rates 2.5–10$\times$ higher than the community network's—show no incremental response to treatment at any threshold. The cross-industry diffusion of Table 1 is therefore carried by non-mutual community ties, consistent with mutual contacts' exposure to the technology being saturated irrespective of the experiment (inframarginality). Second, 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 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 corresponding within-taxi *non-mutual* responses, which enter the structural estimation, are reported in Section 6 and Appendix D.4. The same channel appears in the pre-treatment transaction network (Table B14). Both patterns are unchanged when driver–contact pairs sharing regular customers are excluded (Appendix Table B6). $^{*}p_{\text{RI}} < 0.10$, $^{**} < 0.05$, $^{***} < 0.01$.

Table B6: Network Spillovers Excluding Pairs with Shared Regular Customers

	$\geq 2\%$ (community)	$\geq 5\%$ (strong)	$\geq 10\%$
Cross-industry, Dakar	$+2.88$ [.128] {+41%}	$+4.15^{**}$ [.018] {+39%}	$+5.14^{**}$ [.015] {+68%}
Cross-industry, outside	$+0.68^{***}$ [.010] {+138%}	$+0.68^{***}$ [.009] {+81%}	$+0.73^{**}$ [.030] {+136%}
Within-taxi, Dakar	$+0.08$ [.665] {+14%}	-1.01^{***} [.003] {-31%}	-0.65 [.463] {-12%}
Within-taxi, Thiès	-0.05 [.159] {-49%}	-0.10 [.133] {-24%}	$+0.08$ [.671] {+9%}
N	1,145,898	1,020,292	376,744

Notes: Coefficients $\times 1{,}000$; RI p-values (1,000 permutations) in brackets; effect sizes in braces, computed as $\beta_1/\hat{y}(S{=}0)$ with $\hat{y}(S{=}0) = \bar{y} - \hat{\beta}_1\bar{S}$, predicted adoption at zero treated exposure (controls at their means). 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 and community-tie definition as in Table 1; the strict top-2% trimming is applied *before* the purge, so the estimation universe is exactly that of Table 1 minus the purged pairs and sample sizes are weakly smaller. 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 B5 survives the purge ($+2.31$, $p_{\text{RI}} = 0.003$ at $\geq 5\%$; $+4.18$, $p_{\text{RI}} = 0.009$ 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 B7: Cross-Industry Spillovers Excluding Shared-Customer Pairs and Garage-Exposed Drivers

	$\geq 2\%$ (community)	$\geq 5\%$ (strong)	$\geq 10\%$
Cross-industry, Dakar			
Share treated	$+1.76$ {+23%}	$+3.63^{**}$ {+33%}	$+4.59^{**}$ {+59%}
	(0.47)	(0.30)	(0.43)
	[0.333]	[0.037]	[0.045]
Expected exposure	$+13.01$	-9.10	-47.64
	(15.00)	(10.06)	(14.24)
	[0.386]	[0.366]	[0.001]
Cross-industry, outside Dakar			
Share treated	$+0.48^{*}$ {+78%}	$+0.61^{**}$ {+69%}	$+0.75^{**}$ {+134%}
	(0.15)	(0.09)	(0.13)
	[0.057]	[0.027]	[0.044]
Expected exposure	$+8.62$	-2.63	-2.00
	(5.06)	(3.20)	(4.34)
	[0.089]	[0.410]	[0.645]
Dep. var. mean (Dkr / outside)	8.79 / 0.90	13.41 / 1.29	10.64 / 1.02
N	1,142,867	944,780	347,063

Notes: Specification, sample construction, and cell format exactly as Table 1 (all estimates $\times 1{,}000$; effect sizes in braces; HC1 SEs in parentheses; RI p-values in brackets for Share treated, HC1 for Expected exposure), after two purges applied to the trimmed Table 1 universe: (i) every driver–contact pair sharing a regular customer (≥ 2 transactions with each side; the link definition of Table 5), removing the demand channel; and (ii) every pair of the 285 garage-exposed drivers (self-reported garage-based in the endline survey or matched to a GPS garage roster), removing the population carrying the local-competition channel. The cross-industry learning effects survive with magnitudes close to Table 1 at the strong-tie thresholds ($+4.16 \rightarrow +3.63$ in Dakar and $+0.68 \rightarrow +0.61$ outside at $\geq 5\%$; $+5.16 \rightarrow +4.59$ and $+0.73 \rightarrow +0.75$ at $\geq 10\%$), indicating that cross-industry diffusion operates through social learning rather than shared demand or competitive exposure. Under the same purge, the within-taxi Dakar coefficient at $\geq 5\%$ remains negative and significant (-0.88, $p_{\text{RI}} = 0.025$): competition among drivers is a market-wide phenomenon rather than a garage artifact. $^{*}p_{\text{RI}} < 0.10$, $^{**} < 0.05$, $^{***} < 0.01$.

Table B8: Count Specification: Number of Treated Ties

	$\geq 2\%$	$\geq 5\%$	$\geq 10\%$
Cross-industry, Dakar	$+0.25$ [.114] {+4%}	$+1.67^{**}$ [.015] {+16%}	$+2.01^{***}$ [.008] {+25%}
Cross-industry, outside	$+0.05^{**}$ [.021] {+17%}	$+0.30^{***}$ [.007] {+39%}	$+0.25^{*}$ [.052] {+37%}
Within-taxi, Dakar	$+0.04$ [.214] {+40%}	-0.06 [.840] {-2%}	$+0.39$ [.608] {+8%}
Within-taxi, Thiès	$+0.00$ [.922] {+0%}	$+0.02$ [.786] {+6%}	$+0.29$ [.154] {+52%}
N	1,145,899	1,020,391	376,871

Notes: Coefficients $\times 1{,}000$ on the *number* of treated ties (rather than the treated share), from the specification of Table 1 with NumTreated replacing ShareTreated and the expected-count control replacing expected exposure; community-tie network at the indicated corrected-overlap thresholds, strict top-2% trimming, RI p-values (1,000 permutations) in brackets, effect sizes per additional treated tie in braces, computed as $\beta_1/\hat{y}(\text{NT}{=}0)$, where $\hat{y}(\text{NT}{=}0) = \bar{y} - \hat{\beta}_1\overline{\text{NT}}$ is predicted adoption at zero treated ties with controls at their means ($\overline{\text{NT}} = 11.99$, 1.74, 1.31 at the three thresholds). Cross-industry learning survives the count specification (certified at $\geq 5\%$ and $\geq 10\%$). The within-taxi coefficients—negative and significant in the share specification—are null in counts at every threshold: competition responds to the *share* of one's circle that is treated (market intensity), not to the absolute number of treated contacts, whereas learning operates tie by tie. $^{*}p_{\text{RI}} < 0.10$, $^{**} < 0.05$, $^{***} < 0.01$.

Table B9: 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 B10: 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 B5 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 B11: Robustness: Garage Proximity Spillovers with Onboarding-GPS Assignment

	50 m	75 m	100 m	150 m	300 m	500 m
Panel A: All merchants near garages						
Any treated	-0.0079^{**}	-0.0055	-0.0056	-0.0033	-0.0013	-0.0003
p_{RI}	[0.044]	[0.213]	[0.146]	[0.194]	[0.349]	[0.757]
Effect size	-65%	-61%	-68%	-63%	-46%	-20%
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	-0.0809^{**}	-0.0647^{**}	-0.0671^{**}	-0.0409^{**}	-0.0188^{**}	-0.0091^{**}
p_{RI}	[0.024]	[0.035]	[0.028]	[0.018]	[0.019]	[0.036]
Effect size	-94%	-94%	-95%	-92%	-87%	-78%
N	6,004	7,962	15,277	25,207	56,726	122,764
Panel C: Mutual phone contacts only						
Any treated	$+0.0192$	$+0.0073$	-0.0751	-0.0586	-0.0855^{*}	-0.0589^{*}
p_{RI}	[0.728]	[0.893]	[0.389]	[0.257]	[0.071]	[0.092]
Effect size	$+14\%$	$+5\%$	-37%	-34%	-49%	-45%
N	485	515	658	874	1,226	1,919

Notes: Same design, samples, exposure definitions, and randomization inference as Table 4, but the outcome—within-taxi adoption at the tested garage—assigns the adopting merchant's home garage from the GPS coordinates recorded at merchant *onboarding* (registration), rather than from the modal agent location in the mobility data: the entrant counts as adopting at the tested garage when the onboarding location is nearest to it among the matched garages. Adoption at other garages under this assignment carries effect sizes 1.4 to 1.7 times smaller and is mostly insignificant (not shown). Effect size $= \beta_1/\hat{y}(\text{AnyTreated}{=}0)$, predicted adoption with no treated garage (controls at their means), as in Table 4. RI p-values (1,000 permutations) in brackets. $^{*}p_{\text{RI}} < 0.10$, $^{**} < 0.05$, $^{***} < 0.01$.

Table B12: 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 B5 (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 B17 (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). The structural demand moment and the policy demand exposure use only the pair-specific no-community subset of this network ($N = 4{,}864$ merchants; a shared regular customer has $\geq$ 2 transactions with the specific driver and $\geq$ 2 with the specific merchant); the broad aggregate network is descriptive
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 B14 (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 supports 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 B13: 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 28,720 survey pairs referenced in Section 1.3 come from the drivers who completed the network validation module.

Table B14: Spillovers in the Pre-Experiment Transaction Network ($\geq$ 2 P2P transfers)

	Taxi, Dakar	Cross-ind., Dakar	Cross-ind., other	Taxi, Thiès
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 the Thiès market ($+83\%$ of baseline, $p_{\text{RI}} = 0.001$)—mirroring the mutual-contact channel of Appendix Table B5—and carry no significant local effects. The pattern is identical with a ≥ 1-transfer definition ($N = 60{,}381$; Thiès $+5.23$, $+60\%$, $p_{\text{RI}} = 0.001$; 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 B15: 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	$+82.1\%$	$+186.0\%$	$+43.4\%$	$+0.1\%$	-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	$+38.3\%$	$+82.1\%$	$+28.4\%$	$+8.8\%$	$+13.9\%$	$+8.3\%$
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. Effect size $= \beta_1/\hat{y}(S{=}0)$ with $\hat{y}(S{=}0) = \bar{y} - \hat{\beta}_1\bar{S}$, predicted adoption at zero treated exposure, as in Table 5. Regressions control for expected exposure. $^{*}p_{\text{RI}} < 0.10$, $^{**} < 0.05$, $^{***} < 0.01$.

Table B16: Customer-Network Spillovers by Number of Customer-Linked Drivers

	Number of customer-linked experimental drivers		
	1–5	6–15	16+
Panel A: Within-taxi adoption, Dakar			
Community tie	$+255.43^{***}$	$+336.89^{**}$	$+680.46^{**}$
	(68.18)	(67.96)	(60.99)
	[0.002]	[0.015]	[0.021]
Effect size	$+36.9\%$	$+62.4\%$	$+258.8\%$
N	388	1,092	2,179
No community tie	$+103.52^{**}$	$+348.65^{**}$	$+564.83^{**}$
	(21.82)	(60.51)	(97.76)
	[0.022]	[0.020]	[0.038]
Effect size	$+13.2\%$	$+71.7\%$	$+193.8\%$
N	2,106	1,705	1,053
Panel B: Cross-industry adoption			
Community tie	$+45.84$	$+148.84$	-129.51
	(33.85)	(108.10)	(150.84)
	[0.311]	[0.450]	[0.649]
N	1,321	745	725
No community tie	$+14.29$	$+131.44$	$+116.00$
	(8.08)	(48.71)	(108.66)
	[0.326]	[0.262]	[0.537]
N	20,262	3,571	1,888

Notes: Coefficients, SEs, and dep. var. means multiplied by 1,000; HC1 standard errors in parentheses; randomization-inference p-values (1,000 permutations) in brackets. Each cell is a separate regression of adoption on the treated share of the merchant's customer-linked experimental drivers, controlling for expected exposure, within the indicated bin of the total number of customer-linked drivers. Sample, link definition, and community-tie classification are identical to Table 5: a merchant is linked to a driver through a shared regular customer (≥ 2 transactions with the specific driver and ≥ 2 with the specific merchant, November 2022–November 2023, non-garage drivers); a community tie is a bias-corrected contact-book overlap of at least 2% with any linked driver. Binning on the number of linked drivers holds exposure scale fixed, so the community-tie/no-tie comparison within a column is not driven by community-tied merchants' roughly threefold higher connectivity. Among the lightly exposed (1–5 linked drivers), the community-tie coefficient is $2.5\times$ the no-tie coefficient; the gap closes at 6–15 drivers and is modest at 16+, consistent with social exposure mattering most where treated contacts are scarce. Effect size $= \beta_1/\hat{y}(S{=}0)$ with $\hat{y}(S{=}0) = \bar{y} - \hat{\beta}_1\bar{S}$, predicted adoption at zero treated exposure, as in Table 5. Dep. var. means (per 1,000): Panel A, 845/844 (1–5), 720/679 (6–15), 646/623 (16+) for community-tie/no-tie respectively. $^{*}p_{\text{RI}} < 0.10$, $^{**} < 0.05$, $^{***} < 0.01$.

Table B17: Pre-Intervention P2P Customer Network (≥ 2 fare transactions, Jan 2021–Mar 2022)

	Within-taxi, Dakar	Cross-industry
Share treated	$+36.35$	$+14.79$
	(20.46)	(8.92)
	$[0.136]$	$[0.273]$
Effect size	$+4.6\%$	$+1.9\%$
Dep. var. mean	793.5	762.7
N	3,037	13,000

Notes: Coefficients, SEs, and dep. var. means multiplied by 1,000; HC1 standard errors in parentheses; randomization-inference p-values (1,000 permutations) in brackets. Customers are identified from transaction-level pre-intervention P2P transfers into non-garage experimental drivers (January 2021–March 2022, CFA 500–3,000 per transaction, ≥ 2 qualifying transactions per customer); each customer is then linked to non-experimental merchants through pre-period P2P counterparty pairs (November 2017–February 2022). Merchants who adopted before March 2022 are excluded ($N = 1{,}755$), as are matched (experimental) merchants. Column 1: taxi-category merchants, dep. var. = registration as a taxi merchant in Dakar during the experiment window; column 2: non-taxi merchants, dep. var. = registration as a non-taxi merchant (any region). Regressions control for expected exposure (mean share treated across the 1,000 placebo re-randomizations; coefficients not shown). Estimates are stable across qualification thresholds of 1–2 transactions ($+20$ to $+36$ per 1,000) and attenuate at ≥ 3, where the sample collapses to 828 customers, and are similar using the *number* of treated connections instead of the share, controlling for its expected value ($+20.34$, SE $= 9.11$, $p_{\text{RI}} = 0.110$ within taxi). Because the pre-period fare network is sparse, the two-hop thresholds are looser than in the post-experiment customer network (Table 5), which requires ≥ 2 transactions with the *specific* driver and merchant; imposing that strict dyadic definition here shrinks the within-taxi sample to 670 merchants and yields uninformative estimates ($+6.60$, SE $= 37.58$). 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. $^{*}p_{\text{RI}} < 0.10$, $^{**} < 0.05$, $^{***} < 0.01$.

Table B18: 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$1,000)	$+272.8^{***}$	$+65.9^{***}$
	(23.3)	(25.1)
N	4,450	7,757
$\bar{y}$	0.784	0.744
Panel B: All customer-linked merchants		
ShareTreated ($\times$1,000)	$+281.8^{***}$	$+64.7^{***}$
	(23.3)	(25.0)
N	4,495	7,906

Notes: Each column constructs the shared-customer network G^{cust} from transactions within that window only ($\geq$ 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 B5), 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 and Policy Evaluation: Procedures

This appendix documents the construction of the structural model of Section 6: the moment conditions and estimation samples (D.1), the joint minimum-distance estimator and its weighting (D.2), the baseline-coherence tests that fix the model's structure (D.5), the design-respecting parameter uncertainty procedure (D.6), the identification of the distance decay (D.3), the policy-evaluation machinery (D.7), the channel decompositions (D.8), robustness exercises (D.9), validation against untargeted survey moments (D.10), and the scope of the reported uncertainty (D.11).

D.1 Model, samples, and moment conditions

Adoption probability. For receiver i in auxiliary sample s, the model of equation (4) specifies $P_{is}(A) = 1 - (1 - \Lambda(a_s + x_{is}(A)'\psi))^T$ with $T = 6$, where $\psi = (\beta^s_{\text{learn}}, \beta^w_{\text{learn}}, \beta_{\text{demand}}, \beta_{DI}, \beta_{\text{comp}}, \delta, \beta^{\text{taxi}}_s, c)$ is common across auxiliary samples and a_s is a sample-specific nuisance baseline. For phone learning, exposure is the treated share of i's experimental ties in the strong ($\geq 5\%$) or exclusive weaker (2–5%) overlap band, with treatment constant across all records belonging to the same experimental business. For demand, exposure is the treated share of the exact pair-specific customer-linked experimental taxi businesses; the no-community Table-5 sample identifies β_{demand}, and the full customer-linked sample — where receivers with at least one socially embedded link (a customer link that is also a community phone tie) respond with coefficient $\beta_{\text{demand}} + \beta_{DI}$ — identifies the amplification β_{DI}. For spatial competition, the exposure for a receiver attached to garage g is $1\{\text{community gate}\}\, 1\{\text{AnyTreated}_g\} e^{-\delta d_{ig}}$, where d_{ig} is the exact receiver–garage distance. The eight garage auxiliary moments use the exact Panel-B receiver samples at 50, 75, 100, 150, 300, 500, 750, and 1,000 m. For the within-taxi tie channels, exposure is the treated share of i's mutual (reciprocally saved) or non-mutual community ties in the within-taxi reciprocity samples; mutual exposure enters with coefficient $\beta^{\text{taxi}}_s - c$ for Dakar receivers and β^{taxi}_s for receivers outside Dakar, non-mutual exposure with coefficient $-c$ in Dakar and zero outside (Appendix D.4).

Auxiliary samples. The moments are computed on the exact estimation samples of the experimental tables, with no substitutions: (i) the strict phone-contact network at the

person level, separately for the strong and weak tie-strength bands (the primary, "paper-consistent" specification), and, as a robustness specification, the same network with edges collapsed to the business level; (ii) the post-period pair-specific customer-demand samples (the no-community merchant set and the full customer-linked set with its embedded-link indicator), exactly as in the corresponding table; (iii) the receiver-level garage competition design: the gated Panel-B sample (community gate, non-mutual pairs), with each receiver's exact distance to the treated garage. I use the exact receiver-level design throughout; the uniform-disk approximation is archived as a diagnostic; (iv) the four within-taxi reciprocity samples: mutual and non-mutual community ties at the $\geq 5\%$ threshold, split by receiver geography (Dakar vs. outside), built with the strict top-2% trimming rule and identity-asserted in code against the reciprocity-table rebuild.

Moment conditions. For each auxiliary sample I fit the same auxiliary regression — a linear projection of adoption on a constant, realized exposure, and expected exposure — to the observed outcomes and to the model probabilities $P_i(\psi)$, and match the coefficients. The response moments comprise the exposure coefficients of the phone bands, the demand experiment, and the garage design (including its distance profile, which identifies δ). In addition, each auxiliary sample's mean adoption level is matched *exactly* by its own nuisance baseline a_s: the level moments contribute one parameter each and hence zero net degrees of freedom. The estimated system has fifteen response moments and eight common parameters, leaving seven overidentifying response moments; the minimized criterion Q is reported as a descriptive fit diagnostic, not as a χ^2 test; the moments are not asymptotically Gaussian under the design-based distribution I use, so a χ^2 calibration would be unfounded.

D.2 Joint minimum-distance estimation

Weighting. The primary minimum-distance criterion uses a diagonal weight matrix whose entries are the inverse design variances of the fourteen auxiliary coefficients. These variances are computed from the 991 *unique, jointly complete* stored business-level placebo assignments for which all fourteen coefficients are defined; the same placebo draw index is used for every moment, so cross-moment design correlation is preserved in the draw bank. Nine of the original 1,000 assignments yield an undefined small-radius garage coefficient because the corresponding auxiliary regressor is degenerate; those assignments are retained in the audit but are not used to form the primary joint covariance or inference perturbations. From the same common assignments I also compute the full covariance matrix, which is positive semidefinite and captures the strong correlation of garage coef-

ficients across nested radii. As robustness, I re-estimate using a block-shrinkage inverse-covariance weight and a standardized-identity criterion; the substantive estimates are stable across these choices. Because the exactly identified blocks drive their residuals to zero under any positive weight, the weighting choice matters only for β_{comp} and δ, whose relative weighting across the eight radii changes with the treatment of the cross-radius design covariance. The separability itself is direct to verify: refitting each block alone reproduces the joint estimates to machine precision, and the garage block's standalone criterion equals the minimized joint criterion. Both properties are specific to the preferred specification — under the diagnostic common-intercept restriction the blocks couple and the criterion deteriorates sharply.

Optimization. For each candidate ψ, the nuisance baselines are profiled out exactly: each a_s solves its level-calibration equation numerically. The outer criterion is minimized from a deterministic multistart set. The best basin is reached from a large majority of starts and is confirmed with an independent derivative-free optimization; inferior stationary points are retained as diagnostics rather than discarded. The primary person-edge and business-collapsed phone specifications are estimated separately throughout and are never pooled.

D.3 Identification of the distance decay

Identification. Competition enters receiver i's index as $\beta_{\text{comp}} \cdot \text{AnyTreated}_{g(i)} \cdot e^{-\delta d_i}$, with d_i the exact distance to the garage. The garage moments are the auxiliary coefficients at eight nested radii (50, 75, 100, 150, 300, 500, 750, 1,000 m). Because the model constrains all eight to lie on one exponential profile, the reduced-form coefficient at distance d satisfies approximately

$$\hat{b}(d) \;\approx\; \beta_{\text{comp}}\, e^{-\delta d} \times T\, \bar{p}(1-\bar{p}), \tag{A1}$$

where the second factor is the average marginal effect of the logistic-binomial model. Taking logs, ratios of coefficients across radii identify δ irrespective of the population's adoption level ($\log|\hat{b}(d_1)| - \log|\hat{b}(d_2)| = -\delta(d_1 - d_2)$), while the level at any radius identifies β_{comp}. In the joint estimator this logic operates inside the minimum-distance criterion at the receiver level, using exact distances rather than bin midpoints, with the level scaled by the garage-sample adoption rate — essential, since garage-area adoption is roughly twenty times the network-sample rate.

Transparent two-step diagnostic. A binned version makes the identification visible: regressing $\log|\hat{b}(d)|$ on bin midpoint distance across the six all-merchant (Panel A) bins is log-linear with $R^2 = 0.96$ (predicted-to-actual ratios between 0.83 and 1.22; 4 of six bins within 15%), and yields $\hat{\delta} = 3.73^{***}\,\text{km}^{-1}$ (half-life 186 m). The joint receiver-level estimator implies a flatter decay — $\hat{\delta} = 1.55$, half-life 448 m — with the binned value below the lower edge of the design-respecting interval $[266, 820]$ m. The difference reflects exact receiver distances versus bin midpoints, the exact receiver-level gated sample versus the all-merchant bins, and joint weighting across all moments; the joint estimate is the maintained one, and every policy result uses it.

AnyTreated versus ShareTreated. The reduced-form garage design uses binary AnyTreated for randomization-inference power. Garage rosters are small — of the 36 GPS-matched garages with experimental drivers, the median roster has 2 experimental drivers (mean 3.9; the largest, the airport garage, has 30) — so a treated-share variable is extremely coarse, and the realized variation is essentially binary: 30 garages have AnyTreated $= 1$ and 6 have control-only rosters. The decay identification is unaffected by this choice, because δ comes from coefficient ratios across radii regardless of whether the regressor is binary or continuous; ShareTreated regressions show the same sign and monotone decay with less power. The policy machinery uses the same AnyTreated formulation, so counterfactuals never extrapolate to treatment-intensity variation the design cannot measure.

Decay in data units. Converting each gated (Panel B) garage coefficient to the latent scale, subtracting learning, and dividing by the phone-network baseline yields amplification factors $w(r)$ of roughly 24× at 50 m, 27× at 100 m, 15× at 300 m, and 11× at 500 m: near a garage, the competitive force is an order of magnitude stronger than full community-network exposure. An exponential fit to these factors has a data-side half-life of 400 m, close to the structural 448 m; $w(r)$ combines the exponential kernel with variation in baseline adoption across distance bins, which the marginal-effect conversion translates into additional distance dependence. Setting $|\beta_{\text{comp}}|e^{-\delta d^*} = \beta^s_{\text{learn}}$ at the point estimates gives $d^* \approx 1.3$ km — the distance at which competition has decayed to the size of the strong-tie learning response — placing the crossover at walking-to-driving scale rather than city scale. With the 750 m and 1 km bins targeted, the profile is disciplined by measurement through 1 km, so the crossover involves only a modest extrapolation beyond the design radius. Fitting the six inner bins alone yields a steeper decay (half-life 339 m); the outer bins, individually significant under randomization inference, reject that extrapolation and stretch the half-life by roughly a third.

D.4 Within-taxi identification from the reciprocity cells

The four cells. The within-taxi block is identified by the four within-taxi reciprocity cells at the $\geq 5\%$ threshold (coefficients per 1,000): non-mutual Dakar -0.81 [$p_{RI} = .017$], mutual Dakar $+0.39$ [$.885$], mutual outside Dakar $+2.60$ [$.001$], non-mutual outside Dakar -0.15 [$.015$]. All four are rebuilt from the corrected best-phone network with strict trimming in both layers and identity-asserted in code; Table B5 reports the mutual layer (untrimmed, where reciprocal saves preclude hub contamination), and the trimmed values here are the estimation vintage. The estimation targets the two mutual cells; the non-mutual cells are not targeted, with the outside-Dakar one serving as the held-out test of $\beta_w^{\text{taxi}} = 0$.

Identification argument. Treated drivers operate in Dakar, so receivers outside Dakar face no market-proximate treated competitor: the outside-Dakar mutual cell measures within-industry learning through strong ties free of competition and pins down β_s^{taxi}. The Dakar mutual cell then identifies c as the proximity cost that offsets that learning; under the restriction $\beta_w^{\text{taxi}} = 0$, the same c applies to non-mutual ties, which carry the cost without the learning. The model then predicts the held-out fourth cell at exactly zero (no learning, no competitor); the observed -0.15, though precisely estimated, is resolved by the location diagnostic below as outcome mislabeling rather than transmission.

Rejected one-parameter alternatives. Two natural simplifications fail on the cells' own arithmetic. A common proximity cost imposed directly in reduced-form units forces $\beta_w^{\text{taxi}} = +1.41$ per 1,000 and predicts the held-out cell at $+1.41$ against the observed -0.15. A cost proportional to tie strength implies $\beta_w^{\text{taxi}} = -5.4$ per 1,000 — negative learning. In index units, by contrast, the two estimated costs are statistically indistinguishable: the reduced-form gap between the Dakar cells reflects baseline adoption rates and degree composition, which the logistic index and sample-specific baselines absorb. The economical description of the estimates is one learning channel that requires tie strength and one tie-borne competition cost that requires market proximity.

Location diagnostic: the fourth cell and the geography of the outside margin. The outside-Dakar outcome is defined by the merchant registry's region field, which records where a business registered, not where it operates. Locating the 1,658 window adopters with registry region outside Dakar by their agent-GPS transactions (April–September 2022), 91 percent transact predominantly within 25 km of Dakar — and only 171 have any

transaction presence around Thiès city — so the margin is dominated by peri-urban operators in the Dakar market. Re-estimating the outside cells on genuinely outside adopters: the non-mutual cell falls from -0.147 [.015] to -0.003 [.83] (modal-location rule) or $+0.008$ [.56] (any Thiès-presence rule) — the model's zero prediction, confirmed — while the mutual cell keeps its sign with a comparable proportional effect ($\approx +45\%$ of its baseline) on 53–92 adopters, its significance limited by power alone. Two implications are carried through the text: the $\beta_w^{\text{taxi}} = 0$ restriction holds exactly on correctly measured outcomes, and the targeted mutual cell's population is predominantly the peri-urban ring, so β_s^{taxi} is a lower bound on frictionless within-industry learning (ring-level competition, if any, is absorbed into it). A small residual negative on genuinely outside weak ties would also be consistent with the model: Thiès is some 60 km from Dakar and intercity routes are contested by both cities' taxis, so the Dakar-only competition dummy is an approximation rather than a sharp boundary.

Relation to the spatial competition channel. The tie-borne cost c and the garage kernel $\beta_{\text{comp}} e^{-\delta d}$ are distinct channels with distinct exposure variation: the kernel loads on garage-level treatment at measured receiver–garage distances, the tie costs on the treated share of a receiver's own ties. Reconstruction of the full within-taxi band coefficient from the garage kernel alone under-predicts the observed negative response, which is the direct evidence that part of within-taxi competition travels through ties rather than through space; the tie-cost parameters absorb exactly that component.

D.5 Baseline coherence: what the data impose on the model's structure

Two structural choices are decided by explicit tests rather than convention.

A single common taxi-side baseline is empirically inadequate. As a diagnostic restriction, I require the customer-demand and garage samples to share one taxi-side intercept, so that the model must explain both exposure responses and unconditional adoption levels with a common baseline. The criterion rises from $Q \simeq 0.10$ in the preferred specification to $Q \simeq 35.3$, and the competition parameters move to a suppression boundary: the broad taxi-network baseline rate is far below the adoption rates in the selected garage and customer-linked risk sets. I therefore treat baseline levels as sample-specific nuisance calibrations (each tied to an exactly matched sample-level moment) and interpret the common parameters as structural exposure-response coefficients. The model is structural at the level of the network-exposure-to-adoption mapping used for counterfactual assignments, but these coefficients are not interpreted as deep primitives of preferences,

technology, network formation, or a fully specified equilibrium adoption model. This empirical restriction exercise is not reported as a formal likelihood-ratio or χ^2 test.

Band-specific phone baselines. I test whether the strong- and weak-band phone samples share a baseline. Imposing a shared baseline moves the weak-band learning coefficient by 38% (well beyond my 10% pre-set tolerance), so band-specific baselines are used.

D.6 Design-respecting parameter uncertainty

Procedure. The stored placebo assignments give an empirical distribution of the moment vector under re-assignment of treatment. For each stored placebo assignment I form a pseudo-moment vector $m_b^* = m_{\text{obs}} + (m_b - \bar{m})$, recentering the placebo deviations at the observed moments, and re-estimate the model on m_b^* with the weight matrix held fixed. The re-estimation uses a first-order (GMM-update) step around the point estimate applied to all 991 jointly complete draws, validated against full nonlinear re-estimation with re-profiled nuisance baselines on 150 draws: the maximum absolute discrepancy between the two is about 0.01 for the learning, demand, and tie-cost parameters — negligible against interval widths of roughly 0.4 — and larger for the kernel parameters $(\beta_{\text{comp}}, \eta)$, whose mapping from moments is the most nonlinear; I therefore read the half-life interval as approximate.

Interpretation. The stored placebo draws re-execute the experiment's stratified assignment mechanism, which assigns treatment probabilistically within strata, so the treated-business margin varies across draws as it does under the design: the moment perturbations are draws from the randomization distribution itself. The construction propagates that design uncertainty through the estimator as percentile intervals rather than by test inversion, and I label the results *design-respecting empirical intervals based on the stored placebo assignments*. They quantify joint parameter uncertainty only; assignment variation (Section D.7) and specification uncertainty (the two phone specifications) are reported separately and never combined into a single standard error.

Results. Across all re-estimates: $\beta_{\text{comp}} < 0$ and $\beta_{\text{demand}} > 0$ in every draw; $\beta_s^{\text{taxi}} > 0$ in every draw and both tie-borne costs are jointly positive in 99.5% of draws; the competition half-life $\ln 2/\delta$ has point estimate 448 m with interval $[266, 820]$ m; and the ranking between β_{learn}^s and β_{learn}^w is not identified — the probability that the strong-band response exceeds the weak-band response is 0.32 in the primary specification, so no tie-strength ranking is interpreted anywhere in the paper. On the 150 fully re-estimated draws, the

empirical tail fraction $\Pr(Q_b^* \geq Q_{\text{obs}})$ is 0.66. Thus the observed moment vector sits comfortably within the bulk of the perturbation distribution. The low tail position is consistent with the perturbation distribution overstating moment variability: the stored draws vary slightly in their treated-business margins, and the diagonal weights ignore the strong cross-radius correlation of the garage moments. I treat this only as a design-respecting overidentification diagnostic, not as an exact J-test.

D.7 Policy evaluation

Targetable universe and budget. The policy universe is the 1,454 matched taxi businesses (856 treated and 598 control under the realized assignment). Every policy treats exactly 856 businesses. A business is indivisible: selecting it activates all of its matched stakeholder records (the realized assignment activates exactly 1,256 records). These counts are asserted programmatically for every assignment evaluated.

Receiver populations. Policies are evaluated on a support-preserving receiver master assembled from the same network definitions used in estimation. Phone-network receivers form the cross-industry learning population; the full pair-specific customer-linked merchant universe (8,523 merchants, with the embedded-link indicator fixed at its observed value) forms the demand risk set; the exact gated Panel-B garage receivers not already in the demand set form the spatial-competition risk set; and the within-taxi reciprocity contacts form the tie-channel risk sets — mutual-tie contacts (with the Dakar and outside-Dakar taxi-adoption margins as disjoint outcome events) and non-mutual contacts in Dakar. Within-taxi contacts already counted by the demand or garage groups are excluded from the tie-channel groups, and non-mutual contacts also present in the mutual panel are assigned to the mutual group only, so no receiver's taxi outcome is counted twice. Receivers appearing in more than one layer are reconciled before evaluation and enter the master only once, with their applicable channel indicators retained. The final row-level master contains 3,024,017 unique receivers under the primary specification, assembled from its own trimmed strong/weak receiver unions plus the demand, garage, and tie-channel risk sets. For aggregate policy value each receiver is counted once: cross-industry adoption is defined as adoption in either geographic margin (the two margins are mutually exclusive in the data), and the Dakar and outside-Dakar outcomes are reported separately as descriptive margins. Experimental subjects are excluded. Customer links are defined as in the demand experiment: a merchant–driver pair is linked if they share at least one qualifying customer, i.e., a customer with at least two P2B transactions across non-garage experimental drivers between November 2022 and November 2023.

Receivers reachable by none of the estimated channels have assignment-invariant predicted adoption and therefore drop out of policy differences. No estimated coefficient is applied to a broader *network definition* than the one used to discipline that coefficient. The learning coefficients are nevertheless transported geographically: they are estimated on the Dakar cross-industry phone samples and applied to outside-Dakar receivers on the same strict phone network. I treat this as a maintained transport assumption rather than as additional identification.

Policy baselines. For each policy risk-set group g, a baseline α_g^{pol} is calibrated numerically so that the model's mean adoption probability under the *realized* assignment equals the group's observed adoption rate at the estimated ψ (residuals below 10^{-12}). The primary partition distinguishes cross-industry Dakar and outside-Dakar receivers and, on the taxi side, risk sets defined by demand and garage eligibility. Estimation-sample nuisance baselines are never transported to the policy stage. Under parameter uncertainty, the policy calibration is repeated within each parameter draw before alternative assignments are evaluated.

Policies. Evaluated deterministic rules, all at the 856-business budget: *phone reach* (rank businesses by unique receivers contacted through the strict phone network, union over the business's records and both tie bands, ties broken by business identifier); *greedy coverage* (lazy-greedy maximum coverage of unique receivers); *top mobile-money users* (rank by pre-treatment (Nov. 2017–Feb. 2022) mobile-money activity); and *garage-concentrated* (all garage-member businesses, with the remaining slots filled by 200 documented random draws, reported as the mean over fills). A geographic-reach rule is excluded because no verified receiver-side geography exists; I do not construct proxy geography. No targeting score uses outcome data.

Benchmark and estimand. The benchmark is the *exact-budget random business targeting benchmark*: 2,000 uniform random draws of 856 of the 1,454 businesses, each evaluated exactly. It is not an experimental rerandomization: the budget is held fixed at exactly 856 for comparability across rules, while the experimental mechanism leaves the realized budget random. For any assignment A, the reported estimand is induced adoption $V(A) - V(A_\emptyset)$, the model's expected number of adopters under A minus the no-treatment counterfactual, summed over the policy receiver populations; negative competitive effects are never truncated. Policy gains are stated relative to the mean of exact-budget random business assignments.

Uncertainty propagation. All deterministic policies, the zero assignment, 50 of the 200 garage fills, and a fixed 250-assignment subset of the random bank are re-evaluated at each of the 991 placebo parameter draws, with policy baselines re-calibrated within each draw. Deterministic policies are evaluated directly. For the random subset I use a compressed evaluation based on 101-bin exposure histograms (including one-dimensional histograms for the within-taxi tie exposures, whose indices are linear in a single exposure), exact evaluation of the small demand–garage overlap group, and exact per-garage sufficient statistics. The only approximation is the 0.01-wide exposure binning: against exact evaluation of the same 250 assignments at the point estimate, it changes the implied random mean by 2.6 induced adopters, or 0.078% of the headline reach-targeting gain. The point-estimate random benchmark itself uses all 2,000 assignments evaluated exactly. Assignments are held fixed across parameter draws; policies are never re-optimized within draws, so the intervals describe uncertainty in the value of a fixed rule.

D.8 Channel decomposition

Three complementary decompositions of the reach-targeting gain are computed. *Knockouts:* each channel's coefficient is set to zero, singly and in pairs, and all policies are re-evaluated. For each channel configuration the policy baselines are *re-calibrated under the realized assignment*; thus the decomposition isolates assignment-induced exposure responses while holding each risk set's realized mean adoption fixed, rather than mechanically changing its baseline level. *Shapley decomposition:* the gain of phone-reach targeting over the random mean is attributed to the four channels — cross-industry learning ($\beta^s_{\text{learn}}, \beta^w_{\text{learn}}$), customer demand ($\beta_{\text{demand}}, \beta_{DI}$ jointly), spatial competition (β_{comp}), and the within-taxi tie block (β^{taxi}_s, c) — by averaging marginal contributions over all 24 activation orderings, using the same baseline-recalibration convention. The attribution is exactly additive. Results: the learning contribution alone slightly exceeds the total reach-targeting gain (102%); spatial competition contributes $+2.5\%$ (reach targeting spreads treatment spatially, slightly reducing displacement relative to random); the demand and within-taxi tie contributions are -3.5% and -1.2% — both channels raise adoption under every assignment, so they contribute levels, not gains. Garage concentration's shortfall relative to random decomposes into stronger local displacement and weaker phone reach. Maximizing adoption counts a displaced adoption and a created one symmetrically; the policy ranking does not hinge on this convention, since learning accounts for essentially the entire reach-targeting gain and the competition contribution to that gain is small and positive.

D.9 Robustness

Customer-network density counterfactuals. The measured customer network conditions on digital payment at both ends of each link within the transaction window, so it is a lower bound on the true shared-customer graph. I re-evaluate all policies in a doubled-network world ($2\times$ the measured customer links), treating unobserved linked merchants as responding like observed ones (the per-share response transports under missingness unrelated to treatment) and rescaling the random benchmark identically. At measured density the customer-reach rule is indistinguishable from random ($+2.1\%$) because linked merchants connect to many businesses and their exposure barely depends on placement; in the doubled world it earns $+13\%$, phone reach earns $+44\%$, and the blended phone-plus-customer ranking is best in both worlds ($+59\%$ and $+56\%$). Steeper densities continue the pattern (customer-reach overtakes pure phone reach only beyond $4\times$); the full grid is in `demand_extras_psi9.csv`.

The learning channel's contribution to each rule is invariant to the density scaling — learning operates on phone receivers, whose exposure is unrelated to customer-network density — so its *share* of each rule's total falls with density while its contribution to the targeting *gain* does not:

	Learning-channel induced	Learning share of total induced	
		measured ($1\times$)	doubled ($2\times$)
Random mean (benchmark)	4,912	87%	70%
Phone reach (degree)	8,190	93%	81%
Greedy coverage	8,066	93%	81%
Phone + customer blend	7,587	85%	69%
Top mobile-money users	6,253	90%	76%
Customer reach	4,210	73%	53%

At every density, the degree rule's advantage over random remains the same $\approx 3{,}280$ learning-channel adopters; density scaling changes the level demand contributes under any placement, not the gain learning delivers under a targeted one.

Phone specification. All policy results are recomputed under the business-collapsed phone specification, re-estimated and re-propagated separately, and evaluated on the receiver universe generated by its own business-level collapse and trimming rule (2,121,412 cross receivers, against 1,863,726 under the primary specification). The policy ranking is identical (reach first with probability 0.98; garage last with probability 0.99), with a degree-targeting gain of $+83.9$ percent against $+80.2$ percent under the primary specification. For this specification I emphasize gain counts rather than percentage gains when the

draw-specific random-mean denominator approaches zero. Figure A12 summarizes the mean gains and their 95% design-respecting parameter intervals for all headline policies.

Baseline stratification. Policy baselines are re-calibrated under a coarser stratification that merges the taxi-side risk sets. Induced *levels* shift materially — levels are calibrated objects — but the gains of the headline policies relative to the random benchmark retain their signs, ordering, and magnitudes.

Cost-denominated budget. As a robustness exercise, I denominate the budget in the program's 1,256 matched stakeholder records. Businesses remain indivisible and cost the number of matched records activated when selected. The implemented reach rule uses a deterministic greedy score-per-record packing rule that exhausts the record budget; it selects 803 businesses and gains $+78.1\%$ relative to the mean of documented random feasible cost-fills. I treat this as a budget-robustness heuristic, not as a claim that the packing rule solves a globally optimal knapsack problem.

Support audit. For every policy I report the distribution of each exposure on its receiver population. Reach-targeted policies roughly double the mass of strong-band exposure at full saturation (0.48 of the receiver population, versus 0.27 under the realized assignment); such cells exist in the realized data, so the policies are classified as mild extrapolation along the estimated exposure-response curve rather than evaluation outside the support class. Garage-concentrated targeting stays within the realized support, and no garage exposure exceeds the 500 m design radius by construction. Figure A11 shows the level of the random benchmark across the placebo draws; policy gains are computed against the draw-specific benchmark, so this level variation cancels within each draw.

D.10 Validation against untargeted survey moments

Sign predictions. The model correctly predicts the sign of all four untargeted survey moments and their 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 $+0.098$, effect size $+36.1\%$. This moment is now a *quantitative* out-of-sample prediction. Survey friends sit at the top of the validated overlap distribution — strong ties — and in the customer network taxi friends share regular customers

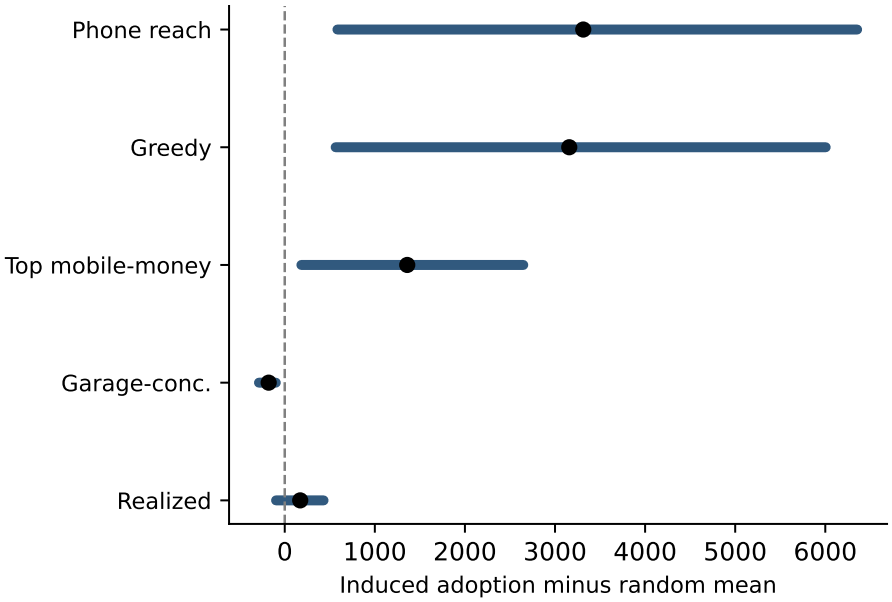


Figure A12: Mean policy gains relative to the mean of exact-budget random business assignments, with 95% design-respecting parameter intervals (1,000 placebo draws).

with the respondent far less often than taxi non-friends (9% vs. 24% of pairs; mean shared customers 4.46 vs. 23.55): market-distant pairs. The model therefore prices a friend tie at β_s^{taxi} with no proximity cost, an effect size of +30.0% $[+12.3,\ +47.4]$ against the observed $+36.1\%$ $[+9.1,\ +63.1]$ — a survey coefficient reproduced from administrative cells, with β_s^{taxi} identified in a different city.

- 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: the strong- and weak-band learning responses are of similar magnitude (their ranking is not statistically identified), so learning operates across the community-tie strength distribution, and 28% of survey "serendipitous" contacts are community ties by the corrected-overlap measure (Table B1); the model therefore 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 for the remaining moments. For m8, m10, and m11, model-implied responses evaluated at representative population-mean exposures remain an order of magnitude smaller than the survey coefficients, so those three are read as sign tests.

Two 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-reported contacts are the right tail of tie strength, not the population average that identifies the band parameters. The taxi-friends moment escapes both problems because the within-taxi block supplies the right comparison: a strong-tie effect size measured on a strong-tie administrative sample.

D.11 What the reported intervals do and do not contain

The reported intervals answer one question: how much the estimated parameters would move had the experiment assigned treatment to different businesses. Everything else is held fixed, and several sources of uncertainty are therefore deliberately outside them. Variation across counterfactual assignments is reported separately, as the random-bank distribution, and the choice of phone specification is not averaged over — the two specifications appear side by side throughout. The functional form is maintained rather than tested: the logistic index, the six-trial horizon, and the exponential shape of the decay. So is the measurement of the networks. Receivers can be mismatched, the customer network is built from post-treatment transactions, and the learning response estimated on cross-industry contacts is assumed to apply wherever those contacts are located. The baseline stratification is examined in the robustness exercises above, and the placebo perturbations approximate the sampling distribution only imperfectly, since the original assignment protocol cannot be fully reconstructed. Finally, the exercise is static. Counterfactuals hold the network and the parameters fixed: adopters induced by a policy do not themselves become senders, and no competitive or general-equilibrium response to the policy is modeled.