Journal of Financial Economics

Customer data access and fintech entry: Early evidence from open banking[☆]

Tania Babina [a,b,c], Saleem Bahaj [d,e,c], Greg Buchak [f], Filippo De Marco [g,c], Angus Foulis [e], Will Gornall [h,*], Francesco Mazzola [i], Tong Yu [j]

[a] Smith School of Business, University of Maryland, 7699 Mowatt Lane, College Park, MD 20740, United States of America
[b] NBER, 1050 Massachusetts Avenue, Cambridge, 02138-5398, MA, United States of America
[c] CEPR, 2nd Floor, 33 Great Sutton Street, London, EC1V 0DX, United Kingdom
[d] University College London, Gower Street, London, WC1E 6BT, United Kingdom
[e] Bank of England, Threadneedle Street, London, EC2R 8AH, United Kingdom
[f] Stanford Graduate School of Business, Stanford University, 655 Knight Way, Stanford, CA 94305, United States of America
[g] Bocconi University, Via Roentgen, 1, Milan, 20136, Italy
[h] Sauder School of Business, University of British Columbia, 2053 Main Mall, Vancouver, BC V6T 1Z2, British Columbia, Canada
[i] ESCP Business School, Turin campus, Via Andrea Doria, 31/a, Turin, 10123, Italy
[j] Imperial College London, Exhibition Road, London, SW7 2AZ, United Kingdom

ARTICLE INFO

Keywords:
Financial regulation
Open banking
Financial innovation
Data access
Data rights
Big data
Artificial intelligence

ABSTRACT

Open banking (OB) empowers bank customers to share their financial transaction data with fintechs and other banks. New cross-country data shows 49 countries adopted OB policies, privacy preferences predict policy adoption, and adoption spurs fintech entry. UK microdata shows that OB enables: (i) consumers to access both financial advice and credit; (ii) SMEs to establish new 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.

The increasing ease with which data is collected, stored, and analyzed has made it a critical input in economic decision-making. Data's growing 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 financial services, 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

☆ Philipp Schnabl was the editor for this article. 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 AFA; Barcelona Summer Forum; Bank of Canada workshop; Bank of Sweden; Baruch College; CKGSB; CMU; Columbia; HBS; 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; IESE (Barcelona); Imperial College London; Maastricht University; MIT; NBER SI Corporate Finance; NBER Economics of Privacy Conference; NFA; NY Fed; NY Fed Fintech Conference; NYU; NYU WAPFIN (Women Assistant Professors in Finance); OCC Conference; OSU; Oxford University; Rome Junior Finance Conference; Queens College School of Business; Stanford; Stockholm School of Economics; Toulouse School of Economics; Wharton; UBC; UBC Winter Finance Conference; Universidad Carlos III de Madrid; University of Maryland; University of Pittsburgh; University of Warwick; University of Washington; UPF; USC; 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.

Received 22 April 2022; Received in revised form 14 September 2024; Accepted 18 September 2024
Available online 5 April 2025

transactions allow firms to learn about a customer's needs and offer tailored financial advice or other products. A small business's transaction data could inform lenders about its health and help a fintech deliver financial management services.

Historically, a customer's financial data has been under her bank's exclusive control, giving that bank an advantage in pricing and customizing financial services.[1] However, banks' exclusive access to this data is being upended by a movement known as open banking (OB). OB empowers bank customers to share their financial transaction data with other financial service providers. For example, OB allows a bank customer 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 are promoting or even mandating it. As of October 2021, regulators in 80 countries have taken at least some steps to encourage the adoption of OB. 49 of the 80 have already adopted their key OB policies. Through OB, policymakers aim to boost 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 and innovate.

In this paper, we explore the causes and consequences of government policies to promote OB. In doing so, we make four 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 drivers. Second, we use microdata from the UK — an early adopter of OB — to provide evidence on how OB policies impact consumers and small- and medium-sized enterprises (SMEs). Third, we examine the global impact of OB policies on financial innovation using our country-level OB policy data. Finally, we provide a quantitative modeling framework for customer data sharing, which measures the overall and distributional welfare effects of OB.

We begin by assembling a comprehensive dataset on government policies to promote OB for the world's 168 largest countries—representing more than 99% of world GDP. We uncover vast heterogeneity in OB policy choices. For example, countries in the European Union (EU) have adopted OB regimes that require sharing only data from transaction accounts (e.g., checking accounts or credit cards). In contrast, countries in other regions typically mandate sharing data from a broader set of financial products (e.g., mortgages or insurance). We also examine the drivers of OB policies and show that consumer trust in sharing data with fintechs predicts OB policy adoption: Intuitively, consumer willingness to share their data increases the potential benefit of these policies. Other country characteristics are less predictive, including economic and financial development, the levels of innovation, or the quality of local institutions.

The prevalence of OB policies motivates further study of the economic effects of these policies. We first focus on the UK where granular microdata offers direct evidence on the adoption and economic impact of OB for both consumers and SMEs. For consumers, we use a survey by the UK Financial Conduct Authority to document two main distinct reasons consumers share data: Financial planning and management (which we term advice OB) and borrowing (credit OB). There is little overlap between users of advice and credit OB, and consumers are more likely to use both OB products 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: Advice OB is associated with greater financial knowledge and credit OB is associated with greater credit access.

For SMEs, UK panel data allows us to estimate the causal impact of OB on borrowers, measure whether banks or non-banks provide new loans, and examine OB's financial inclusion implications. We exploit the fact that the commercial OB-related policy applied only to SMEs with annual sales 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 using a difference-in-differences and event-study designs. We document that SMEs eligible to share data form more new lending relationships with non-bank lenders (e.g., fintechs). 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.

We next provide global evidence on the effect of OB on fintech entry, which regulators regard as a key mechanism through which OB can improve innovation and competition. 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 and event-study designs, we show that VC investment in fintechs surges 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 apps, credit, payments, regtech), consistent with our UK survey evidence that OB data has a wide range of use cases.

While our empirical results offer valuable descriptive and causal evidence regarding OB use, they fall short of addressing several key economic and policy relevant questions related to OB. First, they are largely silent on the mechanisms by which access to OB data increases entry across the two distinct use cases — financial advice and credit — highlighted by our UK consumer results. Second, our differences-in-differences tests have little to say on equilibrium effects, welfare, 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.

We tackle these questions directly using a quantitative model of data usage. This model incorporates consumer data use into a standard industrial organization (IO) model of consumer choice with heterogeneous consumers. In our model, data about a bank's customer — interpreted as either an individual or a business — reveals her preferences (allowing the creation of better products for advice OB) and costliness to serve (allowing to learn default risk for screening in 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 the two use cases using our reduced-form results and pre-OB estimates from the literature. In our credit use case, we calibrate to mortgage products, where data is informative about consumer risk. In our financial advice use case, we calibrate to investment advice, where data is informative about particularized customer needs. 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

[1] As a motivating example, Online Appendix 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.

products, and enabling customers to share their data leads to better-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 goal 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 tailored products. In contrast, the credit OB use case can have negative distributional consequences because OB increases entry precisely by enabling entrants to better exclude unprofitable (higher risk) customers. Users who share unfavorable data lose directly. Users who opt out of sharing are inferred to have strategically hidden unfavorable data, even when opting out due to 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 higher risk. Our quantitative model allows us to weigh this tradeoff, a particularly important question for policymakers concerned with the distributional consequences of OB.

Our model shows that societal preferences for privacy (i.e., unwillingness to share data) not only drive the impact of OB (consistent with our cross-country results), but 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 fintechs enter. However, societal preferences for privacy have a silver lining because they cause customers to opt out of data sharing for privacy reasons, which means that opting out sends only a weak signal about one's riskiness. 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 in lending. The negative inference lenders draw against opt-outs is more than offset by the benefits that these customers derive from increased entry and competition. Consequently, incorporating privacy preferences and the implications of different use cases is an important part of an OB implementation, highlighting a crucial distinction between credit registries and OB data sharing.

To summarize, we document that government policies to promote OB are prevalent: About half of countries have 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 industry, society, and policymakers 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 customers' data and their willingness to share data.

Our paper proceeds as follows. 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 and potentially improve bank customer outcomes.

Second, we engage with the fundamental question, originating with Diamond and Dybvig (1983) and Diamond (1984), over what makes banks special. While fintechs and other non-depository institutions have gained significant market share in transaction-oriented functions like origination and servicing (e.g., Gopal and Schnabl, 2022; Buchak et al., 2024b), 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), suggesting there are significant barriers that limit the growth of new single-product competitors. 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 informational advantages. This explanation dates to Petersen and Rajan (1994), Petersen and Rajan (1995), and, more recently, Granja et al. (2022) and Blickle et al. (2023). Recent empirical work by Ghosh et al. (2024) shows a direct effect of transaction data on screening quality for commercial loans. Berg et al. (2020) and Di Maggio et al. (2022b) show the value of alternative data more generally. Banks' informational advantages are challenged with OB, paving the way for an analysis of how important these advantages are.

Third, we add to the nascent literature on the economic effects of data ownership and access. Theoretical work typically views data as either an input to production or a way to address information asymmetries. Taking the production-input view, Jones and Tonetti (2020) show that a firm may suboptimally hoard product-improving data to prevent entry, motivating the reallocation of data property rights from firms to consumers. Farboodi et al. (2019) model data as valuable for forecasting and suggest that large firms generate more data and benefit from it. Emphasizing the information economics view, 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) emphasize that the economic impact of data sharing is influenced by the extent to which lenders engage in liquidity transformation. He et al. (2024) study how the hardening of soft information, which can result from policies like open banking, affects competition in lending markets. Empirically, Babina et al. (2024) show that larger firms — that naturally generate more data — benefit more from their investments in Artificial Intelligence (AI) and AI use is associated with increased industrial concentration, suggesting that a current status quo (where firms own their customers' data) can stifle entry and weaken competition.

We build on this largely theoretical literature in several ways. OB policies weaken bank data monopolies and give consumers the power to share their data, offering a valuable opportunity to study the economic effects of change in data ownership policies. We document that mandating incumbent firms to share their customers' data spurs new firm entry, but also generates complex competitive interactions that depend on how the consumer data is used (i.e., customer screening in credit products vs. product innovation in financial advice products) with unclear distributional consequences.[2] To the best of our knowledge, we provide the first empirical study on the impact of government policies that open access to rich customer-level financial transaction data. Beyond that, we provide a quantitative framework for studying

[2] More generally, the risks of ever broader data use by firms are not well understood. For example, the use of data in AI applications leads firms to become more systematically risky with unclear impacts on their customers (Babina et al., 2023b).

the 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., 2024a), we connect data to information about consumer heterogeneity around marginal costs and desired product customization. Through these channels, we synthesize both the input-to-production and information economics views of data and highlight their quantitative importance across particular applications. In contrast to, e.g., He et al. (2023) and Parlour et al. (2022), our model emphasizes entry and innovation, which are key policy goals of OB. Moreover, our analysis complements this literature by highlighting the importance of consumer preferences over data privacy (e.g., Acquisti et al. (2016), Tang (2019), Bian et al. (2021), and Chen et al. (2024)) by explicitly incorporating privacy preferences into our structural model.

While conceptually related to data sharing via 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 nonbanks. The richer data that OB covers lends itself to uses beyond screening; however, this very richness creates greater privacy concerns than a standard credit file.[3] We show these aspects of OB are important in driving its effects. Thus, our paper provides evidence of the effects of adopting data-sharing policies more generally, complementing the literature on credit registries.

Fourth, our structural model allows us to 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. (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 interplay between technology and regulation and how they interact with the comparative advantages of depository and non-depository institutions.[4] 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; Philippon, 2019).

Finally, our paper is connected to the literature on the drivers of innovation and entrepreneurship. We document the importance of data access for innovation: 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; Denes et al., 2023; Bai et al., 2022; Babina et al., 2023c).[5]

2. Institutional background and descriptive analysis

This section describes the institutional background of OB policies, details our data collection process, shows the global importance of OB policies, and examines their drivers.

2.1. Institutional background on open banking

OB describes a broad trend where, upon customer request, financial intermediaries (e.g., banks) share — willingly or by regulatory fiat — access to their customers' data with other financial service providers (e.g., fintechs or banks). 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 a bank's individual customers and sometimes also apply to business customers.

While the specifics of government OB efforts vary, the UK's Open Banking Initiative is 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."*[6] Here, third parties refer to both fintechs and other banks. Additionally, banks were required to allow customers to authorize third parties to make payments from their accounts—a practice called payment initiation. OB differs significantly from the UK's existing private sector credit bureaus: It covers richer data (in particular, information on transaction accounts), it gives banks' customers control over their data, it is free to the requester, and banks are forced to participate. These are common features of OB policies around the world and mean that OB goes beyond traditional credit bureaus.

By opening bank data, regulators aim to create an environment where financial intermediaries — both incumbents and 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. With OB, fintechs can access, aggregate, and analyze these separate accounts to provide customized financial advice. She may find it helpful to monitor these accounts in a single place to understand her spending habits and get advice on budgeting, savings, and credit management. Another use case of OB is credit, where potential lenders can access the otherwise private information that a consumer's home bank has about her. For example, with customer permission, a fintech lender could use the data on a bank's customer to query her bank account transactions to help price her a loan. Beyond financial advice and credit, many other use cases have emerged, including identity verification, payments, and insurance. These financial products powered by OB data are typically highly customized to each customer with AI models—consistent with AI, more generally, being used by firms for product innovation so far (Babina et al., 2024).

Data sharing 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 the data provider (e.g., bank) and are a set of standardized, programmatic commands that allow data users (e.g., fintechs) 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

[3] 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.

[4] 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).

[5] 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), VC availability (Kaplan and Lerner, 2010), R&D subsidies (Babina and Howell, 2024), and competition policies (Phillips and Zhdanov, 2017; Babina et al., 2023a).)

[6] Page 11 of "Open Banking, Preparing for Lift off" document. See the 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 3.2 and Online Appendix E.

write access (e.g., payment initiation). In Online Appendix B, we show that in countries that implement OB policies, banks are indeed more likely to provide APIs for customer data sharing.

Even without government OB policies, fintechs have gained access to customer bank data through financial aggregators such as Yodlee and Plaid that collect data via a combination of bilateral agreements and "screen scraping" (web scraping using user-provided passwords). In practice, although these market-based solutions are improving, they are expensive for fintechs and offer incomplete coverage.[7] Incumbent banks' reluctance to voluntarily offer widespread data sharing suggests that they lose monopoly rents — an intuition crystallized in our model — and that there are significant contracting frictions that prevent them from capturing surpluses. For example, bank customer stickiness or a lack of customer sophistication prevents banks from extracting the value of data sharing from customers, and coordination problems around large numbers of (merely hypothetical) fintech entrants prevents a Coasian solution. Importantly, because banks are data monopolists, standard economics predicts that a straightforward arrangement where the bank sells information access to fintechs will lead to markups and an inefficiently low quantity of data access. Thus, government involvement in data sharing appears to be an important force in its widespread adoption.

2.2. Data collection methodology for open banking around the world

We create a comprehensive, hand-collected database of OB government policies (or the lack thereof) for the largest 168 countries (covering over 99% of global GDP). This section describes our methodology; Online Appendix C 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.[8] 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. 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.

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. Third, we reconcile our results against a database of OB regulations maintained by Platformable,[9] 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.[10] As of October 2021, 80 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 statements. 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. There 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, EU Directive 2015/236). PSD2 obligated participating countries to implement its provisions in their respective banking regulations. In the country-level summary statistics in this section, we keep the participating countries separate. For our analyses in Sections 2.4 and 3.3, we weight all countries covered by PSD2 as a single pooled observation.

Implementation status and key dates of government-led 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 towards 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.[11]

Panel (a) of Fig. 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,[12] Brazil is in the process of implementing OB (see here), and the UK has fully implemented its Open Banking Initiative and is considering a follow-on "open finance" regulation.[13] Fig. 1 Panel (b) shows the passage year of countries' major OB government policies.

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 or providing technical standards or 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"). In

[7] For example, financial aggregator pocketsmith.com reports a median connection success rate of 44% for Yodlee among the Canadian banks it claims to cover as of mid-2023. In the US, Fidelity and PNC dropped support for Plaid in late 2023 (see here).

[8] The IMF data is from here. The VC data is from PitchBook and is described later.

[9] Platformable's data is described here.

[10] 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.

[11] 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 building new infrastructure for OB).

[12] 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.

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

Table 1
Open banking government policies summary statistics.

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)

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.

addition to requiring incumbent banks to share data, some OB regimes also require reciprocal sharing by new entrants (e.g., fintechs): Our data shows that only 18% of regimes have data sharing reciprocity (variable "Data reciprocity"). 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: Fig. 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 in what financial products are covered. By definition, all OB regimes cover at least 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. Fewer then 34% of countries 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 data sharing can, in theory, be used both to read data (e.g., pull customer account information) and to write data (e.g., initiate transactions). 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 measures the comprehensiveness of OB policies. The index averages the four key OB policy dimensions discussed above: 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 technical standards for data sharing. This index ranges from 0 (all four dimensions are not yet mandated) to 1 (yes on all four dimensions).

2.4. Drivers of open banking government policies

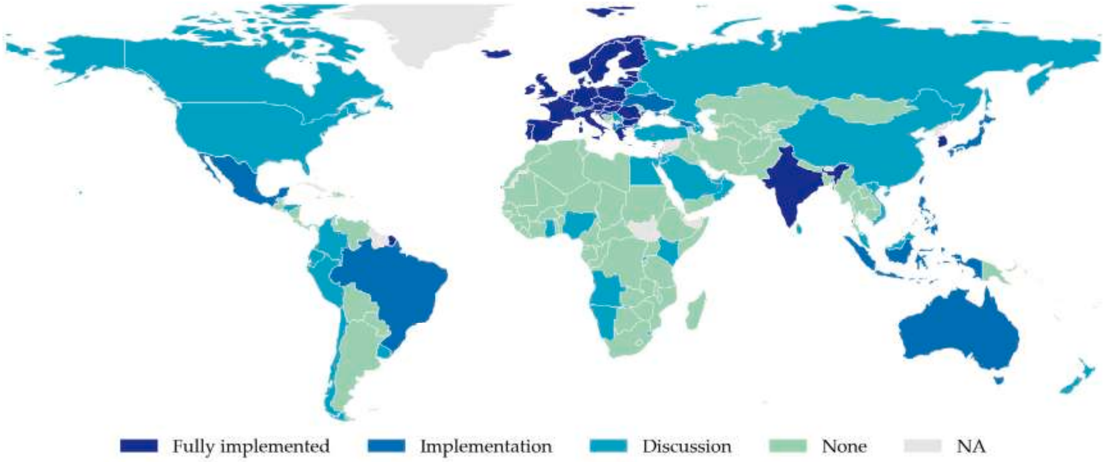
We next examine what factors drive countries to adopt OB policies around the world. In the spirit of Kroszner and Strahan (1999) or (Cornelli et al., 2020), we examine what predicts OB policy adoption using a broad set of country characteristics as summarized in Panel (a) of Table A1 in Online Appendix. We start with basic country-level data, including per capita GDP in thousands of US dollars and population in millions from the World Bank. Given the importance of consumer willingness to share data for OB adoption, we use the measure of consumer trust in sharing data with fintechs from (Chen et al., 2023). 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 (which captures the market power of banks). We take the percentage of banks that are foreign owned from (Claessens and Van Horen, 2013). To capture the quality of institutions, we use the Rule of Law Index from the Cato Institute. Finally, to measure innovation, we add data on VC deals from PitchBook, widely acknowledged as one of the best VC data sources for more recent years.[14]

Using our cross-country data, we then test the association between the time of OB policy implementation and these country characteristics using a Cox proportional hazards model:

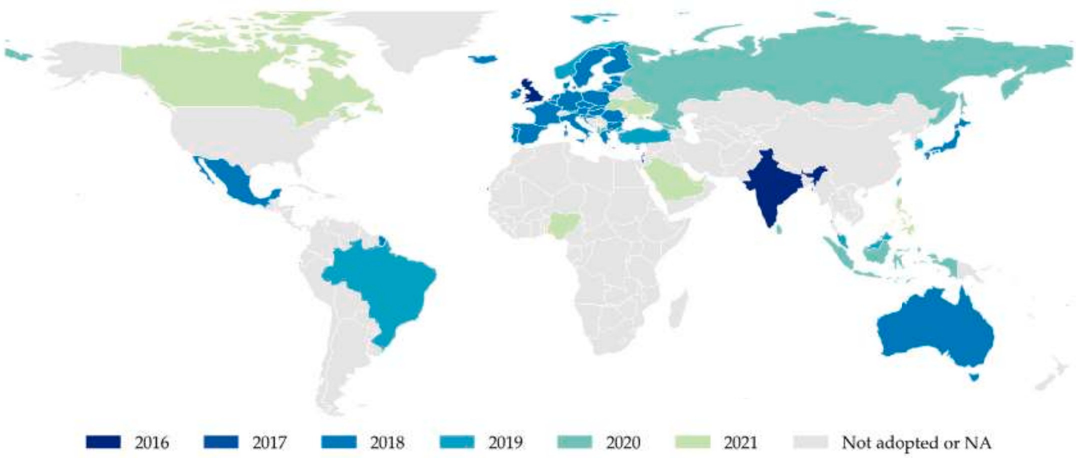
$$h_i(t) = h_0(t)\exp(X_i'\beta + Region_r) \quad (1)$$

where $h_i(t)$ represents the hazard function for the occurrence of the OB outcome (implementation of OB policies through 2021) in year t in country i. This hazard function can be interpreted as the risk of

[14] The data on trust in sharing data with fintechs is based on the survey underlying the EY Global Fintech Adoption Index. Specifically, it measures trust as the portion of survey respondents in each country who "agree" or "strongly agree" that they are comfortable with their main bank to securely share their financial data with fintechs. 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.



(a) Government open banking policy implementation status



(b) Timeline of open banking adoption

Fig. 1. Government-led open banking regimes around the world. 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.

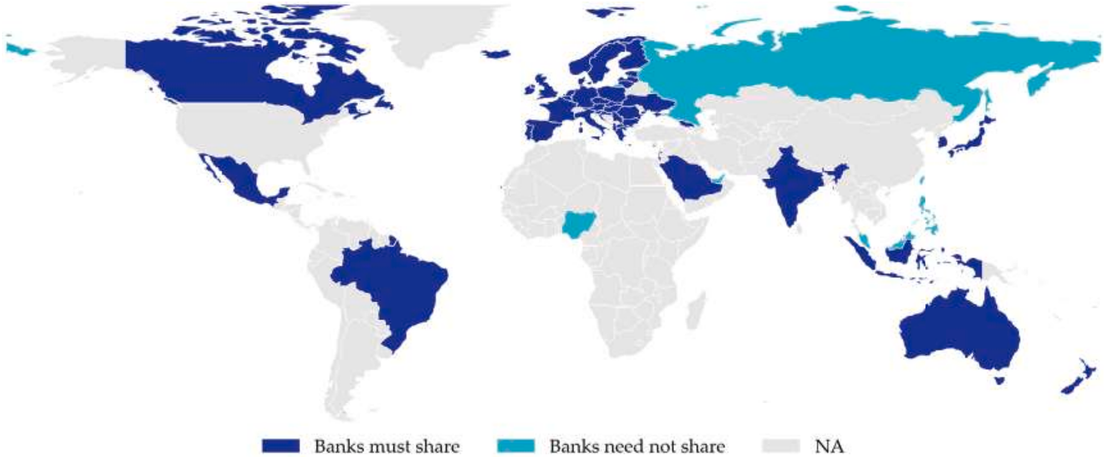
the event happening at time t given it has not yet occurred. X_i' is a vector of country-level characteristics. $Region_r$ are region fixed effects. Data availability causes the number of observations to fluctuate across specifications.

We supplement this regression with a cross-country regression on OB 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) based on key OB policy dimensions. These regressions take the following form, where OB_i denotes the two measures of OB policy for country i as of 2021:

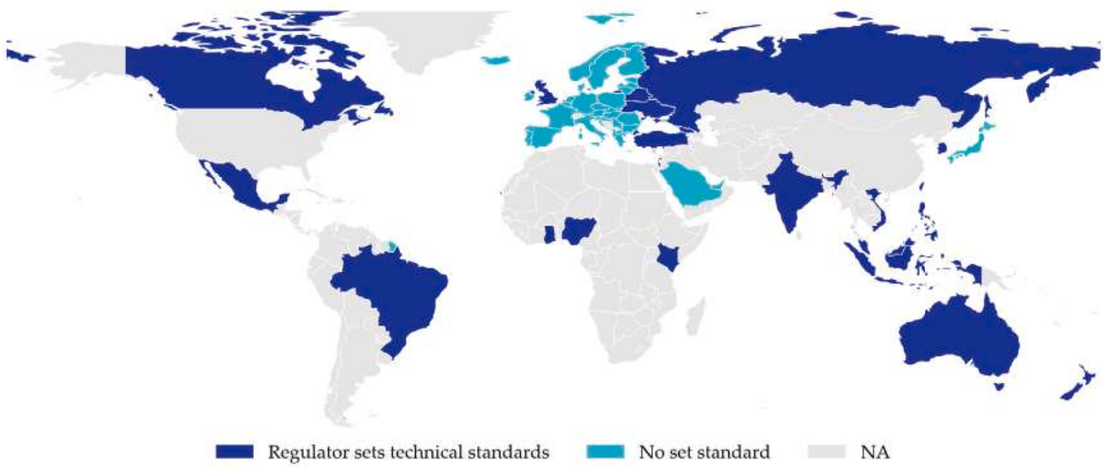
$$OB_i = X_i'\beta + Region_r + \epsilon_i \tag{2}$$

Table 2 presents the determinants of OB adoption speed (columns 1–5), implementation status (columns 6–7), and policy strength (columns 8–9). Columns 1 to 5 use Eq. (1). Since low overall levels of economic development could be associated with the introduction of OB policies in all columns we control for both GDP per capita (and its square) and log population. However, neither a country's GDP nor its population robustly predicts the introduction of OB government policies. 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 (0.15) increase in trust is associated with a significantly higher rate of OB policy adoption, with the hazard (or event occurrence) rate nearly quadrupling.

Other country characteristics are only weakly associated with OB. Column 2 shows that measures of financial development do not predict 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 are not predictive of adoption. In column 4, we find weak and statistically insignificant associations between the adoption of OB policies and both the Rule Law variable and the fraction of foreign-owned banks. 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 largely unchanged and remains statistically significant, while the coefficient on non-fintech VC becomes statistically insignificant.



(a) Banks must share data upon customer request



(b) Regulator sets technical standards

Fig. 2. Open banking government policy dimensions. 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.

Columns 6 to 9 present estimates of Eq. (2). Trust in fintechs is again associated with the OB implementation, with a one standard deviation increase in trust being associated with about two steps of increase on our seven-step scale (column 6). The coefficient is unchanged when we control for non-fintech VC deals (column 7). Columns 8 and 9 show trust in fintechs is associated with our OB Strength Index with borderline significance. Overall, consumer trust in sharing data with 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.[15]

3. The economic effects of open banking

Next, we examine the economic effects of OB. We first focus on the UK (one of the first countries to adopt OB policies): We show that OB enables consumers to access both financial advice and credit (Section 3.1) and leads SMEs to form new lending relationships (Section 3.2). We then examine the global impact of OB policies on financial innovation using our country-level OB policy data (Section 3.3).

3.1. Evidence from UK microdata on consumers

We analyze the use of OB by UK consumers and their financial outcomes using data from the Financial Lives Survey (FLS). The FLS is a 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' demographics, attitudes towards managing their money, financial product usage, and experiences engaging with financial services firms. We use the February 2020 survey which covers usage of OB products for the first time.[16] Online Appendix Table A2 provides summary statistics.

[15] 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, this concern is unlikely to be of first-order importance.

[16] See the survey questionnaire here.

Table 2
Drivers of open banking government policies.

	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

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 open banking (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. 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. 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 policy 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 i) Africa, Middle East & North Africa; ii) Europe & Central Asia; iii) Latin America & the Caribbean; iv) North America; v) 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. The regressions are cross-sectional, where each country in the sample corresponds to a single data point. European Union member states are weighted to count as a single country for estimates and standard errors. Robust standard errors are in parentheses. *** denotes p-value < 0.01, ** denotes <0.05, and * denotes <0.1.

This data has three advantages. First, the survey asks consumers whether they use financial services based on OB, providing novel evidence on uptake. Second, its demographic information allows us to examine what type of consumers adopt OB. Finally, the survey covers consumer financial outcomes so we can examine their association with OB use.

We begin with consumers' uptake of OB. The survey asked 4,310 consumers who report having a day-to-day bank account (necessary to use OB) about their use of OB products. The survey splits 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: Apps that aggregate data from several financial accounts or help users with savings. Credit OB is applications that offer credit, either directly (e.g., lending) or indirectly (e.g., credit ratings or interest rate comparison).[17]

Among consumers who report knowing whether they use these types of services, Online Appendix Table A2 shows that 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 findings in subsequent sections of an OB-led increase in VC fintech investment across a wide range of financial product categories. Surprisingly, we find little association between these two types of OB services. Only 13% of advice OB users also use credit OB, while 20% of credit OB users use advice OB. Overall, the total rate of (unique) OB users is 13%.

[17] 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 interest rate 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 consumers might be 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.

Table 3
Consumers' open banking usage and their financial knowledge and credit access.

		Credit product ownership			
	Financial knowledge (1)	Credit card (2)	Personal loan (3)	Student loan (4)	Pawnbroking loan (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

This table shows the association between financial knowledge, credit product usage, and open banking (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, where each respondent to the survey corresponds to a single data point. 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 OB 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 OB 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 bill 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 younger (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 (reported in parentheses). *** denotes p-value < 0.01, ** denotes <0.05, and * denotes <0.1.

Online Appendix Table A3 shows the cross-sectional association between the use of each type of OB and consumer characteristics. We regress whether a consumer use advice OB (column 1) and credit OB (column 2) on consumer characteristics and location fixed effects. People who have concerns about sharing their OB data are less likely to use both types of OB (variable "Unwillingness to share data"). Employed people are more likely to share their data with both types of OB products, in line with standard models of voluntary disclosure (e.g., Grossman, 1981) as employment status is information absent from credit reports but shareable via OB. People who miss bill payments are also more likely to share, suggesting more demand for both advice and credit products for this financially vulnerable group. Young people are more likely to use advice products (but not credit), while college-educated (variable "Higher education") tend to use credit (but not advice). Other consumer characteristics, such as gender, being white, being married, or risk aversion are not correlated with either OB product uptake.

We next test whether OB usage is associated with consumer financial outcomes. Table 3 relates OB usage to financial knowledge (column 1) and credit product usage (columns 2 to 5). We control for all the consumer characteristics from Online Appendix Table A3 and location fixed effects. 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 that financially savvy people could be more eager to use advice OB. However, another 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 (see here). 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.

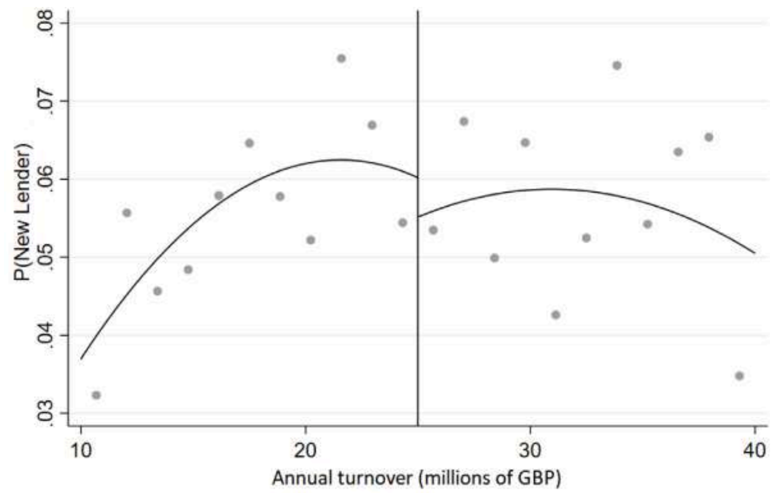
In columns 2 to 5, we look at the link between credit OB use and credit access.[18] We exploit the fact that OB data might be more used for some products than others. Credit cards and personal loans are unsecured credit products likely to benefit from data informative about creditworthiness. Credit OB users are more than 10% more likely to get both credit cards (column 2) and personal loans (column 3).[19] We use student loans and pawnbroking as placebo products that are ex-ante unlikely to benefit from OB. Due to UK regulations, student loan underwriting does not depend on consumer creditworthiness (see here), while pawnbroking is backed by physical collateral and low-tech. As expected, neither student loans (column 4) nor pawnbroking (column 5) are associated with credit OB use.[20] 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, the data on UK consumers shows that OB enables consumers to access financial advice and credit products, and is associated with better consumer financial outcomes.

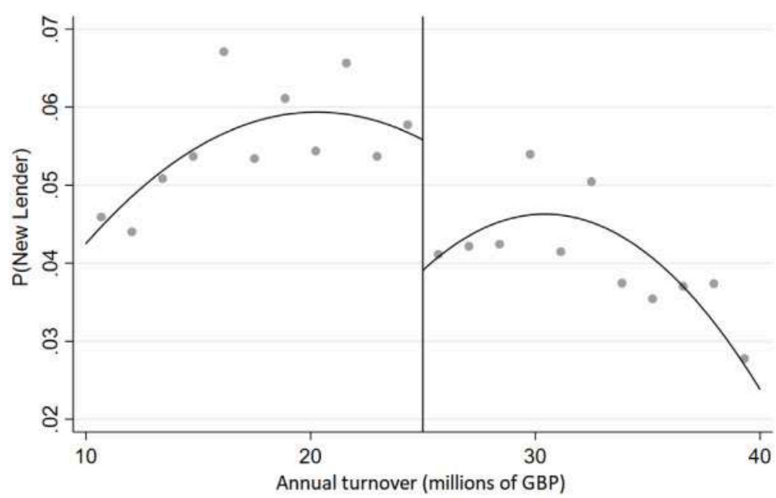
[18] Unfortunately, we cannot observe interest rates on credit products because we do not have this data for the sample of OB respondents in the FLS data.

[19] 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.

[20] Our credit access results could be partially driven by consumers seeking credit from OB lenders signing up for OB too. Although this still shows an active role for OB, we can mitigate this concern by controlling for credit demand. Online Appendix Table A4 shows that our credit effects are robust to controlling for credit use as proxies for demand (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).



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



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

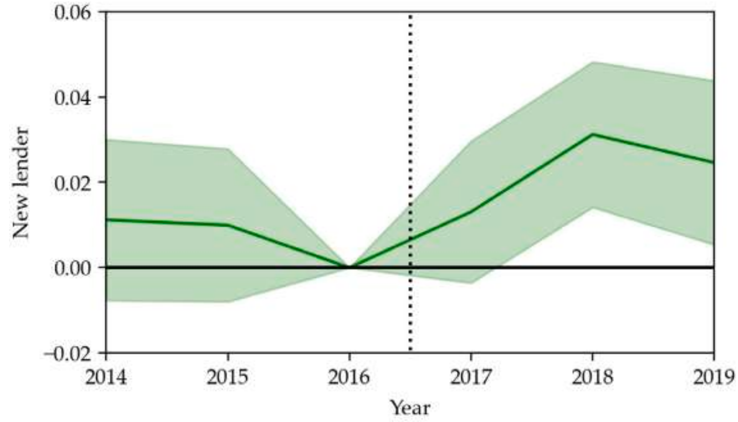
Fig. 3. New SME lending relationships around CCDS eligibility threshold. This figure shows the association between new lending relationship formation and firm sales before and after the implementation of the Commercial Credit Data Sharing (CCDS) policy. The underlying data is company-year data on secured loans for UK firms from Companies House. Panel (a) presents observations from before the implementation of the CCDS (2014–2016) and Panel (b) presents observations after the policy (2017–2019). Each dot is the fraction of firms forming new lending relationships (y-axis) among firms in a given sales bucket (x-axis). We use 22 equally sized buckets from £10 million to £40 million of 2016 firm sales. A firm establishes a new relationship when it gets a loan from a lender that it had not borrowed from in the preceding three years. The vertical line denotes the cutoff firm sales for data sharing under the policy (£25 million), with firms to the left of the line in Panel (b) being treated by the policy and firms to the right of the threshold serving as the control group. The solid curves plot best-fit quadratic polynomials for lending relationship propensity, separately estimated above and below the policy cutoff.

3.2. Evidence from UK microdata on SMEs

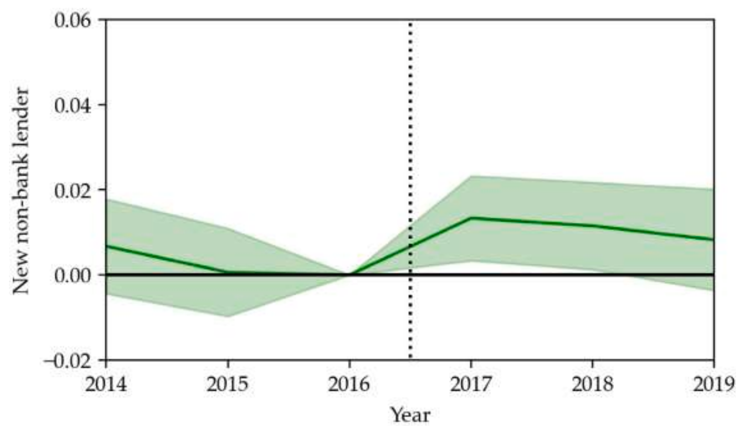
Data on the 2017 launch of the UK's SME-focused OB policy — the "Commercial Credit Data Sharing" (CCDS) — allows us to estimate how OB impacted SMEs' ability to obtain new loans (from banks and non-banks) and test OB's financial inclusion implications. The CCDS is an SME-focused analog of the UK's main OB policy (which covers all bank customers). The CCDS mandated banks to share information on their SME customers, with client approval. Specifically, it required that the nine largest UK banks share detailed information on the transaction accounts, loan repayments, and corporate credit cards of their SME clients with other lenders. Since previous initiatives had made SME credit histories available through credit bureaus, the CCDS principally revealed information about SMEs' transaction accounts (i.e., cash flows). Thus, the information shared on SMEs is analogous to the information individual bank customers share under OB. We briefly describe our analysis of this policy's effect on SME lending, with Online Appendix E providing more detail on the CCDS policy, variable definition, summary statistics, and robustness tests.

The CCDS initiative applied only to SMEs with annual sales below £25 million, which creates quasi-random variation that we exploit for identification. We compare SMEs just below the cutoff (treated) to SMEs just above the cutoff (control) for the three years prior to (2014–2016) and following (2017–2019) the implementation of the policy.[21] We then test how the CCDS policy affects SMEs' ability to form relationships with new lenders. 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

[21] 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.



(a) New lending relationships



(b) New lending relationships with non-banks

Fig. 4. Event-study of SME data sharing and new lending relationships. 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 firms with 2016 sales 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 sales 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 with any lender 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-by-year, region-by-year, and relationship stage-by-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 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. $Any\ New\ Lender_{i,t}$ is an indicator variable equal to one if firm i forms a relationship with a new lender in year t.

Firms in the UK are required to report all claims ("charges") lenders have against their 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).[22] 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.[23] Hence, we observe firms' lending relationships as well as their balance sheet and income statement information over time.

Fig. 3 presents binned scatterplots of new borrowing relationship formation against firm sales 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 in sales threshold before the policy, while Panel (b) shows a discontinuity at that threshold appearing after the policy. Firms below the threshold are more likely

[22] 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.

[23] BvD data is well known for suffering from survivorship bias and various issues with constructing consistent historical panels (Kalemli-Ozcan et al., 2024). 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.

Table 4
SME data sharing and new lending relationships.

	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-by-year FE			Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sector-by-year FE			Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region-by-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	0.000	0.058	0.063	0.064	0.020	0.076	0.021	0.076	0.071

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 firm-year data on secured loans for UK firms with 2016 sales between £10 million and £40 million from Companies House via Bureau van Dijk for the 2014–2019 period. A firm is classified as a Treated SME if its 2016 sales is below the CCDS's £25 million eligibility threshold for data sharing. Post is an indicator variable equal to one after the CCDS was implemented in 2017. Prior CCDS relationship equals one if the firm 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 firm 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 firm takes a loan in the year in question from a lender that it had not borrowed from in the preceding three years. The dependent variable in columns 5 and 7 is an indicator variable equal to one if the firm 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.

to establish a new lending relationship than firms above the threshold after the policy but not before.[24]

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

$$Any\ New\ Lender_{i,t} = \beta \times Treated\ SME_i \times Post_t + \eta X_{i,t-1} + \alpha_i + \gamma_{s,t} + \eta_{g,t} + \nu_{r,t} + \varepsilon_{i,t} \tag{3}$$

We focus on firms with 2016 sales 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 sales 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 measures the focal policy effect. $X_{i,t-1}$ is a lagged vector of firm controls: The log of total assets, cash to total assets, leverage ratio, and credit risk. We include 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 ($\nu_{r,t}$).[25] Regions correspond to the 124 UK postcode areas and industry sectors are based on one-digit SIC codes. Standard errors are clustered at the firm level.

Table 4 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 $Treated SME \times Post$ 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, a 26% 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 the CCDS initially only required the nine largest UK banks to share data. We interact the $Treated\ SME \times Post$ 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 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 ($Prior\ non-CCDS\ relationship_i$). The treatment effect is entirely concentrated among clients of the banks required to share SME data.

Our data allows us to observe whether new lending relationships are with banks (columns 5 and 7) or non-banks (6 and 8). The DiD coefficient is positive but not statistically significant for new relationships formed with banks in column 5, and is both larger and statistically significant in column 6 for new relationships formed with non-banks (e.g., fintechs). The effect is concentrated among clients of the banks required to share their customer data (the triple interaction with $Prior\ CCDS\ relationship_i$ in columns 7 and 8), and these clients are more likely to borrow from non-banks (column 8). Fig. 4 presents

[24] 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.

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

event-study plots for lending relationship formation (Panel (a)) and lending relationship formation with non-bank lenders (Panel (b)) using our main specification in Eq. (3). 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 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. These results show that access to customer bank data leads to non-bank entry into the SME lending market, consistent with the increased fintech entry after OB policy introduction in Section 3.3.

In Online Appendix Table E4, we present an 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 overall borrowing).

Finally, since financial inclusion is a key objective expressed by many policymakers, 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 already had credit or those who did not. Presumably, firms with no prior lending relationship have the most to gain from outside lenders obtaining non-standard data that could be useful in underwriting. 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). We examine the policy's extensive margin effects in column 9 of Table 4 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 in 2016. The rate of relationship formation rises more for firms that had multiple lending relationships prior to the reform. This gives support for the selection mechanism. As we show in Section 4, 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 with non-banks (e.g., fintechs). In terms of distributional effects, treated firms with multiple prior lending relationships are more likely to get new loans. Finally, these new lending relationships are likely beneficial to firms as the SMEs that form new lending relationships with non-banks pay less interest.

3.3. Open banking government policies and financial innovation around the world

Our last set of empirical results focuses on the global consequences of OB policies around the world, broadening our UK-specific analysis. We test whether increased data access spurs financial innovation, which is the most common goal of OB policies. Regulators hope that giving bank customers the ability to share their financial data with fintechs will spark the creation of new firms that offer innovative financial products and increase competition. 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. We use a standard panel event-study design:

$$FintechVC_{i,t} = \sum_{k\neq 0} \beta_k \times OBLag(k)_{i,k,t} + Country_i + Region_r \times Year_t + \epsilon_{i,t}, \tag{4}$$

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 using data from PitchBook.[26] $OBLag(k)_{i,k,t}$ 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.[27] We normalize the year of the OB 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. Standard errors are clustered at the country level.

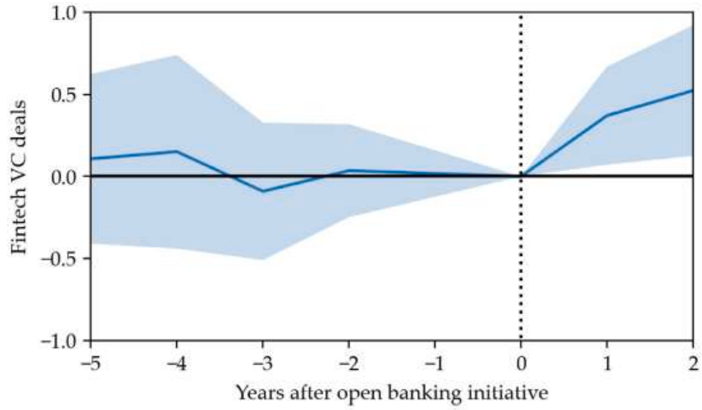
VC data poses two challenges. First, VC activity is skewed, with the US having far more VC investments than any other country. We correct for this using a $\log(1+x)$ transformation of our VC activity measures, which is common in the VC literature (e.g., Gompers and Lerner, 1998 or Li and Zahra, 2012). With this transformation, our tests measure relative increases or decreases in VC activity. Second, the lack of central VC investment registries in most countries makes collecting VC data challenging. Online Appendix Table A5 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.[28] 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 coverage improvements are correlated with the passage of OB government policies.[29] Because we condition on pre-period deals, our results mostly speak to countries that already have developed VC markets.[30] 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.

[26] 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. Since our interest lies in financial innovation, we split the VC 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. 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.

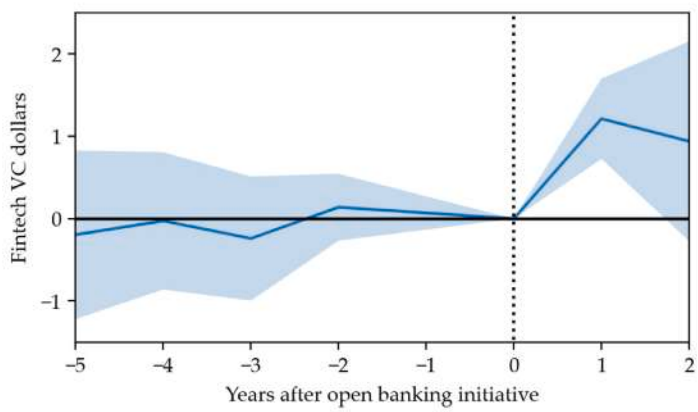
[27] For countries in the sample that never adopt OB, $OBLag(k)_{i,k,t}$ is zero everywhere; these countries help identify region-by-year fixed effects.

[28] 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.

[29] 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.



(a) Log number of fintech VC deals



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

Fig. 5. Event-study of fintech investment after open banking government policies. 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. Regions are (i) Africa, Middle East & North Africa; (ii) Europe & Central Asia; (iii) Latin America & the Caribbean; (iv) North America; (v) 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.

Fig. 5 presents the results from the event-study specification in Eq. (4) and shows a relative absence of pre-trends in fintech VC activity in the number of deals in Panel (a) and the amount invested in 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 sharp increase in fintech VC investments following OB policy adoption is a natural consequence of the uncertainty reduction around the timing of OB policy passage and combined with VCs' fast reactions to new investment opportunities.[31]

Table 5 uses a difference-in-differences design to quantify the relationship between OB policies and fintech VC activity:

$$FintechVC_{i,t} = \beta \times OB_{i,t} + Country_i + \gamma \times Non-fintechVC_{i,t} + Region_r \times Year_t + \epsilon_{i,t}, \qquad (5)$$

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 Eq. (4). 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 ($Trust_i \times Year_t$).

[30] The results in Table 5 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.

[31] 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. VCs target at least 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 investment opportunities (Gompers et al., 2008).

Table 5
Fintech investment after open banking government policies.

	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-by-year FE			Yes	Yes	Yes			Yes	Yes	Yes
Fintech trust-by-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

This table shows changes in fintech venture capital (VC) investment following the implementation of open banking (OB) 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 policy was 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, 9 and 10 control for region-by-year fixed effects. Regions are i) Africa, Middle East & North Africa; ii) Europe & Central Asia; iii) Latin America & the Caribbean; iv) North America; v) 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 (reported in parentheses) are clustered at the country level. *** denotes p-value < 0.01, ** denotes <0.05, and * denotes <0.1.

Table 6
Fintech investment after open banking government policies by fintech product area.

	Alternative lending (1)	Consumer finance (2)	Financial IT (3)	Payments (4)	Regtech (5)	Wealth management (6)	Digital assets (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-by-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

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 Online Appendix D. The independent variable is an indicator variable equal to one if the year in question is after the year major open banking policy was passed in the country in question. We use Equation (5) which controls for the log of one plus the number of non-fintech VC deals, country fixed effects, and region-by-year fixed effects. Regions are (i) Africa, Middle East & North Africa; (ii) Europe & Central Asia; (iii) Latin America & the Caribbean; (iv) North America; (v) 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 (reported in parentheses) are clustered at the country level. *** denotes p-value < 0.01, ** denotes <0.05, and * denotes <0.1.

Across specifications, fintech companies receive significantly more VC investment following the adoption of OB policies. Using our preferred specification from Eq. (5), we find a 0.31 increase in log fintech VC deals (column 4 of Table 5) and a 0.87 increase in log fintech VC dollars (column 9). These estimates are robust to different combinations of controls and fixed effects (columns 1, 3, 6, and 8). The median country-year in this data has 19 fintech VC deals worth \$89 million and so our estimates of β translate into an additional 7 deals and \$125 million 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.[32]

In Section 2.4, we identified consumer trust in sharing their data with fintechs as a potential driver of OB government policies. We next examine if trust in fintechs mediates the effect of OB on VC investments in fintech. We provide suggestive evidence that trust amplifies 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 (column 2) and positive and insignificant for dollars invested (column 7). 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 policies. 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 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

[32] 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).

be passing OB laws around the same time. In Online Appendix D, we present additional robustness tests and show that OB policies that force banks to share their customers' data drive these results.

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 for many financial products or only used for credit underwriting. Since Pitchbook lacks more granular product classifications, we overcome this by using PitchBook's keywords feature to define seven subindustries 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.[33] Details of our classification are in Online Appendix D. Using Eq. (5), Table 6 considers VC investments in companies targeting specific use cases as dependent variables. 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. Overall, we find a broad-based increase in fintech activity following OB policy implementation, which suggests VCs anticipate OB data as offering value not just for credit issuance but for a variety of fintech use cases. This is consistent with our findings in Section 3.1 that UK consumers use many OB-data-reliant products, such as financial advice and credit.

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. First, we model the two use cases of OB identified in our empirical work: Credit and advice. For credit, we use a standard setup where data provides a signal of borrower risk, whereas for advice we adapt and use models with product (e.g., Jones and Tonetti, 2020) or business practice (e.g., Farboodi et al., 2019) improvements to capture data improving financial products. Second, reflecting the goal of OB in promoting financial innovation and competition, we explicitly model new firm entry on the extensive margin together with product and price improvements on the intensive margin. Third, our consumer OB use and fintech entry results confirm that privacy considerations are central to the uptake of the policy. 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.

[33] It is worth mentioning that the size of each of these subindustries is small. Specifically, the median (mean) subindustry-year sees 4 (15) deals worth \$9 million (\$300 million).

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 an individual or an 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. (2024) and our SME analysis in Section 3.2. Alternatively, a financial planning app uses balances and transactions from a consumer's financial accounts to offer her customized financial and tax advice.

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. In 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 $dF(\chi_i|S_i = 0)$.

To account for both advice OB and credit OB use cases, we assume that χ_i provides information on both the consumer-specific marginal cost (mc_i)[34] 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 $v_{ij} \equiv (p_{ij}, g_{ij})$, 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

[34] 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 Online Appendix F.

with consumer i in the prior period. Consumer i receives the following indirect utility from product ij:

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

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 her relationship bank, e.g., due to 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 more relevant financial advice. S_i is an indicator for whether the consumer shares her data with outsiders. When a consumer obtains a product from her relationship bank, she receives both the additional relationship utility θ as well as the customization utility λ. When a consumer obtains a product from an outsider, she only obtains the customization utility and only if she 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 equilibria by obscuring whether a consumer chooses an uninformed offer because she is a high-cost type with high-price offers from insiders, or because she is a low-cost type with a high idiosyncratic preference for the outsider's product (see, 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 $s_j(\boldsymbol{v}_i,\chi_i)$ denote the probability that a consumer with characteristics χ_i chooses firm j's product given all product offerings, including the outside option, $\boldsymbol{v}_i$. This quantity is obtained by integrating the consumer's optimal choice over the consumer-firm taste shocks, $\boldsymbol{\epsilon}_i$:

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

4.1.3. Consumer opt-in to data sharing

Under the OB regime, each consumer chooses whether to opt in to data sharing.[35] 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 $\boldsymbol{v}_i^S$ and $\boldsymbol{v}_i^{\sim S}$ denote the set of offers she receives if she opts in to or out of data sharing, respectively. Let $Eu(\boldsymbol{v}_i)$ denote the consumer's expected utility of the discrete choice problem in Eq. (7) for a given set of offers, with

$$Eu(\boldsymbol{v}_i) = \int \max_j \left\{u(v_{ij},\chi_i)\right\} dF(\boldsymbol{\epsilon}_i). \tag{8}$$

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 privacy and consumer-level heterogeneity.[36] In the same discrete choice framework, we model the consumer's indirect utility of sharing or not sharing her data as follows:

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

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

Here, ϕ represents a society-wide hedonic privacy preference and ϵ_i^S and $\epsilon_i^{\sim S}$ represent consumer-specific i.i.d. privacy preference shocks.[37] 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 the sum of two parts. First, 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. Second, 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 only if data is shared by the consumer:

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

Firms are informed about 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 (v_{ij}) and uninformed firms offering a single product and price to all consumers:

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

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.

[35] 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.

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

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

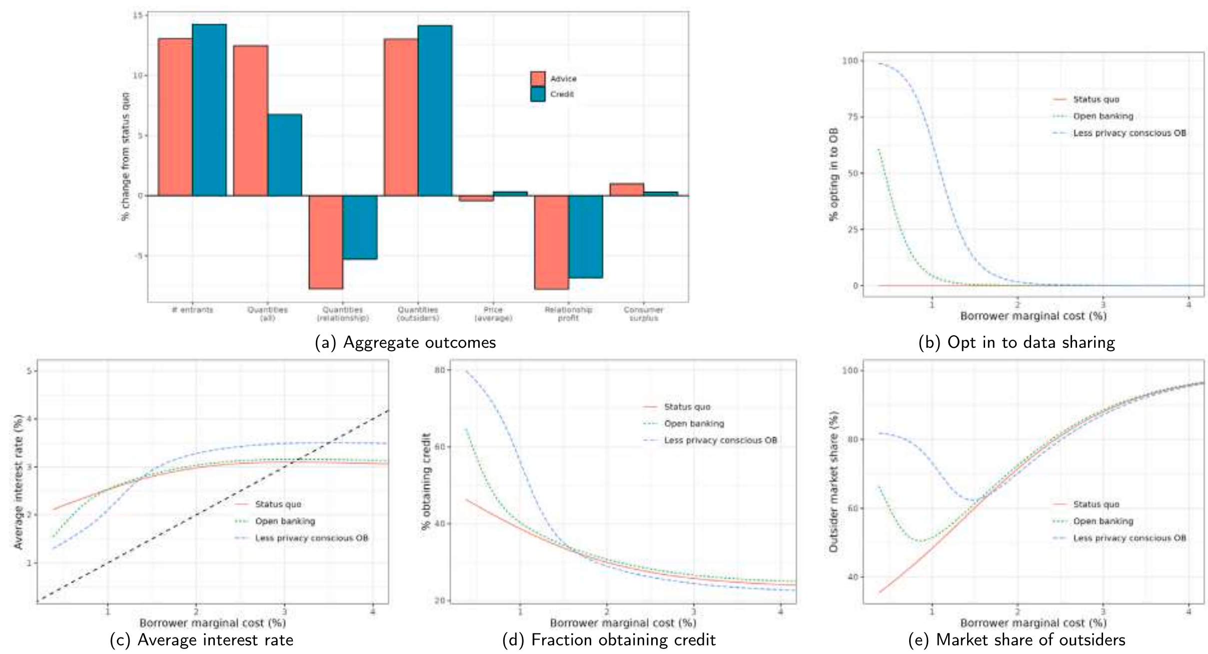


Fig. 6. Aggregate and distributional outcomes of open banking. Panel (a) presents model-implied aggregate changes after open banking (OB), with each bar showing the percentage change in the relevant outcome caused by moving from the status-quo relationship banking regime to the OB regime. Magenta and cyan bars show outcomes for the financial advice and non-GSE residential mortgage calibrations, respectively. # entrants is the number of new entrants. Quantities (all) is the population fraction obtaining the financial service, which we further split into Quantities (relationship), i.e., relationship bank, and Quantities (outsiders), i.e., fintechs. Price (average) is the average fee or rate charged. Relationship profit is relationship banks' profits. Panels (b)–(e) show the distributional outcomes of OB in the credit case. x-axes show borrowers with different marginal costs. Red lines and dotted green lines indicate outcomes for the relationship banking and calibrated OB regime, respectively. Dashed blue lines indicate outcomes in a counterfactual simulation where borrowers' privacy preference is 25% lower. Panel (b) shows the fraction of the population opting into data sharing. Panel (c) shows the average interest rate. Panel (d) shows the fraction of the population obtaining credit. Panel (e) shows the outsiders' market share. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

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 has a similar structure but has an added first stage where consumers choose whether to share their data.

We focus on 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. For a given regime, an equilibrium consists of a set of prices and product customization choices, ν_i, a number of new entrants, consumer product choices, and consumer data-sharing choices. The endogenous choices satisfy the optimal firm entry and profit maximization conditions, optimal consumer product and data sharing choice, and firms' consistent beliefs over consumer choices by type.

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 the credit OB and advice OB use cases described previously. Both calibrations are intended to be quantitatively realistic illustrations of the economic forces affecting two real-world applications of OB.

The non-GSE residential mortgage calibration is an example of where consumer data is useful for underwriting, as the relevant dimension of heterogeneity is in default risk.[38] This market is well studied and there exists estimates for several key parameters in the literature for calibration. The financial advice calibration is an example of the data allowing for a product more tailored to the consumer's needs: The relevant dimension of heterogeneity is what the optimal savings, investment, and tax strategy would be given the consumer's particular financial situation.

We detail our calibration exercise in Online 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 3.3) and consumer adoption of OB from the UK consumer survey (described in Section 3.1). Other parameter estimates, such as consumer price sensitivity and lender marginal costs, are taken from the relevant mortgage (Buchak et al., 2024a) and financial advice (Di Maggio et al., 2022a) literature.

4.2. Consequences of open banking

Using our calibrated model, we first look at the aggregate and distributional effects of OB (Section 4.2.1) before moving on to the role of society-wide privacy preferences (Section 4.2.2). Additional discussion showing the interaction between credit registries and OB is presented in Online Appendix F.3.

4.2.1. Aggregate and distributional consequences of open banking

Fig. 6 Panel (a) compares equilibrium outcomes under OB to those under relationship banking for our financial advice (magenta) and credit (cyan) calibrations. Across both calibrations, entry rises and the quantities of financial services provided increase, although these

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

increases are more dramatic in the advice case. We decompose these aggregate quantity changes into quantity changes from relationship banks (columns "Quantities (relationship)") and other providers (columns "Quantities (outsiders)"). Outsider (e.g., fintech) quantities increase and relationship bank quantities decrease for both products. The substantial increases in quantities offered by outsiders for both financial products are consistent with our findings on large increases in fintech entry in Section 3.3. Average prices in both cases are largely unchanged, although the modest aggregate price changes in the credit case mask dramatic heterogeneity along the distribution of borrower types. Incumbent profits fall and consumer surplus increases in both cases, although the surplus increase is larger for the advice case despite lower entry.

Panels (b) through (e) provide greater insight into the distributional effects of credit OB across the distribution of borrower marginal cost (MC), which reflects default probability. Here, we focus on the comparison between the no-OB status quo in red and the calibrated OB regime in green.[39] Panel (b) shows the fraction of borrowers opting into data sharing. In the no-OB status quo, no consumers share data. Once in the OB regime, the green line shows that the propensity to share data is decreasing in the borrower's unobserved marginal MC: Roughly 60% of borrowers with the lowest MC share their data, while essentially no borrowers with medium or higher MC share data.[40]

Turning to price and quantity outcomes, Panel (c) shows that in the no-OB status quo (red line) average interest rates are only weakly increasing in MC. The relationship is weak because average interest rates are a combination of informed insider rates, which only partially track borrower MC due to information rents, and the uninformed pooled interest rate, which is invariant. In the OB regime (green line), low-MC borrowers opt in to data sharing and reveal their type to outside lenders. These borrowers are offered lower rates from both outsiders and the insider, who now faces greater competition. These lower rates lead to more borrowing (Panel (d)), and to outsiders gaining market share among the lowest-MC borrowers (Panel (e)). In contrast, high-MC borrowers, who choose not to opt in to data sharing, partially reveal to outside lenders their type, although hedonic privacy preferences partially obscure this inference. Therefore, uninformed outsiders charge slightly higher rates as compared to the status quo.

The results in Panels (c) through (d) are consistent with our findings in the SME analysis in Section 3.2: Under OB 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. Importantly, however, because information revelation reduces the adverse selection faced by outsiders, entry rises and in our calibration, entry's positive effect through 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, consider the financial advice calibration. Here, consumer "type" represents the idiosyncratic needs for financial advice. All types are made better off under OB. This arises because customers that share data benefit through outsiders' ability to offer fully customized advice. Furthermore, there is an increase in competition which benefits everyone including customers that do not share their data due to privacy concerns. Since, in contrast to credit OB, no negative information is revealed by sharing data, consumers are more likely to opt in to OB in the advice case than the credit case. Intuitively, all customers benefit from providing more data to their financial advisor, while only customers with low MC directly benefit from providing more data to their loan underwriter. This helps to explain the greater uptake of advice OB than credit OB observed in the UK survey data in Section 3.1.

4.2.2. Consumer attitudes towards privacy

We conclude our analysis of the model by examining how consumer attitudes towards privacy impact the equilibrium outcomes of OB. Fig. 7 shows the impact of varying consumers' mean preference for privacy. The x-axis shows the value of privacy as a multiple of the calibrated value for the UK such that consumers' aversion to data sharing is increasing in the x-value. 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 the financial advice calibration, and the blue lines and marks show outcomes for the credit calibration.

Unsurprisingly, the fraction of consumers opting into data sharing is decreasing in their preference for privacy. However, our finding in Section 3.1 that more consumers opt in to OB for financial advice than for credit is sustained across counterfactual privacy preferences. Next, while in the credit case, high MC borrowers, who do not share their data, do not benefit directly from OB, they do experience two indirect effects: They benefit from increased lender entry and are harmed by their opt-out decision partially revealing their high MC. For societies with weak privacy preferences, the act of not sharing data reveals strong negative information about their type, and so the harm outweighs the benefits of increased entry. In contrast, in societies with privacy preferences similar to the UK, the negative inference from not sharing data is relatively weak, and so the competition and product variety benefits of increased entry outweigh the signaling costs. Our plot showing the fraction of customers made worse off by OB makes this clear. For advice, for the reasons described, all types of consumers are better off at all levels of societal privacy preference. For credit, we see a distinct threshold at about 85% of our calibrated UK privacy preference, below which the signaling cost for high-MC types outweighs the benefit of new entry and a positive number of borrowers are made worse off (marked with x in blue).

We confirm this intuition by revisiting Fig. 6 Panels (b) through (e). 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 Fig. 7. These panels show that as more borrowers opt in to OB (Panel (b)), rates decrease more for low-MC borrowers and increase more for high-MC borrowers (Panel (c)). This leads to greater quantities of credit for low-MC borrowers, but less credit for high-MC borrowers, both overall and from outsiders, relative to the no-OB status quo (Panels (d) and (e)).

4.2.3. Summary

The bottom line from our model is 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.

The complex interplay between use cases, consumer heterogeneity, and societal preferences for privacy leads us to a range of important predictions for the impact of OB on the market for financial products. The advice OB case has little ambiguity. All customers benefit from the option to share their data either directly through better products when sharing data or indirectly through increased competition when they are privacy conscious. In contrast, the results in the credit OB are far more nuanced. Customers with favorable data and low privacy preferences share data and benefit from improved loan terms. Firm entry

[39] We return to a counterfactual with a smaller consumer preference for privacy below (in blue).

[40] 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 (b), 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 $dF(\chi_i)$.

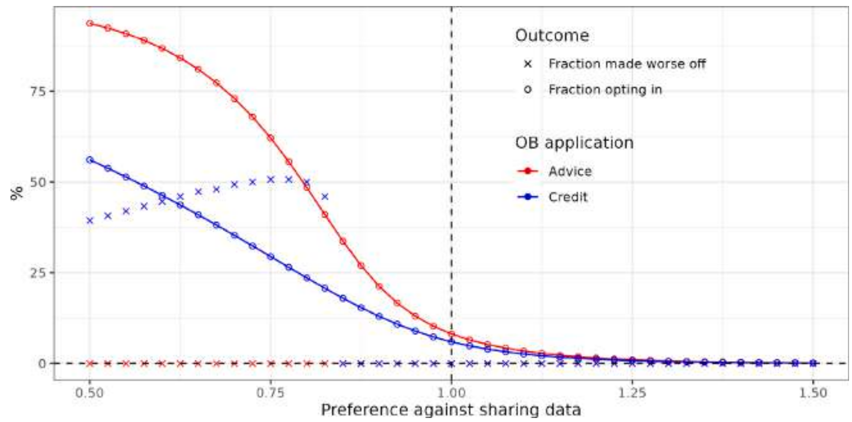


Fig. 7. Effect of Societal Privacy Preferences on Open Banking Equilibria This figure shows how the impact of open banking (OB) varies as societal privacy preferences vary under the model of Section 4. Specifically, it shows outcomes (y-axis) for the advice (red) and credit (blue) OB as population preferences 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. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

increases 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 interest rates. 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.

We find that the societal desire for privacy plays an important role in pinning down distributional consequences of credit OB. More privacy-conscious consumers shield high-MC borrowers from scrutiny: Lenders cannot infer from opting out of OB that the borrower is not sharing because she has a high-MC type. As the preference for privacy decreases, opting out is a more precise signal that the consumer is a costly borrower to serve and lenders charge higher rates. This has the effect of potentially leaving privacy-conscious consumers and those with a high marginal cost worse off.

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. Granular microdata on UK consumers shows that they use OB for credit but also for financial advice, with that usage associated with higher 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 relationship formation is highest for SMEs with prior lending relationships. Large increases in VC fintech investments across different financial products (e.g., financial advice applications, credit, payments, regtech) follow OB policy implementations, suggesting consumer financial transactions data are valuable across many financial applications.

We interpret these results through a general framework of data use and sharing, focusing on the contrasting implications of using data for underwriting (in credit OB) and using data to improve products (in advice OB). 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 to share data 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 high risk consumers are likely to be on the margins of the financial system, and thus precisely those whose financial inclusion policymakers are interested in facilitating. These results are at odds with the financial inclusion goals of OB policies but consistent with our finding that the SMEs who already had credit access benefit the most from OB.

Importantly, these potential negative distributional effects are not present when OB data is used for product improvements rather than for credit 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, and 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.

CRediT authorship contribution statement

Tania Babina: Writing – review & editing, Writing – original draft, Project administration, Formal analysis, Data curation, Conceptualization. **Saleem Bahaj:** Writing – review & editing, Formal analysis, Data curation. **Greg Buchak:** Writing – review & editing, Writing – original draft, Formal analysis, Data curation, Conceptualization. **Filippo De Marco:** Writing – review & editing, Formal analysis, Data curation. **Angus Foulis:** Writing – review & editing, Formal analysis, Data curation. **Will Gornall:** Writing – review & editing, Writing – original draft, Project administration, Formal analysis, Data curation. **Francesco Mazzola:** Writing – review & editing, Formal analysis, Data curation. **Tong Yu:** Writing – review & editing, Formal analysis, Data curation.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Will Gornall reports financial support was provided by Government of Canada Social Sciences and Humanities Research Council Insight Grant 435-2023-0747 "The financial returns to innovation". Filippo De Marco thanks BAFFI for their support

References

Acquisti, A., Taylor, C., Wagman, L., 2016. The economics of privacy. J. Econ. Literature 54 (2), 442–492.
Acs, Z., Åstebro, T., Audretsch, D., Robinson, D.T., 2016. Public policy to promote entrepreneurship: A call to arms. Small Bus. Econ. 47 (1), 35–51.
Aguirregabiria, V., Clark, R., Wang, H., 2019. The geographic flow of bank funding and access to credit: Branch networks, local synergies, and competition. CEPR Discussion Paper No. 13741, Centre for Economic Policy Research, London, United Kingdom.
Allen, F., Gu, X., Jagtiani, J., 2021. A survey of fintech research and policy discussion. Rev. Corp. Finance 1 (3–4), 259–339.
Babina, T., Barkai, S., Jeffers, J., Karger, E., Volkova, E., 2023a. Antitrust enforcement increases economic activity.
Babina, T., Fedyk, A., He, A.X., Hodson, J., 2023b. Artificial intelligence and firms' systematic risk. Available at SSRN 4868770.
Babina, T., Fedyk, A., He, A., Hodson, J., 2024. Artificial intelligence, firm growth, and product innovation. J. Financ. Econ. 151, 103745.
Babina, T., He, A.X., Howell, S.T., Perlman, E.R., Staudt, J., 2023c. Cutting the innovation engine: How federal funding shocks affect university patenting, entrepreneurship, and publications. Q. J. Econ. 138 (2), 895–954.
Babina, T., Howell, S.T., 2024. Entrepreneurial spillovers from corporate R&D. J. Labor Econ. 42 (2).
Bahaj, S., Foulis, A., Pinter, G., 2020. Home values and firm behavior. Amer. Econ. Rev. 110 (7), 2225–2270.
Bai, J., Bernstein, S., Dev, A., Lerner, J., 2022. The dance between government and private investors: Public entrepreneurial finance around the globe. NBER Working Paper No. 28744, National bureau of economic research Cambridge, Mass., USA.
Barth, J.R., Caprio Jr., G., Levine, R., 2004. Bank regulation and supervision: What works best?. J. Financial Intermed. 13 (2), 205–248.
Bartlett, R., Morse, A., Stanton, R., Wallace, N., 2022. Consumer-lending discrimination in the FinTech era. J. Financ. Econ. 143 (1), 30–56.
Beck, T., De Jonghe, O., Schepens, G., 2013. Bank competition and stability: Cross-country heterogeneity. J. Financial Intermed. 22 (2), 218–244.
Ben-Shahar, O., Schneider, C.E., 2011. The failure of mandated discourse. Univ. Pennsylvania Law Rev. 159, 647–749.
Benetton, M., Buchak, G., Robles-Garcia, C., 2022. Wide or narrow? Competition and scope in financial intermediation. CEPR Discussion Paper No. 17497, Centre for Economic Policy Research, London, United Kingdom.
Berg, T., Burg, V., Gombović, A., Puri, M., 2020. On the rise of fintechs: Credit scoring using digital footprints. Rev. Financ. Stud. 33 (7), 2845–2897.
Berg, T., Fuster, A., Puri, M., 2022. Fintech lending. Annu. Rev. Finan. Econ. 14 (1), 187–207.
Bian, B., Ma, X., Tang, H., 2021. The supply and demand for data privacy: Evidence from mobile apps. Available at SSRN 3987541.
Blickle, K., Parlatore, C., Saunders, A., 2023. Specialization in banking.
Boot, A., Hoffmann, P., Laeven, L., Ratnovski, L., 2021. Fintech: What's old, what's new?. J. Financial Stabil. 53, 100836.
Boot, A.W.A., Thakor, A.V., 1997. Financial system architecture. Rev. Financ. Stud. 10 (3), 693–733.
Buchak, G., Hu, J., Wei, S.J., 2021. FinTech as a financial liberator. NBER Working Paper No. 29448, National bureau of economic research Cambridge, Mass., USA.
Buchak, G., Matvos, G., Piskorski, T., Seru, A., 2018. Fintech, regulatory arbitrage, and the rise of shadow banks. J. Financ. Econ. 130 (3), 453–483.
Buchak, G., Matvos, G., Piskorski, T., Seru, A., 2024a. Beyond the balance sheet model of banking: Implications for bank regulation and monetary policy. J. Polit. Econ. 132 (2).
Buchak, G., Matvos, G., Piskorski, T., Seru, A., 2024b. The secular decline of bank balance sheet lending.
Chen, S., Doerr, S., Frost, J., Gambacorta, L., Shin, H.S., 2023. The fintech gender gap. J. Financial Intermed. 54, 101026.
Chen, L., Huang, Y., Ouyang, S., Xiong, W., 2024. Data privacy and digital demand.
Claessens, S., Laeven, L., 2004. What drives bank competition? Some international evidence. J. Money Credit Bank. 36 (3), 563–583.
Claessens, S., Rojas-Suarez, L., 2016. Financial regulations for improving financial inclusion. Task Force Report, Center for Global Development.
Claessens, S., Van Horen, N., 2013. Impact of foreign banks. J. Financial Perspect. 1 (1), 29–42.
Cornelli, G., Frost, J., Gambacorta, L., Rau, R., Wardrop, R., Ziegler, T., 2020. Fintech and big tech credit: A new database. BIS Working Paper No. 887, Bank for International Settlements, Basel, Switzerland.
Crawford, G.S., Pavanini, N., Schivardi, F., 2018. Asymmetric information and imperfect competition in lending markets. Amer. Econ. Rev. 108 (7), 1659–1701.
De Roure, C., Pelizzon, L., Thakor, A.V., 2022. P2P lenders versus banks: Cream skimming or bottom fishing? Rev. Corp. Finance Stud. 11 (2), 213–262.
Denes, M., Howell, S.T., Mezzanotti, F., Wang, X., Xu, T., 2023. Investor tax credits and entrepreneurship: evidence from us states. J. Finance 78 (5), 2621–2671.
Di Maggio, M., Egan, M., Franzoni, F., 2022a. The value of intermediation in the stock market. J. Financ. Econ. 145 (2), 208–233.
Di Maggio, M., Ratnadiwakara, D., Carmichael, D., 2022b. Invisible primes: Fintech lending with alternative data. NBER Working Paper No. 29840, National bureau of economic research Cambridge, Mass., USA.
Di Maggio, M., Yao, V., 2021. FinTech borrowers: Lax screening or cream-skimming?. Rev. Financ. Stud. 34 (10), 4565–4618.
Diamond, D.W., 1984. Financial intermediation and delegated monitoring. Rev. Econ. Stud. 51 (3), 393–414.
Diamond, D.W., Dybvig, P.H., 1983. Bank runs, deposit insurance, and liquidity. J. Polit. Econ. 91 (3), 401–419.
Djankov, S., McLiesh, C., Shleifer, A., 2007. Private credit in 129 countries. J. Financ. Econ. 84 (2), 299–329.
Egan, M., Hortaçsu, A., Matvos, G., 2017. Deposit competition and financial fragility: Evidence from the US banking sector. Amer. Econ. Rev. 107 (1), 169–216.
Egan, M., Lewellen, S., Sunderam, A., 2022. The cross section of bank value. Rev. Financ. Stud. 35 (5), 2101–2143.
Erel, I., Liebersohn, J., 2022. Can FinTech reduce disparities in access to finance? Evidence from the Paycheck Protection Program. J. Financ. Econ. 146 (1), 90–118.
Farboodi, M., Mihet, R., Philippon, T., Veldkamp, L., 2019. Big data and firm dynamics. AEA Papers and Proceedings 109, 38–42.
Fuster, A., Plosser, M., Schnabl, P., Vickery, J., 2019. The role of technology in mortgage lending. Rev. Financ. Stud. 32 (5), 1854–1899.
Ghosh, P., Vallee, B., Zeng, Y., 2024. FinTech lending and cashless payments. J. Finance Forthcoming.
Goldstein, I., Huang, C., Yang, L., 2022. Open banking under maturity transformation. Working Paper.
Gompers, P.A., Gornall, W., Kaplan, S.N., Strebulaev, I.A., 2020. How do venture capitalists make decisions?. J. Financ. Econ. 135 (1), 169–190.
Gompers, P., Kovner, A., Lerner, J., Scharfstein, D., 2008. Venture capital investment cycles: The impact of public markets. J. Financ. Econ. 87 (1), 1–23.
Gompers, P., Lerner, J., 1998. What drives venture capital fundraising?. Brook. Pap. Econ. Activity. Microecon. 1998, 149–204.
Gopal, M., Schnabl, P., 2022. The rise of finance companies and fintech lenders in small business lending. Rev. Financ. Stud. 35 (11), 4859–4901.
Gornall, W., Strebulaev, I.A., 2021. The economic impact of venture capital: Evidence from public companies. Available at SSRN 2681841.
Granja, J., Leuz, C., Rajan, R.G., 2022. Going the extra mile: Distant lending and credit cycles. J. Finance 77 (2), 1259–1324.

Grossman, S.J., 1981. The informational role of warranties and private disclosure about product quality. J. Law Econ. 24 (3), 461–483.
He, Z., Huang, J., Parlartore, C., 2024. Information span in credit market competition.
He, Z., Huang, J., Zhou, J., 2023. Open banking: Credit market competition when borrowers own the data. J. Financ. Econ. 147 (2), 449–474.
Hertzberg, A., Liberti, J.M., Paravisini, D., 2011. Public information and coordination: Evidence from a credit registry expansion. J. Finance 66 (2), 379–412.
Ioannidou, V., Ongena, S., 2010. "Time for a change": Loan conditions and bank behavior when firms switch banks. J. Finance 65 (5), 1847–1877.
Jiang, E., Matvos, G., Piskorski, T., Seru, A., 2020. Banking without deposits: Evidence from shadow bank call reports. NBER Working Paper No. 26903, National bureau of economic research Cambridge, Mass., USA.
Jones, C.I., Tonetti, C., 2020. Nonrivalry and the economics of data. Amer. Econ. Rev. 110 (9), 2819–2858.
Kalemli-Ozcan, S., Sorensen, B., Villegas-Sanchez, C., Volosovych, V., Yesiltas, S., 2024. How to construct nationally representative firm level data from the ORBIS global database: New facts on SMEs and aggregate implications for industry concentration. Am. Econ. J.: Macroecon. 16 (2), 353–374.
Kaplan, S.N., Lerner, J., 2010. It ain't broke: The past, present, and future of venture capital. J. Appl. Corp. Finance 22 (2), 36–47.
Klapper, L., Laeven, L., Rajan, R., 2006. Entry regulation as a barrier to entrepreneurship. J. Financ. Econ. 82 (3), 591–629.
Kroszner, R.S., Strahan, P.E., 1999. What drives deregulation? Economics and politics of the relaxation of bank branching restrictions. Q. J. Econ. 114 (4), 1437–1467.
Laeven, L., Levine, R., 2009. Bank governance, regulation and risk taking. J. Financ. Econ. 93 (2), 259–275.
Li, Y., Zahra, S.A., 2012. Formal institutions, culture, and venture capital activity: A cross-country analysis. J. Bus. Venturing 27 (1), 95–111.
Mullainathan, S., Schnabl, P., 2010. Does less market entry regulation generate more entrepreneurs? Evidence from a regulatory reform in Peru. In: International Differences in Entrepreneurship. The University of Chicago Press, pp. 159–177.
Nam, R.J., 2022. Open banking and customer data sharing: Implications for FinTech borrowers. SAFE Working Paper No. 364.
Ongena, S., Popov, A., Udell, G.F., 2013. "When the cat's away the mice will play": Does regulation at home affect bank risk-taking abroad?. J. Financ. Econ. 108 (3), 727–750.
Parlour, C.A., Rajan, U., Zhu, H., 2022. When fintech competes for payment flows. Rev. Financ. Stud. 35 (11), 4985–5024.
Petersen, M.A., Rajan, R.G., 1994. The benefits of lending relationships: Evidence from small business data. J. Finance 49 (1), 3–37.
Petersen, M.A., Rajan, R.G., 1995. The effect of credit market competition on lending relationships. Q. J. Econ. 110 (2), 407–443.
Philippon, T., 2019. On fintech and financial inclusion. NBER Working Paper No. 26330, National bureau of economic research Cambridge, Mass., USA.
Phillips, G.M., Zhdanov, A., 2017. Venture capital investments and merger and acquisition activity around the world. NBER Working Paper No. 24082, National bureau of economic research Cambridge, Mass., USA.
Puri, M., Robinson, D.T., 2007. Optimism and economic choice. J. Financ. Econ. 86 (1), 71–99.
Puri, M., Zarutskie, R., 2012. On the life cycle dynamics of venture-capital- and non-venture-capital-financed firms. J. Finance 67 (6), 2247–2293.
Ramakrishnan, R.T.S., Thakor, A.V., 1984. Information reliability and a theory of financial intermediation. Rev. Econ. Stud. 51 (3), 415–432.
Stulz, R.M., 2019. Fintech, bigtech, and the future of banks. J. Appl. Corp. Finance 31 (4), 86–97.
Tang, H., 2019. The value of privacy: Evidence from online borrowers. Working Paper.
Thakor, A.V., 2020. Fintech and banking: What do we know?. J. Financial Intermed. 41, 100833.
Vives, X., 2019. Digital disruption in banking. Annu. Rev. Finan. Econ. 11 (1), 243–272.