Bank of England

Customer data access and fintech entry: early evidence from open banking

Staff Working Paper No. 1,059
February 2024

Tania Babina, Saleem Bahaj, Greg Buchak, Filippo De Marco, Angus Foulis, Will Gornall, Francesco Mazzola and Tong Yu

Bank of England

Staff Working Paper No. 1,059

Customer data access and fintech entry: early evidence from open banking

Tania Babina,[(1)] Saleem Bahaj,[(2)] Greg Buchak,[(3)] Filippo De Marco,[(4)] Angus Foulis,[(5)] Will Gornall,[(6)] Francesco Mazzola[(7)] and Tong Yu[(8)]

Abstract

Open banking (OB) empowers bank customers to share transaction data with fintechs and other banks. 49 countries have adopted OB policies. Consumer trust in fintechs predicts OB policy adoption and adoption spurs investment in fintechs. UK microdata shows that OB enables: i) consumers to access both financial advice and credit; and ii) small and medium-sized enterprises to establish new fintech lending relationships. In a calibrated model, OB universally improves welfare through entry and product improvements when used for advice. When used for credit, OB promotes entry and competition by reducing adverse selection, but higher prices for costlier or privacy-conscious consumers partially offset these benefits.

Key words: Open banking, entrepreneurship, fintech, financial innovation, data access, data rights, data portability, Big Data, financial regulation, financial sector, banks.

JEL classification: G21, G28.

---

(1) Columbia University. Email: tania.babina@gsb.columbia.edu
Corresponding author. Columbia University. Email: tania.babina@gsb.columbia.edu
(2) University College London and Bank of England. Email: s.bahaj@ucl.ac.uk
(3) Stanford University. Email: buchak@stanford.edu
(4) Bocconi University. Email: filippo.demarco@unibocconi.it
(5) Bank of England. Email: angus.foulis@bankofengland.co.uk
(6) University of British Columbia. Email: will.gornall@sauder.ubc.ca
(7) ESCP Business School. Email: fmazzola@escp.eu
(8) Imperial College London and Financial Conduct Authority. Email: t.yu18@imperial.ac.uk

We are grateful to Mark Boyd for sharing Platformable's data and Zhaochen Jiang for outstanding research support. For comments, we are grateful to Isha Agarwal, Jonathan Berk, Bo Bian, Charlie Calomiris, Olivier Darmouni, Douglas Diamond, Maryam Farboodi, Zhiguo He, Sean Higgins, Steve Kaplan, Arvind Krishnamurthy, Theresa Kuchler, Luc Laeven, Stefan Nagel, Tomek Piskorski, Ragu Rajan, Amit Seru, Amir Sufi, Ansgar Walther, Luigi Zingales, Yao Zeng, and seminar and conference audiences at the Barcelona Summer Forum; Bank of Canada workshop; Bank of Sweden; CKGSB; Columbia; HEC Lausanne; EFA; Entrepreneurship and Innovation at Nova SBE; Erasmus University in Rotterdam; Esade Spring Workshop (Barcelona); FDIC 21st Annual Bank Research Conference; FIRS; HEC Workshop on Entrepreneurship; Imperial College London; Maastricht University; NBER SI Corporate Finance; NBER Economics of Privacy Conference; NFA; NY Fed Fintech Conference; NYU WAPFIN (Women Assistant Professors in Finance); OCC Conference; Oxford University; Rome Junior Finance Conference; Stanford; Stockholm School of Economics; Wharton; UBC; UBC Winter Finance Conference; Universidad Carlos III de Madrid; University of Warwick; USC; and UTD Finance Conference. Gornall thanks the SSHRC for their support. The FCA Financial Lives Survey data for this research has been provided by the Consumer Data Research Centre, an ESRC Data Investment, under project ID CDRC 1512, ES/L011840/1; ES/L011891/1. The views expressed in this paper are those of the authors and do not necessarily reflect the views of the Bank of England, the MPC, the FPC, or the PRC, nor those of the Financial Conduct Authority.

The Bank's working paper series can be found at www.bankofengland.co.uk/working-paper/staff-working-papers

Bank of England, Threadneedle Street, London, EC2R 8AH
Email: enquiries@bankofengland.co.uk

ISSN 1749-9135 (on-line)

The increasing ease with which data is collected, stored, and analyzed has made data a critical input in economic decision-making. Data's growing economic importance has led to an active discussion about who should control the data generated through private economic activity: A firm or its customers. This issue is particularly salient in the financial services sector, where banks' provision of financial products inherently generates useful customer data. Periodic direct deposits, overdrafts, and late payments help predict a potential borrower's riskiness. Account balances and transactions allow firms to learn about a customer's needs and offer financial advice and other tailored products. A small business's transaction data could inform lenders about its health and help a fintech deliver financial management services.

Historically, customer data has been under the bank's exclusive control and gave banks a comparative advantage in pricing and customizing financial services.[1] However, banks' exclusive access to their customers' financial transaction data is being upended by a movement known as open banking (OB). OB is the trend of empowering bank customers to share their financial transactions data from their bank accounts with other financial service providers. For example, OB allows a bank customer to use a phone application (app) to easily share her bank account history with a potential lender (which can analyze her income and spending habits to underwrite her credit) or with a financial management app (to help her manage her money).

While some banks have implemented OB of their own accord, many governments have taken an active role in promoting or even mandating it. As of October 2021, the regulators of 80 countries have taken steps—some major and others still tentative—to implement policies to promote the adoption of OB with 49 of these having already adopted their key OB policies. In adopting these policies, policymakers want to promote innovative entry, competition, and financial inclusion. Policymakers reason that allowing bank customers to share their financial transaction data will allow fintech entrants and other banks to better compete for business.

In this paper, we explore the causes and consequences of government policies to promote OB. In doing so, we make three key contributions. First, we assemble the first comprehensive, standardized dataset of government-led OB policies. Using this data, we document the ubiquity of OB government policies around the world and examine their political and economic drivers. Second, we provide some early evidence on the effects of OB policies on fintech entry, consumers, and small and medium enterprises (SMEs). Finally, we provide and calibrate a quantitative modeling framework for customer data sharing, which measures the overall and distributional effects of OB.

We begin with careful data collection on government policies to promote OB around the

[1] As a motivating example, Appendix Figure A1 Panel (a) shows non-banks and fintech lenders, which lack such customer data, overwhelmingly use standardized underwriting models such as FICO when originating US residential mortgages. Banks are much more likely to use non-standard credit models, allowing them to exploit their customer data. These non-standard models lead to more individualized pricing: Panel (b) shows that non-standard models lead to more dispersed interest rate residuals than standard models.

world and assemble a comprehensive dataset covering the 168 largest countries—representing more than 99% of world GDP and 98% of world population. We uncover vast heterogeneity in OB policy choices. For example, countries in the European Union (EU) have adopted OB regimes with mandatory data sharing by banks, but without regulator-supplied technical standards for data sharing. In contrast, East Asian countries have favored voluntary bank participation, but spelled out detailed technical standards. We find that consumer trust in sharing data with fintechs predicts OB policy adoption: Intuitively, consumer trust increases the potential benefit of these policies, as people's willingness to share their financial data is crucial for OB's operation. Other country characteristics are less predictive, including economic and financial development, levels of innovation, or the quality of local institutions.

We next provide initial evidence on the economic effects of OB policies. We first examine whether OB policies successfully promote fintech entry, which regulators view as a key mechanism through which OB can improve financial outcomes. We measure fintech entry using data on venture capital (VC) investment in fintech startups. Using the staggered implementation of OB policies across countries in a difference-in-differences design, we show that the number of VC-backed fintech financings increases by a third and the amount of money invested doubles following OB policy adoption. Event studies show a discontinuous increase in fintech activity after the introduction of OB policies, with no pre-trends. Countries whose residents place more trust in sharing data with fintechs see greater post-OB fintech VC investment, suggesting that consumer preferences for data sharing play an important role in OB's impact. Importantly, we observe increases in fintech activity across many financial products (e.g., financial advice applications, credit, payments, regtech), consistent with the hypothesis that OB data is useful for a wide range of financial products beyond credit underwriting.

Granular microdata from the UK—an early adopter of OB—allows us to supplement our cross-country analyses with direct evidence on the adoption and economic impact of OB for both households and SMEs. For households, we use a premier survey on consumers by the UK Financial Conduct Authority to examine two distinct reasons households share data: Borrowing (which we term credit OB) and financial planning and management (advice OB). The prevalence of these non-credit-related uses is consistent with our finding of an OB-led increase in fintech activity across many financial products and sets OB apart from credit registries. There is little overlap between the users of credit and advice OB, although consumers are more likely to use both if they are willing to share their data, are employed, or have missed bill payments. We find suggestive evidence that OB use improves consumer outcomes, with advice OB use being associated with greater financial knowledge and credit OB use being associated with greater access to credit products.

For SMEs, we use panel data from the UK to examine how eligibility for OB impacts SMEs' outcomes. This fills the key empirical gaps in our findings thus far by allowing us to provide a well-identified causal interpretation of the impact of OB on borrowers, to measure whether banks or non-banks provide new loans, and to examine OB's financial inclusion implications. We exploit the fact that the commercial OB-related policy applied only to SMEs with annual turnover below £25 million. This cutoff provides quasi-random variation and allows us to compare outcomes for eligible and non-eligible SMEs following the policy implementation in difference-in-differences and event-study designs. Being eligible to share data makes SMEs more likely to form new lending relationships with non-bank lenders, consistent with increased fintech entry following OB policies. In terms of distributional effects, we find that treated firms *with* prior lending relationships are more likely to get new loans and those SMEs that form new lending relationships with non-banks pay less interest.

Overall, the introduction of OB policies leads to increased fintech entry across a wide range of financial products. Among UK consumers, OB is used for financial advice and credit products, and these uses are associated with greater financial knowledge and credit access, respectively. Among UK SMEs, firms affected by OB are more likely to form new lending relationships, especially with non-banks. These new relationships are driven by the SMEs with prior lending relationships—a finding at odds with the financial inclusion goals of OB policies but fully consistent with the distributional predictions of our model.

While our empirical results offer valuable descriptive and causal evidence regarding OB use, they fall short of addressing several key economic and policy questions related to OB. First, they are largely silent on the economic mechanisms by which access to OB data increases entry across two seemingly distinct data use cases—financial advice and credit—that impact mostly disjoint segments of consumers, as our UK consumer results show. Second, our differences-in-differences tests naturally have little to say on welfare, equilibrium effects, or distributional consequences. Third, while the consumer and SME microdata is informative about the UK case, our cross-country results highlight the importance of customer preferences for sharing data, which raises questions about how OB might look in countries with different social attitudes towards fintechs and privacy. Addressing these questions is not only academically important but also essential for policymakers as OB rule-making continues around the world.

We tackle these questions directly by using a quantitative model of data usage. This model extends standard IO models of consumer choice with heterogeneous consumers by incorporating consumer data use into the model. In our model, data about a bank's customer—interpreted as either a consumer or a business—reveals her preferences (allowing the creation of better products for advice OB) and costliness to serve (allowing screening for credit OB). A relationship bank always sees her data, while other firms see it only if she shares it via OB. We calibrate the model to two financial products using our reduced-form results and pre-OB estimates from the literature. In our credit use case, we calibrate to mortgage products, and consumer data is informative about consumer risk. In our financial advice use case, consumer data is informative about customer needs and allows for the provision of better products. OB spurs innovation and competition in both cases, but through different channels. In the credit OB case, unequal data access discourages entry by giving relationship banks an underwriting advantage and creating adverse selection for entrants. Allowing data sharing reduces this adverse selection, makes entrants more profitable, and, in equilibrium, increases entry. In the advice OB case, unequal data access impairs fintechs' ability to offer customized products, and enabling customers to share their data leads to improved customized products, higher customer demand, and, again, increased entrant profitability and entry.

While OB unambiguously increases competition and innovative entry, our model also shows how these goals can sometimes—but not always—come into conflict with the financial inclusion goals of OB policies. The distributional effects of OB depend critically on how the data is used. All customers benefit in the advice OB use case where the data is used to provide higher-quality or more customized products. In contrast, the credit OB use cases that rely on screening can have negative distributional consequences. OB increases entry for such firms precisely by enabling entrants to better exclude unprofitable (higher risk) customers—and the excluded customers can lose out. Users who share unfavorable data lose directly, while users who opt out of data sharing are inferred to have unfavorable data that they are strategically hiding, regardless of whether it is due to unfavorable financial profiles or strong privacy preferences. Thus, consistent with our reduced-form findings in the SME analysis, the customers who benefit the most may be those who already have credit access. Customers who opt out still gain from increased entry and competition, but lose because they are now inferred to be a higher risk. Our quantitative model allows us to weigh that tradeoff, a particularly important question for policymakers for whom financial inclusion and distributional effects are often first-order concerns.

Our model shows that societal preferences for privacy not only drive the impact of OB (as expected and consistent with our cross-country results), but they also play an important role in explaining these distributional effects. The financial transaction data shared through OB differs from credit registry data not simply because of its utility for generating financial advice, but also because it is by nature more sensitive and many customers are reluctant to share it, as highlighted by our UK consumer data. In our model, strong societal preferences for privacy blunt the impact of OB as few customers opt in to data sharing and so few firms enter. However, preferences for privacy have a silver lining: Customers who opt out of data sharing are more likely to be privacy-conscious than strategically hiding their data, thereby avoiding a strong negative inference about their type. In fact, under reasonable parameters—including those obtained in our UK calibration—OB is welfare-improving for *all* customers even when data is used for screening, because the negative inference lenders draw against opt-outs is more than offset by the benefits that these customers derive from increased firm entry and innovation. Consequently, incorporating privacy preferences and the implications of different use cases is an important part of an OB implementation, highlighting an important distinction between credit registries and OB data sharing.

To summarize, we document that government policies to promote OB are very prevalent, with about half of the countries having some OB efforts. Our empirical analyses and the quantitative model show that OB data can have beneficial economic effects. Our work suggests that the potential implications of OB for academics, policymakers, and industry are large. By giving customers the ability to share their financial transaction data, OB promises to upend the organization of the financial sector. The welfare and distributional effects of this, however, depend crucially on specific uses of customer data as our model shows.

Next, in Section 1, we situate our contribution in the literature. In Section 2, we describe our data. In Section 3, we examine the effects of OB policies, and in Section 4, we provide an economic framework for evaluating our results. Section 5 concludes.

1 Related Literature

Our paper contributes to several strands of literature. First, our research question and methodology connect to the broader literature on cross-country bank regulation. In the wake of the financial crisis, much of this literature focuses on regulation and bank risk, for example, Laeven and Levine (2009), Beck et al. (2013), and Ongena et al. (2013). Our paper is closer to research on regulation and competition, such as that by Claessens and Laeven (2004) who argue contestability and regulation are key drivers of bank competition or Barth et al. (2004) who argue for the role of disclosure and private incentives. We contribute by showing that government policies to promote bank customer data sharing foster entry into the financial sector across many financial products.

Second, we engage with the fundamental question, originating with Diamond and Dybvig (1983) and Diamond (1984), over what, if anything, makes banks special relative to other financial intermediaries. While fintechs and other non-depository institutions have gained significant market share in transaction-oriented functions like origination and servicing, as Gopal and Schnabl (2022) and Buchak et al. (2018) show, they have been slower to replace banks in deeper intermediation roles like underwriting, monitoring, and balance sheet lending. Importantly, banks appear to derive significant value from engaging in multiple intermediation activities simultaneously, as in Egan et al. (2022), Aguirregabiria et al. (2019), or Benetton et al. (2022), which suggests there may exist significant barriers that limit the growth of new single-product competitors in these roles.

Information lies at the heart of relationship banking (Ramakrishnan and Thakor, 1984; Boot and Thakor, 1997) and our paper directly addresses the idea that aggregating data across multiple business lines leads to significant informational advantages. This explanation dates to Petersen and Rajan (1994), Petersen and Rajan (1995), and, more recently, Granja et al. (2022). Recent empirical work by Ghosh et al. (2022) shows, for example, a direct effect of transaction data on screening quality for Indian commercial loans. Berg et al. (2020) and Di Maggio et al. (2022b) show the value of alternative data more generally. OB provides an empirical setting in which banks potentially lose the informational advantage that their wide scope provides to them and paves the way for an analysis of how important these informational advantages are to banks.

Third, we add to the nascent literature examining the implications of data ownership rights. The growing theoretical work on data use typically views data as either an input to production that improves product quality or a way to improve screening or monitoring in settings with information asymmetries. Mandated data sharing generates complex competitive interactions that depend on how the data is used. Taking the production-input view, Jones and Tonetti (2020) show that a firm may hoard product-improving data to prevent entry, and giving data property rights to consumers can generate allocations that are close to optimal. Farboodi et al. (2019) model customer-generated data as valuable in forecasting business conditions and suggest that large firms benefit more from data, a fact confirmed empirically by Babina et al. (2023a) who show that larger firms benefit more from their AI investments. Emphasizing the information economics view, theoretical finance literature like He et al. (2023) and Parlour et al. (2022) highlight how data sharing and portability can increase the quality of lending while having ambiguous effects on consumer welfare and bank profits. Goldstein et al. (2022) emphasizes the theoretical connection between liquidity transformation and lenders' access to information.

We build on this largely theoretical literature in two ways. We provide, to the best of our knowledge, the first empirical study on the impact of government policies that open access to rich customer-level financial and transaction data. While conceptually related to credit registries, e.g., Djankov et al. (2007) and Hertzberg et al. (2011), OB policies differ in important respects. They typically cover consumers regardless of their credit usage and are designed from the outset to facilitate ease of data access by potential bank competitors, including non-banks. The richer data that OB covers (including transactions, income, and savings data) lends itself to uses beyond assessing consumer risk; however, this very richness creates greater privacy concerns than a standard credit file.[2] As we show, these aspects of OB are important in driving its effects. Thus, our paper speaks to these important differences while also providing evidence of the effects of adopting data-sharing policies more generally. Beyond that, we provide a general-purpose quantitative framework for studying the production and use of consumer data in the context of OB. Building on common tools in the IO/finance literature (e.g., Egan et al. (2017), Di Maggio et al. (2022a), Buchak et al. (2023), Benetton et al. (2022)), we connect data to knowledge of consumer heterogeneity around marginal costs and desired customization. Through these channels, we synthesize both the input-to-production

[2] For example, Nam (2022) looks at a German OB fintech and shows that the vast majority of its credit report-sharing applicants are unwilling to also share their OB data.

and information economics views of data and highlight their quantitative importance across particular applications. In contrast to the theoretical models of, e.g., He et al. (2023) and Parlour et al. (2022), our model emphasizes new firm entry and innovation, which is a key policy goal of OB. Because it is quantitative, the model can be easily and credibly calibrated with standard techniques and estimates already in the literature. Finally, our analysis complements this literature by highlighting the importance of consumer preferences over data privacy documented by, e.g., Acquisti et al. (2016), Tang (2019) and Bian et al. (2021). We explicitly incorporate it into our structural model, allowing us to explore novel interactions between privacy preferences, entry and competition, and distributional consequences.

Fourth, our structural model allows us to connect to and broaden the literature around the industrial organization of the financial sector. This literature has studied the role of banks and the increased competition they face from non-depository institutions, e.g., Buchak et al. (2023), Buchak et al. (2018), Fuster et al. (2019), Jiang et al. (2020) (mortgages), Erel and Liebersohn (2022), Gopal and Schnabl (2022) (small business lending in the US), Di Maggio and Yao (2021), De Roure et al. (2022) (personal loans), and Buchak et al. (2021) (deposits). These papers typically highlight the complex interplay between technology and regulation and how they interact with the comparative advantages of depository and non-depository institutions.[3] Our results also connect to the growing literature on financial system structure and financial inclusion (e.g., Claessens and Rojas-Suarez (2016), Bartlett et al. (2022), or Philippon (2019)).

Finally, our paper is connected to the literature on the drivers of entrepreneurship and innovation.[4] We show a large effect of OB policies on innovative entry, which adds to a literature that has shown mixed results on whether policymakers are able to promote high-growth entrepreneurship. Acs et al. (2016) question the general effectiveness of public policies to encourage entrepreneurship, with subsidies of angel investing found to be ineffective (Denes et al., 2023), while Bai et al. (2022) argue government funding of early-stage companies increases local innovation. Other work shows the positive impact of less entry regulation (Klapper et al., 2006; Mullainathan and Schnabl, 2010), more optimistic beliefs (Puri and Robinson, 2007), venture capital availability (Kaplan and Lerner, 2010), weaker competition laws (Phillips and Zhdanov, 2017), lower investor eligibility requirements (Lindsey and Stein, 2019), R&D subsidies (Babina and Howell, 2022), and academic funding (Babina et al., 2023b). Beyond policy impacts, we join the relatively sparse literature connecting data

[3]Literature reviews on the impact of technology in finance can be found in Stulz (2019), Vives (2019), Allen et al. (2021), Thakor (2020), Berg et al. (2022), and Boot et al. (2021).

[4]Entrepreneurs play a crucial role in prominent theoretical explanations for economic growth, including Schumpeter (1911), Lucas Jr (1978), and Baumol (1990). Relative to incumbent firms, new firms have faster productivity and employment growth. This literature includes Kortum and Lerner (2000), Foster et al. (2008), Gennaioli et al. (2013), Haltiwanger et al. (2013), Decker et al. (2014), Glaeser et al. (2015), and Akcigit and Kerr (2018).

access to innovation. Recent work by Ahnert et al. (2022) shows that bank use of information technologies increases employment in new firms. We contribute by showing that government efforts to promote data sharing in the financial sector have fostered investments in VC-backed fintechs.

2 Institutional Background, Data, and Descriptive Analysis

This section describes the institutional background of OB, describes our data collection process, provides high-level summary statistics, and examines the drivers of OB policies.

2.1 Institutional Background on Open Banking

OB describes a broad trend where, upon customer consent, financial intermediaries share—willingly or by regulatory fiat—access to their customers' data with other financial service providers. There are two primary non-mutually exclusive ways in which OB is spreading around the world: Market-led, where banks and fintechs adopt OB without government intervention, and government-led, where regulators institute policies to promote the adoption of OB by the financial sector. This paper focuses on government-led OB policies, which typically apply to both consumer and business bank customers.

While the specifics of government OB efforts vary dramatically, the United Kingdom's (UK) Open Banking Initiative provides an instructive introduction: In 2017, the UK's Competition and Markets Authority (CMA) introduced one of the first OB regulations——commonly known as the CMA Open Banking Order——with the aim of increasing innovation and competition in the retail banking sector. The initiative required that by 2018, banks "*give their personal and business customers the ability to access and share their account data on an ongoing basis with authorized [by the government] third parties.*"[5] Here, third parties refer to both fintechs and other banks. Additionally, banks were required to allow third parties to make payments authorized by banks' customers from their accounts—a practice called payment initiation. OB differs from the UK's existing private sector credit registries in important ways: It covered richer data (in particular, information on transaction accounts, rather than credit only, is shared), it gave banks' customers more control over their data, it was free to the requester, and banks were forced to participate. These differences are common in other countries and mean that OB goes well beyond traditional credit registries.

Data access and payment initiation typically occur through a bank-provided Application Programming Interface (API). APIs are a technology that allows two computer systems (e.g., a bank's and a fintech's) to speak to each other over a network. OB APIs are published by

[5]Page 11 of "Open Banking, Preparing for Lift off" document. See official policy document. A related data-sharing policy focusing exclusively on SME bank customers was introduced in 2015 and implemented in 2017. We discuss this policy in detail in Section 3.3 and Appendix E.

the data provider and are a set of standardized, programmatic commands that allow data users to interact with the provider's customer database and to perform financial services on customers' behalf. The particulars are regime-specific, but API functionality in OB typically allows read access (e.g., querying account data) and sometimes allows write access (e.g., payment initiation).[6] In Appendix B, we show that in countries that implement OB policies, banks are indeed more likely to provide APIs for customer data sharing.

By opening bank data, regulators aim to create an environment where financial intermediaries—both incumbents and fintech entrants—can create new or improved financial services for bank customers and better compete with existing services. The prototypical use case of OB is a financial advice product, such as financial account aggregation, which works as follows. A consumer might have financial accounts scattered across several financial intermediaries: Her bank account, several credit cards, a mortgage, an investment account, and so on. Rather than separately monitoring each of her accounts, she may find it helpful to have this information collected and displayed in a single place. What are her spending habits? Does she have recurring payments or subscriptions she may have forgotten about? Which credit product should she pay down first? How much should she contribute monthly to her retirement account if she wishes to retire by a certain age? With OB, fintech startups can access, aggregate, and analyze these separate accounts to provide customized financial advice.

Other use cases of OB include credit, where potential lenders can access the myriad, and otherwise private, information that a consumer's home bank has about her. For example, with a customer's permission, a fintech lender could use the customer's bank's API to query her bank account transactions and payroll information to help price a loan to that customer. In this way, OB can reduce search costs and level the information playing field between a consumer's home bank and potential competitors. Beyond financial advice and credit, many other use cases have emerged, including identity verification, savings, accounting, automatic overdraft borrowing, and financial product suggestions.

2.2 Data Collection Methodology for Open Banking Around the World

We create a comprehensive and detailed database of OB government policies around the world. Our hand-collected dataset details the OB government policies (or the lack thereof) of the largest 168 countries. This section describes our methodology broadly; Appendix C

[6] While API-enabled OB is currently mainstream, fintechs have historically achieved similar functionality through what is known as "screen scraping" where a customer gives her login credentials for each of her financial institutions to the fintech (e.g., Mint.com). The fintech's software then uses the customer's credentials to log in to each financial institution (as if it were the customer) and extract account data from the financial institution's webpage. Although screen scraping accomplishes similar results to accessing an OB API, screen scraping has numerous weaknesses, including security risks, privacy issues, inefficiency, and unreliability. The API-enabled OB approach addresses these issues.

provides further detail. We base our sample on countries with at least one million people according to the IMF 2018 data or at least 10 VC-backed companies.[7] We aim to be as comprehensive as possible while focusing on a sample of countries for which there is reliable data on OB initiatives, if they exist. In total, we collect data on OB for 168 countries, representing more than 98% of global population and more than 99% of global GDP.

For each country, we manually search for official OB policy documents using Google, and when those are not available, for descriptions of government-led OB initiatives from law firms, research papers, journalists, and industry participants.[8] We classify these policies on multiple dimensions, giving preference to official policy documents (laws, regulations, policy papers, and official statements) to classify the various dimensions of OB policies into standardized categorical variables. Where official policy documents are unavailable, we use other sources.

We ensure accuracy by performing multiple cross-checks. First, two authors independently classify each country's OB regime and jointly reconcile any discrepancies. Second, we use automated news topic searches to uncover any material potentially missed in our manual searches.[9] Third, we reconcile our results against a database of OB regulations maintained by Platformable,[10] an OB advocacy group.

2.3 Summary Statistics on Open Banking Government Policies

Table 1 provides summary statistics on our hand-collected OB data both overall and by region, which we now discuss. As of October 2021, 80—or 43%—of the 168 countries in our sample have at least a nascent government OB effort and 49 have adopted their key OB policies. There is significant heterogeneity by region. 80% of countries in Europe and Central Asia have conducted at least some government OB policies. OB is less present in other regions but all regions in the world have seen at least some government OB effort.

OB regulators frequently cite one or more justifications for implementing OB regimes in their official policy documents and interviews. The three most common are to promote innovation, competition, and financial inclusion. Table 1 shows that 97% of regulators cite innovation as a policy goal; 82% cite competition, and 29% cite financial inclusion.[11] There

[7] The IMF data is from here. The VC data is from PitchBook and are described later in this section.

[8] We use Google as our primary search engine because it has the lion's share of the world search market (88% in June of 2021; see Statista.com). To ensure that using Google does not bias our findings for countries that rely more on other search engines, we also tried using local search engines (e.g., "Yandex" in Russia, "Baidu" in China). We generally found that these alternative search engines did not provide additional relevant articles.

[9] For a given country, a program searches Google for all news articles mentioning ("country name" and "open banking" and ["government" OR "central bank" OR "law" OR "regulation" OR "regulatory framework" OR "supervision")]. This search provides a list of sources of potentially relevant information on government OB interventions to compare to our manual collection. A research assistant then reads the top 10 resulting articles for any discussion on government OB–however small—and flags articles for review by the authors.

[10] Platformable's data is described here.

[11] This variable is missing for countries with no regulatory OB approach and for countries in the early stages

is significant regional heterogeneity in financial inclusion being an OB policy goal: Only 10% of countries in Europe & Central Asia cite financial inclusion, whereas other regions are much more likely to do so.

Finally, we note that the EU adopted and implemented a common OB framework known as the Revised Payment Services Directive (PSD2).[12] PSD2 obligated participating countries to implement its provisions in their respective banking regulations. In the country-level summary statistics later in this section, we keep the participating countries separate. For the purposes of our analyses in Sections 2.5 and 3.1, we weight all countries covered by PSD2 as a single pooled observation.

Implementation Status and Key Dates of Government-led Policies Government-led OB approaches vary both in when they were adopted and how far the implementation has progressed. For countries with some government interest in OB policies, we categorize a country's OB maturity in terms of its implementation status on a 0 to 7 scale, where 0 denotes no effort toward OB, 1–2 correspond to ongoing policy discussions, 3–5 correspond to being in the process of implementation, and 6–7 correspond to full implementation.[13]

Panel (a) of Figure 1 shows the geographical distribution of government-led OB initiatives based on their maturity. As of October 2021, among countries with a government-led approach to OB, 31 (38%) are at the discussion stage, 14 (18%) are in the process of implementation, and 35 (44%) are fully implemented or already seeing follow-on policies. We refer to the 49 countries in the latter two groups as having implemented OB. To provide three examples along the implementation timeline, OB discussion is underway in the US,[14] Brazil is in the process of implementing OB,[15] and the UK has fully implemented its Open Banking Initiative and is considering a follow-on "open finance" regulation.[16] Figure 1 Panel (b) shows the passage year of countries' major OB government policies.

of implementation that have not clearly indicated a policy mandate or objective.

[12]PSD2, Directive (EU) 2015/236.

[13]More specifically, the stages are (1) pre-discussion (some government interest is announced but no actual law or policy implementation is taking place); (2) discussion (the actual law has been discussed or rulemaking is taking place); (3) pre-implementation (the major policy-making has concluded but nothing is yet binding or implemented); (4) early implementation (some data sharing requirements are binding, e.g., bank-level product information, but not personal account/transactions); (5) mid-implementation (personal account/transaction data sharing is binding or OB infrastructure/technical standards have been put in place, but not all planned elements are in place); (6) fully implemented (full implementation as described in the law/rulemaking/policy documents); (7) follow-on regulation or policies (OB is implemented, and regulators are actively working on related policies, such as open finance or open data, or on implementing additional pieces of infrastructure for OB).

[14]The Consumer Financial Protection Bureau (CFPB) is looking into whether to create regulation based on Dodd-Frank's Section 1033 that gives consumers the right to their financial data, but which was never codified into rulemaking and, hence, is not legally binding. See here.

[15]See here.

[16]This policy would broaden data access beyond transaction accounts. See here.

Requirements Set by the Regulator OB government policies differ in what they require of market participants, and indeed, whether they require anything at all. The UK, for example, places explicit de jure legal requirements on banks to participate. Other examples with binding regulatory approaches are Australia, Bahrain, Brazil, the EU, and Israel. In contrast, regulators in Singapore, Malaysia, and Russia do not explicitly mandate data sharing and instead facilitate the adoption of OB by mediating industry discussion, providing technical standards for APIs, or providing infrastructure for data sharing.

As shown in Table 1, among the countries whose OB initiatives have advanced sufficiently for these issues to be decided, we find that 88% require banks to share data (variable "Required data sharing"), while the other 12% do not. In addition to requiring incumbent banks to share data, some OB regimes also require sharing by data users—non-bank financial intermediaries (e.g., fintechs). In other words, some regimes require sharing reciprocity while others do not. Our data shows that only 18% of regimes have data sharing reciprocity (variable "Data reciprocity"), where fintechs that use the data must share. Finally, 39% of countries' regulators lay out technical specifications for data sharing (variable "Regulator provides tech specs"), while the remainder do not. There is significant regional variation in government-led approaches regarding mandatory data sharing and technical specifications: Figure 2 Panels (a) and (b) show these differences graphically for mandatory data sharing and regulator-set technical specifications, respectively.

Open Banking Scope: Covered Services and Functions OB government policies differ dramatically in what financial products and services are covered. OB in its narrowest incarnation covers only transaction accounts: Checking accounts, credit cards, and digital wallets. Some regimes include a broader set of core consumer finance products: Savings accounts, investments, and loans. Still broader regimes, called "open finance", cover all financial services.

By definition, all OB regimes cover at least transaction accounts. Fewer—34%—additionally cover non-transaction accounts (variable "Beyond transaction accts"). Regarding regional heterogeneity, Europe & Central Asia OB policies tend to be very narrow in scope, with only 3% covering non-transaction accounts. In contrast, OB policies in other regions are much broader, with 90% going beyond transaction accounts.

Regarding functionality, OB APIs can, in theory, be used both to read data (e.g., pull customer account information) and to write data (e.g., initiate payments). Some OB regimes focus on data sharing only, and some on both. Our data shows that among those countries where this issue has been decided (variables under "Functionality scope"), only 5% focus on data sharing only, none on payments only, and 95% on both.

Open Banking Strength Index Using our hand-collected data on OB policies, we construct an OB Strength Index, which averages the four key OB policy dimensions discussed in the previous subsection. Namely, these four dimensions reflect whether the regulators have set policies that (i) mandate banks to share data, (ii) require financial service providers (such as fintechs) who use data to share data in return, (iii) cover a wide range of financial products, and (iv) set an API standard. This index ranges from 0 (all four dimensions are no or not yet decided) to 1 (yes on all four dimensions).

2.4 Non-Open Banking Policy Data

Venture Capital Data Spurring innovation is the key objective of OB policies; however, innovative output is notoriously hard to measure. We use data on VC investment into startups as a proxy for innovative entry, as past research has shown that VC-backed startups are generally innovative, fast-growing entrants (Puri and Zarutskie, 2012; Gornall and Strebulaev, 2021). This proxy is a forward-looking measure of profit-motivated investors' expectations, which helps us analyze the effects of still-recent policy interventions.

Using PitchBook data, widely acknowledged as one of the best VC data sources for more recent years, we construct a country-by-year panel of VC deals for the past twenty years, from 2000 to the first half of 2021. We measure VC activity using two standard variables: The number of deals and the investment amount in millions of US dollars. Our interest lies in financial innovation, so we split the deals in each country-year into fintech deals and non-fintech deals, with fintech deals being the deals PitchBook places in the "Financial Software" sub-industry or the "Fintech" vertical. We are interested in measuring the impact of OB on innovation around specific use cases, but Pitchbook lacks more granular product classifications. We overcome this by using PitchBook's keywords feature to define seven sub-industries of fintech: alternative lending, consumer finance, financial IT, payments, regtech (i.e., the use of technology to address regulatory processes), wealth management, and digital assets. Details of our classification are in Appendix D. Because of the cryptocurrency boom and bust cycles and the fact that digital assets are not related to OB, we reclassify digital assets startups as non-fintech for our main analysis, although this does not have any impact on our results.

UK Consumer Data from Financial Lives Survey The Financial Lives Survey is a nationally representative survey of UK consumers conducted by the Financial Conduct Authority (FCA)—one of the main regulators of the UK financial services industry. The survey provides information about consumers' attitudes towards managing their money, the financial products they have and their experiences of engaging with financial services firms, as well as information on consumer demographics. We use the February 2020 survey as this was the first time that consumers were asked about their usage of OB products.[17] See Table

[17]See the survey questionnaire here.

A6 for summary statistics.

UK SME Data from Companies House Firms in the UK are required to report all claims ("charges") lenders have against firms' assets, including lender (bank or non-bank) names, the date the claim commenced, and when the charge ceases, to Companies House (the UK firm Registrar).[18] The information on charges in Companies House is collected by Bureau Van Dijk (BvD) and provided in their FAME database. BvD data also provides annual firm-level financial information matched to charge-holders information.[19] Hence, we observe firms' lending relationships as well as their balance sheet and income statement information over time. See Appendix E for exact details of how our sample is constructed and Table E1 for summary statistics.

Country-level Characteristics We compile a variety of other country-level variables, which are summarized in Panel (a) of Table A1. We start with basic country-level data, including per capita GDP in thousands of US dollars and population in millions from the World Bank. Following Chen et al. (2023), we use the survey underlying the EY Global Fintech Adoption Index to measure trust in fintechs. We measure trust as the portion of survey respondents in each country who "agree" or "strongly agree" that they are comfortable for their main bank to securely share their financial data with fintechs. From the World Bank we also add standard measures of country-level financial sector development, including the quantity of private sector credit to GDP, the number of bank branches per 100k people, and the financial sector's Lerner Index. The Lerner index measures markups over marginal costs, ranges between 0 and 1, and captures the market power of banks, with higher values denoting less competition. We take the percentage of banks that are foreign-owned from Claessens and Van Horen (2013). Finally, to capture the quality of institutions, we use the Rule of Law Index from the Cato Institute, which is on a 0 to 10 scale with higher numbers denoting more favorable conditions.

2.5 Drivers of Open Banking Government Policies

An important preliminary question is what drives countries to adopt OB policies. In

[18]These reports are similar to Uniform Commercial Code (UCC) data on SME lending in the US where lenders make filings on all secured loans to preserve priority in bankruptcy (Gopal and Schnabl, 2022). The charge can be against a specific asset or it can be a charge covering the entirety of the firm's balance sheet or its outstanding invoices in the case of invoice financing. There are strong incentives to ensure this data is accurately reported. Lenders have 21 days to formally register their claim (or face legal barriers to repossessing the assets). Borrowers have an incentive to declare when a charge is satisfied to unencumber their assets. We do not observe unsecured claims. However, the overwhelming majority of loans to UK SMEs are collateralized and hence this data provides a highly representative and timely view of a firm's lending relationships.

[19]BvD data is well known for suffering from survivorship bias and various issues with constructing consistent historical panels (Kalemli-Ozcan et al., 2023). To alleviate this concern and maximize coverage of historical observations, we use annually sampled archived vintages of the FAME database, as in Bahaj et al. (2020), to compile our final panel dataset.

the spirit of Kroszner and Strahan (1999) or Cornelli et al. (2020), we examine what country characteristics predict OB policy adoption.[20] We test the association between the time of OB policy implementation and country characteristics using a Cox proportional hazards model:

$$h_i(t) = h_0(t) \exp(X_i'\beta + Region_r) \tag{1}$$

where $h_i(t)$ represents the hazard function for the occurrence of the OB outcome (implementation of OB policies though 2021) in year t in country i. This hazard function can be interpreted as the risk of the event happening at time t given it has not yet occurred. X_i' is a vector of country-level characteristics. $Region_r$ are region fixed effects that allow us to exploit within-region variation.[21] Data availability causes the number of observations to fluctuate across specifications.

We supplement this regression with a cross-country regression on open banking characteristics. We use both the 0 to 7 OB implementation status (the measure of how far government OB policy has progressed) and the 0 to 1 OB Strength Index (the measure of comprehensiveness of OB policies), which are described in Section 2.3. These regressions take the following form, where OB_i, denoting the two measures of OB for country i as of 2021, is the policy outcome of interest:

$$OB_i = X_i'\beta + Region_r + \epsilon_i \tag{2}$$

Table 2 presents the determinants of open banking adoption speed, implementation status, and policy strength. Columns 1 to 5 use a Cox proportional hazards model. Column 1 shows that consumer trust in sharing their data with fintechs is associated with earlier implementation of OB policies, despite the limited number of observations available for only 27 countries for the trust in fintech data. The effect is economically meaningful: A one standard deviation increase in trust (0.15) almost quadruples the propensity to adopt OB. Other country characteristics are only weakly associated with OB. Column 2 considers three measures of financial development, none of which robustly predicts government-led efforts to promote OB. Column 3 shows that OB policies are somewhat more likely to be adopted in countries with more non-fintech VC deals in 2013, but that fintech VC deals, in particular, are not predictive of adoption. The fact that fintech innovation is not driving OB policies is comforting because it speaks against preexisting financial innovation driving both OB policies and fintech VC deals. In column 4, we find weak and statistically insignificant associations between the adoption of OB policies and both the Cato Institute's Rule Law variable and

[20] The trust in fintechs variable is based on surveys conducted in February and March of 2019, as earlier survey vintages had very low coverage. All other variables are as of 2013, with that year chosen because it predates the earliest OB regimes and because it is the final year that comprehensive Lerner Index data is available from the World Bank.

[21] Following World Bank geographic terms, regions are Africa, Middle East & North Africa; Europe & Central Asia; Latin America & the Caribbean; North America; South Asia, East Asia & Pacific.

the fraction of foreign-owned banks in the country. In column 5, we include both our trust in fintechs measure and non-fintech VC deals as those were the significant predictors: The coefficient on trust in fintechs is unchanged, while the coefficient on non-fintech VC becomes statistically insignificant. Since low overall levels of economic development could be associated with the introduction of OB policies, in addition to our region fixed effects, in all columns we control for both GDP per capita (and its square) and log population to prevent this association from driving the results. However, across columns, neither a country's GDP nor its population robustly predicts the introduction of OB government policies.

Columns 6 to 9 use an OLS specification to test whether trust in fintechs is associated with the maturity (OB implementation) and comprehensiveness (OB Strength Index) of OB policies. Trust in fintechs is again associated with the maturity of OB policies, with a one standard deviation increase in trust being associated with about two steps of increase on our seven-step OB implementation scale (column 6). The coefficient is unchanged when we control for VC deals (column 7). Columns 8 and 9 show trust in fintechs is associated with our OB strength index, controlling for the adoption speed effects previously discussed. We find an association between fintech trust and OB policy strength, although it is of only borderline statistical significance.

Overall, trust in fintechs is associated with the adoption of OB policies. Trust increases the potential benefit of these policies, as people being willing to share their financial data is crucial to the operation of OB.[22] In the next section, we explore the economic consequences of open banking and whether they vary with consumer trust in fintechs.

3 Economic Effects of Open Banking

In this section, we examine the economic effects of OB. We first continue our cross-country analysis to show that fintech entry increases after the adoption of OB policies (Section 3.1). We next use more granular data from the UK to provide evidence that OB use is associated with improved consumer outcomes (Section 3.2) and leads to more fintech lending to SMEs (Section 3.3). These additional analyses both corroborate the more aggregate fintech analysis and provide more causal evidence on the impact of OB.

3.1 Open Banking Policies and Fintech Venture Capital Investment

Financial innovation is the most common goal of OB policies. Regulators hope that giving bank customers the ability to share their financial data with new entrants will spark the creation of new firms that offer innovative financial products and services. We mentioned

[22] A potential concern is reverse causality, as the trust in fintechs was based on a survey conducted in early 2019. However, since consumer trust is likely persistent in nature, we think this concern is unlikely to be of first-order importance.

in Section 2.1 that after OB adoption, banks provide the necessary technology (APIs) for new entrants to access bank customer data (results in Appendix B). We now test whether this data access spurs innovative entry by using data on VC investments in fintechs and a standard panel event-study design:

$$FintechVC_{i,t} = \sum_{k \neq 0} \beta_k \times OBLag(k)_{i,k,t} + Country_i + Region_r \times Year_t + \epsilon_{i,t}, \qquad (3)$$

where $FintechVC_{i,t}$ is a measure of fintech VC activity in country i and year t, measured as either the number of deals or the millions of US dollars invested.[23] $OBLag(k)_{i,k,t}$ is an event time indicator, equal to 1 if country i's adoption of OB government policy occurred k years from time t and zero otherwise.[24] We normalize the year of the policy's passage to zero so that the coefficient β_k measures changes in fintech VC activity k years before or after OB policy passage relative to the year of its passage. $Country_i$ and $Region_r \times Year_t$ are country and region-by-year fixed effects.

VC data poses two key challenges. First, VC activity is very skewed, with the US having far more VC investments than any other country. We correct for this using a $\log(1 + x)$ transformation of our VC activity measures, which means our tests measure relative increases or decreases in VC activity occurring, rather than absolute changes, as is common in the VC literature (e.g., Gompers and Lerner (1998) or Li and Zahra (2012)). Second, the lack of central VC investment registries in most countries makes collecting VC data challenging. Table A2 summarizes our data and shows that PitchBook, despite being one of the best VC databases, has significant gaps in its international coverage. Due to a combination of data collection and low VC activity, only one-quarter of our post-2000 country-years have any fintech VC deals and more than half have no VC deals at all. To reduce the biases created by using log-transformed variables in the presence of zeros and VC data coverage issues, we restrict our attention to countries with active PitchBook coverage. As our first government OB policy passage occurs in 2016 or later and PitchBook coverage improves over time, we restrict our analysis of VC activity to the 2011–2021 period. In addition, we consider only countries that PitchBook already covered before our regression sample period by focusing on countries with five or more fintech deals in the 2000-2010 pre-period, which we refer to as high-coverage countries.[25] Our focus on high-coverage countries and our tests using VC dollars, which load on large and hard-to-miss deals, help attenuate concerns that PitchBook

[23]The staged nature of VC investments means that deal counts tend to measure earlier stage investment and dollar amounts tend to measure later stage investment.

[24]For countries in the sample that never adopt OB, $OBLag(k)_{i,k,t}$ is zero everywhere; these countries help identify region-by-year fixed effects.

[25]Specifically, we consider Australia, Belgium, Brazil, Canada, China, Germany, Denmark, Finland, France, India, Ireland, Israel, Japan, the Netherlands, Norway, Poland, Russia, Spain, Sweden, the United Kingdom, and the United States of America.

coverage improvements are correlated with the passage of OB government policies.[26] Because we condition on pre-period deals, our results mostly speak to countries that already have developed VC markets.[27] Because our filter drops a large number of country-years that *never* had OB, identification in this specification comes chiefly (though not entirely) through the staggered adoption of OB within countries. Intuitively, our regression is comparing VC activity in countries at time t to other countries in the region that will adopt OB but have not adopted it yet. The key identifying assumption is that, absent the treatment, countries within a region would have been on parallel trends.

Figure 3 presents the results from the event-study specification in Equation (3) and shows a relative absence of pre-trends in fintech VC activity, followed by a sharp increase after the passage of OB policy. This pattern holds for the number of deals (Panel (a)) and the amount invested (Panel (b)). In both panels, there is a clear inflection point around the year of the OB policy passage and a change of large economic magnitude: Deals increase by almost half a log point and dollars by about a full log point. The absence of pre-trends is consistent with the parallel trends assumption and OB having a causal impact on country-level fintech VC activity. As we later discuss, these results are robust to a number of alternative specifications.

Our event studies show a sharp increase in fintech VC activity immediately following OB policy passage. While the real effect on consumers and SMEs will not occur until later after policy implementation which typically takes some time, financial investors such as VCs should react to information about future opportunities created by the policy passage. The sharp increase in fintech VC following law passage is a natural consequence of the uncertainty reduction around the timing of OB laws combined with VCs' short time horizons. For example, OB has been in the works in the US since the 2010 Dodd-Frank Act specified that consumers should own their financial transaction data, yet over a decade later, it has not been codified into regulation and hence does not bind on banks.[28] VCs target 30% returns and so see timing as a crucial factor for their financial performance (Gompers et al., 2020). High required returns and a desire to move fast mean that the VC industry is characterized by dramatic year-over-year changes in investment in response to perceived opportunities (Gompers et al., 2008).

Table 3 uses a difference-in-differences design to quantify the relationship between OB

[26] Although only 13% of countries are high-coverage, they include 91% of the VC deals and 94% of the investment value. Thus, our analysis of OB policies on fintech VC activity uses the sample of high-coverage countries in the 2011-2021 period. 99% of these high-coverage country-years have at least one fintech deal, dramatically reducing the econometric issues associated with log-transforming zeros.

[27] The results in Table 3 continue to hold with similar coefficients for the entire sample of countries; however, the large number of zeros makes it hard to interpret the results.

[28] The CFPB has recently announced that a regulatory framework for OB might be coming in 2024, see here.

policies and fintech VC activity:

$$FintechVC_{i,t} = \beta \times OB_{i,t} + Country_i + \gamma \times Non-fintechVC_{i,t} + Region_r \times Year_t + \epsilon_{i,t}, \quad (4)$$

where $OB_{i,t}$ is an indicator variable equal to one if OB was adopted in country i before year t and other variables are as in Equation (3). We are interested in the coefficient β which measures log change in fintech VC activity following the introduction of government OB policies. Alternative specifications remove the control for non-fintech VC, add the interaction of our trust in fintechs measure with OB passage ($\gamma \times OB_{i,t} \times Trust_i$), use year fixed effects instead of region-by-year fixed effects ($Year_t$), or include additional controls for potentially time-varying importance of trust in fintechs ($Trust_i \times Year_t$).

Across specifications, fintech companies receive significantly more VC investment following the adoption of OB policies, whether measured by the number of deals or the dollars invested. Our coefficients are both statistically significant and large in economic magnitude. Using our preferred specification from Equation (4), we find a 0.31 increase in log fintech VC deals (column 4 of Table 3) and a 0.87 increase in log fintech VC dollars (column 9). Our coefficients remain significant in a simple specification using only year and country fixed effects (columns 1 and 6) and in a specification where we keep region-by-year fixed effects but drop our control for non-fintech VC activity (columns 3 and 8). The median country-year in this data has 19 fintech VC deals worth \$89 million and so our effects estimated using Equitation (4) translate into an additional 7 deals and \$125 million dollars annually for the median country. Although these investments are small in absolute terms, small investments in companies with the potential to become large is a defining property of the VC industry.[29]

In Section 2.5, we identified consumer trust in sharing their data with fintechs as a potential driver of OB government policies. A natural question is whether trust in fintechs mediates the effect of OB on VC: Do higher trust societies see a larger post-OB passage increase in fintech VC? In columns 2 and 7, we provide suggestive evidence that trust mediates the effect of OB policies on fintech VC activity, with the coefficient on the interaction between OB passage and trust in fintechs being positive and significant at the 10% level for fintech VC deals and positive and insignificant for dollars invested. These relationships are tentative given that our trust measure is only estimated for a small number of countries and our VC data is inherently noisy. A potential confounder in this setting is that countries that had high trust in fintechs experience both increases in fintech VC activity and the passage of OB laws. However, our country controls absorb a time-invariant relationship between trust and fintechs. Moreover, in columns 5 and 10 we show that our results persist while controlling

[29]For example, less than \$3 billion was invested by US VCs up to 1981 (Gompers et al., 2008), yet that investment included a \$1 million investment in Microsoft and a \$150 thousand investment in Apple (Gornall and Strebulaev, 2021).

for trust-by-year fixed effects: This addresses a concern that trust was more important for fintechs in the later part of the sample and countries happened to be passing OB laws around the same time.

In Table A3, we exploit additional cross-country heterogeneity and its effects on fintech activity: OB policy choices (columns 1–4) and the overall comprehensiveness of OB policies (as measured by our OB Strengths Index in column 5). We find tentative evidence that OB policies that force banks to share their customers data drive our results; however, the power of these tests is limited due to the already-small sample.

We also test if OB spurs fintech entrants offering different financial products. This allows us to shed light on whether the new data made available by OB is used mainly for credit underwriting, or if it is useful for other financial products. Using Equation (4), Table 4 considers VC investments in companies targeting specific use cases as dependent variables.[30] Alternative lending shows a 0.66 log point increase; consumer finance, financial IT, payments, and regtech show increases of between 0.48 and 0.61 log points; and wealth management shows a statistically insignificant 0.43 log point increase. The notable and reassuring exception to this trend is digital assets, where we see an insignificant negative effect. This is intuitive and serves as a placebo test: Digital assets, such as cryptocurrency, are largely unrelated to OB functionality. These results hold (with wealth management gaining statistical significance and payments losing statistical significance) if we remove our non-fintech VC control (not reported). Although the size of each of these subindustries is small,[31] we find a broad-based increase in fintech activity, which suggests VCs anticipate OB data as offering value not just for credit issuance but for a variety of fintech use cases.

Our results continue to hold under numerous alternative specifications. First, following the recent econometric literature on biases created by difference-in-differences regressions with staggered treatment and heterogeneous treatment effects (see, for example, Goodman-Bacon (2021) or Sun and Abraham (2021)), in Figure A2 we show our results are essentially unchanged when we rerun our event studies using the methodology from Gardner (2022) which addresses this concern.

Second, we perform a number of tests to address potential concerns about confounding factors. In Figure A3 and Figure A4 we run two placebo tests: First, shifting the event dates and windows five years earlier, and, second, replacing fintech deals with non-fintech deals as the dependent variable in our main specification, as another precaution against a general rise in innovation causing our results. Reassuringly, we see no effect in either test. In Figure A5 we control for contemporaneous non-fintech VC deals as a proxy for innovation more generally and see the same pattern (following our main specification). This addresses a potential concern that OB adopters enacted broader innovation-promoting policies. In Table

[30] We define these fintech categories based on Pitchbook's fintech industry map (see Appendix D).

[31] Specifically, the median (mean) subindustry-year sees 4 (15) deals worth \$9 million (\$300 million).

A4 we show our effects persist when we rerun our tests while excluding first each country in turn, then Germany and France together (the two countries powerful enough to have an impact on the passage of OB government policy in the EU), and finally the three countries that did not implement OB in our sample (Canada, China, and the United States).

Finally, in Table A5, we show that our results hold under other transformations of fintech VC activity. We first repeat our baseline specification (columns 1 and 2) for comparison purposes. We then follow Jeng and Wells (2000) and consider fintech VC deals (or millions of US dollars invested) scaled by trillions of US dollars of GDP (columns 3 and 4). Next, we consider fintech VC activity divided by total VC activity (columns 5 and 6) and finally the inverse hyperbolic sine transformation of fintech VC activity (columns 7 and 8). Across specifications, we see statistically significant and economically meaningful effects.

3.2 Open Banking Use by Consumers and Their Financial Outcomes

In the previous section, we showed that there is a significant increase in fintech entry following the adoption of OB policies and that this increase is present across many different use cases (e.g., credit, financial advice applications, payments). While these results suggest potential increases in future innovation and competition, they do not provide direct evidence on the adoption and potential impact on OB users, such as consumers and SMEs. VC deals are a forward-looking measure of investment opportunities for high-risk, high-reward ideas. Hence, it is uncertain whether consumers have adopted OB and what effect that adoption has had on their financial outcomes.

We begin to address this gap by examining the use of OB by UK consumers and their financial outcomes using the data from the Financial Lives Survey (FLS) taken in 2020 and described in Section 2.4. This data has three key advantages. First, the survey asks consumers whether they use financial services based on OB, which allows us to provide novel evidence on the uptake of OB. Second, the survey collects a rich set of variables on consumer characteristics, allowing us to examine what type of consumers adopt OB. Finally, the survey collects data on consumer financial outcomes, which we use to examine the association between OB use and consumer financial outcomes.

We start by analyzing the uptake of OB by UK consumers. The survey asked 4,310 consumers who report having a day-to-day bank account (necessary to be able to share financial transaction data used in the provision of OB services) about their use of OB products. We split these products into two broad categories: Advice OB and credit OB. Advice OB is applications that provide information or services to users, such as financial advice apps, tracking apps that aggregate accounts, and apps that help users with savings. Credit OB is applications that offer credit, either directly (e.g., lending) or indirectly (e.g., credit ratings or price comparison).[32]

Table A6 provides summary statistics for our key variables. Among consumers who report knowing whether they use these types of services, 8.6% report the use of advice OB and 5.5% of credit OB. The high use of advice OB shows that OB data is valuable for more than just credit provision, consistent with our finding of an OB-led increase in VC fintech investment across a wide range of financial product categories.[33] Surprisingly, we find little association between the use of these two types of OB services. Only 13% of advice OB users also use credit OB, while only 20% of credit OB users obtain advice OB. This brings the total rate of (unique) OB users in this data to slightly more than 13%.

Table A7 shows the association between the use of each type of OB and consumer characteristics. In a cross-sectional regression, we regress whether a consumer used advice OB (column 1) and credit OB (column 2) on consumer characteristics, while controlling for location fixed effects. Consistent with our cross-country results on trust in fintechs, people who have concerns about sharing their data are less likely to use both types of OB. Employed people are more likely to share, in line with standard models of voluntary disclosure (e.g., Grossman (1981)) as employment status is important positive information absent from standard credit reports but shareable via OB. People who are missing bill payments are also more likely to share, suggesting more demand for both advice and credit for this financially vulnerable group. Finally, we find young people are also more likely to use advice OB and college educated use credit OB, and inconsistent links between OB use and gender, risk aversion, ethnicity, and marital status. Overall, these results suggest that consumer privacy preferences, employment status, and financial circumstances can explain the uptake of OB.

We next test whether OB usage is associated with better consumer financial outcomes. Table 5 shows how reported OB usage relates to financial knowledge (column 1) and credit product usage (columns 2 to 5). While cross-sectional data is poorly suited to deriving causal interpretations, we control for the potential confounders that we can observe in our rich data:

[32] The question we use to proxy for financial advice OB is "RB102c" which asks about the use of financial aggregation apps that allow consumers to see the accounts they hold with different banks in one place (e.g., Money Dashboard, Yolt, MoneyHub) and savings-related apps that help build savings by monitoring consumer current accounts and automatically transferring funds (e.g., Chip, Cleo, Moneybox, Plum). The question we use to proxy for credit OB is "RB102d" which asks about the use of credit products, such as firms offering lending products, credit reference agencies (which use OB to provide alternative credit scores), or price comparison websites (which use OB to prequalify borrowers or match them to lenders). The survey questions ask about specific OB products being used to address the fact that many consumers are unaware of exactly what OB is. In practice, this means OB use will be somewhat under-reported and these rates are a lower bound on the share of consumers using OB services.

[33] Advice OB being the most prevalent OB use case is also consistent with the recent report by the Open Banking Implementation Entity (OBIE), which is in charge of implementing the OB policy in the UK. They find that, among registered Third Party Providers (TPPs) authorized to use consumer transaction data, the largest group of TPPs offer financial decision-making services and these account for the vast majority of OB API calls. Using a sample of international OB fintechs from Platformable, we confirm that financial advice fintechs are more prevalent than credit fintechs.

We include all the consumer characteristics from Table A7 and location fixed effects.[34] In column 1, we find that consumers who use advice OB report 0.16 of a standard deviation higher knowledge about financial matters, potentially suggesting advice OB improves consumers' financial education and awareness. Here, a key concern is reverse causality: Financially savvy people could be more eager to use advice OB. However, a study of UK consumers shows that those who use advice OB are less financially confident ex-ante and report better financial awareness and decision-making ex-post.[35] Interestingly, credit OB use is not associated with improved financial knowledge, potentially because these applications, by design, do not aim to improve consumer financial literacy, but rather mostly help underwrite loans.

In columns 2 to 5, we look at the link between credit OB use and actual credit access.[36] We exploit the fact that OB data might be more used for some products than others. We view credit cards and personal loans as credit products that are ex-ante likely to benefit from consumer data, as both are unsecured credit where consumer transactions and income patterns can inform lenders about consumer creditworthiness. Credit OB users are more than 10% more likely to get both credit cards (column 2) and personal loans (column 3).[37] We use student loans and pawnbroking as a placebo test for OB as these credit products are ex-ante unlikely to be affected by OB. Due to UK regulation, student loan underwriting does not depend on consumer creditworthiness,[38] while pawnbroking is backed by physical collateral and tends to be very low-tech, suggesting it is unlikely to be affected by a fully digital underwriting process in OB lending. Consistent with this, neither student loans (column 4) nor pawnbroking (column 5) are associated with credit OB use.

Our credit card and personal loan results could be partially driven by reverse causality if consumers seeking credit from OB lenders sign up for OB too. Although reverse causality still shows an active role for OB, we can mitigate this concern by controlling for credit demand. Table A8 shows that our credit effects are robust to controls for credit demand using controls for credit use (columns 1 to 4; measured as the number of other credit products a consumer has) or tests with person-level fixed effects (column 5; the specification is run on product-by-person-level data). Together, these results show that credit OB use is robustly associated with access to credit products that are ex-ante expected to benefit from OB underwriting.

Overall, we document significant use of OB products by UK consumers, with slightly greater use of advice OB than credit OB. We also show that advice OB use is associated with

[34]Our key results look very similar if we drop controls for consumer characteristics.

[35]See here.

[36]Unfortunately, we cannot observe interest rates on credit products because we do not have this data for the sample of OB respondents in the FLS data.

[37]We do not provide analysis for the other major credit product—mortgages—because, due to the institutional and regulatory features of the UK mortgage market, it was not ex-ante clear whether this market would benefit from OB. However, in unreported results, we do find that there is an increased probability of getting a mortgage among credit OB users.

[38]See here.

financial knowledge, and that credit OB use is associated with access to the credit. We see these results as a valuable first step in understanding the potential uses and impacts of OB.

3.3 Open Banking Use by SMEs and Their Financial Outcomes

Before turning to the model, we offer an additional empirical analysis that allows us to make stronger causal statements about the effect of OB. To do so, we exploit the 2017 launch of the UK's SME-focused OB policy—the "Commercial Credit Data Sharing" (CCDS)—to estimate its impact on the probability that SMEs obtain new loans from non-banks and test OB's financial inclusion implications.

The CCDS is an SME-focused analog of the UK's main OB policy (which covers all individual and business bank customers), which predated the main OB regulation and mandated information sharing by banks for their SME customers. Specifically, the CCDS required that the nine largest UK banks share detailed information on the current accounts, loan repayments, and corporate credit cards of their SME clients with any other UK lenders if those SMEs approved. Since previous initiatives had made SME borrowers' credit histories publicly available through credit bureaus, the CCDS principally revealed information about SMEs' current accounts (i.e., cash flows) that was previously solely the domain of the banks providing payment services. Thus, the information shared on SMEs is directly analogous to the information consumer and business customers share under OB. Below, we briefly describe our analysis of this policy's effect on SME lending, with Appendix E providing more detail on the CCDS policy (including how it relates to the UK's main OB policy), additional (to Section 2.4) description of data used for this analysis, summary statistics, and robustness tests.

The CCDS initiative applied only to SMEs with annual turnover below £25 million, which creates quasi-random variation that we exploit for identification. We compare SMEs just above the cutoff to SMEs just below the cutoff for the three years prior to (2014–2016) and following (2017–2019) the implementation of the policy.[39] We then test how the CCDS policy affects SMEs' ability to form relationships with new lenders using the firm-level panel data described in Section 2.4. An increased ability to switch or add lending relationships is a direct benefit of greater data sharing and a key channel through which OB is theorized to increase competition and innovation. Following Ioannidou and Ongena (2010), we consider a firm as forming a new relationship if, in a given year, it borrows from at least one lender that is not part of the set of lenders from whom the firm had borrowed in the previous three years.[40] $Any\ New\ Lender_{i,t}$ is an indicator variable equal to one if firm i forms a relationship

[39] While the CCDS was due to go live in April 2016, technical issues meant that data sharing started only in the second half of 2017. Therefore, we include 2016 in the period prior to the reform. We exclude 2020 from the sample because of the potential confounding effects of the COVID-19 pandemic.

[40] Our results are robust to both shorter and longer time windows.

with a new lender in year t.

Figure 4 presents binned scatterplots of new borrowing relationship formation against firm turnover before and after the reform. Panel (a) shows no evidence of a change in the propensity for SMEs to form new lending relationships around the £25 million threshold before the policy, while Panel (b) shows a discontinuity at that threshold appearing after the policy. In particular, firms below the threshold are more likely to establish a new lending relationship than firms above the threshold after the policy but not before the policy.[41]

We formally estimate the effect of the policy on new lending relationships using a difference-in-differences (DiD) design with a linear probability model:

$$Any\ New\ Lender_{i,t} = \beta \times Treated\ SME_i \times Post_t + \eta X_{i,t-1} + \alpha_i + \gamma_{s,t} + \eta_{g,t} + \nu_{r,t} + \varepsilon_{i,t} \quad (5)$$

We focus on firms with 2016 turnover between £10 million and £40 million to cleanly identify the effect of the new data-sharing policy. The treatment indicator variable $Treated\ SME_i$ equals one for firms with turnover below £25 million in 2016. $Post_t$ is an indicator variable equal to one in the years after the policy went live (2017 and later). β—i.e., the difference-in-differences (DiD) coefficient—measures the focal policy effect. $X_{i,t-1}$ is a vector of firm controls lagged by one period: The log of total assets, cash to total assets, leverage ratio, and credit risk. We saturate the model with a rich set of fixed effects, including firm (α_i), sector-by-year ($\gamma_{s,t}$), region-by-year ($\eta_{g,t}$), and lending relationship-stage-by-year ($\nu_{r,t}$).[42] Regions correspond to the 124 UK postcode areas and industry sectors are based on one-digit SIC codes.

Table 6 reports our results. The first four columns show consistently positive effects of the data-sharing policy on SMEs' propensity to borrow from new lenders. In column 1, where we control for year fixed effects only, the Treated SME × Post interaction coefficient is positive and statistically significant, showing the policy increased the probability of SMEs forming new borrowing relationships. We find a 1.36 percentage point increase for treated firms after the policy, which is about a 25% increase from the sample mean relationship formation rate of 5.3%. Adding firm fixed effects (column 2) and our richer set of fixed effects and controls (column 3) slightly increases these estimates.

In column 4, we use an even tighter identification strategy that leverages the fact that only the nine largest UK banks were required to share data under the CCDS. We interact the Treated SME × Post term with both an indicator variable equal to one if SME i had pre-CCDS borrowing relationships with one of the nine banks required to share data under the

[41] The overall downward trend in the new relationship formation rate over time arises mechanically because we fix our sample of firms at the beginning of the period in order to have a balanced panel. This means that the firms in Panel (b) are somewhat older and thus less likely to form new relationships.

[42] Relationship stages are calculated as the deciles of the relationship duration (in months) an SME has with its lenders up to year t.

CCDS ($Prior\ CCDS\ relationship_i$) and an equivalent indicator variable for SMEs that did not have a prior relationship with the nine banks but had a relationship with another lender ($Prior\ non-CCDS\ relationship_i$). Our treatment effect is entirely concentrated among clients of the banks required to share SME data under the CCDS: The triple interaction coefficient for a prior relationship with a bank required to share is positive and statistically significant, while being close to zero for other interactions.

Our data additionally allows us to observe the identity of the new lender. We redefine our outcome variable as new lending relationships with banks (columns 5 and 7) and non-banks (6 and 8). We interpret bank relationships as a proxy for incumbent lenders and non-banks as a proxy for fintechs. Across specifications, the DiD coefficient is positive but not statistically significant for new relationships formed with banks, and becomes both larger and statistically significant for new relationships formed with non-bank lenders. Once again, the effect is concentrated among clients of the banks that the CCDS imposed sharing requirements on (the triple interaction with $Prior\ CCDS\ relationship_i$). This shows that access to customer bank data leads to non-bank entry into the SME lending market, consistent with the fintech entry we document in Section 3.1. In Table E4, we present additional analysis of the policy's effects on SME interest expenses and balance sheets. We find that treated firms with new non-bank relationships see declining interest expenses after the policy (suggesting lower interest rates on new loans), as well as more short-term liabilities and assets (suggesting more borrowing).

Figure 5 presents event-study plots for lending relationship formation (Panel (a)) and lending relationship formation with non-bank lenders (Panel (b)) using our main specification in Equation (5). This figure illustrates that treated and non-treated firms were on approximately parallel trends prior to the policy, with a divergence starting in 2017 (the first year of the data sharing). Post-policy in 2017, the coefficient turns positive and significant, especially for non-banks, and it remains positive for the whole duration after the policy.

Finally, we examine the distributional effects of the data-sharing policy by comparing firms with and without prior lending relationships. It is not ex-ante obvious whether OB will be of greater benefit to those customers who had credit or those who did not. Presumably firms with no prior lending relationship and no credit history have the most to gain from outside lenders obtaining non-standard data that could be useful in underwriting. Indeed, this reflects the commonly expressed view by policymakers that OB data can be useful in increasing financial inclusion. Pushing against this is a countervailing selection mechanism. Customers whose transactions reveal them to be low risk are both more likely to get credit from their relationship lender prior to the policy (because it sees they are low risk) and more able to establish new lending relationships after the policy (because a non-relationship lender can now see they are low risk). Customers whose transaction data indicates them as higher risk face the opposite forces, being both denied credit by their home bank before the policy and potentially being less able to form new relationships after the policy.

We therefore examine the policy's extensive margin effects in column 9 of Table 6 by testing how prior lending relationships mediate the rate of new lending relationship formation. We interact our DiD coefficient with indicator variables for whether a firm had a single or multiple prior lending relationships, with the baseline DiD coefficient capturing the effects for firms with no prior lending relationships. Sharing data makes firms with prior lending relationships more likely to form new lending relationships. We do not find an effect for firms without prior lending relationships. This gives support for the selection mechanism sketched out above rather than the hypothesis that sharing data increases financial inclusion for previously unserved SME borrowers. As we show in the next section, this is consistent with our model's distributional predictions.

Overall, we find that the SMEs affected by the data-sharing policy are more likely to form new lending relationships, mostly with fintechs. In terms of distributional effects, treated firms with prior lending relationships are more likely to get new loans and those SMEs that form new lending relationships with non-banks pay less interest

4 An Economic Framework for Open Banking

We build on the empirical facts documented in the previous section to develop a structural model of how wider access to bank customers' data affects entry, competition, and welfare. Our empirical results illustrate the importance of both credit OB and advice OB: Our model allows us to examine the distinct economic mechanisms that underlie these different data uses. We calibrate our model by linking our novel results on OB firm entry and customer OB adoption with off-the-shelf estimates of financial product markets from the relevant literature. This allows us to assess the welfare and distributional consequences of OB and to extend the insights from our UK microdata to different environments, including countries with different privacy preferences for sharing data. Our model is tailored to speak to three issues.

Advice OB versus credit OB The finance literature has focused on data as a signal of borrower quality (e.g., He et al. (2023)). However, our consumer OB use and fintech entry results show that OB usage is driven roughly equally by advice OB and credit OB. We model these two distinct uses of data, using standard modeling choices for credit and relying on models with product (e.g., Jones and Tonetti (2020)) or business practice (e.g., Farboodi et al. (2019)) improvements for advice. We show how the nature of data use impacts the resulting equilibrium by calibrating our model to two representative applications: Financial advice (for advice OB) and non-agency mortgage loan underwriting (for credit OB). In short, product-improving OB applications (characteristic of advice OB) typically have unambiguously positive impacts on customer welfare, even for customers who opt out of OB, while underwriting applications (characteristic of credit OB) may harm some customers even if they opt out because customers who opt out are inferred to be risky.

Innovation, entry, and unequal data access Financial innovation and fintech entry are the most frequently cited motivations for OB policies, as regulators view fintech entry as a key mechanism through which OB can improve financial outcomes. Therefore, departing from other models of OB, e.g., He et al. (2023) and Parlour et al. (2022), we explicitly model new entry on the extensive margin together with product and price improvements on the intensive margin. This allows us to pinpoint how data access enables innovation, either through directly increasing customer welfare through better products (as is the case in advice OB) or indirectly through encouraging entry by allowing new entrants (with equally good products) to avoid adverse selection (Di Maggio and Yao, 2021) by better-informed incumbents (in credit OB).

Privacy preferences The sensitivity of the information contained in financial transaction data and how this plays into customers' data-sharing decisions is a key difference between OB and traditional credit registries. Our consumer OB use and fintech entry results confirm that many customers are reluctant to share their data and that the reasons for this extend beyond the fear of being offered a less favorable price. Our model builds off this result and considers data-sharing choices as a trade-off between privacy preferences and the better products or lower prices a customer could obtain from revealing her data. Privacy preferences make sharing costly, so some users will choose not to share despite being of a relatively low-risk. As a consequence, the pool of customers choosing not to share includes both customers with unfavorable data (i.e., higher-risk types) making a strategic choice and customers with more favorable data who have strong innate privacy preferences.

As our model illustrates, these issues interact in complex ways. When data is used to improve products, which is a defining feature of advice OB, then all customers benefit from the option to share. Customers who choose to share directly receive better products. In addition, even customers with strong privacy preferences against sharing benefit indirectly from new firm entry and the resulting increase in competition and product variety. In contrast, when data is used to assess customers' cost (e.g., marginal cost representing default probability or alternatively high-risk), a defining feature of credit OB, the results are far more nuanced. Customers with favorable data and low privacy preferences share data, which increases firm entry as potential entrants face less adverse selection. Customers with unfavorable data or strong privacy preferences do not share, which partially signals to outsiders that they are costly to serve, leading to higher prices. These negative effects for non-sharing customers are potentially offset by increased entry and competition. Thus, while on the surface there appears to be an inherent conflict between OB's stated goals of increased competition and innovation with financial inclusion, it is not ex-ante obvious which force dominates for high-cost or privacy-conscious customers.

Our calibrated model allows us to quantitatively assess the importance of these channels. Our baseline calibration uses UK data and shows that OB has a strong positive effect on competition and entry in both cases. This effect is so strong that *all customer types* benefit from OB even in the credit OB case, as it outweighs the costs of being seen to have negative data even for customers applying for credit with the least favorable OB data. In a less privacy-conscious setting, i.e., a country with laxer norms around privacy, our model predicts even more sharing, innovative entry, and competition. However, some customers now lose from OB as increased sharing makes it harder for customers with unfavorable data to pool with customers with strong privacy preferences. In sum, our model shows that a serious quantitative evaluation of OB, and not merely a theoretical one, is necessary for policymakers when thinking about the aggregate and distributional consequences of OB.

Although our calibrations focus on households, our model applies equally well to SMEs. Business-to-business products that use data to deliver advice, such as accounting and receivables management software, should see increased utility, entry, and competition after OB. OB-enhanced SME credit markets will yield similar entry and competition benefits but create potential distributional costs for firms with less favorable data or firms that, for legal or data security reasons, prefer not to share.

4.1 Model

The model extends a standard discrete choice framework by explicitly considering consumer data usage. For expositional purposes, we use the term "consumer," which we interpret generically as applying to either a (human) consumer or a SME. Consumer data allows competing firms to improve their products or pricing by learning about the characteristics of heterogeneous consumers. For example, the pricing of a loan is improved using data from a transaction account that reveals a consumer's credit risk, as shown by Ghosh et al. (2022) and our SME analysis in Section 3.3. Alternatively, a financial planning app uses balances and transactions from financial accounts to offer customized financial and tax advice to a household or SME.

We model two data access regimes that determine which firms can use a consumer's data. In the relationship banking regime, which is the pre-OB status quo, only a consumer's incumbent relationship bank can use her data. Under the OB regime, each consumer chooses whether to opt in to data sharing, and if she does, all firms providing the financial product can use her data regardless of whether they are her relationship bank. If she opts out of data sharing, all firms observe that she opted out, and only the relationship bank can use her data.

4.1.1 Consumer Data and Market Structure

A mass m of heterogeneous consumers, indexed by i, can purchase a financial product. Products are offered by I incumbent firms (i.e., banks) and an endogenous number, N, of new entrants (i.e., fintechs). All firms offer a single product to each consumer, who chooses a single product among the available offerings.

Each consumer is endowed with a vector of characteristics χ_i that is known to the consumer and revealed to firms that can access that consumer's data. Which firms can access the consumer's data depends on the policy regime. Under the relationship banking regime, only a single relationship bank can access the data and learn χ_i, and all other firms only know the unconditional distribution $dF(\chi_i)$. Under the OB regime, the relationship bank still knows χ_i, but, additionally, the consumer decides whether to share her data with all other firms. Let $S_i \in \{0, 1\}$ denote consumer i's (endogenous) choice of whether to opt in to data sharing. If consumer i opts to share data ($S_i = 1$), all firms observe χ_i. If the consumer does not ($S_i = 0$) the non-relationship firms observe only that the consumer opted out of data sharing and consequently infer the endogenous conditional type distribution $dF(\chi_i|S_i = 0)$.

Building off our advice OB and credit OB breakdown, we assume that χ_i provides information on both the consumer-specific marginal cost (mc_i)[43] paid by the lender to provide the product and consumer-specific customization needs (f_i), which if precisely met, provide additional utility to the consumer. Thus, $\chi_i \equiv (f_i, mc_i)$. Marginal cost covers both usage cost (will they exploit credit card bonuses or incur late fees?) and risk (will they default?) and is most linked to credit OB. Customization needs cover product tailoring (how can we set up a financial plan for a particular customer?) and creation (how can we communicate their spending to them or help them save?) and is most linked to advice OB.

4.1.2 Consumer Demand

Consumer i makes a discrete choice of firm j's product from among the $I + N$ competing firms. Product ij is characterized by $\nu_{ij} \equiv (p_{ij}, g_{ij})$, where p_{ij} is price and g_{ij} are non-price characteristics, e.g., whether the offered advice is customized or whether the firm had a relationship with consumer i in the prior period. Consumer i receives the following indirect utility from product ij:

$$u(\nu_{ij}, \chi_i) \equiv -\alpha p_{ij} + (\theta + \lambda)R_{ij} + \lambda(1 - R_{ij})S_i + \epsilon_{ij}. \tag{6}$$

Here, α is the consumer's price sensitivity and p_{ij} is the price. R_{ij} is an indicator for whether firm j is the relationship bank for consumer i, and θ represents the consumer's utility from obtaining the product from its relationship bank due to, e.g., a desire to obtain financial services from a convenient one-stop shop. λ is the extra utility the consumer gets from a financial institution that can provide customization, e.g., by being offered better

[43]We interpret variance in mc_i as the residual *conditional* on observables, e.g., residual variation after controlling for a consumer's publicly available credit score. For example, in countries where credit scores are more informative, we would expect our modeled variance in mc_i to be smaller relative to a country that has no credit scores. We discuss this in more detail in Appendix F.

financial advice. S_i is an indicator for whether the consumer shares her data with outsiders. To provide interpretation, when a consumer obtains a product from its relationship bank, it receives both the additional relationship utility θ as well as the customization utility λ. When a consumer obtains a product from an outsider, it obtains only the customization utility when the consumer shares her data. u is implicitly a function of χ_i because χ_i contains the consumer's desired customization.

Finally, ϵ_{ij} is a horizontal taste shock whose i.i.d. realization is known to the consumer at the time of making the product choice (and only after deciding whether to share her data) but unknown to the firms, creating differentiation and giving individual firms market power. Importantly, these ϵ shocks prevent the unraveling of pure strategy equilbria by obscuring whether a consumer chooses an uninformed offer because she is a high-cost type with high rate offers from insiders, or because she is a low-cost type with a high idiosyncratic preference for the outsider's product (e.g., Crawford et al. (2018)).

Among the offerings and an outside option, u_0, the consumer chooses the product that offers the highest indirect utility. Let $s_j(\boldsymbol{\nu}_i, \chi_i)$ denote the probability that a consumer with characteristics χ_i chooses firm j's product given all product offerings, including the outside option, $\boldsymbol{\nu}_i$. This quantity is obtained by integrating the consumer's optimal choice across the taste shock, $\boldsymbol{\epsilon_i}$:

$$s_j(\boldsymbol{\nu}_i, \chi_i) = \int \mathbb{I}\left\{u(\nu_{ij}, \chi_i) > u(\nu_{ik}, \chi_i),\ \forall k \neq j\right\} dF(\boldsymbol{\epsilon_i}). \tag{7}$$

4.1.3 Consumer Opt-in to Data Sharing

Under the OB regime, each consumer chooses whether to opt in to data sharing.[44] If she shares her data, all $I + N$ firms observe her consumer-specific χ_i. If she does not share her data, her relationship bank observes χ_i, and the other firms observe only that she opted out of data sharing. Let $\boldsymbol{\nu}_i^S$ and $\boldsymbol{\nu}_i^{\sim S}$ denote the set of offers she receives if she opts in or out of data sharing, respectively, and let $Eu(\boldsymbol{\nu}_i)$ denote the consumer's expected utility of the discrete choice problem in Equation (7) for a given set of offers, with

$$Eu(\boldsymbol{\nu}_i) = \int \max_j \left\{u(\nu_{ij}, \chi_i)\right\} dF(\boldsymbol{\epsilon_i}). \tag{8}$$

The consumer makes her data-sharing decision by comparing her expected utility if she shares her data to her expected utility if she does not. We enrich this decision by incorporating a consumer-specific preference for privacy, reflecting both aggregate preferences for

[44] For simplicity, we assume the consumer either shares her data with all the firms or no firms (besides the relationship bank, which already has it). This assumption is nearly without loss of generality because if a consumer is made better off by sharing her data with one extra firm, she is made even better off by sharing her data with all firms. The only exception to this would be if the consumer has increasing hedonic disutility from sharing data with more firms.

privacy and person-level heterogeneity.[45] In the same discrete choice framework, we model the consumer's indirect utility of sharing or not sharing her data as follows:

$$u_i^S = -\phi + Eu(\boldsymbol{\nu}_i^S, \chi_i) + \epsilon_i^S \tag{9}$$

$$u_i^{\sim S} = Eu(\boldsymbol{\nu}_i^{\sim S}, \chi_i) + \epsilon_i^{\sim S} \tag{10}$$

Here, ϕ represents a society-wide hedonic privacy preference and ϵ_i^S and $\epsilon_i^{\sim S}$ represent a consumer-specific privacy preference shock drawn i.i.d.[46] Based on her characteristics and privacy preference, the consumer chooses the greater of these utilities, which yields an endogenous probability of disclosure for each set of consumer characteristics χ_i given by ψ_i:

$$\psi_i = \int \mathbb{I}\left\{u_i^S > u_i^{\sim S}\right\} dF(\epsilon_i^S, \epsilon_i^{\sim S}) \tag{11}$$

Finally, the conditional distribution of types who opt out of data sharing is

$$dF(\chi_i | S_i = 0) = \frac{(1-\psi_i)dF(\chi_i)}{\int_i (1-\psi_i)dF(\chi_i)}. \tag{12}$$

4.1.4 Firms

Entrant firms pay a fixed cost c to enter. Conditional on entry, firms compete in a differentiated Bertrand structure. Firm j's marginal cost for consumer i is equal to the sum of mc_j, a firm-specific cost common to all of j's potential customers, which is known to firms and assumed in our calibration to differ only by incumbent versus new entrant, and mc_i, a consumer-specific cost that is common to all firms selling to consumer i, known by the relationship bank and by new entrants if shared by the consumer:

$$mc_{ij} \equiv mc_j + mc_i. \tag{13}$$

Firms are informed about a consumer i's characteristics, χ_i, if (1) they are consumer i's relationship bank or (2) the economy is in the OB regime and consumer i has opted into data sharing. Uninformed firms know only the distribution of consumer types not sharing data, which in the relationship banking regime is the unconditional consumer distribution, $dF(\chi_i)$, and in the OB regime is the consumer distribution conditional on opting out of data sharing, $dF(\chi_i|S_i = 0)$. Firms set prices and product characteristics to maximize profits, with informed firms setting consumer-specific prices and products (ν_{ij}) and uninformed firms

[45] See, for example, Tang (2019), Bian et al. (2021), and Ben-Shahar and Schneider (2011).

[46] The variance of these shocks being greater than zero precludes a cutoff strategy of opt in versus opt out.

offering a single product and price to all consumers:

$$\Pi_{ij} = \begin{cases} \max_{\nu_{ij}} s_j(\boldsymbol{\nu}_i, \chi_i)(p_{ij} - mc_{ij}) \text{ for firms with data} \\ \max_{\nu_j} \int s_j(\boldsymbol{\nu}_i, \chi_i)(p_j - mc_{ij}) dF(\chi_i) \text{ for firms without data under relationship banking} \\ \max_{\nu_j} \int s_j(\boldsymbol{\nu}_i, \chi_i)(p_j - mc_{ij}) dF(\chi_i | S_i = 0) \text{ for firms without data under OB} \end{cases} \tag{14}$$

Each firm's profit is equal to its profit across all its customers, including both profit from offering targeted products and pricing to customers whose data they know (due to OB data sharing or relationships, if any) and profit from offering an uncustomized product at a single price to the customers whose data they do not know:

$$\Pi_j = \int_i \Pi_{ij} di - c. \tag{15}$$

The entry cost of c implies that in equilibrium, $\Pi_j = c$ for the marginal entrant.

4.1.5 Equilibrium

Events proceed as follows in the relationship banking regime. First, firms choose whether to enter. Second, firms simultaneously set prices and products for both the consumers whose data they have and the consumers whose data they do not have. Third, consumers choose products and consume them. The OB regime follows a similar structure but has an additional stage at the very beginning where consumers choose whether to share their data.

For a given regime, an equilibrium consists of a set of prices and product customization choices $\boldsymbol{\nu}_i$, a number of new entrants, consumer product choices, and consumer data-sharing choices. These endogenous choices satisfy Equation (1) firms maximize profits as given by (14); (2) firms enter as given by Equation (15); (3) consumers make optimal product choices as in Equation (7); (4) consumers optimally choose to opt in to data sharing (Equation (11)); and (5) firms hold consistent beliefs over consumer types as in Equation (12). We restrict our attention to symmetric equilibria within firm types where all informed firms charge the same consumer-specific price, and all uninformed firms charge the same price to observably equivalent consumers.

4.1.6 Model Calibration

We breathe life into the model using simple calibrations based on two products: US non-Government-Sponsored Enterprise (GSE) residential mortgages and financial planning advice. We use these products as representative examples of financial products belonging to the two broader product groups—credit OB and advice OB—where we find significant OB adoption by consumers as described in Section 3.2. We interpret the non-GSE residential mortgage case as a representative example of where the data is useful for underwriting: Where the key consumer-level dimension of heterogeneity is the likelihood of default.[47] Because this particular market is well studied in the literature, there exist off-the-shelf estimates for several key parameters that we can directly import into our model. We interpret the financial advice case as a representative example of where the data is useful for providing a product more tailored to the consumer's needs: Where the key consumer-level dimension of heterogeneity is what the optimal savings, investment, and tax strategy would be given the consumer's particular financial situation. Both applications presented here are intended to be quantitatively realistic illustrations of the underlying economic forces around two important real-world applications of OB.

We detail our calibration exercise in Appendix F. Broadly, our key objects for calibration are the variance of unobserved marginal costs (for mortgages), the value of customized advice (for financial advice), and consumer preferences for privacy (for both cases). We calibrate these parameters through the simulated method of moments, utilizing empirical moments from our earlier reduced-form analysis, including the difference-in-differences estimates of fintech entry (described in Section 3.1) and consumer adoption of OB from the UK consumer survey (described in Section 3.2). Other parameter estimates, such as consumer price sensitivity and lender marginal costs, are taken from the relevant mortgage (Buchak et al. (2023)) and financial advice (Di Maggio et al. (2022a)) literature.

4.2 Consequences of OB

Using our model, we can examine how OB changes financial product markets. We first look at the aggregate effects of OB (Section 4.2.1) before moving on to the distributional consequences (Section 4.2.2) and the role of society-wide privacy preferences (Section 4.2.3). Additional discussion showing the interaction between credit registries and OB is presented in Appendix F.3.

4.2.1 Aggregate Consequences of OB

Figure 6 compares equilibrium outcomes under OB to those under relationship banking for financial advice (magenta) and credit (cyan). Panel (a) focuses on new firm entry, quantities, and prices, while Panel (b) focuses on markups, profits, and consumer surplus. The first bar of Panel (a) shows the change in the number of active firms. Consistent with our reduced-form evidence in Section 3.1, the number of firms increases significantly, and the increase is roughly equal across both use cases, with slightly more entry for credit provision. Despite slightly more entry for credit, our model shows a dramatically larger increase in consumers

[47] We focus in particular on the non-GSE sector because GSEs' guarantees mostly render default risk irrelevant.

obtaining financial advice than credit (columns labeled "Quantities (all)")—which is exactly what we find in UK consumer data in Section 3.2.

We decompose these aggregate quantity changes into quantity changes from relationship banks (columns "Quantities (relationship)") and other providers (columns "Quantities (outsiders)"). Outsider quantities increase and relationship bank quantities decrease for both products, however, the mechanisms are very different. While changes to quantities in the advice context are largely driven by product improvements, quantity changes in the credit case are largely driven by changes in pricing. As OB is adopted, informed outside lenders steal low-marginal cost (MC) consumers from the relationship banks after learning their types and offering lower prices. Thus, average prices offered by outsiders decline dramatically. Facing greater competition from outsiders for low-MC borrowers, relationship bank markups fall, as shown in Panel (b). Additionally, as relationship banks lose low-MC market share, their lending composition shifts towards higher-MC borrowers, and so share-weighted prices increase despite the increase in competition.

Incumbent profits fall and consumer surplus increases in both cases, although the surplus increase is larger for the advice case despite lower entry. These differences are driven by the fact that advice OB provides a better product to *everyone* who uses it, while credit OB gains accrue primarily to low-MC consumers (as we show in the next section), potentially to the detriment of high-MC consumers. The last column shows this effect as the fraction of borrowers receiving below-MC credit falls dramatically with the introduction of OB. In the status quo, uninformed outsiders' single pooling offer is often accepted by borrowers whose MC exceeds the offered interest rate, while following OB implementation, rates rise for borrowers who opt out.

4.2.2 Distributional Consequences of Credit OB

We now examine how the effects of OB vary across consumer types, i.e., their MCs. We focus on the credit example, because in the financial advice context, consumer "type" is synonymous with the idiosyncratically optimal advice that in aggregate makes all types symmetrically better off. That is, in the advice case, *all* borrowers are made better off through outsiders' ability to offer fully customized advice, and OB merely serves to improve the product quality and market competitiveness for everybody.

By contrast, distributional outcomes are more nuanced in the credit case, and the effects of OB vary systematically between low- and high-MC consumers. Figure 7 shows outcomes for consumers versus their MCs. In this section, we focus on the comparison between the no-OB status quo in red and the calibrated OB regime in green. We return to a counterfactual with a smaller consumer preference for privacy in subsequent sections (in blue). Panel (a) shows the fraction of borrowers opting into data sharing in the OB regime. Observe first that in the no-OB status quo, mechanically no consumers share data because, without OB, there is no such option. Once in the OB regime, the green line shows that the propensity to opt in to data sharing is decreasing in the borrower's unobserved MC, to the point where roughly 60% of borrowers with the lowest MC share data, while essentially no borrowers with medium or higher MC share data. Note that this proportion is smoothly decreasing in MC due to borrowers' idiosyncratic preferences for privacy. This smoothness prevents a full (Grossman, 1981) unraveling, and is in contrast to many theoretical signaling models where stark cutoff strategies are common. Importantly, opting out of data sharing does not fully reveal the borrower's type. However, based on the results in Panel (a), it is clear that opting out of data sharing in the OB regime is at least partially revealing, and indeed, the distribution conditional on opting out, $dF(\chi_i|S_i = 0)$ has a higher expected MC than the unconditional distribution $dF(\chi_i)$.

In the no-OB status quo regime (red lines), Panel (b) shows that average interest rates are weakly increasing in MC. Pass-through is incomplete and levels off quickly because while the informed relationship lender charges low-MC borrowers low rates (plus a markup), and high-MC borrowers high rates (plus a markup), shown in Panel (c), high-MC borrowers instead opt for the outsider lenders' flat pooling rates, shown in Panel (d). The outsider's uniform pricing rule results in adverse selection, with low-MC borrowers choosing the relationship bank's offers and high-MC borrowers choosing the outsider's offers, as shown in the outsiders' market share increasing in MCs in Panel (f). As rates increase only modestly in MC, credit quantities, shown in Panel (e), decrease only modestly in borrower MC.

In the OB regime (green lines), low-MC borrowers opt in to data sharing, thus revealing that they are low-MC types to outside lenders, and these opt-ins receive relatively lower rates in equilibrium, as shown in Panel (b). These lower rates come from two sources: First, outsiders directly offer lower rates, as shown in Panel (d), and second, relationship banks now face greater rate competition and reduce their markups, as shown in Panel (c). In effect, competition from OB causes banks to increase pricing differentiation among their customers. These lower rates are reflected in greater borrowing quantities for low-MC borrowers, as shown in Panel (e). Importantly, outsiders gain significant market share among the lowest-MC borrowers who choose to opt in to OB, as shown in Panel (f), resulting in less severe adverse selection, although they still retain significant market share among the highest-MC borrowers who continue to take advantage of the uninformed outsiders' relatively attractive pooling offers.

In contrast, high-MC borrowers, who choose not to opt in to data sharing, partially reveal to outside lenders that they are high MC types, although hedonic privacy preferences partially obscure this inference. Conditional on not sharing data, borrowers' expected MCs are higher, and thus uninformed outsiders charge slightly higher rates as compared to the status quo, as shown in Panel (d). Importantly, the reduced adverse selection faced by outsiders allows for greater entry, and in our calibration, the positive effect of increased entry through greater product variety more than offsets the negative effect of higher prices, *even among the highest-MC borrowers*. Thus, the quantity of credit provided increases for *all* borrowers under the OB regime. Finally, while all borrowers in our calibration benefit, the largest benefits accrue to the best types who were ex-ante receiving the most credit. This is consistent with our distributional findings in the reduced-form SME analysis in Section 3.3, where we find that in the OB regime, firms with prior lending relationships are more likely to get new loans and those firms that form new lending relationships with non-banks pay less interest.

4.2.3 Consumer Attitudes Towards Privacy

We conclude our analysis of the model by examining how consumer attitudes towards privacy impact equilibrium outcomes of OB. We begin this analysis by varying consumers' mean preference for privacy and show the results in Figure 8. The x-axis shows the value of privacy as a multiple of the calibrated value. Moving left-to-right along the x-axis increases consumers' hedonic aversion to sharing data, perhaps similar to moving from a country with a high trust in fintechs (e.g., Korea) to a country with low trust in fintechs (e.g., Japan). The lines with circle markers show the fraction of consumers opting into data sharing. The x marks show the fraction of consumers (regardless of whether they opt in to OB) who are made *worse* off by OB in a utility sense. The red lines and marks show outcomes for financial advice, and the blue lines and marks show outcomes for credit.

Observe first that, unsurprisingly, the fraction of consumers opting into data sharing is decreasing in their preference against sharing data. Additionally, despite having identical preferences against sharing data in both use cases, our calibration successfully matches our calibration targets discussed in Section 3.2 where more consumers opt in to data sharing for financial advice than data sharing for credit, and these differences in opt-in rates persist across counterfactual privacy preferences. Intuitively, differences in opt in rates arise because sharing data with outside financial advice providers makes all consumers weakly better off with no drawbacks. In contrast, while opting into credit OB data sharing benefits low-MC borrowers, high-MC consumers are potentially hurt.

Our plot showing the fraction of customers made worse off by OB makes this mechanism clear. For advice, all types of consumer are better off at all levels of societal preference because they all benefit from increased entry and competition. For credit, we see a distinct threshold at about 85% of our calibrated UK privacy preference. If the average privacy preference is above that level, all consumers benefit from OB because the increased profitability for outside lenders among the small fraction of low-MC borrowers who opt in to OB leads to increased lender entry, and the benefits of increased product variety from entry for high-MC borrowers more than offsets the small price increases that come from (mildly) negative inferences about their types from their opt-out choice. These negative inferences are small because when many borrowers opt out due to privacy concerns, the act of not sharing data reveals little about the consumer's MC type. In contrast, with weaker social preferences against sharing data, the act of not sharing data reveals more negative information about the borrower's MC type. Thus, they have a harder time masking that they are high-MC types and suffer the consequences of higher rates that dominate any benefit they receive from increased product variety.

We confirm this intuition by revisiting Figure 7. Here, the dashed blue lines reflect a counterfactual where the privacy preference is decreased by 25%—corresponding to 0.75 on the x-axis of Figure 8. These panels show that as more borrowers opt in to OB (Panel (a)), rates decrease more for low-MC borrowers and increase more for high-MC borrowers (Panel (b)). This is primarily driven by changes in outside lenders' pricing behavior (Panels (c) and (d)). This leads to greater quantities of credit for low-MC borrowers, and less credit for high-MC borrowers, both overall and from outsiders, relative to the no-OB status quo (Panels (e) and (f)).

In summary, a greater societal desire for privacy shields high-MC borrowers from scrutiny: Lenders cannot infer from opting out of data sharing that the borrower is not sharing because she has a high-MC type. As the hedonic preference for privacy decreases, choosing not to share becomes more indicative of being a high-MC type. Lenders correctly make this inference and charge them higher rates. In the credit case, this has the effect of potentially harming privacy-conscious consumers. In the advice case, there are no negative inferences to be drawn about OB opt-outs and so even privacy-conscious individuals benefit from increased entry.

5 Conclusion

Our paper examines the dramatic rise of OB, which is now present in some form in roughly 80 countries. Using a hand-collected dataset of OB government policies around the world, we document significant heterogeneity in these policies' timing, purpose, and implementation. Large increases in VC fintech activity across different financial products (e.g., financial advice applications, credit, payments, regtech) follow OB policy implementations. Granular micro-data on UK consumers shows they use OB for credit but also for financial advice, with that usage associated with credit use and greater financial knowledge, respectively. Data on UK SMEs affected by OB shows they form more new lending relationships, especially with non-banks. These new relationships are driven by the SMEs with prior lending relationships—a finding at odds with the financial inclusion goals of OB policies but fully consistent with our model's distributional predictions.

We interpret these results through a general framework of data use and sharing, focusing on the contrasting implications of using data for underwriting and using data to improve products. OB increases entry in both use cases through very different channels: For credit, data allows entrants to underwrite more effectively and reduce adverse selection; while for product improvements, data allows entrants to improve their product quality. Although our results suggest OB is achieving its innovation-promotion goals, our framework highlights how OB-enabled credit underwriting can harm consumers whose data would indicate their riskiness. Being able to opt out offers only partial protection to these consumers, as the act of opting out itself sends a signal from which lenders draw a negative inference. Moreover, these consumers are likely to be on the margins of the financial system, and thus precisely those whose financial inclusion policymakers are interested in facilitating. For example, OB data is increasingly used to screen potential renters via the screening service Tink, and customers who are unwilling to share their data risk being cut out of basic housing markets.

Importantly, these potential negative distributional effects are not present when OB data is used for product improvements rather than for screening, and preliminary evidence suggests that product improvements are an equally—if not more—quantitatively relevant OB application. Additionally, social privacy preferences can ameliorate some of the worst distributional effects and prevent a stigma from non-sharing. In our quantitative calibration on UK data, the benefits of entry and innovation more than offset the losses from information revelation for even the riskiest borrowers, with many borrowers seeing major benefits. This result is specific to our calibration and our estimates of UK privacy preferences, highlighting the importance of quantitative models like ours for evaluating the impacts of OB.

As policymakers set the path of future banking regulation, our paper helps put these tradeoffs in perspective. Data lies at the heart of relationship banking, and large financial institutions benefit from their special ability to aggregate huge amounts of customer data. Because of that, removing banks' monopoly on customer data has the potential to transform the very nature of relationship banking. If opening data reduces banks' economies of scope, the entire banking ecosystem could reorganize around more specialized and interconnected firms. The large reaction of fintech investment to OB policy implementations shows the potential for disruption and just how valuable innovators perceive this data to be, while our results on non-bank SME borrowing document real disruption to an important market.

More generally, the role that data ownership and access plays in endogenously creating and maintaining market power is a first-order question in an increasingly data-driven economy, sectors that are dominated by a small number of data-intensive firms. Opening data to potential competitors and innovators in order to spur innovation, increase competition, and ultimately raise welfare is a natural policy response, and our paper is the first to provide a global comparative analysis of such policy initiatives. Our work aims to set the stage for future research on OB and the use of data in finance and beyond by highlighting why it matters and the key tradeoffs it raises. However, this potentially profound disruption and restructuring of the financial system is still in its infancy. Important empirical and theoretical questions remain about how these policies will impact the behavior and outcomes of consumers, businesses, and financial firms.

References

Acquisti, A., C. Taylor, and L. Wagman (2016). The economics of privacy. *Journal of Economic Literature 54*(2), 442–492.

Acs, Z., T. Åstebro, D. Audretsch, and D. T. Robinson (2016). Public policy to promote entrepreneurship: A call to arms. *Small Business Economics 47*(1), 35–51.

Aguirregabiria, V., R. Clark, and H. Wang (2019). The geographic flow of bank funding and access to credit: Branch networks, local synergies, and competition. *CEPR Discussion Paper No. 13741*.

Ahnert, T., S. Doerr, N. Pierri, and Y. Timmer (2022). Does IT help? Information technology in banking and entrepreneurship. *BIS Working Paper No. 998*.

Akcigit, U. and W. R. Kerr (2018). Growth through heterogeneous innovations. *Journal of Political Economy 126*(4), 1374–1443.

Allen, F., X. Gu, and J. Jagtiani (2021). A survey of fintech research and policy discussion. *Review of Corporate Finance 1*(3-4), 259–339.

Babina, T., A. Fedyk, A. He, and J. Hodson (2023a). Artificial intelligence, firm growth, and product innovation. *Journal of Financial Economics forthcoming*.

Babina, T., A. X. He, S. T. Howell, E. R. Perlman, and J. Staudt (2023b). Cutting the innovation engine: How federal funding shocks affect university patenting, entrepreneurship, and publications. *Quarterly Journal of Economics 138*(2), 895–954.

Babina, T. and S. T. Howell (2022). Entrepreneurial spillovers from corporate R&D. *Journal of Labor Economics forthcoming*.

Bahaj, S., A. Foulis, and G. Pinter (2020). Home values and firm behavior. *American Economic Review 110*(7), 2225–2270.

Bai, J., S. Bernstein, A. Dev, and J. Lerner (2022). The dance between government and private investors: Public entrepreneurial finance around the globe. *NBER Working Paper No. 28744*.

Barth, J. R., G. Caprio Jr, and R. Levine (2004). Bank regulation and supervision: What works best?. *Journal of Financial Intermediation 13*(2), 205–248.

Bartlett, R., A. Morse, R. Stanton, and N. Wallace (2022). Consumer-lending discrimination in the fintech era. *Journal of Financial Economics 143*(1), 30–56.

Baumol, W. J. (1990). Entrepreneurship: Productive, unproductive, and destructive. *Journal of Political Economy 98*(5), 893–921.

Beck, T., O. De Jonghe, and G. Schepens (2013). Bank competition and stability: Cross-country heterogeneity. *Journal of Financial Intermediation 22*(2), 218–244.

Ben-Shahar, O. and C. E. Schneider (2011). The failure of mandated discourse. *University of Pennsylvania Law Review 159*, 647–749.

Benetton, M., G. Buchak, and C. Robles-Garcia (2022). Wide or narrow? Competition and scope in financial intermediation. *CEPR Discussion Paper No. 17497*.

Berg, T., V. Burg, A. Gombović, and M. Puri (2020). On the rise of fintechs: Credit scoring using digital footprints. *Review of Financial Studies 33*(7), 2845–2897.

Berg, T., A. Fuster, and M. Puri (2022). Fintech lending. *Annual Review of Financial Economics 14*(1), 187–207.

Bian, B., X. Ma, and H. Tang (2021). The supply and demand for data privacy: Evidence from mobile apps. *Available at SSRN 3987541*.

Boot, A., P. Hoffmann, L. Laeven, and L. Ratnovski (2021). Fintech: What's old, what's new?. *Journal of Financial Stability 53*, 100836.

Boot, A. W. A. and A. V. Thakor (1997). Financial system architecture. *Review of Financial Studies 10*(3), 693–733.

Buchak, G., J. Hu, and S. J. Wei (2021). Fintech as a financial liberator. *NBER Working Paper No. 29448*.

Buchak, G., G. Matvos, T. Piskorski, and A. Seru (2018). Fintech, regulatory arbitrage, and the rise of shadow banks. *Journal of Financial Economics 130*(3), 453–483.

Buchak, G., G. Matvos, T. Piskorski, and A. Seru (2023). Beyond the balance sheet model of banking: Implications for bank regulation and monetary policy. *Journal of Political Economy forthcoming*.

Chen, S., S. Doerr, J. Frost, L. Gambacorta, and H. S. Shin (2023). The fintech gender gap. *Journal of Financial Intermediation 54*, 101026.

Claessens, S. and L. Laeven (2004). What drives bank competition? Some international evidence. *Journal of Money, Credit and Banking 36*(3), 563–583.

Claessens, S. and L. Rojas-Suarez (2016). Financial regulations for improving financial inclusion. *Center for Global Development Task Force Report*.

Claessens, S. and N. Van Horen (2013). Impact of foreign banks. *Journal of Financial Perspectives 1*(1), 29–42.

Cornelli, G., J. Frost, L. Gambacorta, R. Rau, R. Wardrop, and T. Ziegler (2020). Fintech and big tech credit: A new database. *BIS Working Paper No. 887*.

Crawford, G. S., N. Pavanini, and F. Schivardi (2018). Asymmetric information and imperfect competition in lending markets. *American Economic Review 108*(7), 1659–1701.

De Roure, C., L. Pelizzon, and A. V. Thakor (2022). P2P lenders versus banks: Cream skimming or bottom fishing? *Review of Corporate Finance Studies 11*(2), 213–262.

Decker, R., J. Haltiwanger, R. Jarmin, and J. Miranda (2014). The role of entrepreneurship in US job creation and economic dynamism. *Journal of Economic Perspectives 28*(3), 3–24.

Denes, M. R., S. T. Howell, F. Mezzanotti, X. Wang, and T. Xu (2023). Investor tax credits and entrepreneurship: Evidence from U.S. states. *Journal of Finance forthcoming*.

Di Maggio, M., M. Egan, and F. Franzoni (2022a). The value of intermediation in the stock market. *Journal of Financial Economics 145*(2), 208–233.

Di Maggio, M., D. Ratnadiwakara, and D. Carmichael (2022b). Invisible primes: Fintech lending with alternative data. *NBER Working Paper No. 29840*.

Di Maggio, M. and V. Yao (2021). Fintech borrowers: Lax screening or cream-skimming?. *Review of Financial Studies 34*(10), 4565–4618.

Diamond, D. W. (1984). Financial intermediation and delegated monitoring. *Review of Economic Studies 51*(3), 393–414.

Diamond, D. W. and P. H. Dybvig (1983). Bank runs, deposit insurance, and liquidity. *Journal of Political Economy 91*(3), 401–419.

Djankov, S., C. McLiesh, and A. Shleifer (2007). Private credit in 129 countries. *Journal of Financial Economics 84*(2), 299–329.

Egan, M., A. Hortaçsu, and G. Matvos (2017). Deposit competition and financial fragility: Evidence from the US banking sector. *American Economic Review 107*(1), 169–216.

Egan, M., S. Lewellen, and A. Sunderam (2022). The cross section of bank value. *Review of Financial Studies 35*(5), 2101–2143.

Erel, I. and J. Liebersohn (2022). Can fintech reduce disparities in access to finance? Evidence from the paycheck protection program. *Journal of Financial Economics 146*(1), 90–118.

Farboodi, M., R. Mihet, T. Philippon, and L. Veldkamp (2019). Big data and firm dynamics. *AEA Papers and Proceedings 109*, 38–42.

Foster, L., J. Haltiwanger, and C. Syverson (2008). Reallocation, firm turnover, and efficiency: Selection on productivity or profitability?. *American Economic Review 98*(1), 394–425.

Fuster, A., M. Plosser, P. Schnabl, and J. Vickery (2019). The role of technology in mortgage lending. *Review of Financial Studies 32*(5), 1854–1899.

Gardner, J. (2022). Two-stage difference-in-differences. *Available at arXiv 2207.05943*.

Gennaioli, N., R. La Porta, F. Lopez-de Silanes, and A. Shleifer (2013). Human capital and regional development. *Quarterly Journal of Economics 128*(1), 105–164.

Ghosh, P., B. Vallee, and Y. Zeng (2022). Fintech lending and cashless payments. *Available at SSRN 3766250*.

Glaeser, E. L., S. P. Kerr, and W. R. Kerr (2015). Entrepreneurship and urban growth: An empirical assessment with historical mines. *Review of Economics and Statistics 97*(2), 498–520.

Goldstein, I., C. Huang, and L. Yang (2022). Open banking under maturity transformation. *Working Paper*.

Gompers, P., A. Kovner, J. Lerner, and D. Scharfstein (2008). Venture capital investment cycles: The impact of public markets. *Journal of Financial Economics 87*(1), 1–23.

Gompers, P. and J. Lerner (1998). What drives venture capital fundraising?. *Brookings Papers on Economic Activity. Microeconomics 1998*, 149–204.

Gompers, P. A., W. Gornall, S. N. Kaplan, and I. A. Strebulaev (2020). How do venture capitalists make decisions?. *Journal of Financial Economics 135*(1), 169–190.

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

Gopal, M. and P. Schnabl (2022). The rise of finance companies and fintech lenders in small business lending. *Review of Financial Studies 35*(11), 4859–4901.

Gornall, W. and I. A. Strebulaev (2021). The economic impact of venture capital: Evidence from public companies. *Available at SSRN 2681841*.

Granja, J., C. Leuz, and R. G. Rajan (2022). Going the extra mile: Distant lending and credit cycles. *Journal of Finance 77*(2), 1259–1324.

Grossman, S. J. (1981). The informational role of warranties and private disclosure about product quality. *Journal of Law and Economics 24*(3), 461–483.

Haltiwanger, J., R. S. Jarmin, and J. Miranda (2013). Who creates jobs? Small versus large versus young. *Review of Economics and Statistics 95*(2), 347–361.

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

Hertzberg, A., J. M. Liberti, and D. Paravisini (2011). Public information and coordination: Evidence from a credit registry expansion. *Journal of Finance 66*(2), 379–412.

Ioannidou, V. and S. Ongena (2010). "Time for a change": Loan conditions and bank behavior when firms switch banks. *Journal of Finance 65*(5), 1847–1877.

Jeng, L. A. and P. C. Wells (2000). The determinants of venture capital funding: Evidence across countries. *Journal of Corporate Finance 6*(3), 241–289.

Jiang, E., G. Matvos, T. Piskorski, and A. Seru (2020). Banking without deposits: Evidence from shadow bank call reports. *NBER Working Paper No. 26903*.

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

Kalemli-Ozcan, S., B. Sorensen, C. Villegas-Sanchez, V. Volosovych, and S. Yesiltas (2023). How to construct nationally representative firm level data from the ORBIS global database: New facts on SMEs and aggregate implications for industry concentration. *American Economic Journal: Macroeconomics forthcoming.*

Kaplan, S. N. and J. Lerner (2010). It ain't broke: The past, present, and future of venture capital. *Journal of Applied Corporate Finance 22*(2), 36–47.

Klapper, L., L. Laeven, and R. Rajan (2006). Entry regulation as a barrier to entrepreneurship. *Journal of Financial Economics 82*(3), 591–629.

Kortum, S. and J. Lerner (2000). Assessing the contribution of venture capital to innovation. *RAND Journal of Economics 31*(4), 674–692.

Kroszner, R. S. and P. E. Strahan (1999). What drives deregulation? Economics and politics of the relaxation of bank branching restrictions. *Quarterly Journal of Economics 114*(4), 1437–1467.

Laeven, L. and R. Levine (2009). Bank governance, regulation and risk taking. *Journal of Financial Economics 93*(2), 259–275.

Li, Y. and S. A. Zahra (2012). Formal institutions, culture, and venture capital activity: A cross-country analysis. *Journal of Business Venturing 27*(1), 95–111.

Lindsey, L. A. and L. C. Stein (2019). Angels, entrepreneurship, and employment dynamics: Evidence from investor accreditation rules. In *Sixth Annual Conference on Financial Market Regulation*.

Lucas Jr, R. E. (1978). On the size distribution of business firms. *Bell Journal of Economics 9*(2), 508–523.

Mullainathan, S. and P. Schnabl (2010). Does less market entry regulation generate more entrepreneurs? Evidence from a regulatory reform in Peru. In *International Differences in Entrepreneurship*, pp. 159–177. The University of Chicago Press.

Nam, R. J. (2022). Open banking and customer data sharing: Implications for fintech borrowers. *SAFE Working Paper No. 364*.

Ongena, S., A. Popov, and G. F. Udell (2013). "When the cat's away the mice will play": Does regulation at home affect bank risk-taking abroad?. *Journal of Financial Economics 108*(3), 727–750.

Parlour, C. A., U. Rajan, and H. Zhu (2022). When fintech competes for payment flows. *Review of Financial Studies 35*(11), 4985–5024.

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

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

Philippon, T. (2019). On fintech and financial inclusion. *NBER Working Paper No. 26330*.

Phillips, G. M. and A. Zhdanov (2017). Venture capital investments and merger and acquisition activity around the world. *NBER Working Paper No. 24082*.

Puri, M. and D. T. Robinson (2007). Optimism and economic choice. *Journal of Financial Economics 86*(1), 71–99.

Puri, M. and R. Zarutskie (2012). On the life cycle dynamics of venture-capital- and non-venture-capital-financed firms. *Journal of Finance 67*(6), 2247–2293.

Ramakrishnan, R. T. S. and A. V. Thakor (1984). Information reliability and a theory of financial intermediation. *Review of Economic Studies 51*(3), 415–432.

Schumpeter, J. A. (1911). *Theorie Der Wirtschaftlichen Entwicklung.* Leipzig: Duncker & Humblot.

Stulz, R. M. (2019). Fintech, bigtech, and the future of banks. *Journal of Applied Corporate Finance 31*(4), 86–97.

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

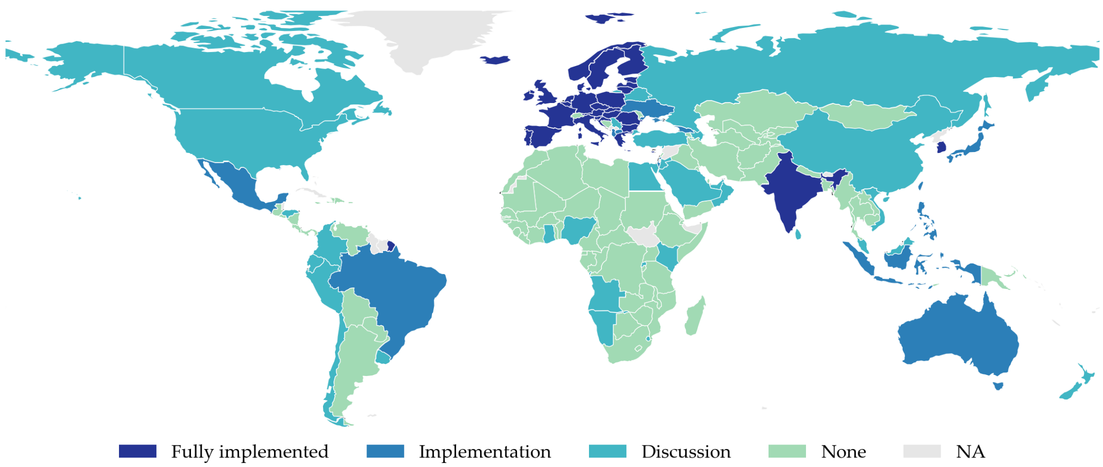
Tang, H. (2019). The value of privacy: Evidence from online borrowers. *Working Paper*.

Thakor, A. V. (2020). Fintech and banking: What do we know?. *Journal of Financial Intermediation 41*, 100833.

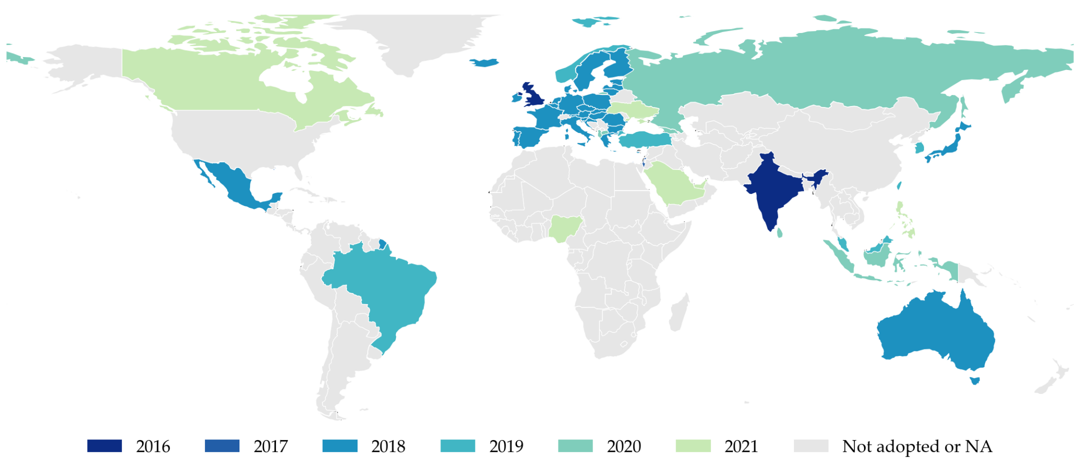
Vives, X. (2019). Digital disruption in banking. *Annual Review of Financial Economics 11*(1), 243–272.

Figure 1: Government-led Open Banking Regimes Around the World

Note: These maps show the current implementation status of government-led open banking policies and the year in which the major open banking policy was passed. Panel (a) shows the implementation status of their government open banking policies. Fully implemented corresponds to countries that have implemented open banking government policies; Implementation to those that have determined the specifics of the open banking approach and are currently implementing it; Discussion to those either considering implementing open banking policies or discussing that implementation; None to those with no government open banking approach; and NA to those where we have not collected data. Panel (b) shows the passage year of countries' major open banking policies. Data on government open banking policies is current as of October 2021.



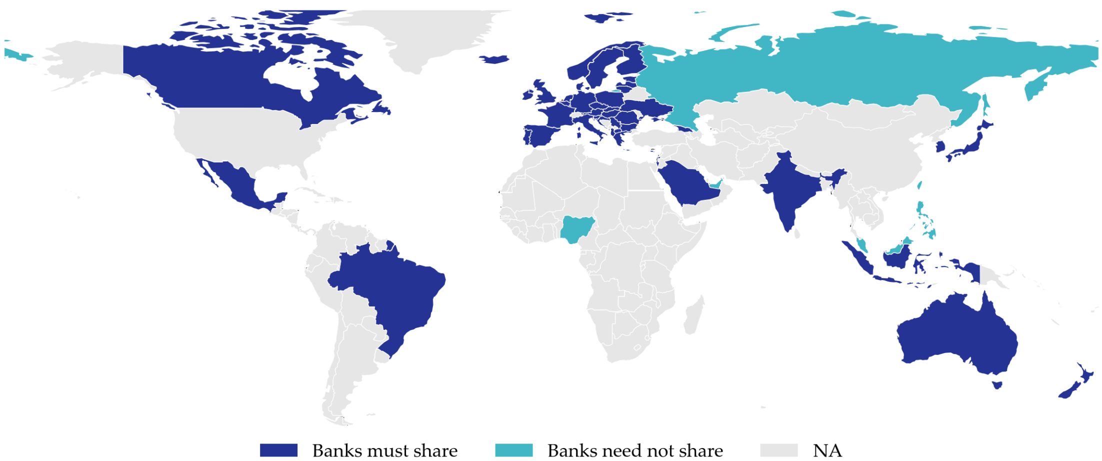
(a) Government open banking policy implementation status



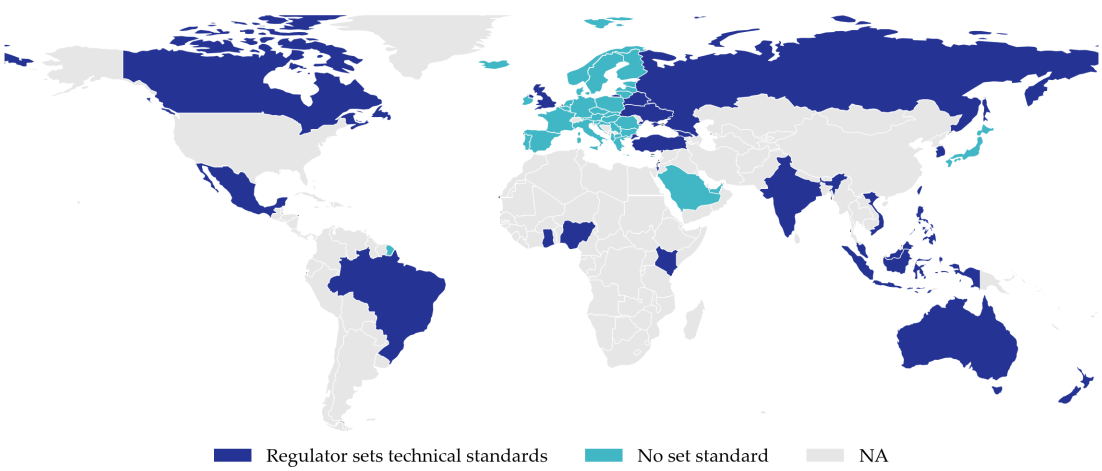
(b) Timeline of open banking adoption

Figure 2: Open Banking Government Policy Dimensions

Note: These maps show mandated data sharing and technical specifications among countries with government-led open banking efforts developed enough to specify those policy dimensions. Panel (a) shows whether the current or proposed policy requires banks to share data upon customer request. Panel (b) shows whether the regulator sets a technical standard for open banking application programming interfaces—the technology used to share bank customer data. Countries marked NA either have no government-led open banking regime, are too early in discussion for the issue to be decided, or were excluded from our data collection. Data on government open banking policies is current as of October 2021.



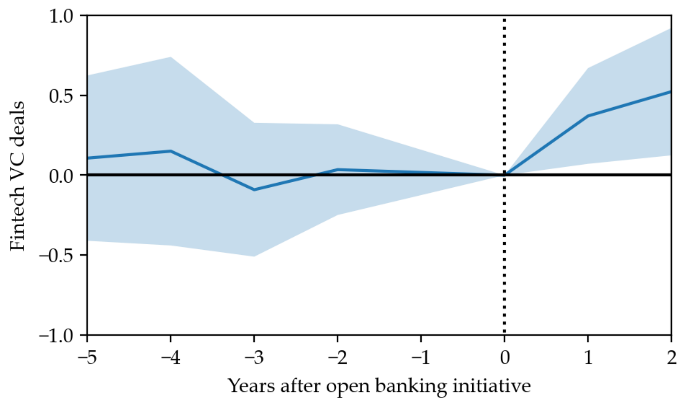
(a) Banks must share data on customer request



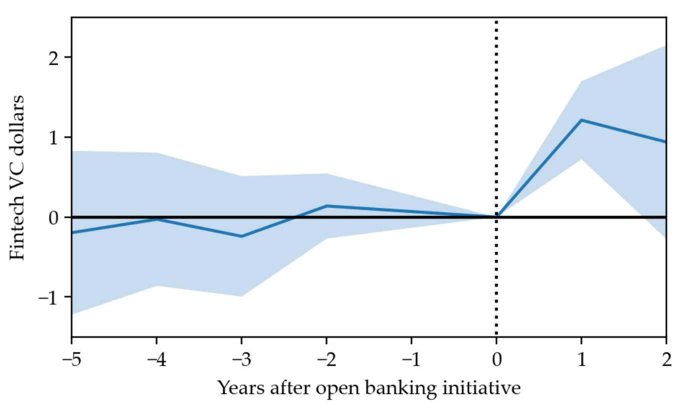
(b) Regulator sets technical standards

Figure 3: Event-study of Fintech Investment After Open Banking Government Policies

Note: This figure shows changes in fintech venture capital (VC) activity around the passage of open banking government policies using a panel event-study analysis. We perform this analysis on our high-coverage Pitchbook panel of 2011-2021 data for the 21 countries with at least five fintech VC deals in the 2000–2010 period. Panel (a) shows an event study on the log of one plus the number of fintech VC deals, and Panel (b) shows an event study on the log of one plus the millions of US dollars invested in fintech VC deals. Year 0 is the passage year of each country's major open banking initiative. The coefficient for year 0 is set to zero and other coefficients are presented net of country fixed effects and region-by-year fixed effects for Africa, Middle East & North Africa; Europe & Central Asia; Latin America & the Caribbean; North America; South Asia, East Asia & Pacific, following World Bank geographic terms. European Union member states are weighted to count as a single country for estimates and standard errors. The shaded regions denote 95% confidence intervals calculated using standard errors clustered at the country level.



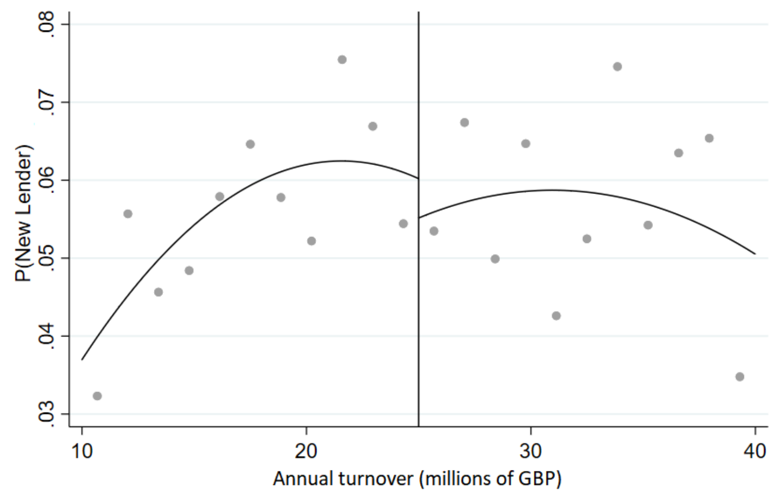
(a) Log number of fintech VC deals



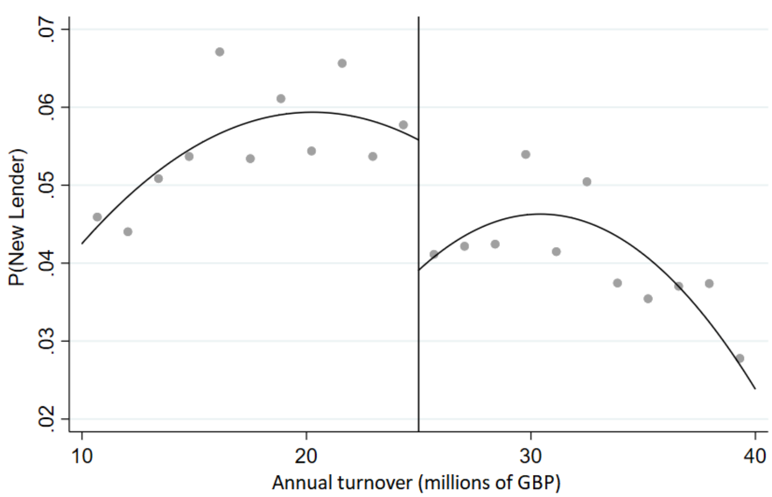
(b) Log amount of fintech VC investment in millions of US dollars

Figure 4: New SME Borrowing Relationships around CCDS Eligibility Threshold

Note: This figure shows the association between new lending relationship formation and firm turnover before and after the implementation of the Commercial Credit Data Sharing (CCDS) policy. The underlying data is company-year data on secured loans for UK SMEs from Companies House. Panel (a) presents firm-years from before the implementation of the CCDS (2014-2016) and Panel (b) presents firm-years after the policy (2017-2019). Each dot is the portion of SMEs forming new lending relationships (y-axis) among firms in a given turnover bucket (x-axis). An SME establishes a new relationship when they get a loan from a lender they had not borrowed from in the preceding three years. We use 22 equally-sized buckets from £10 million to £40 million of 2016 firm turnover. The solid curves plot best-fit quadratic polynomials for lending relationship propensity, separately estimated above and below the cutoff. The vertical line denotes the cutoff firm turnover for data sharing under the policy (£25 million), with firms to the left of the line in Panel (b) being treated by the policy.



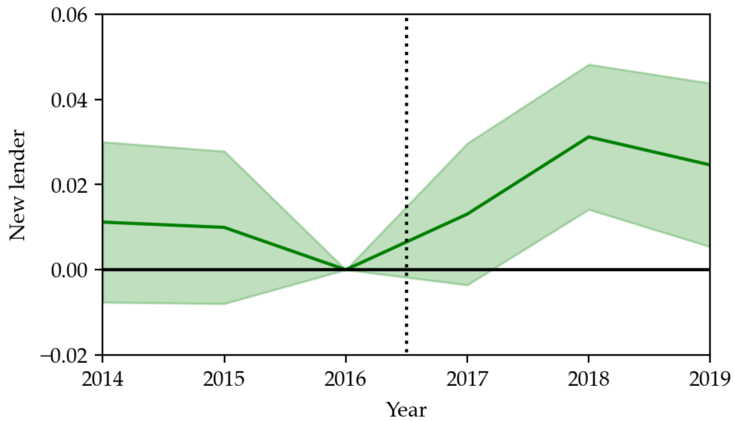
(a) New lenders by SME turnover before data-sharing policy (2014-2016)



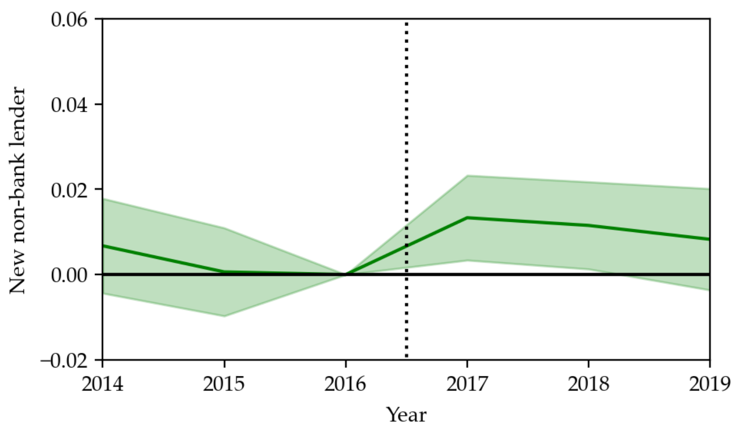
(b) New lenders by SME turnover after data-sharing policy (2017-2019)

Figure 5: Event-study of SME Data Sharing and New Borrowing Relationships

Note: This figure shows changes in new lending relationship formation for SMEs treated by the UK Commercial Credit Data Sharing (CCDS) policy using a panel event-study analysis. The underlying data is company-year data on secured loans for UK SMEs with 2016 turnover between £10 million and £40 million from Companies House via Bureau van Dijk for the 2014–2019 period. Firms are classified as treated if their 2016 turnover is below the CCDS's £25 million eligibility threshold, with firms above the threshold serving as the control group. Panel (a) shows an event study on the rate of new lending relationships for treated firms, while Panel (b) shows an event study on the rate of new lending relationships with non-banks. The event study specification is estimated using one period lagged firm-level control variables of the log of total assets, a low credit risk dummy, cash to total assets, and leverage ratio, as well as firm, sector-year, region-year, and relationship stage-year fixed effects. Low credit risk is defined as a QuiScore above 80, sectors are defined based on 1-digit 2003 UK SIC codes, regions are the 124 postcode areas, and relationship stage is the decile of the average relationship length the firm has with its lenders. The shaded regions denote 95% confidence intervals calculated using standard errors clustered at the firm level.



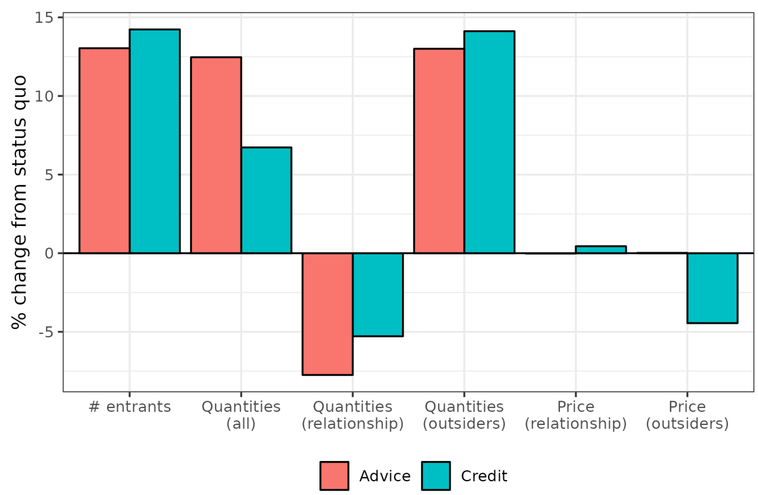
(a) New lending relationships



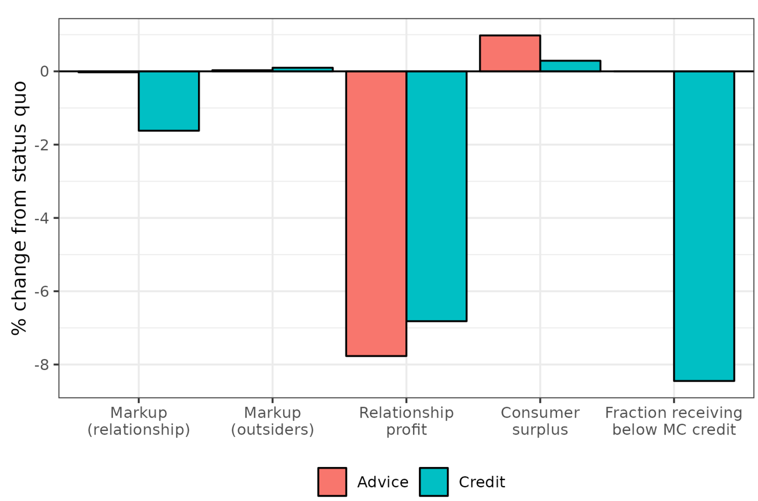
(b) New lending relationships with non-banks

Figure 6: AGGREGATE OUTCOMES OF OPEN BANKING

Note: This figure shows how open banking (OB) impacts economic outcomes in the model of Section 4. Panel (a) presents changes in entry, quantities, and prices after OB and Panel (b) presents changes in markups, profits, and consumer surplus. Each bar shows the percentage change in the relevant outcome caused by moving from the status-quo regime to the OB regime. For example, a bar with a height of 10% denotes that the number in question increased by 10% under the OB calibration, not 10 percentage points. Magenta bars show outcomes for the financial advice calibration, roughly following Di Maggio et al. (2022a). Cyan bars show outcomes for the non-GSE residential mortgage calibration, roughly following Buchak et al. (2023). # entrants is the number of new entrants. Quantities (all) is the fraction of the population obtaining the financial service, which we further split into the quantity provided by relationship banks, Quantities (relationship), and the quantity provided by other firms, Quantities (outsiders). Price (relationship) is the average fee or rate charged by relationship lenders and Price (outsiders) is the average price charged by other firms. Markup (relationship) is the average rate or fee markup over marginal cost earned by the relationship lender and Markup (outsiders) is the average for other firms. Relationship profit is profits earned by relationship banks (new entrants earn zero profit in equilibrium). Fraction receiving below MC credit is the fraction of the population that obtains credit for a price below their marginal cost.



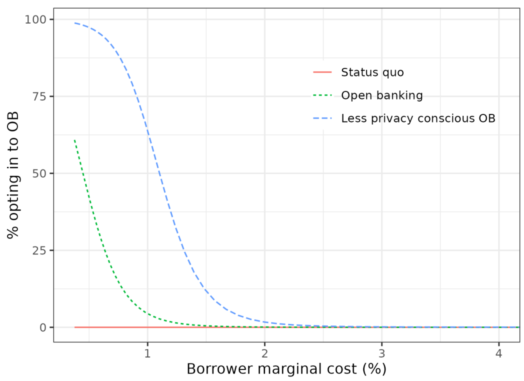
(a) Entry, quantities, and prices



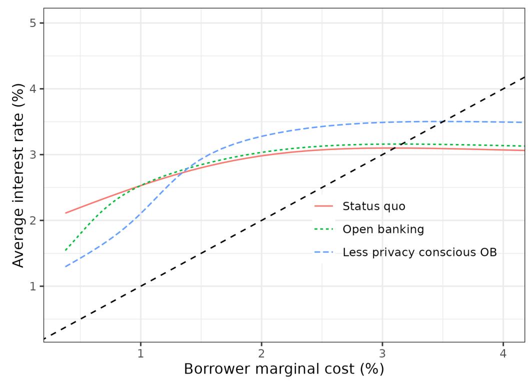
(b) Markups, profits, and surplus

Figure 7: Heterogeneous Effects of Open Banking in Underwriting by Marginal Costs

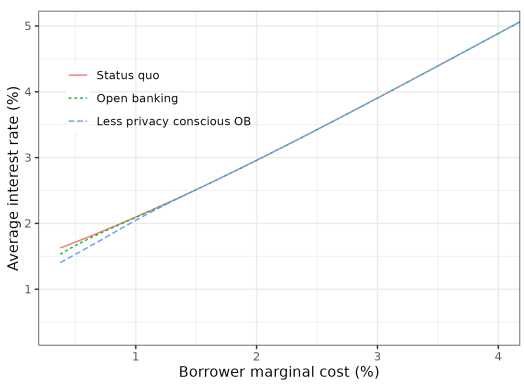
Note: This figure shows how open banking differentially impacts consumers with different marginal costs in the model of Section 4. In each panel, the x-axis shows borrowers with different marginal costs and the y-axis shows a particular outcome. Red lines indicate outcomes for the relationship banking regime. Dotted green lines indicate outcomes in the calibrated open banking regime. Dashed blue lines indicate outcomes in a counterfactual open banking simulation where borrowers' privacy preference is decreased by 25%. Panel (a) shows the fraction of the population opting into data sharing. Panel (b) shows the average interest rate paid for credit. Panel (c) shows the average interest rate charged by the relationship lender. Panel (d) shows the average interest rate charged by outsiders, firms other than the relationship bank. Panel (e) shows the fraction of the population obtaining credit. Panel (f) shows the market share of outsiders.



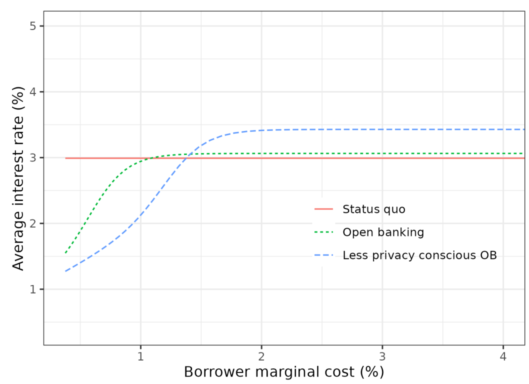
(a) Opt in to data sharing



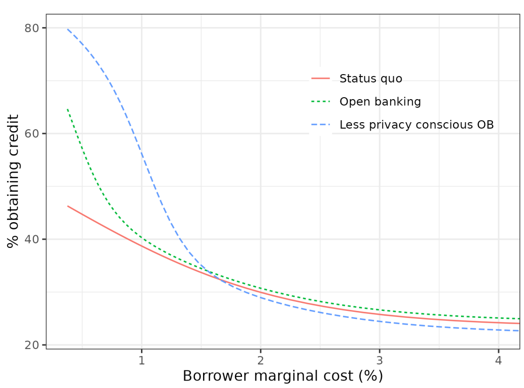
(b) Average interest rate



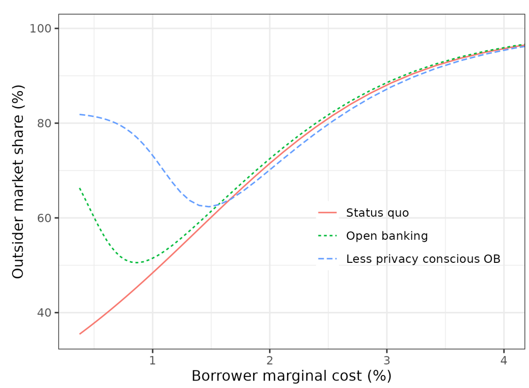
(c) Relationship interest rate



(d) Outsider interest rate



(e) Fraction obtaining credit



(f) Market share of outsiders

Figure 8: Effect of Societal Privacy Preferences on Open Banking Equilibria

Note: This figure shows how the impact of open banking varies as societal privacy preferences vary under the model of Section 4. Specifically, it shows outcomes (y-axis) for the advice (red) and credit (blue) OB as population preferences for privacy vary (x-axis). The solid lines with circle indicators show the fraction of the population opting into open banking. The x markers show the fraction of the population made worse under open banking. Privacy preferences are presented as a multiple of the baseline calibration, with a lower value corresponding to individuals being more willing to share data.

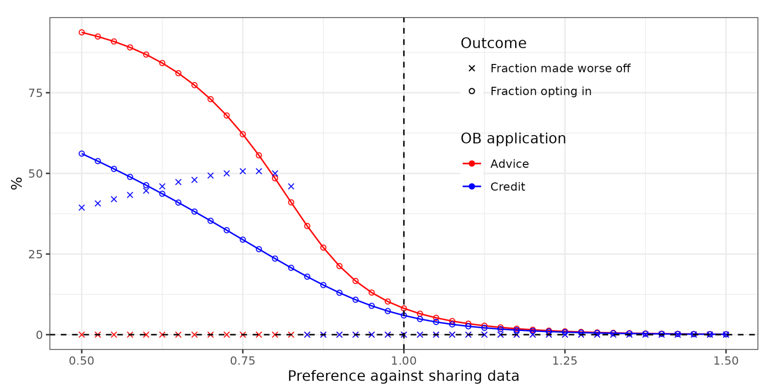


Table 1: OPEN BANKING GOVERNMENT POLICIES SUMMARY STATISTICS

Note: This table presents summary statistics on open banking government policies for 168 countries. The first number in each column is the percentage of countries fitting the given criteria and the number in parentheses is the number of countries under consideration. Government-led open banking presence considers all countries in the respective region for which data were collected, while the other categories (policy mandates, status, participation, product scope, and functionality scope) consider only those countries with a government-led open banking approach that has advanced far enough that the issue in question has been (at least preliminarily) decided. Columns split the sample into regions, with geographic terms following the World Bank definitions.

Variable	Worldwide	Africa, Middle East & North Africa	Europe & Central Asia	Latin America & the Caribbean	North America	South Asia, East Asia & Pacific
Government-led open banking presence	48% (168)	25% (65)	80% (50)	32% (25)	67% (3)	56% (25)
Policy justification						
Innovation	97% (65)	100% (9)	97% (39)	100% (3)	100% (1)	92% (13)
Competition	82% (65)	67% (9)	87% (39)	100% (3)	0% (1)	77% (13)
Inclusion	29% (66)	40% (10)	10% (39)	100% (3)	100% (1)	54% (13)
Status						
Discussion	38% (80)	75% (16)	12% (40)	75% (8)	100% (2)	36% (14)
Mid-implementation	18% (80)	6% (16)	12% (40)	25% (8)	0% (2)	43% (14)
Implemented	44% (80)	13% (16)	75% (40)	0% (8)	0% (2)	21% (14)
Policy strength						
Required data sharing	88% (57)	67% (6)	97% (37)	100% (2)	100% (1)	64% (11)
Data reciprocity	18% (56)	33% (6)	0% (36)	100% (2)	100% (1)	45% (11)
Regulator provides tech specs	39% (62)	63% (8)	15% (39)	100% (2)	100% (1)	83% (12)
Beyond transaction accts	34% (56)	80% (5)	3% (36)	100% (3)	100% (1)	91% (11)
Functionality scope						
Data sharing only	5% (58)	0% (6)	0% (38)	50% (2)	100% (1)	9% (11)
Payments only	0% (58)	0% (6)	0% (38)	0% (2)	0% (1)	0% (11)
Both	95% (58)	100% (6)	100% (38)	50% (2)	0% (1)	91% (11)

Table 2: DRIVERS OF OPEN BANKING GOVERNMENT POLICIES

Note: This table shows whether ex-ante country characteristics predict the implementation of open banking government policies. Columns 1–5 consider Cox proportional hazards models testing the adoption year of open banking based on the period up to October 2021. Columns 6–7 consider a cross-country OLS regression of the status of a country's OB regulation, expressed as a zero-to-seven score of each country's open banking implementation progress as of October 2021, with 0 being no action, 1–2 being increasingly serious levels of discussion, and 3–7 being levels of implementation progress. Columns 8–9 consider our OB Strength Index, a zero-to-one measure of the strength of each countries' open banking regime equal to the average of four indicators of policy strength (banks needing to share data, data-using firms needing to share data, regulators setting technical standards, and coverage of financial products beyond transaction accounts). The independent variables are country characteristics. Trust in fintechs is the portion of survey respondents who report being willing to share their financial data with fintechs, as reported by Chen et al. (2023). Bank branches per 100k people, Private sector credit to GDP, and Financial sector Lerner index are from the World Bank. Non-fintech VC deals and Fintech VC deals are from PitchBook and are used after taking the log of one plus the number of VC deals. Foreign-owned banks is the share of banks that are foreign-owned and are from the Claessens and Van Horen (2013) foreign bank ownership data. The Rule of Law Index is from the Cato Institute and is on a zero-to-ten scale with higher numbers denoting more favorable conditions. OB adoption year is the calendar year of OB adoption. All specifications include GDP per capita in thousands of US dollars, the square of GDP per capita in hundreds of thousands of US dollars, and the log of population based on World Bank data, as well as region fixed effects for Africa, Middle East & North Africa; Europe & Central Asia; Latin America & the Caribbean; North America; South Asia, East Asia & Pacific, following World Bank geographic terms. All independent variables are as of 2013, except for the trust in fintechs measure which is from early 2019. European Union member states are weighted to count as a single country for estimates and standard errors. The regressions are cross-sectional, where each country in the sample corresponds to a single data point. *** denotes p-value < 0.01, ** denotes <0.05, and * denotes <0.1.

	Open banking adoption					OB implementation (0-7)		OB Strength Index (0-1)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Trust in fintechs	9.755*				8.803**	15.130***	14.006***	5.403*	4.836*
	(5.284)				(4.044)	(3.939)	(4.079)	(2.471)	(2.436)
Branches per 100k people		-0.007							
		(0.020)							
Private sector credit to GDP		-0.001							
		(0.005)							
Financial sector Lerner index		0.246							
		(1.301)							
Non-fintech VC deals			0.311**		0.540		0.488		0.144
			(0.135)		(0.438)		(0.659)		(0.148)
Fintech VC deals			0.098						
			(0.271)						
Foreign owned banks				-0.150					
				(0.463)					
Rule of Law Index				0.049					
				(0.126)					
OB adoption year								0.231	0.256
								-0.207	(0.164)
Per capita GDP ($k)	0.202***	0.067	0.021	0.047**	0.139***	0.442***	0.330*	0.072	0.056
	(0.065)	(0.042)	(0.021)	(0.023)	(0.053)	(0.096)	(0.180)	(0.069)	(0.071)
Per capita GDP ($100k) squared	-20.218***	-6.516	-2.869	-4.696**	-14.530***	-43.110***	-32.951*	-4.409	-3.234
	(5.707)	(4.100)	(2.120)	(2.235)	(4.725)	(8.154)	(16.022)	(5.957)	(6.116)
Log population	-0.086	0.011	-0.095*	0.027	-0.232**	0.044	-0.159	0.216**	0.121
	(0.133)	(0.042)	(0.057)	(0.092)	(0.106)	(0.176)	(0.308)	(0.085)	(0.145)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	27	86	163	130	27	27	27	19	19
Concordance index	0.893	0.882	0.924	0.895	0.917				
Adjusted R^2						0.698	0.696	0.789	0.800

Table 3: Effect of Open Banking Government Policy on Fintechs

Note: This table shows changes in fintech venture capital (VC) investment following the implementation of open banking government policies. The table uses a difference-in-differences design on our high-coverage Pitchbook panel of country-year data spanning 2011-2021 for the 21 countries with at least five fintech deals in the 2000–2010 period. The dependent variable in columns 1 to 5 is the log of one plus the number of fintech deals in a country-year, and in columns 6 to 10 it is the log of one plus the amount invested in millions of US dollars. The main independent variable is After OB initiative which is an indicator variable equal to one if the year in question is after the year major open banking laws were passed in the country in question. Columns 2 and 7 include After OB initiative × trust in fintechs which interacts that term with country-level trust in fintechs variable equal to the portion of survey respondents who report being willing to share their financial data with fintechs, as measured for the EY Global Fintech Adoption Index and reported by Chen et al. (2023). Columns 4 and 9 include a control for non-fintech VC activity using Pitchbook data, transformed the same way as fintech VC activity. All specifications control for country fixed effects; columns 1, 2, 6, and 7 contain controls for year fixed effects; and columns 3, 4, 5, 8, and 9 control for region-by-year fixed effects for Africa, Middle East & North Africa; Europe & Central Asia; Latin America & the Caribbean; North America; South Asia, East Asia & Pacific, following World Bank geographic terms. Columns 5 and 10 additionally offer time-varying controls for the trust-in-fintech measure, with the coefficient on the control variable being estimated separately for each calendar year. European Union member states are weighted to count as a single country for estimates and standard errors. Standard errors are clustered at the country level. *** denotes p-value < 0.01, ** denotes <0.05, and * denotes <0.1.

	Fintech VC deals					Fintech VC dollars				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
After OB initiative	0.214*	0.101	0.416**	0.308**	0.432*	0.746**	0.802*	1.058**	0.874**	0.949*
	(0.111)	(0.164)	(0.159)	(0.125)	(0.205)	(0.267)	(0.390)	(0.415)	(0.368)	(0.487)
After OB initiative × Trust in fintechs		0.594*					0.041			
		(0.280)					(0.784)			
Non-fintech VC deals				0.498***						
				(0.139)						
Non-fintech VC dollars									0.338***	
									(0.105)	
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes				Yes	Yes			
Region-Year FE			Yes	Yes	Yes			Yes	Yes	Yes
Fintech trust-Year FE					Yes					Yes
Observations	231	176	231	231	176	231	176	231	231	176
Adjusted R^2	0.919	0.918	0.930	0.937	0.925	0.877	0.869	0.894	0.898	0.888

Table 4: EFFECT OF OPEN BANKING GOVERNMENT POLICY ON FINTECHS BY PRODUCT AREA

Note: This table shows changes in fintech venture capital (VC) investment by different product areas following the implementation of government open banking policies. The table uses a difference-in-differences design on our high-coverage Pitchbook panel of country-year data spanning 2011-2021 for the 21 countries with at least five fintech deals in the 2000–2010 period. The dependent variable in each specification is the log of one plus the number of VC deals in a country-year and given subsector of fintech, where subsectors are defined based on Pitchbook keywords as described in Appendix D. The independent variable is an indicator variable equal to one if the year in question is after the year major open banking laws were passed in the country in question. We use Equation (4) which controls for the log of one plus the number of non-fintech VC deals, country fixed effects, and region-by-year fixed effects for Africa, Middle East & North Africa; Europe & Central Asia; Latin America & the Caribbean; North America; South Asia, East Asia & Pacific, following World Bank geographic terms. European Union member states are weighted to count as a single country for estimates and standard errors. Standard errors are clustered at the country level. *** denotes p-value < 0.01, ** denotes <0.05, and * denotes <0.1.

	Alternative lending	Consumer finance	Financial IT	Payments	Regtech	Wealth management	Digital assets
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
After OB initiative	0.656**	0.480***	0.608***	0.409*	0.503**	0.432	-0.136
	(0.296)	(0.140)	(0.140)	(0.209)	(0.187)	(0.293)	(0.259)
Non-fintech VC control	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	231	231	231	231	231	231	231
Adjusted R^2	0.872	0.835	0.882	0.871	0.882	0.887	0.835

Table 5: Consumers OB Usage and their Financial Knowledge and Credit Access

Note: This table shows the association between financial knowledge, credit product usage, and OB usage using person-level responses to the Financial Conduct Authority's 2020 Financial Lives Survey in the UK. We use a cross-sectional OLS specification. The dependent variable in column 1 is the respondent's answer to the question "How knowledgeable would you say you are about financial matters?" on a 0 (not at all knowledgeable) to 10 (very knowledgeable) scale. The dependent variables in columns 2 to 5 are indicator variables equal to one if the respondent currently holds the credit product in question or held it in the last 12 months. Advice OB is an indicator variable equal to one if the respondent uses open banking for financial advice products, i.e., answers yes to using financial aggregation apps that allow consumers to see the accounts they hold with different banks in one place (e.g., Money Dashboard, Yolt, MoneyHub) or savings-related apps that help build savings by monitoring consumer current accounts and automatically transferring funds (e.g., Chip, Cleo, Moneybox, Plum). Credit OB is an indicator variable equal to one if the respondent uses open banking for credit products, i.e., answers yes to using firms offering customized lending products, credit reference agencies (which can personalize credit reports), or price comparison websites (e.g., rates offered by different lenders). Respondent controls are indicator variables for being unwilling to share data (respondent gives a score of 3 or below on a 0-to-10 scale to the question "Thinking about Open Banking, how willing would you be to give your bank permission to securely access your banking information?"), being employed (working full- or part-time), missing bill payments (reports missing bills payments in at least three of the last six months or finds keeping up with domestic bills and credit commitments to be a heavy burden), having high risk aversion (gives a score of 3 or below on a 0-to-10 scale to the question "Are you a person who is generally willing to take risks?"), having at least some post-secondary education, being aged 18–39 years, being male, being of white ancestry, and being married or in a registered civil partnership. All specifications control for county (UK local authority) fixed effects and estimate robust standard errors. *** denotes p-value < 0.01, ** denotes <0.05, and * denotes <0.1.

		Credit product ownership			
	Financial knowledge	Credit card	Personal loan	Student loan	Pawnbroking loan
	(1)	(2)	(3)	(4)	(5)
Advice OB	0.370***	0.039	0.020	-0.030	0.006
	(0.143)	(0.034)	(0.026)	(0.026)	(0.004)
Credit OB	0.019	0.126***	0.108***	0.002	0.001
	(0.197)	(0.040)	(0.035)	(0.034)	(0.005)
Respondent controls	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
Observations	3,098	3,104	3,104	3,104	3,104
Adjusted R^2	0.158	0.167	0.089	0.325	0.025

Table 6: SME Data Sharing and New Lending Relationships

Note: This table shows changes in new lending relationship formation for SMEs treated by the UK Commercial Credit Data Sharing (CCDS) policy. The table uses a difference-in-differences design on company-year data on secured loans for UK SMEs with 2016 turnover between £10 million and £40 million from Companies House via Bureau van Dijk for the 2014–2019 period. An SME is classified as a Treated SME if its 2016 turnover is below the CCDS's £25 million eligibility threshold. Post is an indicator variable equal to one after the CCDS was implemented in 2017. Prior CCDS relationship equals one if the SME had an existing lending relationship in 2016 with one of the nine banks required to share SME data under the CCDS, while Prior non-CCDS relationship is an equivalent indicator variable for SMEs that did not have a prior relationship with the nine banks but had a relationship with another lender. Single relationship and multiple relationships are indicator variables equal to one if in 2016 the SME had loans from one lender or loans from multiple lenders, respectively. The dependent variable in columns 1–4 and 9 is an indicator variable equal to one if the SME takes a loan in the year in question from a lender they had not dealt with in the preceding three years. The dependent variable in columns 5 and 7 is an indicator variable equal to one if the SME similarly takes a new loan and that loan is from a bank, while in columns 6 and 8 the indicator variable is equal to one if the loan is from a non-bank. Firm controls are the log of total assets, a low credit risk dummy, cash to total assets, and leverage ratio, all lagged one year. Low credit risk is defined as a QuiScore above 80, sectors are defined based on 1-digit 2003 UK SIC codes, regions are the 124 postcode areas, and relationship stage is the decile of the average relationship length the firm has with its lenders. Standard errors are clustered at the firm level and are in parentheses. *** denotes p-value < 0.01, ** denotes <0.05, and * denotes <0.1.

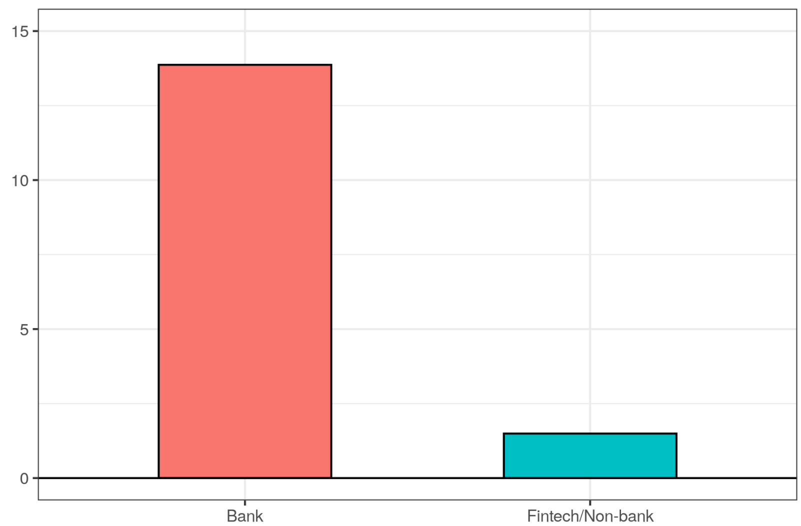
	Any new lender				New bank	New non-bank	New bank	New non-bank	Any new lender
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated SME × Post	0.0136***	0.0156***	0.0153***	0.0003	0.0061	0.0093***	0.0008	0.0005	0.0003
	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.004)
Treated SME	-0.0021								
	(0.004)								
Treated SME × Post × Prior CCDS relationship				0.0228***			0.0067	0.0146***	
				(0.009)			(0.007)	(0.006)	
Treated SME × Post × Prior non-CCDS relationship				0.0064			0.0046	0.0017	
				(0.013)			(0.010)	(0.009)	
Treated SME × Post × Single relationship									0.0129*
									(0.008)
Treated SME × Post × Multiple relationships									0.0279**
									(0.012)
Firm controls			Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes							
Firm FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Relationship stage-year FE			Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sector-year FE			Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region-year FE			Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	39,089	39,089	39,089	39,089	39,089	39,089	39,089	39,089	39,089
Adjusted R^2	0.00	0.058	0.063	0.064	0.020	0.076	0.021	0.076	0.071

Appendix

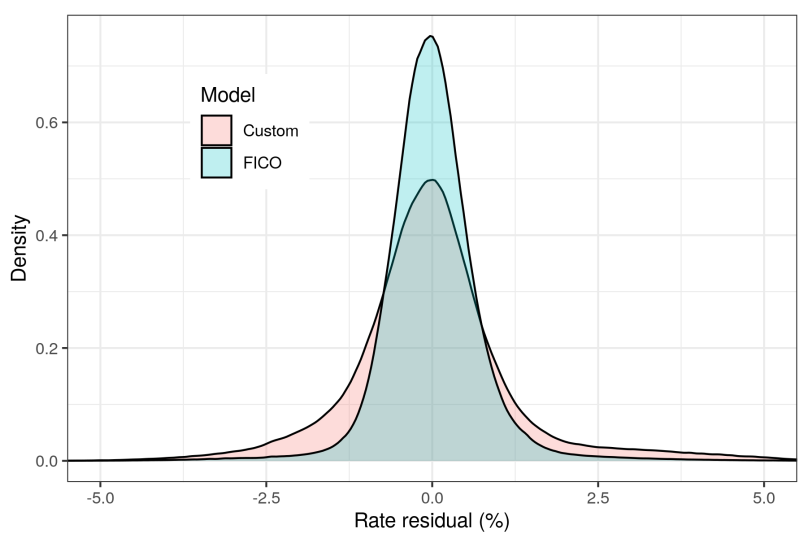
A Additional Tables and Figures

Figure A1: Data Use by Banks and Non-banks/Fintechs in the US Mortgage Market

Note: This figure shows the use of credit-scoring models by banks and non-banks and interest rate residuals in the US residential mortgage market. Panel (a) shows the fraction of mortgages originated using a credit scoring model besides standardized Equifax, Experian, FICO, or Vantage Score for depository (red) and non-depository (blue) institutions. Panel (b) shows the distribution of interest rate residuals for custom (red) and standardized (blue) credit scoring models after controlling for interacted LTV, loan purpose, lien status, loan type, debt-to-income ratio, whether the loan is a reverse mortgage, open-end line of credit, made for a business purpose, HOEPA status, construction method, occupancy type, and conforming status fixed effects, plus year-MSA fixed effects. Data is from HMDA for 2018 and 2019, merged with the Avery file to identify lender type.



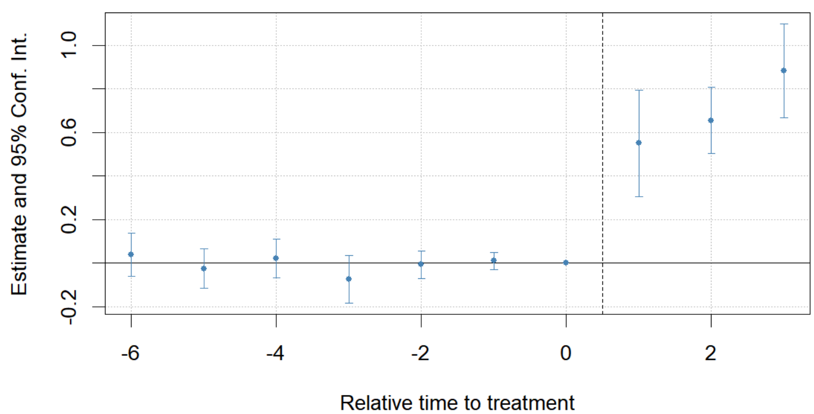
(a) Percentage of mortgages originated using alternate credit scoring methods



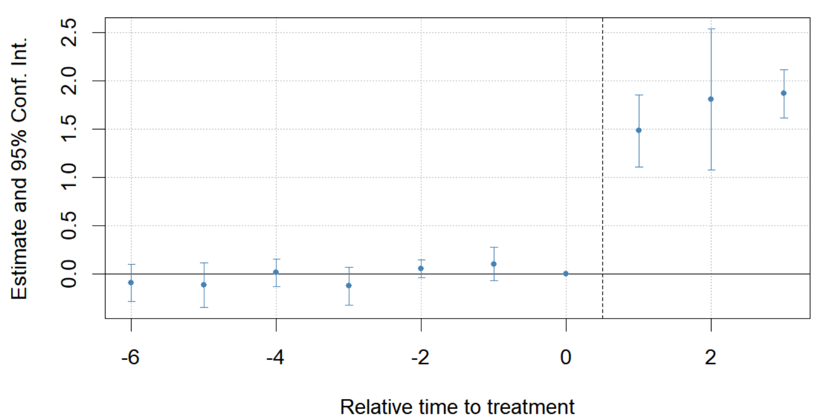
(b) Mortgage interest rate residuals by credit scoring method

Figure A2: Event-study of Fintech VC and Open Banking Using Gardner (2022) TWFE

Note: This figure replicates our main event study (Figure 3) of changes in fintech venture capital (VC) activity around the passage of open banking (OB) government policies but follows the two-way fixed-effect (TWFE) specification of Gardner (2022). We use the default specification provided by the R did2s package available here. We perform this analysis on our high-coverage Pitchbook panel of 2011-2021 data for the 21 countries with at least five fintech VC deals in the 2000–2010 period. Panel (a) shows an event study on the log of one plus the number of fintech VC deals, and Panel (b) shows an event study on the log of one plus the millions of US dollars invested in fintech VC deals. Year 0 is the passage year of each country's major open banking initiative. The coefficient for year 0 is set to zero and other coefficients are presented net of country fixed effects and region-by-year fixed effects for Africa, Middle East & North Africa; Europe & Central Asia; Latin America & the Caribbean; North America; South Asia, East Asia & Pacific, following World Bank geographic terms. European Union member states are weighted to count as a single country for estimates and standard errors. The bars indicate 95% confidence intervals calculated using standard errors clustered at the country level.



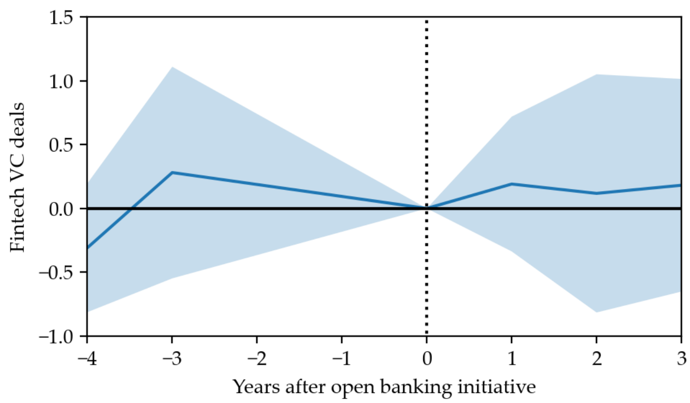
(a) Fintech VC deals following Gardner (2022)



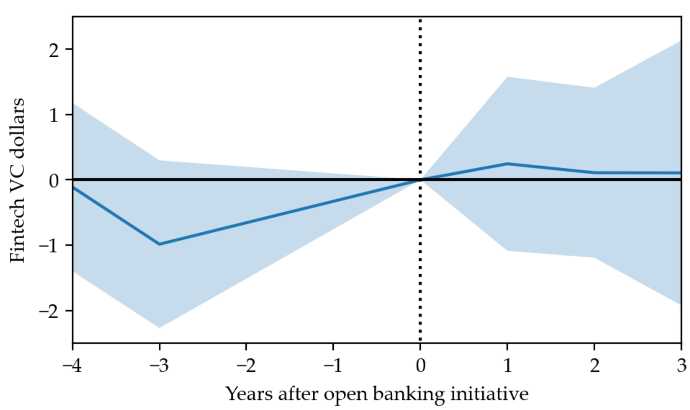
(b) Fintech VC dollars following Gardner (2022)

Figure A3: Placebo Test Event-study of Fintech VC Five Years Before Open Banking

Note: This figure conducts a placebo test of our main event study (Figure 3) of changes in fintech venture capital (VC) activity around the passage of open banking (OB) by shifting all events and windows forward five years. We perform this analysis on a Pitchbook panel of 2006-2016 data for the countries with at least five fintech VC deals in the 1995–2005 period. Panel (a) shows an event study on the log of one plus the number of fintech VC deals, and Panel (b) shows an event study on the log of one plus the millions of US dollars invested in fintech VC deals. Year 0 is five years before the passage year of each country's major open banking initiative. The coefficient for year 0 is set to zero and other coefficients are presented net of country and region-by-year fixed effects for Africa, Middle East & North Africa; Europe & Central Asia; Latin America & the Caribbean; North America; South Asia, East Asia & Pacific, following World Bank geographic terms. European Union member states are weighted to count as a single country for estimates and standard errors. The shaded regions denote 95% confidence intervals calculated using standard errors clustered at the country level.



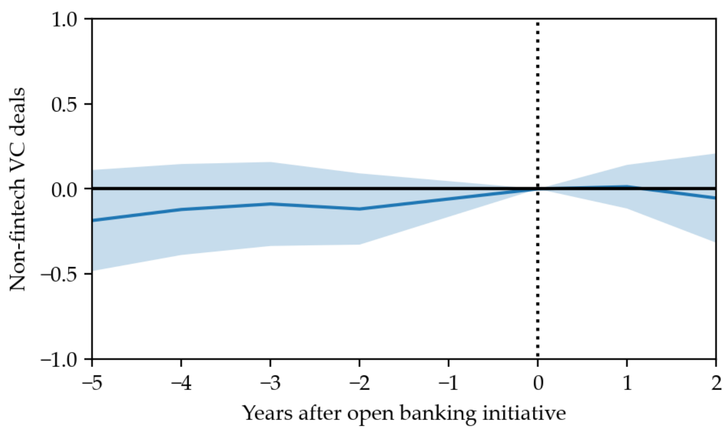
(a) Placebo test of fintech VC deals five years before passage of OB



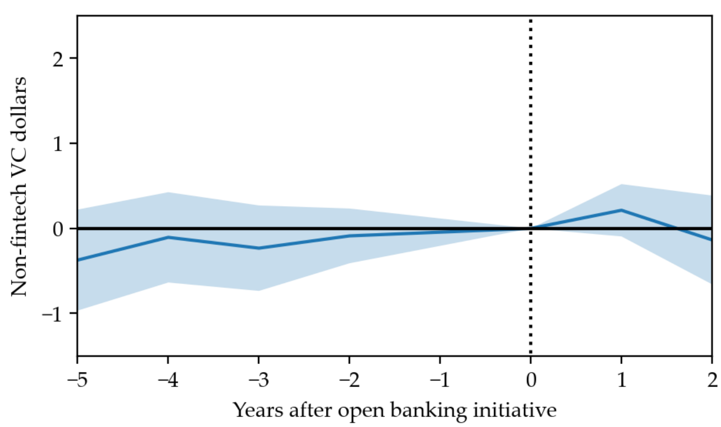
(b) Placebo test of fintech VC dollars five years before passage of OB

Figure A4: Placebo Test Event-study of Non-fintech VC After Open Banking

Note: This figure conducts a placebo test of our main event study (Figure 3) by switching non-fintech VC deals and fintech VC deals (and likewise for dollars) for our regression specification and sample construction. We perform this analysis on a Pitchbook panel of 2011-2021 data for the countries with at least five non-fintech VC deals in the 2000–2010 period. Panel (a) shows an event study on the log of one plus the number of non-fintech VC deals, and Panel (b) shows an event study on the log of one plus the millions of US dollars invested in non-fintech VC deals. Year 0 is five years before the passage year of each country's major open banking initiative. The coefficient for year 0 is set to zero and other coefficients are presented net of country and region-by-year fixed effects for Africa, Middle East & North Africa; Europe & Central Asia; Latin America & the Caribbean; North America; South Asia, East Asia & Pacific, following World Bank geographic terms. European Union member states are weighted to count as a single country for estimates and standard errors. The shaded regions denote 95% confidence intervals calculated using standard errors clustered at the country level.



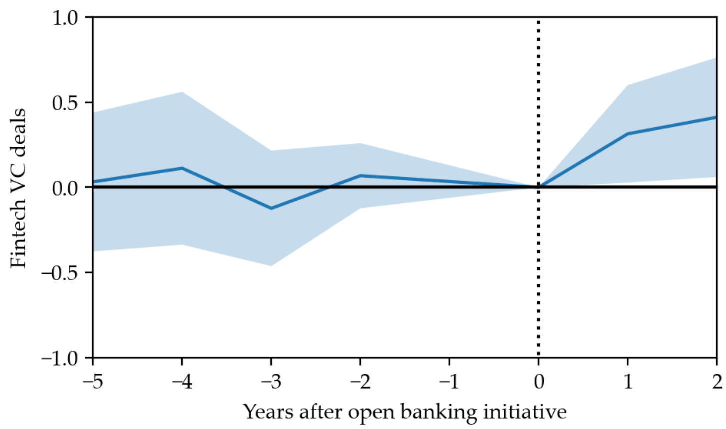
(a) Placebo test of non-fintech VC deals



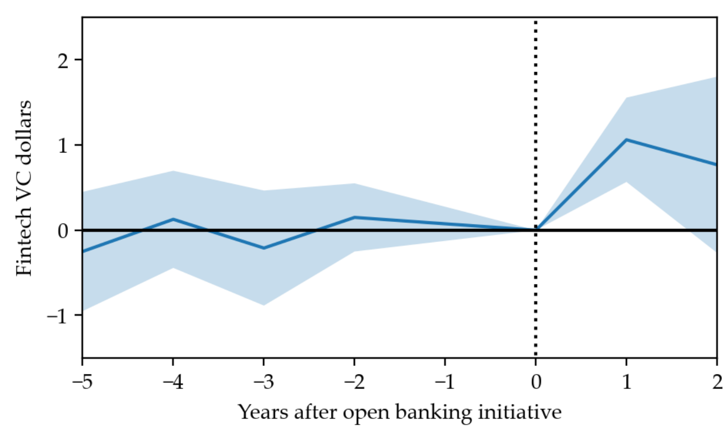
(b) Placebo test of non-fintech VC dollars

Figure A5: Event-study of Fintech Investment Controlling for Non-fintech VC deals

Note: This figure provides a robustness check to our main event study (Figure 3) by adding a control for non-fintech VC activity. We perform this analysis on a Pitchbook panel of 2011-2021 data for the countries with at least five fintech VC deals in the 2000–2010 period. Panel (a) shows an event study on the log of one plus the number of fintech VC deals while controlling for non-fintech VC deals. Panel (b) shows an event study on the log of one plus the millions of US dollars invested in fintech VC deals while controlling for the log of one plus the millions of dollars invested in non-fintech VC deals. Year 0 is five years before the passage year of each country's major open banking initiative. The coefficient for year 0 is set to zero and other coefficients are presented net of country and region-by-year fixed effects for Africa, Middle East & North Africa; Europe & Central Asia; Latin America & the Caribbean; North America; South Asia, East Asia & Pacific, following World Bank geographic terms. European Union member states are weighted to count as a single country for estimates and standard errors. The shaded regions denote 95% confidence intervals calculated using standard errors clustered at the country level.



(a) Fintech VC deals after controlling for non-fintech VC



(b) Fintech VC dollars invested after controlling for non-fintech VC

Table A1: Country Data Summary Statistics

Note: This table presents summary statistics on our country-level variables. Panel (a) reports values for a cross-section of country characteristics for all countries for which we have collected OB data; we use this sample to examine which country characteristics predict OB policy adoption. Panel (b) reports values for 2011–2021 panel data for our high-VC-coverage sample of countries that have at least five fintech venture capital (VC) deals in the 2000–2010 period; we use this sample for panel regressions of open banking's impact on fintech VC activity. For each variable, we present the number of observations, the average value, the standard deviation, and assorted percentiles. The first set of variables (under "Open banking variables") concern the status of open banking policies as of October 2021. After open banking (OB) initiative equals one in country-years after a major open banking policy was passed. The next three variables are set at the country level based on that country's OB policies: OB implemented is an indicator variable equal to one if the open banking policy was implemented or is in the pre-implementation stage, OB implementation is a 0-7 rating of the open banking policy progress where higher numbers denote more progress toward regulation, and the OB Strength Index is our 0-1 measure of open banking policy strength. All other variables, except "Trust in fintechs", are measured in 2013 year, which we use for pre-open banking country characteristics for our cross-country regressions; Trust in fintechs is measured in 2019—the earliest available year with comprehensive data. VC deals, non-fintech VC deals, and fintech VC deals are presented next and are from PitchBook and used after taking the log of one plus the number (and are hence different from Table A2). Per capita GDP in thousands of US dollars, the square of per capita GDP in hundreds of thousands of US dollars, the log of population (in millions), private sector credit to GDP, bank branches per 100k people, and the financial sector Lerner index are from the World Bank. Trust in fintechs is the proportion of survey respondents who report being willing to share their financial data with fintechs, as reported by Chen et al. (2023) The Lerner index ranges between 0 and 1 and measures the market power of banks, with higher values denoting less competition. Foreign-owned banks are from the Claessens and Van Horen (2013) foreign bank ownership data. The Rule of Law Index is from the Cato Institute and is on a 0 to 10 scale with higher numbers denoting more favorable conditions.

Panel (a) Cross sectional data for 168 country sample

	Count	Mean	Std. dev.	10th pct.	25th	50th.	75th	90th
Observations	168							
Open banking variables								
After open banking initiative	168	0.00	0.00	0.00	0.00	0.00	0.00	0.00
OB implemented	168	0.29	0.45	0.00	0.00	0.00	1.00	1.00
OB implementation	168	1.83	2.46	0.00	0.00	0.00	4.00	6.00
OB Strength Index	168	0.15	0.26	0.00	0.00	0.00	0.25	0.50
Venture capital variables								
VC deals	168	1.58	1.97	0.00	0.00	0.69	2.79	4.79
Non-fintech VC deals	168	1.54	1.95	0.00	0.00	0.69	2.66	4.68
Fintech VC deals	168	0.56	1.03	0.00	0.00	0.00	0.69	1.99
Other explanatory variables								
Per capita GDP ($k)	163	14.65	20.66	0.75	1.38	5.58	18.35	45.54
Per capita GDP ($100k) squared	163	0.06	0.17	0.00	0.00	0.00	0.03	0.21
Log population	156	2.48	2.09	0.60	1.35	2.33	3.39	4.33
Trust in fintechs	27	0.23	0.15	0.09	0.12	0.16	0.33	0.42
Private sector credit to GDP	149	54.19	47.37	11.08	18.70	39.79	70.53	124.82
Branches per 100k people	155	17.30	15.63	2.72	4.94	12.48	23.54	37.15
Financial sector Lerner index	94	0.31	0.13	0.17	0.23	0.29	0.37	0.47
Foreign-owned banks	134	0.43	0.28	0.03	0.20	0.42	0.66	0.80
Rule of Law Index	146	5.18	1.56	3.43	3.94	4.77	6.47	7.56

Table A1: Country Data Summary Statistics (continued)

Panel (b) Panel data for 21 countries with high-VC-coverage sample

	Count	Mean	Std. dev.	10th pct.	25th	50th.	75th	90th
Observations	231							
Open banking variables								
After open banking initiative	231	0.32	0.47	0.00	0.00	0.00	1.00	1.00
OB implemented	231	0.81	0.39	0.00	1.00	1.00	1.00	1.00
OB implementation	231	4.81	1.89	1.00	4.00	6.00	6.00	6.00
OB Strength Index	231	0.43	0.34	0.25	0.25	0.25	0.50	1.00
Venture capital variables								
VC deals	231	5.75	1.22	4.47	4.92	5.53	6.31	7.46
Non-fintech VC deals	231	5.67	1.21	4.36	4.82	5.42	6.21	7.29
Fintech VC deals	231	3.19	1.38	1.61	2.30	3.00	3.88	5.10
Other explanatory variables								
Per capita GDP ($k)	231	40.26	22.00	6.78	27.13	44.27	52.62	62.73
Per capita GDP ($100k) squared	231	0.21	0.18	0.00	0.07	0.20	0.28	0.39
Log population	231	4.40	4.00	1.69	2.27	3.64	4.85	5.80
Trust in fintechs	176	0.19	0.16	0.07	0.10	0.14	0.15	0.56
Private sector credit to GDP	199	107.33	44.47	51.88	65.25	105.49	141.13	167.65
Branches per 100k people	197	24.86	13.72	8.92	14.79	21.86	32.93	38.36
Financial sector Lerner index	72	0.27	0.12	0.11	0.18	0.28	0.34	0.41
Foreign owned banks	231	0.27	0.24	0.02	0.08	0.20	0.40	0.58
Rule of Law Index	126	7.08	1.44	4.35	6.78	7.45	8.18	8.61

Table A2: PitchBook Data Summary Statistics

Note: This table presents summary statistics on our PitchBook venture capital (VC) deal data for 168 countries over 2000–2021. The first column presents statistics on the entire dataset, the next two columns present data for 2000–2010 and 2011–2021 for low-coverage countries, and the final two columns present data for 2000-2010 and 2011-2021 for high-coverage countries. High-coverage countries are those with five or more fintech VC deals in the 2000–2010 period, while countries with fewer than five are low-coverage countries. The first set of rows presents the number of countries in each sample, both those with open banking implemented or in the pre-implementation stage as of October 2021 and those that have not reached that stage. The second set of rows presents the number of country-year observations in each sample, both those that are after an open banking policy was passed in that country and other observations. The third set of rows presents statistics on country-year VC investment: any VC deals indicates the percentage of country-years with a VC deal, mean and median raw VC deals present the average number of deals in country-years, and mean and median raw VC dollars ($m) presents the average value of VC deals in a country-year in millions of US dollars. The fourth set of rows presents similar statistics on country-year fintech VC investment.

	All countries	Low-coverage countries		High-coverage countries	
	2000-2021	2000-2010	2011-2021	2000-2010	2011-2021
Countries					
Count of countries	168	147	147	21	21
Countries with open banking implemented	49	32	32	17	17
Countries without open banking implemented	119	115	115	4	4
Country-year observations					
Count of country-year observations	3,696	1,617	1,617	231	231
Country-years after open banking passed	139	0	84	0	55
Country-years before open banking passed	3,557	1,617	1,533	231	176
Country-year VC activity					
Any VC deals (%)	44.6	23.4	50.1	99.1	100.0
Mean raw VC deals	74.5	1.1	13.0	212.7	880.6
Median raw VC deals	0.0	0.0	1.0	38.0	251.0
Mean raw VC dollars ($m)	718.1	6.3	85.8	1,725.2	9,119.4
Median raw VC dollars ($m)	0.0	0.0	0.0	170.6	1,109.9
Country-year fintech VC activity					
Any fintech VC deals (%)	25.3	3.2	31.4	64.1	98.7
Mean raw fintech VC deals	6.0	0.0	1.8	8.4	74.8
Median raw fintech VC deals	0.0	0.0	0.0	1.0	19.0
Mean raw fintech VC dollars ($m)	81.2	0.1	20.9	66.8	1,085.9
Median raw fintech VC dollars ($m)	0.0	0.0	0.0	1.3	88.6

Table A3: Effect of Open Banking Government Policy Characteristics on Fintechs

Note: This table shows changes in fintech venture capital (VC) investment activity following the implementation of different types of open banking policies by governments around the world. The table uses a difference-in-differences design on our high-coverage Pitchbook sample of country-year data spanning 2011-2021 for the 21 countries with at least five fintech deals in the 2000–2010 period. The dependent variable in each specification is the log of one plus the number of fintech VC deals in each country-year. The independent variables are different characteristics of open banking government policies interacted with an indicator variable equal to one if the year in question is after the year major open banking laws were passed in the country in question. In column 1 we indicate whether banks are mandated to share the data with other financial service providers upon consumer request; in column 2 whether there is data reciprocity between banks and other financial service providers (e.g., if fintechs have to share customer data with banks); in column 3 whether regulators set technical standards for open banking implementation; and in column 4 whether, in addition to bank payment accounts, open banking policies cover other financial products and services (e.g., mortgages, insurance). In column 5, we interact with the Open Banking Strength Index, which we define as the average of those four policy dimensions used in columns 1 to 4. All specifications have a control for non-fintech VC activity, country fixed effects, and region-by-year fixed effects for Africa, Middle East & North Africa; Europe & Central Asia; Latin America & the Caribbean; North America; South Asia, East Asia & Pacific, following World Bank geographic terms. European Union member states are weighted to count as a single country for estimates and standard errors. Standard errors are clustered at the country level. *** denotes p-value < 0.01, ** denotes <0.05, and * denotes <0.1.

	Fintech VC Deals				
	(1)	(2)	(3)	(4)	(5)
Banks must share × after OB	0.284**				
	(0.124)				
Users must share × after OB		0.264			
		(0.173)			
Regulated specifications × after OB			0.179		
			(0.142)		
Beyond transactions × after OB				0.141	
				(0.308)	
OB Strength Index × after OB					0.488
					(0.370)
After OB initiative	0.037	0.220	0.208	0.247	0.030
	(0.156)	(0.146)	(0.184)	(0.171)	(0.273)
Non-fintech VC control	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes
Region-Year FE	Yes	Yes	Yes	Yes	Yes
Observations	231	231	231	231	231
Adjusted R^2	0.937	0.937	0.937	0.936	0.937

Table A4: Leave-one-out Country Effect of Open Banking Government Policy on Fintechs

Note: This table shows how our estimate of the effect of open banking on fintech VC deals estimated using Equation (4) varies when we exclude countries from our data. Each row corresponds to a data sample that is equal to our high-coverage Pitchbook panel data of country-year data spanning 2011-2021 for the 21 countries with at least five fintech deals in the 2000–2010 period, but excluding each country one at a time in the first 21 rows, excluding France and Germany (the two largest EU countries) together in the next row, and excluding Canada, China, and the USA together (the three countries that did not pass OB laws in our sample period) in the final row. The Coefficient column presents the coefficient on post-open banking (parameter on After OB initiative) estimated using a difference-in-differences design on that sample, with the Standard error, t stat. and p-value columns similarly presenting their respective statistics. All specifications have a control for non-fintech VC activity, country fixed effects, and region-by-year fixed effects for Africa, Middle East & North Africa; Europe & Central Asia; Latin America & the Caribbean; North America; South Asia, East Asia & Pacific, following World Bank geographic terms. European Union member states are weighted to count as a single country for estimates and standard errors. Standard errors are clustered at the country level. *** denotes p-value < 0.01, ** denotes <0.05, and * denotes <0.1.

	Coefficient	Standard error	t stat.	p-value
Excluding AUS	0.294***	0.092	3.201	0.009
Excluding BEL	0.300**	0.114	2.637	0.023
Excluding BRA	0.299**	0.111	2.685	0.023
Excluding CAN	0.301**	0.114	2.630	0.025
Excluding CHN	0.172*	0.094	1.826	0.098
Excluding DEU	0.301**	0.115	2.626	0.024
Excluding DNK	0.300**	0.115	2.617	0.024
Excluding ESP	0.296**	0.115	2.570	0.026
Excluding FIN	0.298**	0.115	2.588	0.025
Excluding FRA	0.298**	0.114	2.603	0.025
Excluding GBR	0.329*	0.156	2.112	0.061
Excluding IND	0.357**	0.131	2.717	0.022
Excluding IRL	0.304**	0.115	2.646	0.023
Excluding ISR	0.299**	0.111	2.685	0.023
Excluding JPN	0.310*	0.144	2.162	0.056
Excluding NLD	0.297**	0.116	2.559	0.027
Excluding NOR	0.264*	0.123	2.135	0.058
Excluding POL	0.297**	0.117	2.543	0.027
Excluding RUS	0.355*	0.169	2.102	0.062
Excluding SWE	0.295**	0.115	2.570	0.026
Excluding USA	0.301**	0.114	2.630	0.025
Excluding DEU and FRA	0.301**	0.115	2.608	0.024
Excluding CAN, CHN, and USA	0.176*	0.095	1.856	0.096

Table A5: Effect of Open Banking Government Policy on Fintechs Using Other Transformations

Note: This table shows changes in fintech venture capital (VC) investment following the implementation of open banking government policies. The table tests our main difference-in-differences regression specification (Equation 4) on our high-coverage Pitchbook panel data of country-year data spanning 2011-2021 for the 21 countries with at least five fintech deals in the 2000–2010 period. Odd columns consider dependent variables based on the number of fintech deals, while even columns consider dependent variables based on the value of fintech deals in millions of US dollars. Each pair of columns considers different transformations of these measures, with columns 1 to 2 considering the log of 1 plus the measure of fintech activity (our main specification), columns 3 and 4 considering the ratio of fintech activity to trillions of dollars of GDP, columns 5 and 6 considering the ratio of fintech VC activity to total VC activity in the country-year, and columns 7 and 8 considering the inverse hyperbolic sine (IHS) of fintech VC activity in a country-year. The independent variable is a dummy variable equal to one if the year in question is after the year major open banking laws were passed in the country in question. The mean of the dependent variables, the standard deviation of the dependent variables, and Cohen's d values are presented to aid in the interpretation of the effect sizes under the different transformations. Cohen's d is the ratio between the regression coefficient on After OB initiative and the standard deviation of the dependent variable. All specifications have a control for non-fintech VC activity, country fixed effects, and region-by-year fixed effects for Africa, Middle East & North Africa; Europe & Central Asia; Latin America & the Caribbean; North America; South Asia, East Asia & Pacific, following World Bank geographic terms. European Union member states are weighted to count as a single country for estimates and standard errors. Standard errors are clustered at the country level. *** denotes p-value < 0.01, ** denotes <0.05, and * denotes <0.1.

	Log (1 + Fintech VC)		Fintech VC / GDP		Fintech VC / Total VC		IHS Fintech VC	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Deals	Dollars	Deals	Dollars	Deals	Dollars	Deals	Dollars
After OB initiative	0.308**	0.874**	17.758*	298.962***	0.027**	0.079**	0.303*	0.868*
	(0.125)	(0.368)	(9.438)	(24.498)	(0.009)	(0.026)	(0.146)	(0.403)
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Non-fintech VC control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	231	231	200	200	231	231	231	231
Adjusted R^2	0.937	0.898	0.752	0.452	0.754	0.621	0.923	0.89
Std. of dependent variable	1.38	2.33	23.52	328.21	0.04	0.11	1.47	2.44
Mean of dependent variable	3.19	4.58	25.51	194.18	0.08	0.11	3.80	5.20
Cohen's d	0.22	0.38	0.75	0.91	0.62	0.74	0.21	0.36

Table A6: FLS Data Summary Statistics

Note: This table presents summary statistics from the Financial Conduct Authority's 2020 Financial Lives Survey For each variable, we present the number of observations, the average value, the median value, and the standard deviation. Observation counts vary as we exclude "don't know" and/or "prefer not to say" responses. Advice OB is an indicator variable equal to one if the respondent uses open banking for financial advice products, i.e., answers yes to using financial aggregation apps that allow consumers to see the accounts they hold with different banks in one place (e.g., Money Dashboard, Yolt, MoneyHub) or savings-related apps that help build savings by monitoring consumer current accounts and automatically transferring funds (e.g., Chip, Cleo, Moneybox, Plum). Credit OB is an indicator variable equal to one if the respondent uses open banking for credit products, i.e., answers yes to using firms offering customized lending products, credit reference agencies (which can personalize credit reports), or price comparison websites (e.g., rates offered by different lenders). Financial knowledge is the respondent's answer to the question "How knowledgeable would you say you are about financial matters?" on a 0 (not at all knowledgeable) to 10 (very knowledgeable) scale. Credit card, personal loan, student loan, and pawnbroking loan are indicator variables equal to one if the respondent currently holds the credit product in question or held it in the last 12 months. The number of other credit products variables are the count of credit products the respondent reports owning, excluding the product in question. Unwillingness to share data equals one if the respondent gives a score of 3 or below on a 0 to 10 scale to the question "Thinking about Open Banking, how willing would you be to give your bank permission to securely access your banking information?". Employed equals one if the respondent reports working full- or part-time. Missing bill payments equals one if the respondent reports missing bills payments in at least three of the last six months or finds keeping up with domestic bills and credit commitments to be a heavy burden. Risk aversion equals one if the respondent gives a score of 3 or below on a 0 to 10 scale to the question "Are you a person who is generally willing to take risks?". Higher education equals one if the respondent has completed at least some post-secondary education. Young equals one if the respondent is aged 18–39. Male equals one if the respondent reports being male. White equals one if the respondent is of white ancestry. Married equals one if the respondent is married or in a registered civil partnership.

	Count	Mean	Median	Std. dev.
Open banking usage				
Advice OB	3,923	0.086	0	0.281
Credit OB	3,943	0.055	0	0.227
Financial knowledge and credit product ownership				
Financial knowledge	4,266	6.545	7	2.272
Credit card	4,310	0.536	1	0.499
Personal loan	4,310	0.107	0	0.309
Student loan	4,310	0.182	0	0.386
Pawnbroking loan	4,310	0.004	0	0.066
Number of other credit products				
Excluding credit cards	4,310	1.054	1	1.302
Excluding personal loans	4,310	1.483	1	1.368
Excluding student loans	4,310	1.408	1	1.456
Excluding pawnbroking	4,310	1.585	1	1.49
Respondent characteristics				
Unwillingness to share data	3,940	0.524	1	0.499
Employed	4,281	0.453	0	0.498
Missing bill payments	4,234	0.141	0	0.348
Risk aversion	4,257	0.371	0	0.483
Higher education	3,963	0.518	1	0.5
Young	4,310	0.406	0	0.491
Male	4,253	0.451	0	0.498
White	4,188	0.916	1	0.277
Married	4,172	0.452	0	0.498

Table A7: Determinants of Consumer Open Banking Usage in the UK

Note: This table shows the association between OB usage and demographic variables using person-level responses to the Financial Conduct Authority's 2020 Financial Lives Survey in the UK. We use a cross-sectional OLS specification. The dependent variable in column 1 indicates if the respondent uses open banking for financial advice products, i.e., answers yes to using financial aggregation apps that allow consumers to see the accounts they hold with different banks in one place (e.g., Money Dashboard, Yolt, MoneyHub) or savings-related apps that help build savings by monitoring consumer current accounts and automatically transferring funds (e.g., Chip, Cleo, Moneybox, Plum). The dependent variable in column 2 indicates if the respondent uses open banking for credit products, i.e., answers yes to using firms offering customized lending products, credit reference agencies (which can personalize credit reports), or price comparison websites (e.g., rates offered by different lenders). Other variables are dummy variables measuring person-level characteristics. Unwillingness to share data equals one if the respondent gives a score of 3 or below on a 0 to 10 scale to the question "Thinking about Open Banking, how willing would you be to give your bank permission to securely access your banking information?". Employed equals one if the respondent reports working full- or part-time. Missing bill payments equals one if the respondent reports missing bills payments in at least three of the last six months or finds keeping up with domestic bills and credit commitments to be a heavy burden. Risk aversion equals one if the respondent gives a score of 3 or below on a 0 to 10 scale to the question "Are you a person who is generally willing to take risks?". Higher education equals one if the respondent has completed at least some post-secondary education. Young equals one if the respondent is aged 18–39. Male equals one if the respondent reports being male. White equals one if the respondent is of white ancestry. Married equals one if the respondent is married or in a registered civil partnership. All specifications control for county (UK local authority) fixed effects and estimate robust standard errors. *** denotes p-value < 0.01, ** denotes <0.05, and * denotes <0.1.

	Advice OB (1)	Credit OB (2)
Unwillingness to share data	-0.022**	-0.037***
	(0.010)	(0.009)
Employed	0.039***	0.024***
	(0.011)	(0.009)
Missing bill payments	0.043**	0.049***
	(0.017)	(0.015)
Risk aversion	-0.000	-0.010
	(0.010)	(0.008)
Higher education	-0.000	0.022***
	(0.010)	(0.009)
Young	0.047***	0.017
	(0.013)	(0.011)
Male	-0.002	-0.013
	(0.010)	(0.008)
White	-0.014	0.021
	(0.023)	(0.017)
Married	0.001	0.005
	(0.011)	(0.009)
County FE	Yes	Yes
Observations	3,217	3,232
Adjusted R^2	0.035	0.025

Table A8: Consumers OB Usage and their Financial Knowledge and Credit Access with Credit demand Controls

Note: This table shows the association between credit product ownership and open banking usage using person-level responses to the Financial Conduct Authority's 2020 Financial Lives Survey in the UK. We use a cross-sectional OLS specification. The dependent variables in columns 1 to 4 are indicator variables equal to one if the respondent currently holds the credit product in question or held it in the last 12 months. In column 5, we run a similar specification at the person-by-product-level with each person entering the data four times, once for each potential credit product, and the omitted category is student loans. Credit OB is an indicator variable equal to one if the respondent uses open banking for credit products, i.e., answers yes to using firms offering customized lending products, credit reference agencies (which can personalize credit reports), or price comparison websites (e.g., rates offered by different lenders). Number of other credit products is the total number of all credit products that the respondent holds other than the credit product in the respective column. Columns 1 to 4 control for county (UK local authority) fixed effects and demographic controls. Column 5 controls for the interaction between the credit product and respondents' (demographic) controls as well as person fixed effects. Respondent controls are indicator variables for unwillingness to share, employed, missing bill payments, risk aversion, higher education, young, male, white, and married. Unwillingness to share data equals one if the respondent gives a score of 3 or below on a 0 to 10 scale to the question "Thinking about Open Banking, how willing would you be to give your bank permission to securely access your banking information?". Employed equals one if the respondent reports working full- or part-time. Missing bill payments equals one if the respondent reports missing bills payments in at least three of the last six months or finds keeping up with domestic bills and credit commitments to be a heavy burden. Risk aversion equals one if the respondent gives a score of 3 or below on a 0 to 10 scale to the question "Are you a person who is generally willing to take risks?". Higher education equals one if the respondent has completed at least some post-secondary education. Young equals one if the respondent is aged 18–39. Male equals one if the respondent reports being male. White equals one if the respondent is of white ancestry. Married equals one if the respondent is married or in a registered civil partnership. All specifications estimate robust standard errors. *** denotes p-value < 0.01, ** denotes <0.05, and * denotes <0.1.

	Credit card (1)	Personal loan (2)	Student loan (3)	Pawnbroking loan (4)	Product level (5)
Credit OB	0.071*	0.068**	0.006	-0.002	
	(0.039)	(0.033)	(0.033)	(0.005)	
Credit card × Credit OB					0.119**
					(0.047)
Personal loan × Credit OB					0.089**
					(0.044)
Pawnbroking loan × Credit OB					-0.010
					(0.038)
Number of other credit products	0.101***	0.061***	-0.004	0.005**	
	(0.008)	(0.005)	(0.005)	(0.002)	
Respondent controls	Yes	Yes	Yes	Yes	
County FE	Yes	Yes	Yes	Yes	
Product-Respondent controls FE					Yes
Person FE					Yes
Observations	3,232	3,232	3,232	3,232	12,928
Adjusted R^2	0.223	0.146	0.32	0.043	0.412

B Government-led Open Banking and Incumbent Banks' Data Sharing

Government OB policies either force or encourage banks to allow other financial service providers to access their customer data upon customer request. The basic threshold question is whether banks indeed share their customer data following OB government policies. Since APIs are the main technology for data sharing in OB, we examine whether the presence of OB government policies is associated with bank API offerings. We use bank API data from Platformable, which is a global leader in data on OB APIs.[48] To our knowledge, this data provides the best global coverage of banks' APIs, and allows us to examine whether government efforts to encourage OB adoption are actually effective at opening up banks by leading them to introduce APIs.

Table B1 shows the results of a cross-sectional, cross-country regression of the prevalence of APIs in banks of each country against our measures of that country's government OB implementation:

$$BankAPIs_i = \beta \times OB_i + X_i'\gamma + Region_r + \epsilon_i, \tag{16}$$

where $BankAPIs_i$ is the log-transformed number of banks with APIs (columns 1 to 3) or the percentage of the top 10 banks in each country that offer APIs (columns 4 to 6). OB_i is one of three types of OB outcomes. First, we use a 0/1 indicator for whether the government has already implemented OB policies in a country as of October 2021 (columns 1 and 4). Second, we use a continuous measure of how far the implementation of government OB policy has progressed, with 0 denoting none and 7 denoting fully implemented with follow-on regulation (columns 2 and 5). Third, we use the interaction between our 0/1 OB policy indicator and our 0 to 1 OB Strength Index (columns 3 and 6), which is described at the end of Section 2.3. $Region_r$ are region fixed effects, and X_i' is a vector of ex-ante basic economic country characteristics (GDP per capita and population).

There is a strong positive association between OB policies and bank API offerings. Column 1 shows that countries with OB policies have about twice as many banks offering APIs, with columns 2, 4, and 5 yielding qualitatively similar numbers. Columns 3 and 6 show that these effects are driven by more comprehensive OB policies. These results provide the first systematic evidence that government policies to promote OB might have already had a significant effect on data sharing in the financial service industry, and that counties that have more comprehensive OB policies (as measured by our OB Strength Index) are likely to see

[48]Platformable collects industry data on OB and open finance by systematically identifying API providers and consumers using bank and fintech website sources, fintech registers such as EUCLID (EU) and FCA (UK), assessing API consumers and providers from fintech association membership lists, and by surfacing new initiatives from newsletters and industry alerts. Data is collected on a rolling basis, with each entity assessed at least once every three months.

more data sharing. These results also suggest that banks are not voluntarily sharing data, a result consistent with the model we later present.

Table B1: Open Banking Government Policy and Bank API Offerings

Note: This table shows the association between government open banking policies and banks' open application programming interfaces (APIs) using an OLS specification. The sample includes the sample of countries for which we have collected OB data (168 countries) that we could also obtain bank data (158 countries). The dependent variable in columns 1 to 3 is the log of one plus the number of banks offering APIs and in 4 to 6 it is the percentage of the top 10 banks in each country (as ranked by 2020 assets in Bureau van Dijk) that offer APIs as of the end of 2021. APIs are the technology used to share bank customer data under open banking. The independent variable of interest in columns 1 and 4 is Open banking implemented (0/1) which is an indicator variable equal to one if open banking was implemented in the country in question as of October 2021; in columns 2 and 5 it is Open banking implementation (0–7) which is a 0–7 rating of the extent of open banking government policy implementation progress as of October 2021, with 0 being no action, 1–2 being increasingly serious levels of discussion, and 3–7 being levels of implementation progress; and in columns 3 and 6 it is the interaction of the open banking implemented (0/1) indicator variable with our Open Banking Strength Index which is a measure of policy strength. The open banking implemented indicator corresponds to being in or after the pre-implementation stage or equivalently to a level of 3 or above. All specifications include GDP per capita in thousands of US dollars, the square of GDP per capita in hundreds of thousands of US dollars, and the log of population, all based on World Bank data as of 2013. Region fixed effects are for Africa, Middle East & North Africa; Europe & Central Asia; Latin America & the Caribbean; North America; South Asia, East Asia & Pacific, following World Bank geographic terms. European Union member states are weighted to count as a single country for estimates and standard errors. All specifications estimate robust standard errors. *** denotes p-value < 0.01, ** denotes <0.05, and * denotes <0.1.

	Banks with APIs			% of top 10 banks with APIs		
	(1)	(2)	(3)	(4)	(5)	(6)
Open banking implemented (0/1)	0.794***		-0.286	0.218***		-0.005
	(0.245)		(0.224)	(0.072)		(0.053)
Open banking implementation (0-7)		0.233***			0.058***	
		(0.046)			(0.013)	
OB Strength Index × OB implemented			1.238***			0.256***
			(0.341)			(0.087)
Per capita GDP ($k)	0.031**	0.016	0.027**	0.004	0.000	0.003
	(0.012)	(0.011)	(0.012)	(0.003)	(0.003)	(0.003)
Per capita GDP ($100k) squared	-0.893	0.404	-0.612	-0.023	0.266	0.035
	(1.705)	(1.475)	(1.610)	(0.447)	(0.394)	(0.433)
Log population	0.164***	0.145**	0.162***	0.025**	0.021**	0.025**
	(0.060)	(0.056)	(0.059)	(0.010)	(0.009)	(0.010)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	158	158	158	158	158	158
Adjusted R^2	0.577	0.629	0.596	0.461	0.509	0.476

C Open Banking Data Collection and Variable Definitions

This appendix describes the construction of our OB government policies dataset and defines variables. Each observation in the dataset corresponds to a country's OB approach as of the collection date.[49]

C.1 What is an "approach" and what makes an approach "open banking"?

A government-led OB approach does not need to be a single law or policy; many countries' OB approaches in fact are composed of several separate policies. Rather, an approach encompasses the totality of the country's OB government efforts.

The line between OB policies and related but non-OB policies can be unclear, and a single simple definition cannot encompass all cases. For our purposes, there are two reasons for us to classify a regulatory approach as OB:

1. **Functional**: Does the regulator's approach have the key functional elements of OB? Specifically, does it facilitate programmatic access (e.g., through an API) to financial intermediaries' customers' data for the purposes of data sharing or payments?

2. **Nominal**: Do regulators, journalists, or industry groups refer to the regulation as "open banking"?

The functional approach is more objective and can be applied to countries that have progressed sufficiently far down the pathway of discussing and implementing OB policies. The nominal approach is useful in cases where regulators have only recently been discussing OB but none of the functional elements have yet been formalized. The following two regulations may be similar to OB but we do not consider to be OB policies and we list them as illustrative counterexamples:

1. General Data Protection Regulation (GDPR): This EU law grants consumers certain privacy rights over their data. However, GDPR is not an OB law because it does not mandate that commercial entities (specifically, banks) in possession of the data share it upon customer request. Note, however, that the EU does have an OB law, the PSD2.

2. Regulation related to central bank digital currencies (CBDC): Movements to create payment systems utilizing CBDC are payments regulations but are not open payments regulations, as they do not mandate open data sharing between market participants. There have been many payments-related regulations (CBDC and other) that modernize payments but are not "open" in any sense, aside from, for example, reporting requirements to regulators.

[49]Most recently, October 2021.

Having defined what constitutes an "approach" and what makes an approach an "open banking" approach, we now define in detail the variables we collect and the classification decision rules. With each data category, we provide notes to clarify decision rules and address common questions.

C.2 Data categories and variable definitions

C.2.1 Open banking approach and regulatory mandate

- **government_led_initiative**: Is there a government-led initiative around OB?
  - **Yes.**
  - **No.**
- **regulatory_entity_type**: Which type of regulator is leading the OB effort?
  - **Monetary authority**: A financial regulator, e.g., a central bank.
  - **Competition authority**: A regulator tasked with anti-trust or other competition-related enforcement, e.g., the Competition and Markets Authority in the UK.
  - **Consumer protection authority**: A regulator tasked with consumer protection, e.g., the Consumer Financial Protection Bureau in the US or a data privacy authority.
- **innovation_mandate**: Is increasing innovation a proffered policy mandate?
  - **Yes**: Spurring the creation or adoption of new financial products or technologies is either discussed or explicitly stated as policy goals.
  - **No**: Otherwise.
- **competition_mandate**: Is increasing competition a proffered policy mandate?
  - **Yes**: Increasing entry, increasing competition, decreasing markups, or related issues are either discussed or explicitly stated as policy goals.
  - **No**: Otherwise.
- **inclusion_mandate**: Is increasing financial inclusion a proffered policy mandate?
  - **Yes**: Increasing access to the financial system, serving the unbanked, fighting inequality, or related issues are either discussed or explicitly stated as policy goals.
  - **No**: Otherwise.

How do we denote efforts coordinated between both regulators and market participants?

We define these as government-led efforts. The justification for this is that almost all major government policies involve some level of collaboration or input from industry. In the US, for example, there are open comment periods and meetings with industry and lobbyists. Fundamentally, however, these initiatives work through the government, and so to the extent that the government has any authoritative hand in leading the regulation, we consider it as government-led.

Which agency type do we select in cases where several are responsible?

We select the regulator most aligned with the proffered mandate or rationale for OB. For example, in the case of Australia, we select the Australian Competition and Consumer Commission because the country's OB policy mandate is most closely aligned with that of a "competition authority".

C.2.2 Timeline and initiative

- **initiative_name**: Name of the government-led policy initiative.
- **initiative_passed_date**: Date that the OB legislation is signed into law, or date when the first non-regulation government major effort to promote OB goes into effect (e.g., for Singapore we use November of 2016—the date when the Monetary Authority of Singapore (MAS) published a comprehensive roadmap: API Playbook—which, in effect, set the gold standard for regulatory advice on the topic in Asia: see here). For efforts that have not yet been signed into law or resulted in a major government policy, this field is TBD.
- **data_sharing_date**: First date at which the legal mandate on customers' data sharing begins to bind, or (in cases of non-legally-binding policies) when the government sets up the infrastructure that allows customer data sharing.
- **oct_21_status**: Implementation status as of October 2021.
  - **Nothing**: No government-led OB.
  - **Pre-discussion**: Some government interest but no actual law or implementation is taking place.
  - **Discussion**: The actual law has been introduced or passed and rulemaking is taking place.
  - **Pre-implementation**: The law is passed and rules have been set, but nothing is yet binding.

  - **Early implementation**: Some data sharing requirements are binding (e.g., bank-level product information), but not personal account/transactions.
  - **Mid implementation**: Personal account/transaction data sharing is binding, but not all planned elements are in place (e.g., not all planned API functionality exists.)
  - **Fully implemented**: Full implementation as described in the law/rulemaking.
  - **Follow-on regulation**: OB is implemented, and regulators are actively working on related regulation such as open finance or open data more broadly.

Which government effort do we focus on when there are several?

We focus on the first major government OB effort.[50] For example, in the United States, several regulatory bodies have expressed interest in OB (e.g., the Treasury/OCC and the Consumer Financial Protection Bureau (CFPB)). The CFPB's effort through Dodd-Frank Section 1033 is the most important US regulatory effort. In the UK, the 2016 CMA9 order was the first major open banking law, although it is subject to pending follow-up regulation to broaden its scope.

What if the precise date is unavailable?

In cases where the precise date cannot be found or is ambiguous for some reason, we use the most precise date that can be inferred from the data. For example, if the best information for a country that can be located says the law passed in "the second half of 2020," we will assign the date as July 1, 2020.

What event defines the data sharing date?

In cases where data sharing is mandated, this is the date. In cases where data sharing is not mandated but, for example, the regulator sets API standards, we use the date at which the API standards go into force. In cases where the regulation initially applies only to a subset of later planned entities (e.g., the UK Open Banking Initiative applies to 9 large banks), we use the date at which the requirements first apply to any entity.

C.2.3 Standards

- **regulatory_technical_specifications**: Does the regulator set technical specifications for data sharing / payments?
  - **Yes.**

[50] Given the recency of the OB trend, this is almost always also the latest OB approach with the exceptions being the United Kingdom and Sweden. These two countries had earlier, abortive OB attempts that we exclude due to their limited implementation.

- **No.**

What happens when regulators and industry collaborate on technical specifications?

This field is "Yes" if technical standards are either developed internally by the regulator, arrived at through collaboration of the regulator with industry participants, or mandated by the regulator to be developed by industry participants.

C.2.4 Open banking scope

- **financial_services_scope**: How wide is the set of financial products covered under OB?
  - **Narrow**: Transaction accounts only.
  - **Broad**: Transaction accounts and other "core" financial products (e.g., loans).
  - **Very broad**: Above products plus "non-core" financial products (e.g., insurance).
- **transaction_accounts_covered**: Does the regulation cover transaction accounts?
  - **Yes.**
  - **No.**
- **nontransaction_accounts_covered**: Does the regulator cover financial products aside from transaction accounts?
  - **Yes.**
  - **No.**
- **share_account_data**: Does the regulator either require or facilitate the sharing of customers' transaction account data?
  - **Yes.**
  - **No.**
- **payment_initiation**: Does the regulator require or facilitate technology to allow the initiation of customer payments by third parties?
  - **Yes.**
  - **No.**

What do we include in transaction accounts?

Any financial account that allows for cash-like transactions, e.g., checking accounts, debit cards, credit cards, and digital wallets.

What are core and non-core financial products?

Core products are consumer financial products that banks typically offer, including, e.g., loans or investment services. Non-core products are either consumer finance products that banks do not typically offer, e.g., insurance, or financial products that are not "consumer" finance products, such as small business loans.

Is a payment service like Venmo or Alipay an OB transaction service?

No, these services do not rely on open APIs interfacing with banks. See the definition of an OB approach above.

C.2.5 Sharing scope and reciprocity

- **data_holders_share**: Do data holders (e.g., banks) have to share their customers' data (upon customer request)?
  - **Yes.**
  - **No.**
- **data_users_share**: Do data users (e.g., fintechs) have to share their customers' data (upon customer request)?
  - **Yes.**
  - **No.**

C.2.6 Miscellaneous

- **PSD2**: Is this country a party to Europe's PSD2?
  - **Yes.**
  - **No.**

C.3 Miscellaneous notes

How do we define scope, sharing rules, and so on in cases where the regulators have not yet decided on an approach?

We denote these cases as "TBD" and exclude them from sections of the analysis where we split or condition on these variables.

D Classification of Fintech Startups

PitchBook groups tens of thousands of startups into the "Financial Software" subindustry and the "Fintech" vertical, but does not offer a more granular industry definition. We overcome this using PitchBook's keywords feature with categories from PitchBook's 2021Q1 fintech market map and keywords derived from those startups. PitchBook's fintech market map divides recent fintech financing rounds into eight broad categories: alternative lending, capital markets, consumer finance, digital assets, financial services IT, payments, regtech, and wealthtech. Importantly, these categories were designed around use cases and without OB in mind.

Although innovative startups are by nature often hard to classify, these categories roughly span the current fintech market. Alternative lending includes retail and commercial lending. Capital markets includes institution-focused capital market applications, including trading, data, and capital management. Consumer finance encompasses digital banking, rewards programs, and credit cards. Digital assets covers cryptocurrency and related applications. Financial services IT includes both APIs and enterprise architecture. Regtech includes risk management and compliance startups. Wealth management includes investment advisory and brokerage services.

For each of those categories, we derive a list of keywords used by the startups in that category. These keywords were assigned by PitchBook analysts covering the company, with the typical company having four keywords. Keywords range from general to specific, for example, the most frequently used keywords for companies in the regtech segment of the market map are regtech vertical, fraud detection, fraud detection platform, regulatory compliance, fintech, artificial intelligence, and risk management.

We find the relative frequency of each keyword within each category. For example, the regtech vertical keyword accounts for 3% of the keywords used by startups in the regtech category and less than 1% for all the other categories. A keyword is distinctive to a category if it is in the top 25 keywords for that category and its usage rate in that category is twice the sum of its usage rates in the other categories. Regtech vertical, fraud detection, fraud detection platform, and regulatory compliance are all distinctive keywords for the regtech category. Fintech, artificial intelligence, and risk management are not because they are commonly used across categories. The capital markets category focuses on institutional services and lacks distinctive keywords (its top keywords are financial technology, financial software, financial platform, and financial services) and so we drop it.

We assign fintech startups into categories using the distinctive keywords for each category. A startup is classified as a regtech startup if it is marked with regtech vertical, fraud detection, fraud detection platform, regulatory compliance, or other distinctive keywords for the regtech category. Fintech companies often offer a broad scope of services and can be hard to assign to a single category. Our keyword-based classification system accommodates this by allowing companies to be in multiple categories. For example, the US company SeedFi offers packages of borrowing and saving to lower-income customers placing it in both the alternative lending and consumer finance categories. The resulting categories are relatively balanced, with the largest categories (wealth management, financial IT) being about two-and-a-half times as large as the smallest category (consumer finance).

E Open Banking Use by SMEs and Their Financial Outcomes

Institutional Background for the SME Data-Sharing Policy The UK's headline open banking policy was introduced in 2017 by the UK's Competition and Markets Authority (CMA) as one of the first national OB policy initiatives. This policy is what we refer to as the UK's OB policy in our cross-country analysis and required that, by 2018, banks provide their personal and business customers with the ability to access and share their current account data on an ongoing basis with authorized third parties, such as fintechs and other banks.

However, the UK had started to set the stage for a related policy targeting SMEs two years earlier. In November 2015, concerned with the high concentration of its national banking market, the UK Government enacted the "Small Business Enterprise & Employment Act 2015" (the Act). The intention of the Act was to lower entry barriers for alternative credit providers in the SME credit segment, thereby stimulating competition. The Act's initiative related to OB-related data sharing is the Commercial Credit Data Sharing (CCDS) scheme.[51] We describe this initiative in detail next.

The CCDS regulation required nine (large) UK banks to share current account data (i.e., data from transaction accounts with a bank), as well as the up-to-date performance of loans and corporate credit cards, of all their SME customers with other lenders (including banks and non-banks) via four designated Credit Reference Agencies (CRAs).[52] When other lenders join the scheme, they commit to sharing their own SME portfolio data with the designated CRA within one year under a reciprocity rule. SMEs were affected if they had annual turnover below £25 million. As of 2017,[53] each CRA receives raw financial transaction data on a monthly basis and consolidates that data into a common format such that the data could be easily delivered to any credit provider that wants to lend to a prospective borrower. Since both SME consent to the CCDS program and a loan application are required for lenders to receive data about a specific SME from CRAs, this scheme can be regarded as credit OB. While SMEs' credit histories were widely available through credit bureaus even prior to the CCDS, the new policy supplemented these credit files with data on SME cash flows. Specifically, credit files could now be enriched with monthly snapshots of SMEs' transactions and current account balances.

The CMA Open Banking Order (OBO) and the CCDS policies can be thought of as twin OB policies: They both allowed bank customers to share their current account data (with bank customers' approval) with third parties. The main differences between the two policies are the following. First, the CCDS only applied to SMEs, while the OBO applied to both businesses and consumers. Second, the technical implementation of data sharing is different.

[51] A full summary of all the initiatives related to credit market access is available here.

[52] These are Experian, Equifax, Dun & Bradstreet, and Creditsafe.

[53] While the CCDS was due to go live in April 2016, technical issues meant that data sharing started only in 2017.

The CCDS mandated banks to share SME data with CRAs (who then shared that data onward), while the OBO mandated banks to share customer data directly with third parties via APIs. Third, the type of data being shared is slightly different. In principle, the OBO provides more detail as it is real-time transaction-level current account information, unlike the CCDS, which is a month-end summary of the current account (e.g., max/min/average balance, credit/debit turnover, and rejected payments). Fourth, while both policies offer bank customers a choice of sharing their data, the CCDS offers a bank-account-level option to opt out of sharing, while the OBO gives a more granular data-using-application-level option to opt in. Since among business customers, anecdotally SMEs benefited most from the main OB initiative, our results could be interpreted as the combined effect of both the CCDS and the OBO.

In terms of SMEs that were affected by this policy, the CCDS covers firms with up to £25 million in turnover. This threshold can be regarded as quasi-exogenous. It differs from the typical UK definition of an SME that is used in official statistics and that determines Companies House reporting standards (fewer than 250 employees and less than £36 million in turnover). It also does not match with key thresholds in the tax system (e.g., VAT is only payable on turnover over £80k).

Sample and Summary Statistics As we describe in Section 2.4, our UK firm panel data comes from Bureau Van Dijk (BvD), which offers firm financial data as well as data on all claims ("charges") against firms' assets by lenders collected from Companies House. We match charges to firm financials via the Companies House ID number. Charge holder names (lenders) are recorded as strings of text and we match those lenders. Finally, we classify lenders as banks (including foreign banks) or non-banks using the Bank of England's Historical Banking Regulatory Database (HBRD) and the Financial Conduct Authority's Financial Services Register (FSR). Unmatched lenders are classified based on their name.[54]

These matched data allow us to observe firms' lending relationships as well as their balance sheet and income statement information over time, although we do not observe the interest rates or amounts borrowed for individual loans. To construct our final sample, we restrict our attention to limited liability firms and exclude both subsidiaries and companies whose primary industry is mining (UK SIC codes 1010-1450), utilities (4011-4100), finance and insurance (6511-6720), public administration (7511-7530), or education and health (8010-8540). This allows us to focus on non-financial firms with their own balance sheets in sectors with limited government ownership. We focus on firms with 2016 turnover between £10 million and £40 million to cleanly identify the effect of the new data-sharing policy. We also condition the sample on firms reporting turnover as well as information to compute our baseline control variables: total assets, non-equity liabilities, cash holdings, and QuiScore (a measure of credit

[54] Information on the HBRD can be found here and the FCA FSR here. We classify unmatched lenders with "bank" in their name as banks and other unmatched lenders as non-banks.

risk) for at least one year in each of the pre- and post-treatment time windows. This leaves us with a sample of 39,089 observations on 6,886 unique firms.

Table E1 presents the summary statistics of the variables. All ratios are Winsorized at the 1% level. As described in the main paper, we consider a firm as having established a new lending relationship if it gets a loan from a lender with whom it did not have a relationship with in the preceding three years. The probability that an SME establishes a new lending relationship in a given year is 5.3%. Over the entire sample period, SMEs switch more often to bank lenders than to non-bank lenders. Note that banks were responsible for 72% of observed relationships in 2016 in our sample so this may simply reflect that there are more banks than non-banks in the SME lending market. An indicator variable "Prior CCDS relationship" shows that 57% of our firms have a prior loan from one of the 9 banks required to share under the data-sharing policy, and this increases to 73% when conditioning on the firm having an outstanding loan (unreported). A significant number of firms (21%) have no secured credit relationship (i.e., there is no lien against their assets in the data), and 44% have only one lender at the time of the policy introduction. The average firm in our sample is 24 years old and has a leverage ratio (defined as the ratio of total non-equity liabilities to total assets) of 59%. Additionally, the average QuiScore in our sample is 90.3 (where larger values mean lower risk), and the largest proportion of firms comes from the manufacturing, services, and retail sectors, a mix broadly in line with the aggregate economy.[55]

We next assess the comparability of treated and control firms in 2016, the year in which the policy was passed (noting that, as discussed in the main paper, actual data sharing only started in 2017). Table E2 tests for differences between the sample means of our variables for the treated and the control groups. In the first three rows, we can see that both groups have a similar rate of switching to a new lender, both overall and to bank and non-bank lenders in particular. When it comes to financial characteristics, treated firms are smaller than control firms, but that is mechanically driven by the definition of treatment, which is based on turnover size. Control firms (above £25 million in turnover) hold slightly more cash and have higher leverage ratios. These differences are statistically significant but small in economic magnitude. Our baseline specification controls for these potentially important differences. However, the two groups are almost identical in terms of credit risk, firm age, and lending relationship characteristics (number and length). Finally, the distribution of the two groups in terms of industry slightly differs, but these differences are absorbed by sector-times-year fixed effects in our main specification.

Robustness We conduct several robustness tests focusing on our main SME specification

[55] We define sectors following the 2003 UK Standard Industrial Classification (SIC) of Economic Activities. Specifically, manufacturing firms are firms with UK SIC codes between 1511 and 3720; real estate between 7011 and 7032, or between 4511 and 4550 (construction); wholesale between 5010 and 5190; retail 5211-5274; transport 6010-6340, or 6410-6412; services 7411-7490, or 5510-5552 (hotel and restaurants), or 7210-7260 (computer), 9211-9310 (other services), or 6420 (telecommunications); and everything else is other sector.

(column 3 of Table 6), which we present in Table E3. First, the CCDS policy does not specify which year the turnover eligibility threshold refers to. Although firm turnover is persistent over time and we include firm fixed effects in the regression, the cutoff may be measured with some error in 2016. Thus, in the first column of Table E3, we assign treatment based on firm turnover in 2017 instead of 2016. We observe that our key test coefficient remains positive and statistically significant, with a comparable magnitude to our baseline results.

Second, although our identification strategy helps alleviate concerns about the comparability between treated and control firms, we redo our baseline results using a matching strategy. As shown in the balance test in Table E2, the treated and control groups differ slightly from one another in terms of some observable characteristics (cash and leverage, other than obvious differences in size). Therefore, in column 2 we reestimate our test using a matching procedure. We match each firm in the control group to at most four firms in the treated group based on 2016 values of lagged total assets, leverage, cash-to-asset ratio, a credit risk indicator, sector, and location. The difference-in-differences coefficient remains similar in magnitude and statistically significant.

In the next two columns, we shrink the turnover window (£10-£40 million) we used for the firm sample to £15-£35 million (column 3) or £20-£30 million (column 4). While standard errors increase due to the reduced sample size, the point estimates remain very similar to the baseline coefficient, confirming the positive effect of the policy on the probability of SMEs adding new borrowing relationships.

Finally, the last two columns change the window used to identify the existing lending relationships a firm has in previous years. We move from the baseline window (3 years) to 1 year (column 5) or 5 years (column 6). In both cases, our coefficient of interest remains statistically significant and similar in magnitude to our baseline result.

Real effects After determining the effect of the policy on the SMEs' new borrowing relationships, we analyze whether and how firm real and financial outcomes are affected in Table E4. In the first column, we explore the loan pricing effect of the policy.[56] We find that only loans from new non-bank lenders, rather than those from bank lenders, are associated with lower total firm interest expenses.

In the next two columns of Table E4, we investigate the effects on firm liabilities. We focus on the triple interaction between treated SMEs meeting the revenue cutoff, the post-period, and firms that took a new non-bank loan, as that is where the policy has stronger effects (i.e., see Table 6). We show that short-term liabilities increase for treated firms after the policy, especially when switching to non-banks because of the CCDS scheme. The coefficient is positive and statistically significant for short-term liabilities, but it is not significant for long-term liabilities. These results suggest that new non-banks use new data to gain customers

[56] Interest-paid is not a well-reported item in BvD. For each firm, we address this issue by replacing missing values with the average interest expenses over its pre- or post-reform period.

and provide short-term credit. This could be consistent with the policy opening access to non-traditional short-term sources of finance (e.g., factoring). Finally, we see that the asset side of the firm balance sheet expands, as new credit through data sharing likely leads to an increase in total assets (column 4).

Table E1: SME Lending Data Summary Statistics

Note: This table presents summary statistics on our sample of UK SMEs with £10m–£40m in 2016 turnover. Each data point is an SME-year for the 2014–2019 period (with variables denoted as $t-1$ being lagged one year). Our data is from Bureau Van Dijk and covers secured borrowing and company financials. Any new lenders is an indicator variable equal to one if the SME received a new loan in the year in question from a lender they did not have a loan from in the preceding three years. New bank lenders and New non-bank lenders are similarly equal to one if the SME has received a new loan from a new bank or non-bank lender, respectively. Log Total assets, Cash / Total assets, and Leverage are the respective accounting variables. Low risk is a dummy variable equal to one if the firm has a BvD QuiScore above 80. For other characteristics, the 2016 value is used for all firm years. Turnover is the SME's revenue in 2016 in millions of British pounds. Treated SME is an indicator variable equal to one if the SME's 2016 turnover was below £25 million. No relationship, single relationship, and multiple relationships are indicator variables equal to one if, in 2016, the SME had no loans, loans from one lender, or loans from multiple lenders, respectively. Prior CCDS relationship equals one if the SME had an existing lending relationship in 2016 with one of the nine banks required to share SME data under the CCDS, while Prior non-CCDS relationship is an equivalent indicator variable for SMEs that did not have a prior relationship with the nine banks but had a relationship with another lender. Relationship length is the log of one plus the number of months the SME's average lending relationship has lasted, as of 2016. Firm age is the SME's age in years as of 2016. Manufacturing, Services, Real estate, Retail, Wholesale, Transportation, and Other sector are indicator variables defined based on UK SIC codes as described in Appendix E. All variables are Winsorized at the 1% level. The number of observations, mean, standard deviation, and percentiles are presented for each variable.

	Count	Mean	SD	10th pct.	25th	50th.	75th	90th
New lending relationships								
Any new lender$_{i,t}$	39,089	0.053	0.224	0	0	0	0	0
New bank lender$_{i,t}$	39,089	0.036	0.186	0	0	0	0	0
New non-bank lender$_{i,t}$	39,089	0.020	0.140	0	0	0	0	0
Accounting variables								
Log Total assets$_{i,t-1}$	39,089	9.259	0.817	8.400	8.738	9.152	9.652	10.239
Cash / Total assets$_{i,t-1}$	39,089	0.685	0.258	0.288	0.511	0.738	0.911	0.977
Leverage$_{i,t-1}$	39,089	0.587	0.272	0.242	0.391	0.584	0.762	0.904
Low risk$_{i,t-1}$	39,089	0.915	0.279	1	1	1	1	1
Firm characteristics								
Turnover$_i$	39,089	19.500	7.337	11.534	13.506	17.496	24.255	30.675
Treated SME$_i$	39,089	0.771	0.420	0	1	1	1	1
No relationships$_i$	39,089	0.214	0.410	0	0	0	0	1
Single relationship$_i$	39,089	0.436	0.496	0	0	0	1	1
Multiple relationships$_i$	39,089	0.350	0.477	0	0	0	1	1
Prior CCDS relationship$_i$	39,089	0.571	0.495	0	0	1	1	1
Prior non-CCDS relationship$_i$	39,089	0.215	0.411	0	0	0	0	1
Relationship length$_i$	39,089	8.860	9.677	0	1.499	6.086	12.630	21.847
Firm age$_i$	39,089	23.550	20,767	5	10	17	31	50
Firm sectors								
Manufacturing$_i$	39,089	0.198	0.399	0	0	0	0	1
Services$_i$	39,089	0.305	0.461	0	0	0	0	1
Real estate$_i$	39,089	0.141	0.348	0	0	0	0	1
Retail trade$_i$	39,089	0.049	0.216	0	0	0	0	0
Wholesale trade$_i$	39,089	0.205	0.404	0	0	0	0	1
Transportation$_i$	39,089	0.053	0.225	0	0	0	0	0
Other sector$_i$	39,089	0.047	0.212	0	0	0	0	0

Table E2: SME TURNOVER THRESHOLD BALANCE TESTS

Note: This table presents the results of a balance test on our sample of UK SMEs with £10m–£40m in 2016 turnover. We compare treated (2016 turnover≤£25 million) with control (2016 turnover>£25 million) SMEs as of 2016. Our data is from Bureau Van Dijk and covers secured borrowing and company financials. Any new lenders is an indicator variable equal to one if the SME received a new loan in the year in question from a lender they did not have a loan from in the preceding three years. New bank lenders and New non-bank lenders are similarly equal to one if the SME has received a new loan from a new bank or non-bank lender, respectively. Log Total assets, Cash / Total assets, and Leverage are the respective accounting variables for 2015. Low risk is a dummy variable equal to one if the firm has a BvD QuiScore above 80 for 2015. Turnover is the SME's 2016 revenue in millions of British pounds. No relationship, single relationship, and multiple relationships are indicator variables equal to one if, in 2016, the SME had no loans, loans from one lender, or loans from multiple lenders, respectively. Prior CCDS relationship equals one if the SME had an existing lending relationship in 2016 with one of the nine banks required to share SME data under the CCDS, while Prior non-CCDS relationship is an equivalent indicator variable for SMEs that did not have a prior relationship with the nine banks but had a relationship with another lender. Relationship length is the log of one plus the number of months the SME's average lending relationship has lasted, as of 2016. Firm age is the SME's age in years as of 2016. Manufacturing, Services, Real estate, Retail trade, Wholesale trade, Transportation, and Other sector are indicator variables equal to one for firms in these sectors defined based on UK SIC codes as described in Appendix E. All variables are Winsorized at the 1% level. The number of observations, mean, and standard deviation are presented for each group and cross-group differences are tested using a t-test.

	Control			Treated			
	Count	Mean	SD	Count	Mean	SD	Difference
New lending relationships							
Any new lender$_i$	1,656	0.061	0.239	5,230	0.054	0.226	-0.007
New bank lender$_i$	1,656	0.042	0.201	5,230	0.038	0.191	-0.004
New non-bank lender$_i$	1,656	0.019	0.138	5,230	0.018	0.133	-0.001
Accounting variables							
Log Total assets$_{i,t-1}$	1,656	9.598	0.818	5,230	9.085	0.766	-0.513***
Cash / Total assets$_{i,t-1}$	1,656	0.697	0.256	5,230	0.679	0.259	-0.018**
Leverage$_{i,t-1}$	1,656	0.614	0.268	5,230	0.591	0.266	-0.023***
Low risk$_{i,t-1}$	1,652	0.913	0.282	5,219	0.907	0.291	-0.007
Firm characteristics							
Turnover$_i$	1,656	30.801	4.205	5,230	16.287	4.019	-14.515***
No relationships$_i$	1,656	0.222	0.416	5,230	0.209	0.406	-0.014
Single relationship$_i$	1,656	0.432	0.495	5,230	0.441	0.497	0.009
Multiple relationships$_i$	1,656	0.346	0.476	5,230	0.350	0.477	0.004
Prior CCDS relationship$_i$	1,656	0.575	0.494	5,230	0.572	0.495	-0.003
Prior non-CCDS relationship$_i$	1,656	0.202	0.402	5,230	0.219	0.414	0.017
Relationship length$_i$	1,656	8.431	9.476	5,230	8.824	9.561	0.393
Firm age$_i$	1,656	23.367	21.148	5,230	23.113	20.302	-0.254
Firm sectors							
Manufacturing$_i$	1,656	0.171	0.377	5,230	0.203	0.403	0.032***
Services$_i$	1,656	0.332	0.471	5,230	0.306	0.461	-0.026**
Real estate$_i$	1,656	0.146	0.353	5,230	0.140	0.347	-0.006
Retail trade$_i$	1,656	0.050	0.217	5,230	0.050	0.217	0.000
Wholesale trade$_i$	1,656	0.219	0.414	5,230	0.200	0.400	-0.019*
Transportation$_i$	1,656	0.043	0.204	5,230	0.054	0.226	0.010*
Other sector$_i$	1,656	0.039	0.193	5,230	0.047	0.212	0.008

Table E3: Alternative Tests on SME Data Sharing and Lending

Note: This table shows changes in new lending relationship formation for SMEs treated by the UK Commercial Credit Data Sharing (CCDS) policy tested using alternative specifications. Each specification builds off of our baseline test that uses a difference-in-differences design on company-year data on secured loans for UK SMEs with 2016 turnover between £10 million and £40 million from the UK Companies House via Bureau van Dijk for the 2014–2019 period. An SME is classified as a Treated SME if its 2016 turnover is below the CCDS's £25 million eligibility threshold and, therefore, is potentially affected by the data-sharing policy. Post is an indicator variable equal to one after the CCDS was implemented in 2017. The dependent variable is an indicator variable equal to one if the SME takes a loan in the year in question from a lender they had not dealt with in the preceding three years. In column 1, we vary the baseline specification by defining treatment based on 2017 turnover rather than 2016 turnover. In column 2, we match each firm in the control group to at most four firms in the treated group based on 2016 values of lagged total assets, leverage, cash-to-asset ratio, a credit risk indicator, sector, and location. In columns 3 and 4, we restrict the sample to firms with 2016 turnover in the given range. In columns 5 and 6, we classify relationships as new based on a shorter or longer lookback time, respectively. Firm controls are the log of total assets, a low credit risk dummy, cash to total assets, and leverage ratio, all lagged one year. Low credit risk is defined as a QuiScore below 80, sectors are defined based on 1-digit 2003 UK SIC codes, regions are the 124 postcode areas, and relationship stage is the decile of the average relationship length the firm has with its lenders. Standard errors are clustered at the firm level and are in parentheses. *** denotes p-value < 0.01, ** denotes <0.05, and * denotes <0.1.

	2017 Turnover		Turnover Window		Relationship Window	
	for Threshold	Matched	£15-35m	£20-30m	1-year	5-year
	(1)	(2)	(3)	(4)	(5)	(6)
Treated SME × Post	0.0199***	0.0138**	0.0129**	0.0126	0.0137***	0.0148***
	(0.005)	(0.006)	(0.006)	(0.009)	(0.005)	(0.005)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm controls	Yes	Yes	Yes	Yes	Yes	Yes
Relationship stage-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Sector-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Region-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	39,089	22,638	23,587	11,052	39,089	39,089
Adjusted R^2	0.064	0.069	0.069	0.078	0.067	0.063

Table E4: Real Effects of SME Data-sharing

Note: This table shows changes in accounting measures for SMEs treated by the UK Commercial Credit Data Sharing (CCDS) policy. We use a difference-in-differences design on company-year data on secured loans for UK SMEs with 2016 turnover between £10 million and £40 million from the UK Companies House via Bureau van Dijk for the 2014–2019 period. An SME is classified as a Treated SME if its 2016 turnover is below the CCDS's £25 million eligibility threshold and, therefore, is potentially affected by the data-sharing policy. Post is an indicator variable equal to one after the CCDS was implemented in 2017. New non-bank relationship and New bank relationship are indicator variables equal to one if the SME has a new borrowing relationship in the year in question and that relationship is with a non-bank or a bank, respectively. New refers to relationships with lenders the SME did not have a loan from in the preceding three years. The dependent variable in column 1 is the ratio of interest expenses to total assets. The dependent variables in columns 2 to 4 are the logarithm of short-term (ST) liabilities (2), long-term (LT) liabilities (3), and total assets (4). Firm controls are the log of total assets, a low credit risk dummy, cash to total assets, and leverage ratio, all lagged one year. Low credit risk is defined as a QuiScore above 80, sectors are defined based on 1-digit 2003 UK SIC codes, regions are the 124 postcode areas, and relationship stage is the decile of the average relationship length the firm has with its lenders. Standard errors are clustered at the firm level and are in parentheses. *** denotes p-value < 0.01, ** denotes <0.05, and * denotes <0.1.

	Interest / Assets (1)	Log (ST liabilities) (2)	Log(LT liabilities) (3)	Log(Total assets) (4)
Treated SME × Post × New non-bank relationship	-0.0123*	0.1409**	0.2891	0.1756***
	(0.007)	(0.071)	(0.285)	(0.050)
Treated SME × Post × New bank relationship	0.0042	0.0678	0.2040	0.0074
	(0.003)	(0.049)	(0.209)	(0.031)
Treated SME × Post	0.0005	0.0255**	0.0402	0.0214***
	(0.000)	(0.010)	(0.046)	(0.005)
Firm FE	Yes	Yes	Yes	Yes
Firm controls	Yes	Yes	Yes	Yes
Pairwise interactions	Yes	Yes	Yes	Yes
Relationship stage-year FE	Yes	Yes	Yes	Yes
Sector-year FE	Yes	Yes	Yes	Yes
Region-year FE	Yes	Yes	Yes	Yes
Observations	30,530	35,101	34,710	35,115
Adjusted R^2	0.528	0.837	0.831	0.955

F Model Solution, Calibration, and Additional Comparative Statics

This appendix section details the solution method and calibration of our structural models of (1) US non-GSE residential mortgages (primarily jumbo mortgages) and (2) consumer financial advice. In both cases, we impose common structural assumptions on consumer heterogeneity. The horizontal taste shocks ϵ_{ij} follow a type-one extreme value distribution.[57] We assume the privacy taste shocks ϵ_i^S follow the same distribution and normalize their variance to 1.

F.1 Model Solution

The model solution requires solving for equilibrium interest rates, firm entry, and each consumer type's probability of sharing data. Optimal firm pricing gives a series of first-order conditions that pin down rates. Firm entry is pinned down by the zero-profit condition. Data sharing probabilities are pinned down through consumer privacy preferences. While each equilibrium object to solve for impacts all others, we find the fixed point by solving a large system of non-linear equations given by the preceding equilibrium conditions. Code to solve the model is available upon request.

F.2 Model Calibration

In both contexts, the parameters to calibrate are as follows: (1) the value of customization λ, (2) distributional parameters of marginal costs, which we assume are log-normally distributed with a mean and variance to be calibrated, (3) preferences for the incumbent relationship lender, (4) the outside option utility of not buying a product, (5) the entry cost of new firms, (6) the MC advantage (or disadvantage) for incumbents, (7) the sensitivity to interest rates and prices, and (8) the hedonic value of privacy. This yields 17 parameters: 8 in each context plus a common value of privacy. The sources for these are reported in Table F2.

We first assume that the unobserved marginal cost variance is zero for financial advice and the value of customization is zero in the mortgage context. Next, where possible, we take estimates available from the literature or industry reports for these parameters. In our case, the means of the marginal cost distributions and incumbent advantages come from Buchak et al. (2023) and industry reports,[58], while price sensitivity comes from Di Maggio et al. (2022b) and Buchak et al. (2023) for the respective applications. The other 9 parameters

[57] This is a common distributional assumption in models of discrete choice and yields highly tractable market share and pricing equations. See, e.g., in the finance context, Buchak et al. (2023).

[58] For advice marginal costs, with $mc_{bank} = 1.5\%$ and $mc_{fintech} = 0.35\%$, we use reported fees on JPMorgan's website for automated versus particularized financial advice, respectively.

(the hedonic value of privacy, the value of customization for financial advice, the unobserved marginal cost variance for mortgages, the preference for service from an incumbent for both applications, the outside option utility for both applications, and the entry cost for both applications) are calibrated through our SMM procedure.

We use 10 moments to discipline our overidentified SMM procedure. As a starting point, we assume a market has 30 pre-OB firms, which reflects the typical mortgage market in the US, Buchak et al. (2023), and that the same number of firms provide advice. We assume that markets for financial advice share common overlap. Next, assuming a pre-OB steady-state number of firms, from our summary statistics and event studies, we can infer the change in the number of firms following OB introduction. Incumbent shares come from the respective financial advice and mortgage papers. Quantities come from industry reports in the case of financial advice,[59] and facts on aggregate outstanding mortgage debt for credit underwriting. Finally, we use data from the UK FLS 2020 consumer survey to measure OB takeup rates. Table F1 shows the targeted moments and their values.

Our SMM approach searches for 9 parameters to match these 10 moments as closely as possible. While all parameters are determined simultaneously, it is instructive to detail the connection between the parameters and the key moments that discipline them. First, hedonic preference for privacy is closely tied to the OB takeup rate. As the preference for privacy becomes smaller, fewer people opt in to data sharing. Second, the value of customization in the financial advice case most directly determines the amount of new firm entry. When the advice is more valuable, more firms enter. Third, the unobserved MC variance in the credit underwriting case jointly determines both the amount of new entry post-OB as well as the overall number of firms. Other things equal, more unobserved MC variance reduces the number of firms pre-OB, and also impacts firm entry post-OB. Fourth and fifth, across both cases, the preference for the incumbent/relationship lender most directly drives the incumbent market share: as this preference increases, incumbents gain more market share. Sixth and seventh, across both applications, the outside option utility most directly drives overall quantities: as the outside option utility increases, quantities decrease. Finally, eighth and ninth, the entry cost most directly drives the number of firms operating in the market. The results of our calibration are shown in Table F2.

[59] See here.

Table F1: TARGET MODEL MOMENTS

Note: This table details the moments used for calibration of the open banking model described in Section 4. It provides the description of the target, the target moment, the moment from the calibrated model, and the data source for the target. Event study refers to the reduced-form analysis on fintech entry in this paper from Table 3, FLS refers to the Financial Conduct Authority's 2020 Financial Lives Survey and represents our results on consumer update of OB products described in Section 3.2. The relevant industry reports used for quantities in financial advice OB can be found here. Pre-OB number of firms assumes that lenders (e.g., incumbent banks) also provide financial advice.

Moment	Target	Model	Data source
Financial advice			
Δ firms	0.13	0.13	Event study
Incumbent share	0.01	0.01	Di Maggio et al. (2022b)
Quantities	0.35	0.39	Industry reports
Pre-OB number of firms	29.95	30.00	Buchak et al. (2023)
OB takeup share	0.086	0.081	FLS
Mortgage origination			
Δ firms	0.15	0.14	Event study
Incumbent share	0.12	0.13	Buchak et al. (2023)
Quantities	0.40	0.33	US outstanding mortgage debt
Pre-OB number of firms	29.89	30.00	Buchak et al. (2023)
OB takeup share	0.055	0.057	FLS

Table F2: CALIBRATED MODEL PARAMETERS

Note: This table shows the calibrated parameters for the open banking model described in Section 4. It provides the parameter, its description, the calibrated value, and how it was calibrated. We cite the relevant industry reports in footnotes. SMM refers to the simulated method of moments, targeting those moments described in Table F1. The relevant industry reports used for average marginal cost in financial advice OB can be found here. The relevant industry reports used for Incumbent MC advantage in financial advice OB can be found here.

Parameter	Description	Calibrated value	Source
Common parameters			
ϕ	Value of privacy	0.75	SMM
Financial advice			
λ^a	Value of customization	0.98	SMM
σ_a^{MC}	Unobserved MC variance	0.00	Assumption
θ^a	Preference for incumbent	0.37	SMM
u_0^a	Outside option utility	2.39	SMM
c^a	Entry cost	0.91	SMM
μ^{MC}	Average marginal cost	0.35	Industry reports
μ_r^{MC}	Incumbent MC advantage	-1.15	Industry reports
α^u	Rate sensitivity	1.38	Di Maggio et al. (2022b)
Mortgage origination			
λ^u	Value of customization	0.00	Assumption
σ_u^{MC}	Unobserved MC variance	0.73	SMM
θ^u	Preference for incumbent	2.44	SMM
u_0^u	Outside option utility	1.17	SMM
c^u	Entry cost	0.60	SMM
μ^{MC}	Average marginal cost	1.62	Buchak et al. (2023)
μ_r^{MC}	Incumbent MC advantage	0.00	Buchak et al. (2023)
α^u	Rate sensitivity	1.14	Buchak et al. (2023)

F.3 Comparative Statics on the Informativeness of Data

An important model comparative static is the informativeness of borrower information. In the case of credit, the main model parameter governing the value of information is the variance of the unobserved marginal cost. This corresponds to the informativeness of borrower information because when the unobserved variance is high, revealing the information eliminates more variance; when the informativeness of borrower information is low, revealing that information does not. One interpretation of this is that a low variation corresponds to an information environment where a credit registry has already revealed most of the useful information on borrower creditworthiness.

In the case of financial advice, the main model parameter governing the value of information is the value of customization, λ. This corresponds to the value of information because it governs the size of the effective product improvement for service providers with access to customer data. Similar to the above, one interpretation of a high λ is a regime where advice customization has a large impact on customer outcomes and OB data is very useful in customization.

Figure F1 Panels (a) and (b) show how several model outcomes vary as a percent change relative to the status quo relationship regime. Panel (a) corresponds to altering the variance of the unobserved marginal cost for credit, and Panel (b) corresponds to altering the value of customization for financial advice. In both panels, the x-axis corresponds to the scale applied to the baseline, i.e., 0.85 in Panel (a) means the variance of unobserved marginal cost is set to 0.85 times its calibrated value. Moving left-to-right along the x-axis corresponds to information being more revealing. In each figure, the black line shows the fraction of the population opting into data sharing, the red line shows the percentage increase in the number of operating firms, the blue line shows the percentage change in relationship bank profits, and the green line shows the percentage change in service provision. The comparison is done at each point on the x-axis between the status quo world with that parameter value and the OB world with that parameter. For example, at an x-value of 0.85, we are looking at how entry changes in going from a world with 0.85 of the MC variance with no OB to a world with 0.85 of the MC variance with OB.

Beginning with Panel (a), we find that customer take-up into data sharing is *decreasing* with the variance of borrower marginal cost. While this is initially counterintuitive, it is driven by the fact that increased MC dispersion reduces, in an absolute sense, the number of outside lenders issuing loans because they are more severely affected by adverse selection. This is still true—and self-reinforcing—in the OB regime because of the incomplete take-up of data sharing. When there are fewer outside lenders, it is less beneficial for borrowers to share their information with them, and thus fewer borrowers share their data.

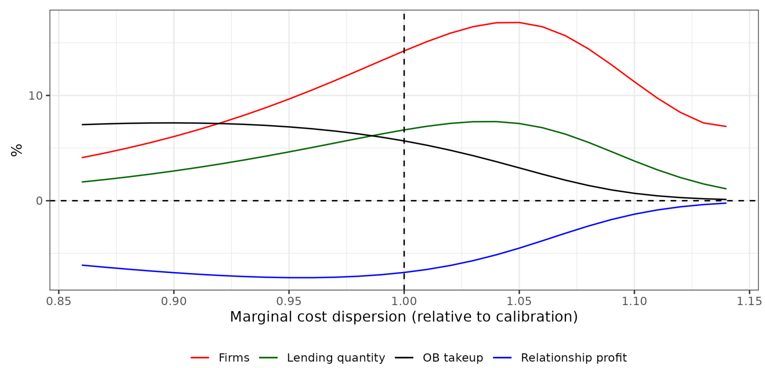
We find an inverse-U-shaped pattern in the percent increase in outside lenders, shown in red. For low marginal cost dispersion, while more borrowers opt in to OB, the informational content of the data is lower, and hence there is relatively modest entry. As marginal cost dispersion increases, there are two opposing forces: on one hand, data becomes more revealing, and so OB has a larger impact on firm entry. On the other hand, data sharing take-up declines, and so the number of potential OB-enabled consumers declines. The former force dominates at first but is ultimately swamped by the latter. The increase in lending quantities tracks new firm entry, and relationship bank profit has the inverse effect, following the same logic.

The effect of borrower information is more straight-forward in the case of financial advice, and the relationships are all monotonic, as shown in Panel (b). As the value of customized advice increases from left to right, data sharing takeup increases. This occurs due to the direct effect of it being more beneficial to share data when it allows firms to offer better products, and additionally due to a compounding effect of new firm entry. As data becomes more useful in customizing advice, more outsider providers enter because they can offer (and charge for) an increasingly superior product. This forms a beneficial feedback loop where more firms offering services draw more customers into data sharing, and more customers sharing their data draws more firms in to serve them. The quantity of advice provided increases following the same logic, while new entrants steal an increasing share of customers from the relationship advisor, reducing their profits.

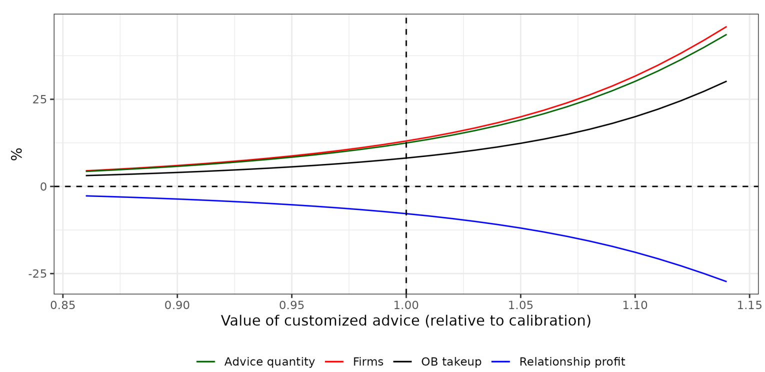
These counterfactuals illustrate the stark economic differences arising from how data is used across OB applications. Incentives around data used for product improvements are straightforward. More revealing data means firms with access provide better products, thus inducing more data sharing and firm entry in a virtuous cycle. In the credit application, these two forces sometimes operate in opposition. When data is more useful, uninformed outsiders face greater adverse selection and the quantity of outsider firms shrinks. With few outsiders and optional data sharing, the benefits of sharing data are small, so few consumers opt in, meaning new entry opportunities for firms are limited. The mediator in this vicious cycle is customers' endogenous opt-in choice, and so policymakers designing OB policies need to consider consumer incentives to share data, particularly when the data may be used against them.

Figure F1: Comparative Statics on Data Informativeness

Note: This figure shows comparative statics on the informativeness of consumer data for the open banking (OB) model described in Section 4. The plots show, compared to the no-OB status quo, the number of service providers (red), relationship bank profit (blue), service quantity (green), and OB takeup rate (black). Panel (a) shows results for the mortgage underwriting case and varies the dispersion of unobserved marginal cost. The x-axis shows the marginal dispersion relative to the baseline calibration. Panel (b) shows the results for the financial advice case and varies the value of customized advice, with the x-axis showing the value of customization relative to the baseline calibration.



(a) Credit



(b) Advice