Open Banking and Credit Allocation*

Emily Breza† Bernardus Van Doornik‡ Maryam Farboodi§
Dimas Fazio¶ Janis Skrastins‖

This version: June 16, 2026
Preliminary — please do not circulate

Abstract

We study how giving customers control of their financial data reallocates credit. Brazil's Open Finance lets households and firms share the cash-flow records that digital payments leave in their bank accounts with any lender. In Central Bank of Brazil administrative data, exploiting institutions' staggered entry into the regime, individuals gain access to credit when their bank joins, while firms—which rarely consent—lose it. Within borrower, credit shifts toward the bank that stands to learn the most about her cash flows—by 4–7 percentage points of her total credit—while incumbents that already observed those flows lose retail share: data sharing erodes informational rents. The reallocation concentrates in unsecured credit, grows with the informativeness of the payment trail, and accrues to lenders with cheap funding and the technology to use data. Open finance can broaden access and competition, but its gains are gated—by consent, by prior information, and by the capacity to act on data.

JEL codes: G21, G28, G51, O16, D82, L86.
Keywords: open banking, open finance, instant payments, Pix, information sharing, credit allocation, relationship lending, fintech, financial inclusion.

*This draft is preliminary and incomplete; results are work in progress and subject to revision. We thank seminar and conference participants for helpful comments. The views expressed here are those of the authors and do not necessarily reflect those of the Central Bank of Brazil. All errors are our own.
†Harvard University and NBER.
‡Central Bank of Brazil.
§Massachusetts Institute of Technology and NBER.
¶National University of Singapore.
‖University of Utah.

1 Introduction

Banking runs on private information. A long tradition in financial economics holds that what makes a bank special is what it knows: by observing the flows through a customer's accounts, an incumbent lender learns what no outsider can see, prices accordingly, and earns rents on the relationship (Black, 1975; Fama, 1985; Sharpe, 1990; Rajan, 1992). Checking-account activity is among the strongest predictors of default (Mester et al., 2007; Norden and Weber, 2010), and because this information arrives as a byproduct of holding the account, it has belonged, de facto, to the bank. Credit bureaus and registries share slices of it—repayment histories, balances outstanding—and reshape lending where they operate (Pagano and Jappelli, 1993; Djankov et al., 2007; Hertzberg et al., 2011; Sutherland, 2018). But the underlying cash-flow record itself has stayed put.

Two pieces of digital infrastructure are now unbundling that information from the account. Instant payment systems—India's UPI, Brazil's Pix, and dozens of counterparts—have pulled the daily transactions of entire populations into bank accounts, converting cash income into standardized electronic records. Open-banking regimes, adopted by more than forty jurisdictions, give customers the legal right to port those records to any institution they choose (Babina et al., 2025). Together they reassign property rights over banking's core asset: the borrower, not her bank, now owns the data trail her economic life generates (Jones and Tonetti, 2020; Farboodi and Veldkamp, 2025). This paper asks the question that reassignment poses: when borrowers can hand any lender the information that incumbents once monopolized, who lends, to whom, and on what terms?

Theory does not settle the question. Information sharing can erode incumbents' rents, intensify competition, and draw rationed borrowers into the market (Stiglitz and Weiss, 1981; Petersen and Rajan, 1995; Hauswald and Marquez, 2003); it can also soften competition, advantage lenders with the wrong incentives, or stall on consent (He et al., 2023; Parlour et al., 2022). Adjudicating empirically is demanding because it requires observing not only credit but the *information itself*—what each lender knew about each borrower, and how that changed when sharing began. Existing evidence, drawn largely from the United Kingdom and from cross-country adoption, observes neither (Babina et al., 2025).

Brazil lets us observe both. Pix, launched by the Central Bank of Brazil in November 2020, reached near-universal adoption within a few years; together with card settlements, bank wires, and *boleto* collections, it leaves a granular, account-level record of electronic cash flows for essentially the entire adult population and most firms. Brazil's Open Finance regime, phased in from 2021, lets customers port exactly these records across institutions. Participation was staggered: the largest banks were mandated to join in early 2021, while

smaller banks and fintechs certified their data-sharing APIs on a schedule stretching to the present. We obtain the certification dates from the third-party vendor that certifies every participant before it goes live, and we verify on the Central Bank's Open Finance dashboard that data begin to flow within weeks of certification—so certification marks the moment an institution can transmit and receive customer data.

Crucially, we can also measure the information at stake in each lending relationship. Before Open Finance, a bank observed a customer's electronic cash flows only insofar as they passed through its own accounts; afterward, the customer can grant any participant access to the flows observed by *all* of her institutions. For each borrower–bank pair and month we therefore construct a *potential information gain*: the share of the borrower's economy-wide electronic cash flows—aggregated across the income- and revenue-bearing rails—that Open Finance can make visible to the bank if the customer shares, net of the share the bank already saw. Because individual consents are unobserved, this object measures what each lender *stands to learn* about each borrower, and estimates based on it carry an intent-to-treat interpretation. It is the centerpiece of our design. We complement the administrative records with original survey questions on digital-payment use that we fielded in collaboration with SEBRAE, Brazil's micro- and small-business support service.

Adoption is endogenous on every margin: institutions choose when to join, customers whether to consent, and credit comoves with the macroeconomic forces driving digital adoption. We therefore proceed through designs of increasing stringency—bank-level event studies around API certification, within-bank contrasts across products and segments, and a staggered borrower-level difference-in-differences—before settling on a *within-customer* design that exploits multi-homing. Many Brazilian households and firms borrow from several banks, so we can compare, for the same borrower in the same month, the banks that stand to learn much about her to the banks that stand to learn little. Borrower $\times$ month fixed effects absorb her total credit demand, her risk, any common shock—and, most importantly, her endogenous decision of whether and how much to share at all; conglomerate $\times$ month fixed effects absorb each bank's supply conditions, including its decision to join Open Finance and its overall push to use shared data; and borrower $\times$ conglomerate fixed effects absorb the level of each relationship. Identification comes only from *differential* potential information gains across a borrower's banks, holding the borrower fixed. This is the strongest specification in the paper: it controls for the most obvious confounders—demand, risk, consent, and bank supply—and identifies the effect from clean, relationship-level variation in what each bank stands to learn.

The results unfold from the market down. At the bank level, lending expands once an institution certifies its Open Finance API: total credit and the number of client relationships

rise together—by roughly 12 log points in the quarter of certification and 30 log points within five to six quarters—while the value-weighted lending rate falls, all on flat pre-trends. Quantities up and prices down is the signature of an outward shift in credit supply, consistent with joining banks both using the data they receive and competing harder for customers they can now screen. Because banks choose when to certify, we treat these event studies as descriptive; they set the aggregate stage, and identification comes from the designs below.

At the borrower level, the gains are real but uneven. When a borrower's own primary bank joins Open Finance, individuals become more likely to hold credit—an extensive-margin expansion of roughly 4–5 percentage points within six months, on flat pre-trends—and, conditional on borrowing, they borrow more. Firms move the other way: the probability that a firm holds credit *falls* by 2–3 percentage points. The asymmetry tracks the consent architecture. Roughly 30% of retail customers consent to share their data, against about 2% of firms—a gap reflecting more demanding consent processes, weaker perceived benefits, and strategic concerns about revealing revenues. Whatever the mix of causes, the opt-in design of open finance determines whom the regime reaches. All in all, the evidence suggests that banks tilt their lending portfolios away from corporate and toward retail credit, where information acquisition has become cheaper—a shift that may carry unintended consequences for productive investment in the economy.

Our central finding comes from the within-customer design, which asks *which* bank captures a borrower's credit. Holding the borrower fixed, credit shifts toward the bank that stands to learn more about her cash flows: moving a relationship from no potential information gain to a full gain raises the bank's share of the borrower's credit by 4.7 percentage points for individuals, 6.8 for small (*Simples*-regime) firms, and 4.3 for larger firms over seven quarters, on flat pre-trends. Because the borrower $\times$ month fixed effects absorb total demand, these estimates isolate *reallocation across lenders*, not changes in total borrowing: within a borrower, information moves credit from the banks that knew her to the banks that now can.

To interpret our empirical finding, we develop a model of credit-market competition with data sharing. A borrower of uncertain quality spreads her borrowing across differentiated banks. Each bank's screening precision increases with the share of her cash flows it observes—not necessarily handles, scaled by the informativeness of the payment trail and by the bank's information technology; better screening lowers the cost of lending, a channel that is strongest for unsecured credit and vanishes for fully collateralized products. Open finance raises every participating bank's observable share of the borrower's flows to the pooled total, so each bank's cost falls with its potential information gain. The model delivers insights that organize the analysis: information gains reallocate the borrower's credit toward high-gain banks;

effects concentrate in unsecured credit, scale with the informativeness of the trail, and accrue only to lenders with the funding and technology to act on data.The framework is deliberately stylized, a richer treatment of strategic consent and equilibrium pricing is the subject of ongoing work, but it is enough to map each empirical exercise to a comparative static.

The reallocation sits exactly where theory says information is most valuable. The effect is concentrated in relationships where the bank previously observed *none* of the borrower's flows; as the bank's prior information rises, the effect declines monotonically; and for individuals it turns *negative* at high levels—banks that already saw most of a retail customer's flows lose share once those flows become visible to rivals. An incumbent's advantage rests not on what it knows but on knowing more than its competitors, and open finance competes that advantage away. This is the erosion of informational rents at the heart of relationship-lending theory (Sharpe, 1990; Rajan, 1992; Hauswald and Marquez, 2003), measured here directly at the relationship level; Sutherland (2018) documents the analogous pattern when lenders join a credit bureau.

The composition of the reallocation pins down the mechanism. The credit that moves is overwhelmingly *unsecured*, where screening matters most; effects on strongly collateralized products, where collateral substitutes for information, are negligible. The response is larger where the payment trail is more informative: firms in Pix-intensive industries, whose revenues the data capture best, see smaller contractions and larger intensive-margin gains—direct evidence that the payment system and the sharing regime are complements (Ghosh et al., 2026). And information pays only for lenders equipped to act on it: tracing potential flows from data senders to receivers, lending-share gains are largest for deposit-taking institutions, with their low cost of funds, and for institutions that had invested in information technology before Open Finance, while non-banks—mainly fintechs—receive far more data through Open Finance than they transmit.

We probe the leading threats to this interpretation. The borrower's endogenous consent decision and any borrower-wide shock are absorbed by borrower $\times$ month fixed effects; a bank's decision to join, and its contemporaneous appetite for lending, are absorbed by conglomerate $\times$ month effects; and comparisons are made within join-timing cohorts. The remaining variation—which of a borrower's banks stands to learn more, determined by the historical geography of her payment flows—would have to correlate with relationship-specific demand shocks timed to the bank's certification to confound the estimates; flat pre-trends offer no sign of this. The results are also robust to alternative flow measures and to heterogeneity-robust event-study estimators.

Taken together, the evidence suggests a consistent answer to the opening question. Information sharing broadens access and intensifies competition—but its incidence is governed

by three gates: who consents (individuals far more than firms), who previously held the information (informed incumbents lose their rents), and who can act on it (lenders with cheap funding and the technology to use data).

We contribute, first, to the empirical literature on *open banking*. Theory shows that borrower-owned data can either intensify or soften credit-market competition depending on how they shift screening advantages (He et al., 2023), and existing evidence links open-banking policies to fintech entry using cross-country and UK variation (Babina et al., 2025; Ma, 2024). We provide loan-level, economy-wide evidence from a major emerging market, and—unusually—we measure the information itself: the interaction of open finance with a comprehensive payment record lets us observe how much each lender stands to learn about each borrower, not merely that a regime exists, which in turn shifts the empirical question from *whether* open banking matters to *how*.

Second, we contribute to work on *payments, data, and credit*. Digital-payment and digital-footprint data predict repayment and expand access (Berg et al., 2020; Björkegren and Grissen, 2020; Alok et al., 2024; Ghosh et al., 2026), and payment technologies have real effects on households and firms (Suri and Jack, 2016; Higgins, 2024; Crouzet et al., 2023; Dalton et al., 2024; Agarwal et al., 2025; Burga et al., 2025). Recent work on Pix examines bank liquidity (Ding et al., 2025) and deposit competition (Sarkisyan, 2024). We show that the payment system and the information-sharing regime are *complements* in the credit market: the payment system creates the data, the sharing regime allocates them, and the credit response scales with the informativeness of the trail. Where that literature measures what payment data *predict*, we measure what happens when the data *change hands*.

Third, we speak to the theory of *information in lending* (Sharpe, 1990; Rajan, 1992; Petersen and Rajan, 1994; Pagano and Jappelli, 1993; Hauswald and Marquez, 2003; Liberti and Petersen, 2019; Parlour et al., 2022) and to evidence on information-sharing institutions (Djankov et al., 2007; Hertzberg et al., 2011; Sutherland, 2018). Where credit bureaus share repayment outcomes, open finance shares the cash-flow record itself— information closer to the soft knowledge that once defined relationship lending, now rendered hard and portable (Liberti and Petersen, 2019), and central to how uncollateralized credit is underwritten (Lian and Ma, 2021). Our within-customer estimates provide, to our knowledge, some of the first direct evidence that changes in a specific lender's information about a specific borrower reallocate credit toward the lender that learns—and away from the incumbent that no longer knows more.

Section 2 describes the institutions, Section 3 the data and our information measures, and Section 4 documents who uses Pix and how intensively. Section 5 lays out the empirical designs and Section 6 presents the results. Section 7 develops a model to interpret the

empirical findings and Section 8 discusses implications, and Section 9 concludes.

2 Institutional Background

2.1 Cash-flow data: the rails behind a bank account

Money moves into and out of a Brazilian bank account over a small set of electronic rails, each of which leaves a time-stamped, counterparty-identified record at the account-holding institution. Pix, the instant payment system, settles account-to-account transfers in seconds through the BCB's SPI. Bank wires (TED) settle through the Reserve Transfer System (STR). *Boleto* (bank-slip) collections—the workhorse of bill and invoice payment—and legacy DOC transfers clear through the private clearing house (CIP). Credit- and debit-card sales reach the merchant's account as card settlements. Together these rails carry essentially all electronic income and revenue: wages, sales receipts, invoices, transfers. Our information measures aggregate *all* of them—not Pix alone—into a single ledger of each agent's electronic cash flows.

Pix nonetheless deserves its prominence, because it is the rail that made the ledger nearly universal. Launched by the Central Bank of Brazil (BCB) in November 2020, Pix enables free (for individuals) or low-cost, 24/7 transfers between any two participating institutions, using simple aliases ("keys") such as a phone number or tax ID; participation by large institutions is mandatory, which delivered immediate interoperability. Within a few years Pix reached near-universal coverage of the adult population and a large majority of firms, displacing cash and—more slowly—cards (Argente et al., 2025; Higgins, 2024; Crouzet et al., 2023). By pulling formerly unrecorded cash transactions onto an electronic rail, Pix converted the cash economy into data.

The result is a granular, economy-wide ledger of income and spending—exactly the inputs underwriting models value, and data of the kind shown to predict repayment and expand access for thin-file borrowers (Berg et al., 2020; Björkegren and Grissen, 2020; Ghosh et al., 2026; Alok et al., 2024). The rails thus did more than move payments: they created an information asset, held piecemeal by the institutions where each agent's accounts happen to sit.

2.2 Open Finance: who can see the data

A payment record is only as useful as the set of parties that can act on it. By default, the cash-flow records of Section 2.1 are observed only by the institutions holding the accounts on either side of a transaction. Brazil's Open Finance regime (originally "Open Banking," rebranded

"Open Finance" as its scope widened) governs how that information can be shared more broadly. The regulatory framework was issued in May 2020 and implemented in phases. In February 2021, the largest institutions were *mandated* to expose standardized APIs through which their customers could, upon consent, share registration data, account and transaction data, and credit-product data with other institutions. Smaller institutions and fintechs could join voluntarily, and many did, on a staggered schedule that continues to the present; payment initiation was added in late 2021–2022, and the data perimeter has since expanded toward investments, insurance, and foreign exchange. By 2026, hundreds of institutions participate.

Three design features matter for our analysis. First, sharing is *customer-permissioned*: data move only when the customer—an individual or a firm—grants consent, and the customer chooses what to share and with whom. This makes Open Finance a richer but more demanding regime than a traditional credit bureau, and it places the consent decision—and the frictions around it—at the center of incidence. Second, participation is *interoperable and standardized*: certified APIs make the shared data comparable across institutions, lowering the cost of using "hard" information (Liberti and Petersen, 2019). Third, adoption is *staggered across institutions*, which is the source of our identifying variation.

Individual adoption has been substantial: the conversion rate of initiated consents rose from roughly 30% in 2023 to over 55% by early 2026, with more than one hundred million unique individual consents. Firm adoption, by contrast, has lagged dramatically—on the order of one to two million unique firm consents—reflecting a more demanding consent process, weaker awareness and perceived benefits, privacy and strategic-reporting concerns, and ambiguity about who may consent on behalf of a multi-owner firm. An ongoing randomized evaluation finds that simple informational treatments raise firm adoption by several percentage points, concentrated among younger, more educated owners and those concerned with privacy (Barboni et al., 2025). This retail–firm asymmetry in consent will recur throughout our results.

2.3 Join dates and the timing of information transmission

Identifying the effect of Open Finance requires knowing *when* each institution actually began sharing and receiving data. We define an institution's join date as the date on which its APIs were certified by Raidiam, the third-party vendor that all participants must clear before going live; Raidiam provided these dates to us directly. We cannot observe individual consents or the content of transmitted data in BCB administrative records. However, using the BCB's Open Finance dashboard (available from 2023), we verify that aggregate data

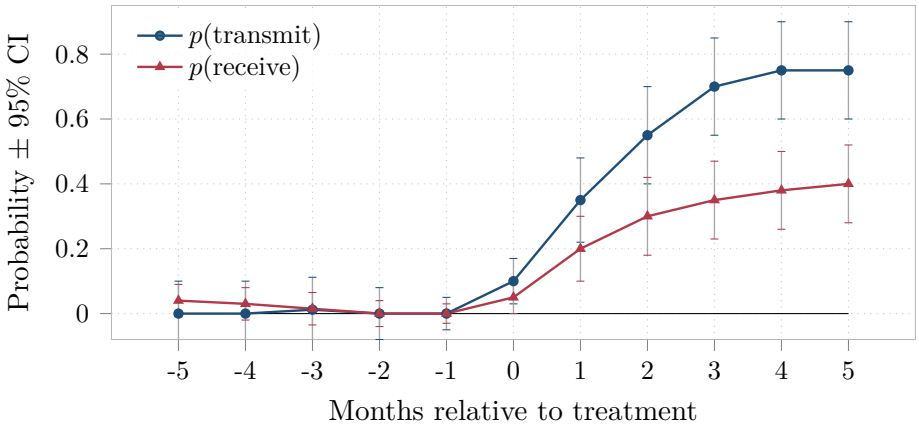


Figure 1: Data transmission begins shortly after API certification. Event-study of Open Finance data transmission/receipt around the month of API certification (normalized at -1), constructed from the Central Bank of Brazil's Open Finance dashboard. Data flows rise sharply once an institution certifies its API, validating certification dates as the onset of participation.

transmission and receipt for an institution rise sharply within weeks of API certification (Figure 1). Certification dates are therefore a sharp, administratively recorded proxy for the onset of an institution's participation. Following the empirical pattern of joining, we group institutions into waves—Wave 1 (early 2021, dominated by mandated banks), Wave 2 (late 2021–2023, including large fintechs), and later joiners—and we use both the wave structure and the exact certification month in our designs.

3 Data and Measurement

3.1 Sources

We combine several administrative datasets maintained by the Central Bank of Brazil (BCB), linked at the level of anonymized individual (CPF) and firm (CNPJ) identifiers and aggregated to the level of *financial conglomerates* (we treat each banking group as a single decision-maker, consolidating entities under a common parent).

Credit registry (SCR). Credit outcomes come from the credit registry (*Sistema de Informações de Crédito*, SCR), the BCB's loan-level supervisory dataset. For every credit operation in the country above a reporting threshold of R$200—effectively, every meaningful credit contract—the SCR records monthly the outstanding balance, the product type and its collateralization (from unsecured products to fiduciary-lien instruments), interest rate, maturity, arrears, and the identities of borrower and lender—so we observe, for each borrower, the full portfolio of her relationships and whether a lender is a new or an incumbent

one. Our principal outcome is the active portfolio balance; we also use the value-weighted interest spread, new originations, and the secured/unsecured split.

Electronic cash flows. To capture the information that Open Finance makes shareable, we use a consolidated cash-flow base recording, by account holder and month, *all* electronic flows passing through a bank account, by rail (Section 2.1): instant payments (Pix, received and sent, drawn from transaction-level SPI records), credit- and debit-card settlements, bank wires (TED, settled through the STR), interbank clearing-house transfers including legacy DOC (CIP), and *boleto* collections. Aggregating these rails—rather than Pix alone—captures the income- and revenue-bearing flows of each agent and lets us measure both the informativeness of a customer's cash-flow trail and which institution observes which flows. Transaction-level Pix records additionally provide counterparty counts and the split between business and personal accounts used in Section 4.

Banking relationships (CCS). The registry of client relationships (*Cadastro de Clientes do Sistema Financeiro*, CCS) records, in current and historical files, which institutions each customer holds accounts with and since when. From it we define the set of potential lending relationships, incumbency, multi-homing, and each customer's "primary bank."

Open Finance participation. Our Open Finance variables come from the API-certification dates of each conglomerate, provided by Raidiam—the third-party vendor that certifies every participant—together with the aggregate data transmission and receipt series published on the BCB Open Finance dashboard (from 2023), which we use to validate the certification dates and for the sender–receiver information-flow decomposition.

Individual and firm characteristics. Individual and firm registries provide demographics (such as owner age), firm attributes including *MEI* (individual micro-entrepreneur) and *Simples* (simplified-tax-regime) status and sector, and ownership links between firms and their individual owners, which we use to classify firms by size and formality. For employment outcomes and workforce characteristics we use RAIS, the Ministry of Labor's matched employer–employee registry.

Firm survey (in collaboration with SEBRAE). Finally, we designed survey questions that SEBRAE—Brazil's micro- and small-business support service—fielded on our behalf in its *Pulso das Empresas* survey of micro and small enterprises (SEBRAE, 2025). The questions cover how firms collect their sales (Pix, cards, cash), whether Pix is accepted

through a business account or the owner's personal account, and owner characteristics such as age, gender, and education. These data underlie the adoption facts of Section 4.

3.2 Sample and panel

Because the linked universe is very large, we draw random samples of individuals and firms and build a borrower × conglomerate × month panel covering the pre- and post-Open Finance period. The unit of observation is a potential lending relationship in a month. We also collapse the panel to the borrower × month and conglomerate × month levels for the borrower-level and bank-level designs, respectively. We split all analyses into three borrower types: individuals, *Simples* (small) firms, and larger firms.

3.3 Measuring the potential information gain

Our key innovation is to measure, for each borrower–bank pair, how much information about the borrower's electronic cash flows a bank *stands to gain* when Open Finance comes online. For borrower i, conglomerate b, and month t, define s_{ib}, the *prior information share*, as the fraction of i's economy-wide electronic cash flows that were received through accounts at b around the time b joined Open Finance, aggregating the same electronic rails described above—a comprehensive measure of the income- and revenue-bearing flows a bank could observe directly, not merely its Pix activity. We define C_{it}, the *open-finance information stock*, as the cumulative fraction of i's electronic inflows that is observable through the set of institutions that have joined Open Finance (and could therefore be shared). The *potential information gain* of relationship ib is

$$\Delta_{ib,t} \;=\; \Big(C_{it} - s_{ib}\Big) \cdot \mathbf{1}\{i\text{'s first bank and } b \text{ have both joined by } t\},$$

the incremental share of i's flows that b *can* see through Open Finance beyond what it already observed, $\Delta_{ib,t} \in [0, 1]$. The gain is *potential* because BCB administrative records do not reveal whether, or with which banks, i actually shares: $\Delta_{ib,t}$ is what b would learn if she did. Estimates based on $\Delta_{ib,t}$ therefore have an intent-to-treat interpretation and, if anything, understate the effect of realized information transfers. For brevity we refer to $\Delta_{ib,t}$ as the information gain. Our baseline measure aggregates all electronic rails, including credit- and debit-card settlements; we obtain similar results from Pix-only and sent-flow variants.

To test which lenders benefit the most from Open Banking, we bin relationships by the prior information share s_{ib} into ordered groups: *never-joiner*; *no prior information* ($s_{ib} = 0$);

Table 1: Summary statistics *(placeholder)*

	Individuals	*Simples* firms	Larger firms
1(any credit)	—	—	—
Total credit (log)	—	—	—
Credit share (primary bank)	—	—	—
Prior information share s_{ib}	—	—	—
Information gain Δ_{ib}	—	—	—
Pix inflow share of electronic inflows	—	—	—
# relationships per borrower	—	—	—
Observations	183,194,949	33,004,879	3,627,026

Notes: Borrower × conglomerate × month panel from Central Bank of Brazil administrative data. Observation counts correspond to the within- customer estimation sample of Table 2. Remaining cells to be completed.

and increasing levels (e.g., $(0, 0.3]$, $(0.3, 0.6]$, $(0.6, 1)$, and $= 1$). We interact information gain with these prior-information groups.

3.4 Outcomes

At the relationship level our main outcome is the borrower's *credit share* at the bank, $share_{ib,t} = \text{credit}_{ib,t} / \sum_{b'} \text{credit}_{ib',t}$, computed monthly. Because the denominator is the borrower's total credit, the credit share is, by construction, a statement about *allocation across lenders*. We also use product-specific shares (unsecured; fiduciary-lien). At the borrower level we use an indicator for any credit (extensive margin) and log total credit conditional on borrowing (intensive margin), plus the value-weighted spread. At the bank level we use log total credit and the number of client relationships.

3.5 Summary statistics

Table 1 reports summary statistics. [*Placeholder: to be populated from the estimation sample—means and standard deviations of credit outcomes, information shares and gains, Pix usage, and borrower/relationship characteristics, by borrower type.*] Two patterns motivate the design. First, multi- homing is common: a large share of borrowers maintain relationships with more than one conglomerate, which is what makes the within-customer comparison feasible. Second, Pix-borne flows constitute a large and rising share of electronic inflows, especially for individuals and small firms, so that Open Finance consent conveys substantial incremental information.

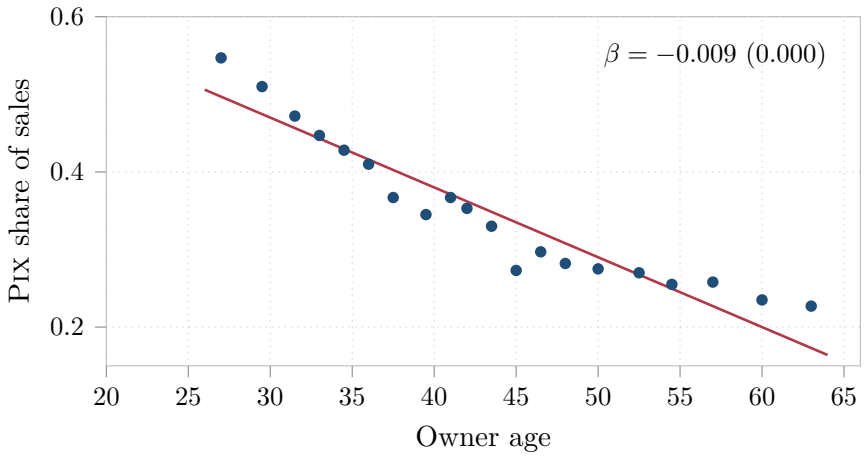


Figure 2: Older owners use Pix less: a generational gradient. Binned scatter of the share of a micro-entrepreneur's sales collected via mobile (Pix) payments against owner age, with a linear fit. The slope is -0.009 per year of age (standard error ≈ 0.000). Source: authors' survey questions, fielded in collaboration with SEBRAE (*Pulso das Empresas*).

4 Stylized Facts: Pix Adoption and Its Incidence

Before turning to credit, we document a set of facts about *who* uses Pix and *how intensively*, because the value of Open Finance to a borrower is bounded by how much of her economic life the payment system records. In the aggregate, Pix adoption is near-universal—more than nine in ten Brazilian adults transact with it, and it has overtaken cards in transaction volume—but its intensity and incidence vary systematically across agents in ways that shape the information content of the data trail. We draw these facts from the credit-registry-linked payment data and from the survey questions we fielded in collaboration with SEBRAE (Section 3).

A generational gradient in adoption. The share of a firm's sales collected through mobile (Pix) payments falls steeply with the age of the owner. Figure 2 is a binned scatter of the Pix sales share against owner age across micro-entrepreneurs: it declines from roughly 55% for owners in their late twenties to about 23% for those in their early sixties—a gradient of about 0.9 percentage points per year of age (the fitted slope is -0.009, precisely estimated). This is the firm-side counterpart of the consumer pattern that the propensity to use new payment technologies falls with age, which Crouzet et al. (2026) show governs how fast businesses adopt: firms serving younger customers face stronger incentives to accept digital payments, generating the strategic adoption complementarities emphasized by Alvarez et al. (2023) and Higgins (2024).

Smaller and women-owned firms use Pix more, but accept it less formally. Within firms, the Pix sales share is *higher* for smaller (MEI, individual micro-entrepreneur) firms and for women-owned firms—consistent with Pix substituting for cash among the smallest sellers. Yet acceptance through a dedicated *business* account, rather than a personal one, is far from universal: only about half of MEI firms accept Pix in a business account, and this fraction is lower for women, especially those with less formal education. The distinction is consequential for Open Finance: revenue that arrives in a personal account is recorded against the owner's individual taxpayer ID (CPF) rather than the firm's (CNPJ), blurring the firm's verifiable cash-flow record precisely for the smallest and most informal businesses—a formalization gap we return to when interpreting the weaker credit response of firms.

Incidence across sectors. Pix usage is most intensive among independent contractors, freelancers, and service providers, and lowest—relative to other instruments—in retail, a sector with high pre-existing card penetration. Consistent with this, Pix appears to displace cash more than cards; indeed, Pix acceptance *crowds in* card acceptance, complementing rather than cannibalizing existing electronic instruments (Sampaio and Ornelas, 2024). The income-informativeness of a firm's Pix trail is therefore highest in cash-displacing, service-oriented sectors and lower in card-intensive retail—heterogeneity we exploit directly in Section 6.5. Relatedly, Burga et al. (2025) find that Pix raised wages in small, cash-intensive establishments, compressing the skill premium.

Why these facts matter for credit. Two implications carry into the credit analysis. First, because Pix records a larger share of activity for younger, smaller, and service-sector firms, i.r. agents for whom Pix (and other electronic) inflows capture a larger share of true income, they benefit more from Open Banking. Second, the incidence of Pix adoption foreshadows the incidence of Open Finance benefits: the agents whose flows are best captured (individuals and small service firms) are precisely those for whom we later find the largest information-driven credit reallocation, while the business-account formalization gap is one channel behind the retail–firm asymmetry that recurs throughout our results.

5 Empirical Strategy

Adoption of Open Finance is staggered across institutions, and consent is endogenous. We therefore use a sequence of designs, each addressing a different threat to identification. The within-customer specification of Section 5.4 is the strongest: it absorbs the most obvious

confounders and identifies from clean within-borrower variation, so we place the most weight on it.

5.1 Bank-level design

Staggered adoption around API certification. We estimate event studies of bank outcomes around each conglomerate's certification month,

$$y_{b,t} = \sum_{k \neq -1} \theta_k \, \mathbf{1}\{t - \tau_b = k\} + \mu_b + \lambda_t + \varepsilon_{b,t},$$

where τ_b is b's certification month. Later (or never) certifiers serve as controls for early joiners.[1] The threat is that certification timing is endogenous to other bank decisions; we therefore treat these estimates as descriptive and lean on the within-bank and within-customer designs for identification.

5.2 Within-bank design

Within a bank, Open Finance consent rates differ sharply across the bank's own segments—about 30% for retail customers versus roughly 2% for firms. The gap partly reflects the mechanics of consent. An individual owns her information and can share it in a few taps; for a firm, the bank must verify that the person granting consent has the *legal authority* to share the firm's data, and firm accounts are routinely operated by financial managers and staff who lack that authority. Moreover, information is differentially relevant across products (unsecured vs. secured). Comparing the evolution of a bank's unsecured vs. secured portfolio, and its retail vs. corporate lending, after it joins holds the bank's overall supply conditions fixed.

[1] Under staggered adoption with heterogeneous treatment effects, two-way fixed-effects event studies can be contaminated by "forbidden" comparisons that use already-treated units as controls (Goodman-Bacon, 2021; de Chaisemartin and D'Haultfœ uille, 2020; Sun and Abraham, 2021). Our event studies use not-yet-treated and never-treated institutions as the comparison group, and we confirm robustness to heterogeneity-robust estimators (Callaway and Sant'Anna, 2021; Borusyak et al., 2024). The within-customer design of Section 5.4 sidesteps the issue entirely: identification there comes from differences *across a given borrower's banks within the same month*, so no cross-cohort or already-treated comparisons enter.

5.3 Borrower-level staggered difference-in-differences

We ask what happens to a borrower's access to and use of credit when her own primary bank joins Open Finance. In a borrower × month panel,

$$y_{i,t} = \sum_{k \neq -1} \beta_k \, \mathbf{1}\{t - \tau_{b(i)} = k\} + \alpha_i + \eta_{g(i),t} + \varepsilon_{i,t},$$

where $\tau_{b(i)}$ is the certification month of i's primary bank, α_i is a borrower fixed effect, and $\eta_{g(i),t}$ is a bank-segment × month fixed effect. Outcomes are the extensive margin (any credit) and the intensive margin (log credit). We estimate individuals and firms separately.

5.4 Within-customer information-gain design

Our strongest specification exploits multi-homing: it controls for the most obvious confounders—borrower demand, risk, and consent; bank supply and joining—and identifies the effect from clean, relationship-level variation in what each bank stands to learn. Consider a borrower i who banks with several conglomerates. When Open Finance comes online, different banks of i obtain different information gains $\Delta_{ib,t}$, depending on how much of i's flows each already observed. We estimate

$$share_{ib,t} \;=\; \beta\, \Delta_{ib,t} \;+\; \underbrace{\alpha_{i,t}}_{\text{borrower}\times\text{month}} \;+\; \underbrace{\delta_{b,t}}_{\text{conglomerate}\times\text{month}} \;+\; \underbrace{\gamma_{ib}}_{\text{borrower}\times\text{conglomerate}} \;+ \varepsilon_{ib,t}, \quad (1)$$

and a dynamic analog that replaces $\Delta_{ib,t}$ with interactions of the (pair-level, time-invariant) information gain with event-time indicators relative to b's join. The borrower × month fixed effect $\alpha_{i,t}$ is the workhorse: it absorbs *everything* about the borrower in that month—her total credit demand, her risk, her income shocks, and any aggregate condition—so that β is identified purely from differences in information gains *across the same borrower's banks at the same time*. The conglomerate × month effect $\delta_{b,t}$ absorbs each bank's time-varying supply (cost of funds, risk appetite, technology), and the borrower × conglomerate effect γ_{ib} absorbs the persistent level of each relationship. We saturate $\alpha_{i,t}$ further by interacting it with the bank's join cohort, so that comparisons are within borrower–month *and* join-timing cohort. Standard errors are clustered by conglomerate. Recall that primarily, we think of Open Banking as an information sharing mechanism. Following this logic, we identify the lenders who benefit the most from Open Bankin group lenders base on how much prior information they had about the borrowers. In order to do so, we replace $\beta\Delta_{ib,t}$ with $\sum_g \beta_g \left(\Delta_{ib,t} \times \mathbf{1}\{\text{prior group} = g\}\right)$.

Equation (1) addresses the central confounders. The borrower's endogenous consent

decision—whether, when, and how much of her data she chooses to share in aggregate—together with any borrower-level demand or risk shock, is differenced out by $\alpha_{i,t}$, because consent operates at the customer level while $\Delta_{ib,t}$ varies across her banks. The bank's endogenous decision to join Open Finance, and its overall drive to acquire and use shared cash-flow data, along with all other bank-level supply shocks, are absorbed by $\delta_{b,t}$. What remains is the *relationship-specific* potential information gain, which is determined by the historical geography of the borrower's flows across her banks—plausibly unrelated to contemporaneous, relationship-specific credit-demand shocks once the borrower × month and bank × month margins are removed. The identifying assumption is that, absent differential information, a borrower's credit share at her high-gain and low-gain banks would have evolved in parallel; the flat pre-trends in the dynamic estimates support this.

6 Results

6.1 Bank-level: lending expands and rates fall after joining

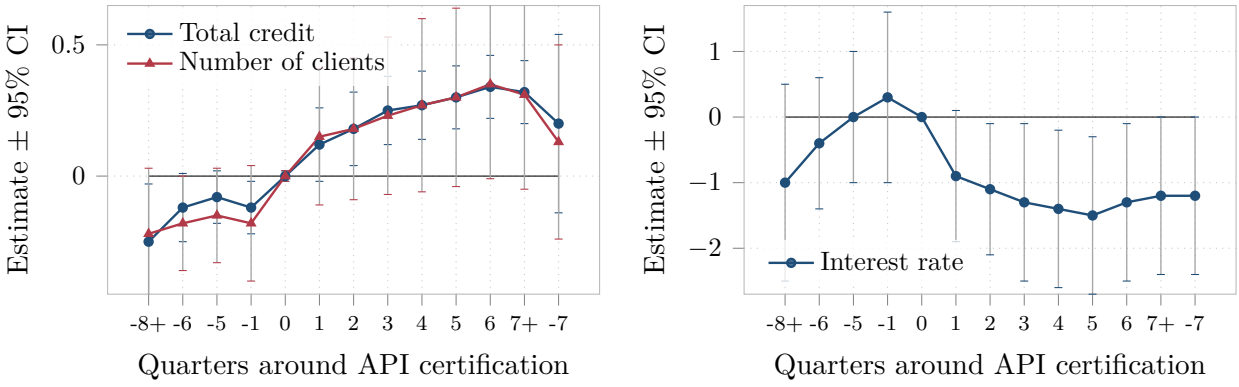
We begin at the bank level. Figure 3 plots event studies of bank outcomes around the certification of a conglomerate's Open Finance API. After an institution joins, its total credit and its number of client relationships rise together—by roughly 12 log points in the quarter of certification to 30–34 log points within five to six quarters (panel a)—while its value-weighted lending rate declines by an economically meaningful amount (panel b), with pre-trends close to flat. The joint expansion of quantities and contraction of prices is the signature of an outward shift in the effective supply of credit, consistent with better screening and intensified competition once data become portable.

Our empirical findings also point to products benefiting for which cash-flow information is most decision- relevant,i.e., *unsecured* credit. We find that the expansion is tilted toward unsecured products: the unsecured share of the portfolio rises after joining (Section 6.4), consistent with prediction of the model developed in Section 7.

These bank-level patterns are only suggestive: the timing of certification may itself respond to other bank decisions, so we do not lean on them for identification and turn instead to the disaggregated within-bank and within-customer designs below, which hold the bank or the borrower fixed.

6.2 Borrower-level: new retail borrowers, tighter access for firms

We next ask what happens to a borrower when her own primary bank joins Open Finance. Figure 4 shows the extensive margin. For *individuals*, the probability of holding (non-trivial)



(a) Total credit and number of clients

(b) Value-weighted interest rate

Figure 3: Bank-level outcomes around Open Finance API certification. Staggered event studies of conglomerate-level outcomes relative to the certification quarter (normalized at −1). Total credit and the number of client relationships rise after joining; the value-weighted lending rate falls. Bands are 95% confidence intervals. These estimates are descriptive: certification timing may be endogenous to other bank decisions.

credit rises steadily after the bank joins, reaching roughly +4–5 percentage points within five to six months, on a flat pre-trend—open finance is associated with new retail borrowers entering credit. For *firms*, the extensive margin moves in the opposite direction, declining by about 2–3 percentage points. The intensive margin (log credit conditional on borrowing) rises for both groups (Figure 5). The contrast is consistent with two non-exclusive mechanisms that we probe below: a reallocation by banks away from firm and toward retail lending, and information revealed about firms that is, on net, less favorable than for individuals. It also mirrors the stark difference in consent rates (retail $\approx$ 30% vs. firms $\approx$ 2%): the population of firms that consents is small and selected. All in all, this evidence suggests that banks shift their lending portfolios away from corporate and toward retail credit, where information acquisition has become cheaper—a shift that may carry unintended consequences for productive investment in the economy (Section 8).

6.3 Within-customer: information reallocates credit toward the bank that learns

We now turn to the within-customer design—the most stringent of our specifications—which holds the borrower fixed and asks whether the bank that stands to gain more information through Open Finance captures more of the borrower's credit. Figure 6 reports the dynamic version of equation (1): the response of a bank's credit *share* to its information gain, by event time relative to joining. Pre-trends are flat. After joining, the credit share rises monotonically with information gain, reaching, for a relationship that moves from zero to full information

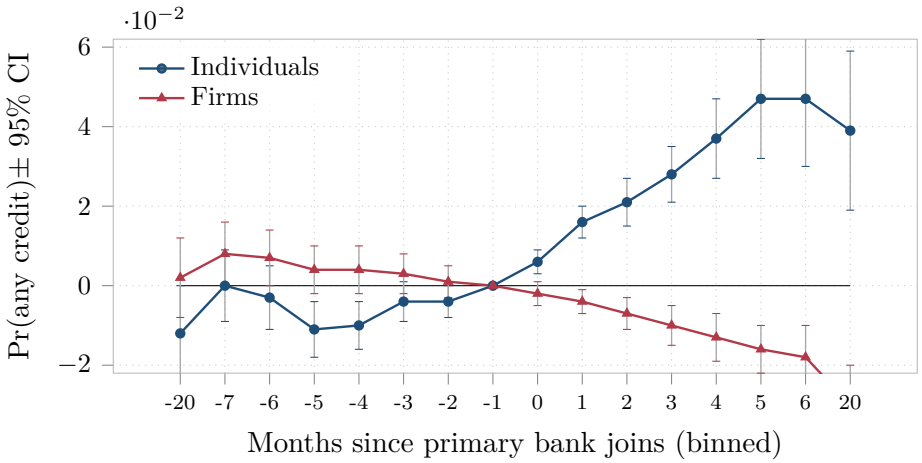


Figure 4: Extensive margin after the borrower's primary bank joins Open Finance. Event study of an indicator for holding credit, estimated separately for individuals (blue) and firms (red), with borrower and bank-segment×month fixed effects. Individuals become more likely to borrow; firms less likely. Bars are 95% confidence intervals.

gain, about +4.7 percentage points for individuals, +6.8 for *Simples* (small) firms, and +4.3 for larger firms seven quarters out. Because borrower × month fixed effects absorb the borrower's total demand, these magnitudes describe *reallocation across a borrower's lenders driven by information*. Prediction **??**.

Table 2 reports the level specification and, crucially, its interaction with the bank's *prior* information about the borrower. Two results stand out. First, the average effect of information gain on credit share is positive and precise for all three borrower types (columns 1, 3, 5): a full information gain raises the credit share by 3.3 percentage points for individuals, 5.6 for *Simples* firms, and 3.6 for larger firms. Second—and this is the sharpest test of the mechanism—the effect is concentrated in relationships where the bank *previously had no information* (the "No info" row): +2.6, +5.6, and +4.1 percentage points, respectively. As prior information rises, the marginal effect falls, and for individuals it turns significantly *negative* at high prior information (−5.1 percentage points at the top group). Banks that already observed most of a retail customer's flows *lose* share when those flows become visible to everyone, consistent with the predictions of the model that we develop in Section 7.

6.4 Product composition: the action is in unsecured credit

Table 3 decomposes the credit-share response by product type. Information gain raises the share of *unsecured* credit for all borrower types (panel A), with a large effect for *Simples* firms (+8.5 percentage points). By contrast, the effect on semi-secured credit is negative—significantly so for firms, which *substitute away* from semi-secured products (panel B)—and

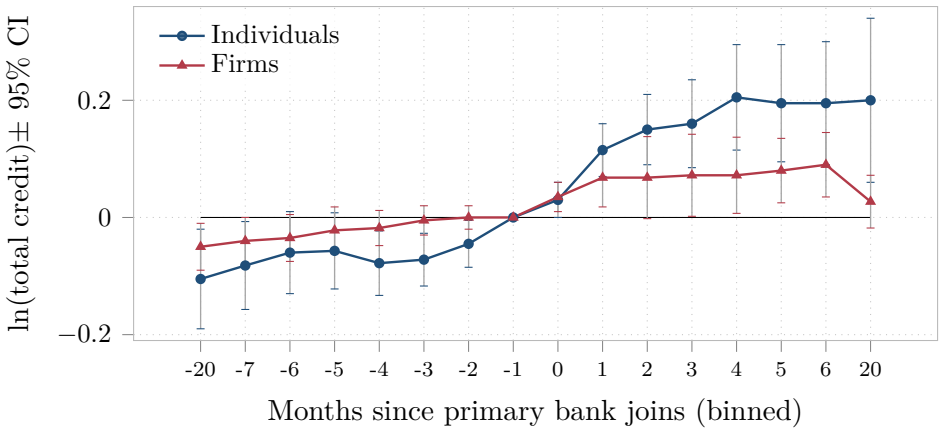


Figure 5: Intensive margin after the borrower's primary bank joins Open Finance. Event study of log total credit (conditional on holding credit above R$1,000), estimated separately for individuals (blue) and firms (red), with borrower and bank-segment×month fixed effects. Conditional on access, both groups borrow more after their bank joins, with a larger increase for individuals. Bars are 95% confidence intervals.

the effect on the most strongly collateralized (fiduciary-lien) products is small (panel C). This is exactly Prediction **??**: cash-flow information matters for credit that is not backed by collateral, and matters little where collateral already substitutes for information.

6.5 The payments–information complementarity

Prediction **??** holds that the credit response to information sharing should be larger where the underlying payment trail is more informative. Table 4 tests this for firms, interacting the treatment (the firm's bank joins Open Finance) with an indicator for operating in a *high-Pix-usage* industry, where a larger share of revenue flows through Pix and is therefore captured by the shared data. While the average firm sees its probability of credit fall after its bank joins (−2.5 percentage points), firms in high-Pix industries experience a partial offset on the extensive margin (+0.6 percentage points) and a significantly larger intensive-margin expansion (+1.4 percentage points). The value of open finance to a firm rises with how much of its economic activity Pix records —the clearest evidence that the instant-payment system and the information-sharing regime are complements in credit markets (Ghosh et al., 2026).

6.6 Who turns information into lending?

An important input to policy design is the beneficiaries of Open Finance. Three pieces of evidence, drawn from the post-2023 Open Finance dashboard and the within-customer design, point the same way. First, Figure 7 shows that non-banks (mainly fintechs) are

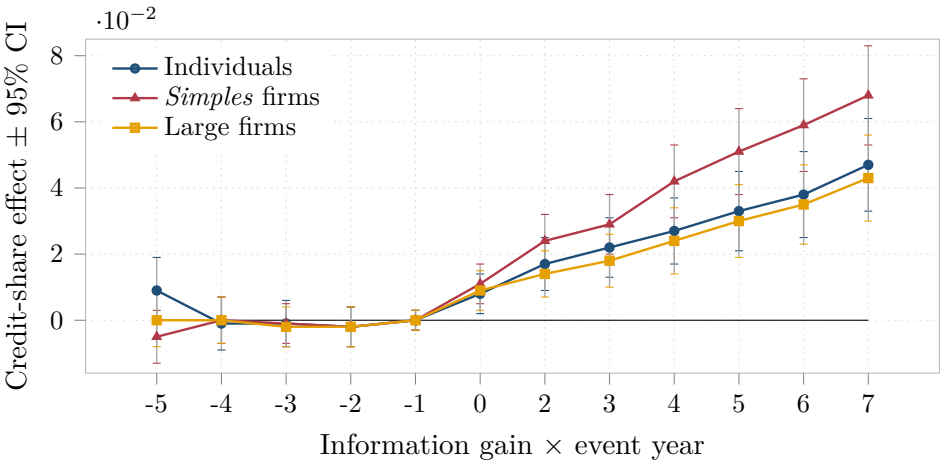


Figure 6: Credit share responds to information gain (within borrower × month). Dynamic estimates of the effect of a bank's information gain Δ_{ib} on its share of the borrower's credit, by event quarter relative to joining, for individuals, *Simples* firms, and larger firms. Specification as in equation (1) with borrower×month×cohort, conglomerate×month, and borrower×conglomerate fixed effects; standard errors clustered by conglomerate. The horizontal axis is information gain interacted with relative year.

disproportionately data *consumers*: their ratio of data received to data transmitted exceeds one throughout, while deposit-taking banks transmit at least as much as they receive. Open finance moves information from incumbent deposit-takers toward fintech entrants. Second, the lending-share gains from a given information flow are larger when the receiving institution is a *deposit-taker*, which enjoys a lower cost of capital and can therefore act on favorable information by lending (Table 5, panel A). Third, the gains are larger when the receiving institution had invested more in *information technology* before Open Finance and is thus better positioned to ingest and use the data (panel B). Information is necessary but not sufficient: turning it into credit requires cheap funding and the capacity to use it.

These decompositions are estimated within the same within-customer specification: we interact the relationship's information gain with the deposit-taking status and pre-Open Finance IT intensity of the *sending* and *receiving* conglomerates, so the comparison runs across the banks of the same borrower in the same month, and the heterogeneity is identified from differences in who receives the data, not from differences across borrowers. The interpretation is correspondingly sharp. A common fear at the launch of open-banking regimes was that incumbents would be "expropriated" by non-banks cherry-picking their best customers. The evidence gives a more textured answer: non-banks are indeed net receivers of information (Figure 7), and incumbents that already knew a retail customer well do lose share (Table 2); but the lending-share gains concentrate among deposit-takers, which fund loans cheaply, and among institutions that had invested in IT before Open Finance (Table 5).

Table 2: Within-customer: information gain and credit share, by prior information

Dep. var.: credit share	Individuals		*Simples* firms		Larger firms	
	(1)	(2)	(3)	(4)	(5)	(6)
Info gain	0.0330***		0.0564***		0.0361***	
	(0.0111)		(0.0080)		(0.0087)	
Info gain × prior-information group:						
× No info ($s_{ib} = 0$)		0.0264***		0.0557***		0.0408***
		(0.0089)		(0.0071)		(0.0096)
× Level 1 $(0, 0.3]$		0.0088		0.0528***		0.0358***
		(0.0081)		(0.0061)		(0.0105)
× Level 2 $(0.3, 0.6]$		−0.0217*		0.0467**		−0.0079
		(0.0114)		(0.0194)		(0.0117)
× Level 3 $(0.6, 1)$		−0.0506***		0.0476		0.1637
		(0.0129)		(0.0460)		(0.1014)
Observations	183,194,949		33,004,879		3,627,026	
R^2	0.778		0.873		0.873	
Borrower×Conglomerate FE			Yes			
Conglomerate×Month FE			Yes			
Borrower×Month×Cohort FE			Yes			

Notes: Borrower×conglomerate×month panel. The dependent variable is the bank's share of the borrower's total credit. "Info gain" is the incremental share of the borrower's economy-wide electronic cash flows that the bank can observe through Open Finance, net of what it already observed; the baseline aggregates all electronic rails, including credit- and debit-card settlements, with Pix-only and sent-flow variants yielding similar results (Section 3.3). Prior-information groups bin the relationship by the bank's pre-Open Finance observed share s_{ib}. Standard errors clustered by conglomerate in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

The competitive consequences of open finance thus depend jointly on data access, the cost of capital, and technological capacity—not on data access alone (He et al., 2023).

7 Conceptual Framework

In this section, we develop a deliberately stylized model of competition among differentially informed lenders. We start by holding the borrower's total demand fixed and asks how information sharing reallocates her borrowing across banks, exactly the object our within-customer design identifies. We plan to extent the analysis to the extensive margin of financial inclusion. The model builds on Sharpe (1990), Rajan (1992), and Hauswald and Marquez (2003), adapted to customer-permissioned data portability as in He et al. (2023).

Table 3: Within-customer credit-share response by product type

Dep. var.: product-specific credit share	Individuals (1)	*Simples* firms (2)	Larger firms (3)
Panel A: Unsecured			
Info gain	0.0258**	0.0854***	0.0305***
	(0.0121)	(0.0116)	(0.0088)
Panel B: Semi-secured			
Info gain	−0.0030	−0.0519***	−0.0177**
	(0.0023)	(0.0104)	(0.0076)
Panel C: Fiduciary lien			
Info gain	0.0010	0.0048*	0.0075
	(0.0008)	(0.0025)	(0.0047)
Observations	140,907,578	25,606,730	2,783,568
Fixed effects	Borrower×Congl.,	Congl.×Month,	Borrower×Month×Cohort

Notes: As in Table 2; the dependent variable is the bank's share of the borrower's credit *within* the indicated product category. Standard errors clustered by conglomerate. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

7.1 Setup

Let θ denote borrower's repayment probability which we refer to as her latent quality. A borrower with latent quality θ can borrow from $J \geq 2$ banks, indexed by b. Banks share a common prior over θ and differ in how precisely they can screen her: bank b's screening precision is

$$\tau_b \;=\; \bar{\tau} \;+\; \kappa_b \, \nu \, m_b, \tag{2}$$

where $m_b \in [0,1]$ is the share of the borrower's economy-wide electronic cash flows that b observes, $\nu \geq 0$ is the informativeness of those flows about θ (the quality of the payment trail, common across all borrowers), $\kappa_b \geq 0$ is b's capacity to convert data into credit assessments (bank b's information technology), and $\bar{\tau}$ is baseline precision from public sources such as the credit bureau. We assume $\kappa_b \nu > 0$ for at least one bank.

To simplify the analysis, we assume that the expected unit cost at which b can lend is given by

$$c_b \;=\; \rho_b \;+\; \ell(\tau_b; \phi), \tag{3}$$

where ρ_b is b's marginal cost of funds (low for deposit-takers) and $\ell(\tau; \phi)$ is the expected credit loss on a marginal loan, which depends on screening precision τ and on the product's collateralization $\phi \in [0,1]$. The simplifying assumption is that the functional form of $\ell(\tau; \phi)$ is given exogenously and satisfy the following restriction.

Assumption 1. *$\ell(\tau; \phi) = (1 - \phi)\, \lambda(\tau)$, where λ is differentiable with $\lambda' < 0$ and $\lambda'' \geq 0$.*

Table 4: Firm credit and Pix-usage intensity (payments–information complementarity)

	(1) Any credit	(2) ln(credit amount)
Treated bank	−0.0245***	0.0039
	(0.0032)	(0.0069)
Treated bank × high-Pix industry	0.0063**	0.0142***
	(0.0024)	(0.0039)
Observations	34,000,000	1,395,217
Borrower FE / Bank FE / Month FE	Yes	Yes

Notes: Firm sample. "Treated bank" indicates the firm's bank has joined Open Finance. "High-Pix industry" indicates above-median Pix usage in the firm's industry (in firms' own accounts). Standard errors clustered by conglomerate. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

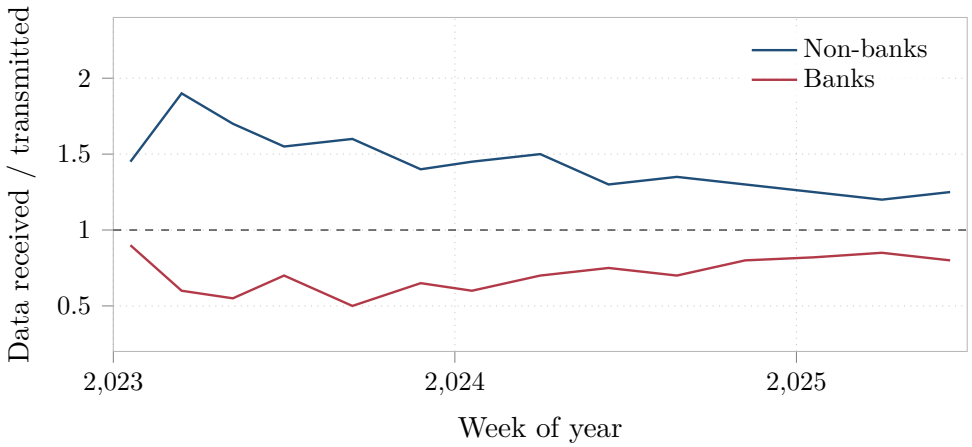


Figure 7: Non-banks receive more information than they transmit. Ratio of Open Finance data received to data transmitted, by institution type, weekly from 2023 (BCB Open Finance dashboard). Non-banks (blue) consistently exceed one; deposit-taking banks (red) remain at or below one. The plotted lines smooth the weekly dashboard series.

Assumption 1 captured a screening problem with two properties in reduced form: 1) collateral recovers the fraction ϕ of the exposure regardless of θ, 2) expected losses on the uncollateralized residual fall as screening improves. Thus, a better-informed lender funds fewer bad risks and prices the rest more accurately ($\lambda' < 0$), with weakly diminishing returns to precision ($\lambda'' \geq 0$).[2] Importantly, information is irrelevant for fully collateralized credit ($\ell_\tau = 0$ at $\phi = 1$).

The borrower views lenders as imperfect substitutes: different borrowers have a comparative advantage in convenience, switching costs, product fit, etc. As such, she spreads her

[2]For instance, let the bank observe a signal $x = \theta + \varepsilon$ with noise variance $1/\tau$, extend credit when posterior quality clears a hurdle, and lose the fraction $1 - \phi$ of the exposure in default. Expected losses per unit lent then take the form $(1 - \phi)\lambda(\tau)$ with λ decreasing in τ; $\lambda'' \geq 0$ encodes diminishing returns to precision.

Table 5: Information flows: who gains lending share (sender → receiver)

Dep. var.: credit share; coef. on info gain ×...	Individuals (1)	*Simples* firms (2)	Larger firms (3)
Panel A: by deposit-taking status			
No deposits → No deposits	0.0062*	0.0095**	0.0092
	(0.0026)	(0.0033)	(0.0060)
No deposits → Deposits	0.0115**	0.0126**	0.0095
	(0.0039)	(0.0037)	(0.0084)
Deposits → No deposits	0.0061*	0.0076**	0.0082**
	(0.0024)	(0.0028)	(0.0026)
Deposits → Deposits	0.0113**	0.0121**	0.0129**
	(0.0039)	(0.0039)	(0.0043)
Panel B: by pre-Open Finance IT intensity			
Low IT → Low IT	0.0071**	0.0094**	0.0106**
	(0.0021)	(0.0032)	(0.0033)
Low IT → High IT	0.0126**	0.0122**	0.0141**
	(0.0047)	(0.0040)	(0.0050)
High IT → Low IT	0.0092**	0.0112**	0.0103**
	(0.0031)	(0.0040)	(0.0035)
High IT → High IT	0.0145*	0.0132*	0.0139*
	(0.0058)	(0.0053)	(0.0059)
Observations	129,314,574	23,298,729	2,560,401
Fixed effects	Borrower×Congl., Congl.×Month, Borrower×Month		

Notes: The coefficient on information gain is interacted with the deposit-taking status (panel A) or pre-Open Finance IT intensity (panel B) of the *sending* and *receiving* institutions. Receivers that take deposits or are IT-intensive show the largest lending-share gains. Standard errors clustered by conglomerate. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

borrowing across banks rather than concentrating it at the cheapest one. We abstract away from modeling the detail of borrower preference and summarize differentiated competition with the logit allocation

$$s_b = \text{share}_b \;=\; \frac{e^{-\beta c_b}}{\sum_{b'} e^{-\beta c_{b'}}}, \qquad \beta > 0, \tag{4}$$

where β indexes how strongly the allocation responds to cost differences. Note that Equation (4) emerges endogenously from a differentiated Bertrand competition in which each of the borrower's marginal borrowing decisions carries an i.i.d. extreme-value taste shock and lenders price at a common markup over cost.[3]

[3]Under Bertrand–logit competition the equilibrium markup is $1/[\beta(1-\text{share}_b)]$, approximately constant when individual shares are modest; with a common markup, prices inherit cost differences one-for-one and (4) holds as written. None of the results below depend on the logit form: any allocation with shares decreasing in own cost, increasing in rivals' costs, and summing to one delivers Propositions 1–2 qualitatively.

7.2 Information sharing via Open Finance

Before Open Finance, bank b sees only the flows that land in its own accounts: $m_b = s_b$, the relationship's *prior information share*. Once the borrower's institutions join Open Finance and she consents, any participating bank can access the flows observed across *all* of her institutions—the pooled share $C \geq s_b$; so, m_b rises to C. Bank b's precision rises from $\tau_b^{\text{pre}} = \bar{\tau} + \kappa_b \nu s_b$ to $\tau_b^{\text{post}} = \bar{\tau} + \kappa_b \nu C$, and its lending cost falls by

$$\delta_b \;\equiv\; \ell\Big(\tau_b^{\text{pre}}; \phi\Big) - \ell\Big(\tau_b^{\text{post}}; \phi\Big) \;=\; (1-\phi)\,\kappa_b\,\nu \int_{s_b}^{C} \Big[-\lambda'(\bar{\tau} + \kappa_b \nu m)\Big]\, dm \;\geq\; 0. \tag{5}$$

We call $\Delta_b \equiv C - s_b \in [0,1]$ the relationship's *information gain*: it is the length of the integration range in (5). In the model the gain is realized; empirically we observe its *potential* value, since realized learning requires the customer's consent.[4] Estimates based on the measured gain of Section 3.3 therefore carry an intent-to-treat interpretation.

7.3 Equilibrium lending shares

Proposition 1 (Information reallocates credit)**.** *Following Open Finance, fixing the borrowing amount, bank b's credit share is given by*

$$\text{share}_b^{\text{post}} = \frac{\text{share}_b^{\text{pre}}\, e^{\beta \delta_b}}{\sum_{b'} \text{share}_{b'}^{\text{pre}}\, e^{\beta \delta_{b'}}}, \qquad \textit{so} \qquad \Delta\,\text{share}_b \;\approx\; \beta\,\text{share}_b\Big(\delta_b - \bar{\delta}\Big), \tag{6}$$

to first order in $\beta\delta$, with $\bar{\delta} \equiv \sum_{b'} \text{share}_{b'} \delta_{b'}$; bank b gains share if and only if δ_b exceeds the certainty-equivalent average $\tilde{\delta} \equiv \beta^{-1} \ln \sum_{b'} \text{share}_{b'}^{\text{pre}} e^{\beta \delta_{b'}}$. The cost reduction δ_b in (5):

(i) is strictly increasing in the information gain Δ_b and, holding the pooled share C fixed, strictly decreasing in prior information s_b;

(ii) is strictly decreasing in collateralization ϕ whenever $\delta_b > 0$, with $\delta_b = 0$ at $\phi = 1$;

(iii) is strictly increasing in informativeness ν and in IT capacity κ_b for relationships with no prior information ($s_b = 0$), and in general whenever $s_b\Big[-\lambda'(\tau_b^{\text{pre}})\Big] < C\Big[-\lambda'(\tau_b^{\text{post}})\Big]$;

(iv) in the leading case of locally linear screening returns ($\lambda' \equiv -\gamma$), takes the multiplicative form $\delta_b = (1-\phi)\,\gamma\,\kappa_b\,\nu\,\Delta_b$.

Proof. The exact update follows from substituting $c_b - \delta_b$ into (4); expanding in $\beta\delta$ gives the approximation, and $\text{share}_b^{\text{post}} > \text{share}_b^{\text{pre}}$ iff $e^{\beta\delta_b} > \sum_{b'} \text{share}_{b'}^{\text{pre}} e^{\beta\delta_{b'}} = e^{\beta\tilde{\delta}}$. For (i), the integrand in (5) is strictly positive, and $\partial\delta_b/\partial s_b = -(1-\phi)\kappa_b\nu[-\lambda'(\tau_b^{\text{pre}})] < 0$. For (ii),

[4] We plan to extend our analysis to explicitely model consumer;s consent, as we explain below.

$\partial\delta_b/\partial\phi = -\delta_b/(1-\phi)$. For (iii), differentiating (5) in ν (or κ_b) gives $(1-\phi)\kappa_b\{C[-\lambda'(\tau_b^{\text{post}})] - s_b[-\lambda'(\tau_b^{\text{pre}})]\}$ (up to the positive factor ν/κ_b), which is positive under the stated condition and always at $s_b = 0$. Part (iv) is immediate. □

Two features of Proposition 1 discipline the empirical work. First, fixing the total borrowing, because shares sum to one, reallocation is *zero-sum within a borrower*: the within-customer design estimates precisely the cross-bank differences $\delta_b - \tilde{\delta}$, purged of anything common to the borrower. Second, the model separates the *value of the data*—$\nu\Delta_b$, governed by the payment trail and the bank's prior overlap with it—from the *capacity to act on the data*—κ_b, ρ_b, and ϕ. This separation maps one-for-one into our heterogeneity tests. Finally, by (6) the share response scales with share$_b$: a lender priced out of the borrower's market by a high funding cost ρ_b (share$_b \approx 0$) converts even a large δ_b into little additional lending.

Proposition 2 (Erosion of informational rents)**.** *Suppose at least one rival of bank b has a strictly positive cost reduction. If b already observed everything Open Finance can reveal ($s_b = C$, hence $\Delta_b = 0$ and $\delta_b = 0$), then b strictly loses credit share even though its own cost is unchanged.*

Proof. With $\delta_b = 0$ and $\delta_{b'} > 0$ for some active b', the denominator of (6) strictly exceeds one, so b's share strictly falls. Raising s_b shortens the integration range in (5) and, when $-\lambda'$ is decreasing ($\lambda'' > 0$), removes its largest values; share losses follow from $\delta_b < \tilde{\delta}$ for s_b close enough to C. □

Proposition 2 restates a classic idea in modern form: an incumbent's advantage rests not on what it knows but on knowing *more than its rivals*; equalizing information competes the rents away (Sharpe, 1990; Rajan, 1992; Hauswald and Marquez, 2003). Open Finance compresses cross-bank information differences within each borrower, so informationally thin relationships expand while informationally rich incumbents cede share.

More generally, δ_b declines in s_b through two reinforcing channels: the increment Δ_b is mechanically smaller, and, when $\lambda'' > 0$, the marginal value of additional precision is lower for an already well-informed lender. The share response to Open Finance therefore declines in s_b and turns negative for s_b high enough.

The most important implication of the model involves the identity of lenders who benefit the most from Open Finance. The model implies that the reallocation is largest for relationships in which the lender previously knew little ($s_b \approx 0$), where the increment Δ_b is largest and the marginal value of precision highest (Proposition 2). Incumbents that already observed most of a borrower's flows (s_b high) gain little and may *lose* share, as competitors close the information gap and bid away borrowers on whom the incumbent earned informational rents (Sharpe, 1990; Rajan, 1992; Sutherland, 2018).

A second prediction of the model involves the products that are most impacted by Open Banking. The model implies that the effects are concentrate in products for which cash-flow information is most decision- relevant—*unsecured* credit—rather than in collateralized lending, where the collateral substitutes for information.

Finally, the model implies that the credit response to information sharing is increasing in the informativeness of the underlying payment trail. Where Pix (and other electronic) inflows capture a larger share of true income—e.g., in high-Pix-usage sectors—the same consent conveys more, and credit responds more (Ghosh et al., 2026).

Extensive margin The propositions hold total borrowing fixed. To speak to the borrower-level evidence, we plan to incorporate a reservation rate into borrowers problem: the borrower takes credit if and only if her best offer is better than her reservation rate, $\min_b p_b \leq \bar{r}$. Information sharing moves $\min_b p_b$ through the posterior about θ. When the revealed cash flows are *favorable*, every recipient's expected loss falls, the best offer improves, and previously rationed borrowers enter—an extensive-margin expansion. When they are *unfavorable*, sharper screening prices the borrower out—the flip side of relaxing the information friction that rations credit in the first place (Stiglitz and Weiss, 1981). Sharing thus reallocates access from revealed-bad toward revealed-good borrowers, and its net effect on a customer class depends on the composition of what the data reveal.

Consumer consent Consent gates the entire mechanism of information sharing among finnacial institutions. Data move only when the borrower opts in, so $\delta_b > 0$ requires consent, and borrowers who expect favorable revelation (and have low privacy costs) consent first. This implies a positive selection that predicts gains for consenters on average. Where consent is rare, most borrowers generate no δ_b at all. It follows that banks may redirect attention toward the data-rich retail segment, and, if lenders draw inferences from non-consent itself, non-consenters can be priced *as if* unfavorable information had been revealed. Put differently, giving consent is a a*Bayesian persuasion* problem for borrowers. The model thus rationalizes the retail extensive-margin expansion we document. may contribute.

8 Discussion

What do we learn? Three lessons organize the evidence. First, *information itself moves credit*: the reallocation runs toward the banks that stand to learn the most, is largest where prior information was absent, and operates in the products where screening matters—the mechanism that information-based theories of lending highlight, rather than a generic

"fintech" effect. Second, *the consent architecture determines incidence.* Households, who can consent in a few taps, gain access—most strongly exactly where lenders previously knew nothing, where information sharing can do most to overcome adverse selection. Firms, for whom consent requires verified legal authority, remain largely outside the regime, and banks appear to tilt their portfolios toward the retail segment where information acquisition has become cheaper. Third, *data alone are not enough*: the gains accrue to lenders with cheap funding and the technology to use the data, so open finance redistributes lending toward funded, IT-capable institutions rather than toward data-hungry entrants per se.

The portfolio tilt carries a warning. Because the credit that moves is overwhelmingly unsecured (Table 3) and access shifts from firms toward households, open finance may redirect credit from productive corporate uses toward consumption. Our firm-side results—larger intensive-margin gains in high-Pix industries (Table 4)—indicate a productive component tied to revenue information, but the aggregate shift away from corporate credit could have unintended consequences for productive investment in the economy. Tracing the real effects on investment, employment, and household welfare is a priority for future work.

For policy, two implications stand out. Consent design is first-order, and on the firm side the friction is legal as much as technological: clarifying who may authorize sharing on a firm's behalf—or creating standing authorizations for firm representatives—could unlock the corporate segment, and even light-touch interventions raise firm adoption (Barboni et al., 2025). And because information sharing pays off where the payment trail is rich and where receiving lenders can act on it, open-finance regimes are complements to—not substitutes for—payment digitization and a competitive, well-capitalized lending sector.

9 Conclusion

For most of banking history, the information generated by a borrower's accounts belonged, in practice, to her bank. Brazil's combination of a comprehensive electronic cash-flow record—crowned by the near-universal instant payment system Pix—and a customer-permissioned sharing regime (Open Finance) hands that information to the borrower, and we have used the reassignment to ask how information allocates credit. The answer is direct: information moves credit. Holding the borrower fixed, the bank that gains more visibility into her electronic cash flows captures a larger share of her borrowing, and does so most where it previously knew least, while informationally rich incumbents lose retail share. The reallocation is concentrated in unsecured credit, is larger where the payment trail is more informative, and accrues to lenders with the funding and technology to use the data. At the borrower level, new retail borrowers gain access even as firm access tightens—an asymmetry that

tracks the large consent gap between households and firms.

Information-sharing regimes built on digital payments can therefore broaden access and intensify competition, but their incidence is governed by who consents, who previously held the information, and who can act on it. Much remains to be done—tracing real effects, sharpening the bank-level identification, and understanding the firm-consent gap—but the evidence already marks the payments-plus-information architecture now spreading across emerging markets as a powerful, and unevenly distributed, force in credit allocation.

References

Agarwal, Sumit, Wenlan Qian, Yuan Ren, Hsin-Tien Tsai, and Bernard Yeung. The real impact of fintech: Evidence from mobile payment technology. *Management Science*, 2025. Articles in Advance.

Alok, Shashwat, Pulak Ghosh, Nirupama Kulkarni, and Manju Puri. Breaking barriers to financial access: Cross-platform digital payments and credit markets. NBER Working Paper No. 33259, 2024.

Alvarez, Fernando E., David Argente, Francesco Lippi, Esteban Méndez, and Diana Van Patten. Strategic complementarities in a dynamic model of technology adoption: P2p digital payments. NBER Working Paper No. 31280, 2023.

Argente, David, Paula González Alvarez, Esteban Méndez, and Diana Van Patten. Drivers of digital payment adoption: Lessons from brazil, costa rica, and mexico. NBER Working Paper No. 34280, 2025.

Babina, Tania, Saleem Bahaj, Greg Buchak, Filippo De Marco, Angus Foulis, Will Gornall, Francesco Mazzola, and Tong Yu. Customer data access and fintech entry: Early evidence from open banking. *Journal of Financial Economics*, 169:103950, 2025.

Barboni, Giorgia et al. Open finance adoption among micro and small enterprises in brazil. Ongoing randomized evaluation, Innovations for Poverty Action and Central Bank of Brazil, 2025. Work in progress.

Berg, Tobias, Valentin Burg, Ana Gombović, and Manju Puri. On the rise of fintechs: Credit scoring using digital footprints. *Review of Financial Studies*, 33(7):2845–2897, 2020.

Björkegren, Daniel and Darrell Grissen. Behavior revealed in mobile phone usage predicts credit repayment. *The World Bank Economic Review*, 34(3):618–634, 2020.

Black, Fischer. Bank funds management in an efficient market. *Journal of Financial Economics*, 2(4):323–339, 1975.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess. Revisiting event-study designs: Robust and efficient estimation. *Review of Economic Studies*, 91(6):3253–3285, 2024.

Burga, Carlos, Jacelly Cespedes, Carlos Parra, and Bernardo Ricca. Financial technologies, labor markets, and wage inequality: Evidence from instant payment systems. Working paper (SSRN 5026874), 2025.

Callaway, Brantly and Pedro H.C̃. Sant'Anna. Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230, 2021.

Crouzet, Nicolas, Apoorv Gupta, and Filippo Mezzanotti. Shocks and technology adoption: Evidence from electronic payment systems. *Journal of Political Economy*, 131(11):3003–3065, 2023.

Crouzet, Nicolas, Pulak Ghosh, Apoorv Gupta, and Filippo Mezzanotti. Demographics and technology diffusion: Evidence from mobile payments. NBER Working Paper No. 34885, 2026.

Dalton, Patricio S., Haki Pamuk, Ravindra Ramrattan, Burak Uras, and Daan van Soest. Electronic payment technology and business finance: A randomized, controlled trial with mobile money. *Management Science*, 70(4):2590–2625, 2024.

de Chaisemartin, Clément and Xavier D'Haultfœ uille. Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–2996, 2020.

Ding, Ding, Rodrigo Gonzalez, Yiming Ma, and Yao Zeng. The effect of instant payments on the banking system: Liquidity transformation and risk-taking. Banco Central do Brasil Working Paper Series No. 619, 2025.

Djankov, Simeon, Caralee McLiesh, and Andrei Shleifer. Private credit in 129 countries. *Journal of Financial Economics*, 84(2):299–329, 2007.

Fama, Eugene F. What's different about banks? *Journal of Monetary Economics*, 15(1): 29–39, 1985.

Farboodi, Maryam and Laura Veldkamp. A model of the data economy. *Review of Economic Studies*, 2025. Forthcoming.

Ghosh, Pulak, Boris Vallée, and Yao Zeng. Fintech lending and cashless payments. *Journal of Finance*, 81:1053–1101, 2026.

Goodman-Bacon, Andrew. Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277, 2021.

Hauswald, Robert and Robert Marquez. Information technology and financial services competition. *Review of Financial Studies*, 16(3):921–948, 2003.

He, Zhiguo, Jing Huang, and Jidong Zhou. Open banking: Credit market competition when borrowers own the data. *Journal of Financial Economics*, 147(2):449–474, 2023.

Hertzberg, Andrew, José María Liberti, and Daniel Paravisini. Public information and coordination: Evidence from a credit registry expansion. *Journal of Finance*, 66(2):379–412, 2011.

Higgins, Sean. Financial technology adoption: Network externalities of cashless payments in mexico. *American Economic Review*, 114(11):3469–3512, 2024.

Jones, Charles I. and Christopher Tonetti. Nonrivalry and the economics of data. *American Economic Review*, 110(9):2819–2858, 2020.

Lian, Chen and Yueran Ma. Anatomy of corporate borrowing constraints. *Quarterly Journal of Economics*, 136(1):229–291, 2021.

Liberti, José María and Mitchell A. Petersen. Information: Hard and soft. *Review of Corporate Finance Studies*, 8(1):1–41, 2019.

Ma, Xinchen. Open banking and competition in banks and fintech: Evidence from mobile apps. Job Market Paper, London School of Economics, 2024.

Mester, Loretta J., Leonard I. Nakamura, and Micheline Renault. Transactions accounts and loan monitoring. *Review of Financial Studies*, 20(3):529–556, 2007.

Norden, Lars and Martin Weber. Credit line usage, checking account activity, and default risk of bank borrowers. *Review of Financial Studies*, 23(10):3665–3699, 2010.

Pagano, Marco and Tullio Jappelli. Information sharing in credit markets. *Journal of Finance*, 48(5):1693–1718, 1993.

Parlour, Christine A., Uday Rajan, and Haoxiang Zhu. When fintech competes for payment flows. *Review of Financial Studies*, 35(11):4985–5024, 2022.

Petersen, Mitchell A. and Raghuram G. Rajan. The benefits of lending relationships: Evidence from small business data. *Journal of Finance*, 49(1):3–37, 1994.

Petersen, Mitchell A. and Raghuram G. Rajan. The effect of credit market competition on lending relationships. *Quarterly Journal of Economics*, 110(2):407–443, 1995.

Rajan, Raghuram G. Insiders and outsiders: The choice between informed and arm's-length debt. *Journal of Finance*, 47(4):1367–1400, 1992.

Sampaio, Matheus C. and José Renato Haas Ornelas. Payment technology complementarities and their consequences in the banking sector: Evidence from brazil's pix. Banco Central do Brasil Working Paper Series No. 600, 2024.

Sarkisyan, Sergey. Instant payment systems and competition for deposits. Working paper, 2024.

SEBRAE. Pulso das empresas: Survey of micro and small enterprises. Serviço Brasileiro de Apoio às Micro e Pequenas Empresas, 2025.

Sharpe, Steven A. Asymmetric information, bank lending, and implicit contracts: A stylized model of customer relationships. *Journal of Finance*, 45(4):1069–1087, 1990.

Stiglitz, Joseph E. and Andrew Weiss. Credit rationing in markets with imperfect information. *American Economic Review*, 71(3):393–410, 1981.

Sun, Liyang and Sarah Abraham. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199, 2021.

Suri, Tavneet and William Jack. The long-run poverty and gender impacts of mobile money. *Science*, 354(6317):1288–1292, 2016.

Sutherland, Andrew. Does credit reporting lead to a decline in relationship lending? evidence from information sharing technology. *Journal of Accounting and Economics*, 66(1):123–141, 2018.

A Additional Figures

This appendix collects supporting figures referenced in the text: the within-bank shift toward unsecured and corporate lending after joining Open Finance (Figures 8–9) and the classification of relationships by prior information (Figure 10).

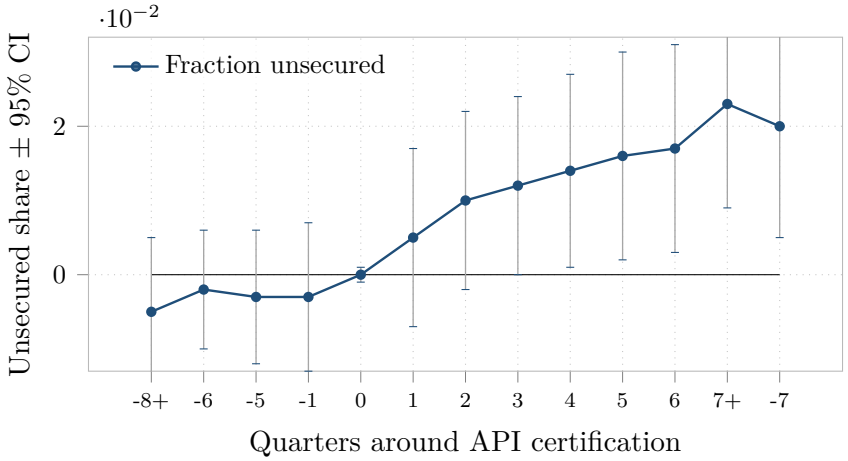


Figure 8: The unsecured share of the portfolio rises after a bank joins Open Finance. Within-bank event study of the fraction of the portfolio that is unsecured, around the quarter of API certification (normalized at -1). The unsecured share rises by one to two percentage points over the post-period, consistent with cash-flow information substituting for collateral. Bands are 95% confidence intervals.

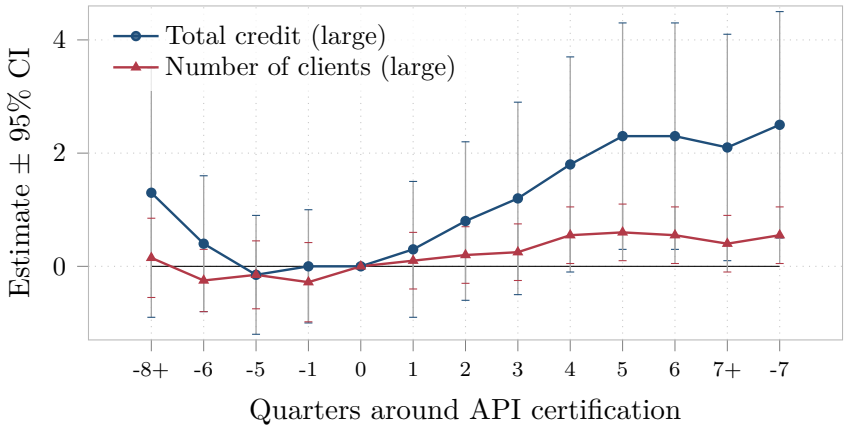


Figure 9: Corporate lending after a bank joins Open Finance. Within-bank event study of (log) total credit and the number of large-firm client relationships around API certification. Both rise in the post-period. Bands are 95% confidence intervals.

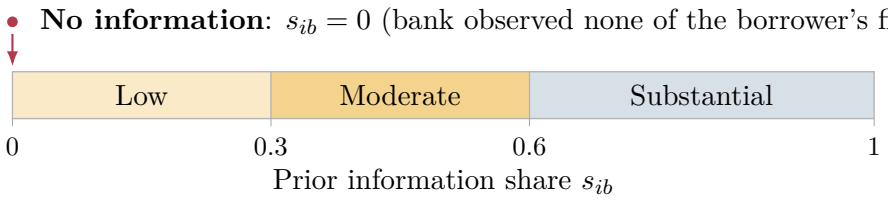


Figure 10: Classifying bank–borrower relationships by prior information. Each relationship is grouped by the share s_{ib} of the borrower's electronic cash flows that the bank already observed before Open Finance: no information ($s_{ib} = 0$), low $(0, 0.3]$, moderate $(0.3, 0.6]$, and substantial $(0.6, 1)$. These groups underlie the heterogeneity in Table 2.