Data as Collateral: Open Banking for Small Business Lending

Tong Yu*
Imperial College London

Job Market Paper |Latest version|

December 2, 2024

Abstract

Open banking enables small businesses to share their bank financial data with potential lenders. I examine the effect of open banking on collateralization in small business lending. For identification, I exploit institutional features of the UK's open banking policy that creates a discontinuity in firms' eligibility to share data. Using a novel loan-level dataset covering the entire UK secured business loan market, I document that open banking eases the pledge of assets like accounts receivable and inventory. Firms eligible to share data are more likely to pledge such assets as collateral, thereby improving their access to credit. These effects are more pronounced for firms facing greater information asymmetry and those with greater information available to share. These findings highlight the role of open banking in reducing collateral constraints by mitigating information asymmetry.

Keywords: Data, Collateral, Open Banking, Small Business Lending, Open Data Economy

*Tong Yu is with Imperial College London. I am indebted to Ramana Nanda, Franklin Allen, and Rajkamal Iyer for their continuous guidance and support. I would also like to thank Tania Babina, Bruno Biais, Jamie Coen, Claudia Custodio, Marco Di Maggio, Matthias Efing, Itay Goldstein, Ziang Li, José-Luis Peydró, Manju Puri, Tarun Ramadorai, Magdalena Rola-Janicka, Philip Strahan, Per Strömberg, Basil Williams; and seminar participants at the Financial Conduct Authority, HEC Paris Finance PhD Workshop, and Imperial College London for their helpful discussions and comments that improved this paper. I am grateful to the data teams at the Office for National Statistics, Companies House, and the Financial Conduct Authority for their assistance with data access. I also thank the open banking policy team, the bank supervision team, the economic research team, and the market research team at the Financial Conduct Authority for sharing their valuable expertise, in particular, Gareth Thomas, Margaret Watmough, Robert Cross, and Tim Burrell. This research was funded in part by the Brevan Howard Centre Dissertation Fellowship at Imperial College London. At the time of writing, Tong Yu was an employee of the Financial Conduct Authority. The views expressed in this paper are those of the author and do not necessarily reflect the views of the Financial Conduct Authority. All errors are my own.

Contact information: t.yu18@imperial.ac.uk, www.tong-yu.com

1 Introduction

Data and collateral play a crucial role in small business lending. Lenders typically assess a small business's creditworthiness by evaluating its historical financial performance. However, this financial data is often held by the bank that provides the business's current account and is not widely shared. This data monopoly creates an information asymmetry between small businesses and potential new lenders. To mitigate this asymmetry, lenders require small businesses to pledge collateral. Yet, unlike larger firms, small businesses often lack physical assets, such as real estate, that can serve as collateral. This results in collateral constraints that limit their access to credit[1] and hinder their productivity and growth (see, e.g., Schmalz, Sraer and Thesmar (2017); Catherine et al. (2022)).

Recently, a digital revolution known as *Open Banking*, has begun reshaping the financial landscape, particularly for small businesses.[2] For instance, Babina et al. (2024) show that open banking helps small businesses build new lending relationships. Many small businesses rely on a single bank for financial services, making the bank account financial data a rich source for understanding a business's operations, financing, and investments. Open banking enables small businesses to share their bank account data directly with potential lenders, weakening the bank's data monopoly and shifting the ownership of financial data from banks to the businesses themselves.[3] This transformation raises an important question: What is the impact on loan collateralization when borrowers can share their bank account data?

I address this question by examining the impact of the UK's open banking policy on small business collateralization. In the UK, most small business loans are collateralized with two main types of security interests, or liens: floating liens and fixed liens. A floating lien is a general security interest that covers general categories of assets, which are not individually identified and may fluctuate in quantity and value. For example, current assets like accounts receivable and inventory can be used as collateral under a floating lien, allowing the business to sell or replace these assets as needed. As a result, the actual value in the claim may change due to the turnover in accounts receivable and inventory. In contrast, a fixed lien is tied to a specific asset–such as land, property, or equipment–that cannot be sold or disposed of without the lender's permission and generally has a stable, predictable value.

[1]The financing gap for small businesses is enormous. According to the National Small Business Association, more than 25% of small businesses in the US are unable to secure adequate financing. The UK Government estimates that the financing gap for UK small businesses is approximately £22 billion. The World Bank estimates that small businesses in developing countries have an unmet financing need of $5.2 trillion, equivalent to 19% of the GDP of these countries.

[2]See Nam (2023); Babina et al. (2024); Alok et al. (2024) for the impact of open banking on consumers.

[3]Ghosh, Vallee and Zeng (2023) show that an increased use of cashless payments generates transferable and verifiable information, enabling FinTech lenders to screen borrowers more efficiently and thereby improving borrowers' access to credit. This finding provides an economic rationale for open banking.

Despite having a high proportion of current assets, small businesses face challenges in using floating liens as collateral. In my data, the average current asset ratio for small businesses is 75%, yet only 6% of these businesses manage to secure loans with floating liens in a given year. Figure 1 shows an almost zero correlation between the likelihood of using floating liens to secure loans from new lenders and the size of current assets, in contrast to a significantly positive correlation between the use of fixed liens and the size of fixed assets. In other words, the likelihood of using floating liens does not necessarily increase as current assets grow. One possible reason is that the fluctuating nature of these assets makes their values difficult to predict. Firms with lower-quality or highly volatile assets may be more likely to apply for credit, which can lead to an adverse selection problem. Additionally, floating lien collateral–such as accounts receivable or inventory–can be sold or diverted without the lender's consent, creating a moral hazard if these assets are "tunneled out." As a result, without access to detailed information on the firm's operations and transactions, lenders may hesitate to extend credit secured by floating liens.[4]

Similar to Babina et al. (2024), I exploit the UK's small business-focused open banking policy–the Commercial Credit Data Sharing (CCDS) scheme, but I depart from this work by studying loan collateralization–how loans are collateralized–in the small business lending market, which has important policy implications. The CCDS scheme mandated nine leading UK banks to share their small business customer data–including business current account data, corporate credit card data, and business loan performance data–with other UK lenders upon the customer's request. Importantly, the CCDS scheme applies only to small businesses with annual revenue up to £25 million, creating a quasi-random assignment that determines whether a business can share its bank financial data. This exogenous threshold in the data sharing eligibility allows for a sharp regression discontinuity (RD) design suitable for causal inference about the impact of open banking on collateralization.[5] To the best of my knowledge, this paper is the first to examine the causal effect of data sharing policy on loan collateralization.

I rely on two UK administrative datasets, including loan-level data from the Company Charge Register and firm-level data from the Inter-Departmental Business Register between 2015 and 2018. This novel and comprehensive dataset offers three main advantages. First, it covers nearly the entire universe of UK firms and captures the entire secured business loan market in the UK. Second, it specifies the type of collateral used–whether floating liens, fixed liens, or a combination–on each loan, allowing me to study the impact on loan collateralization. Third, it includes firm-level information, such as revenue from tax returns,

[4] This is consistent with Udell (2004), which refers to current assets like accounts receivable and inventory as "information-intensive" for lenders.

[5] The CCDS threshold of £25 million can be considered exogenous, as it differs from the primary UK definitions of an SME: €50 million, according to the UK Government, and £36 million, as defined by the UK Companies Act 2006.

which enables me to determine a firm's eligibility for data sharing under open banking.

In the first part of the paper, I develop a parsimonious model of bank lending to illustrate the role of open banking. In the model, the borrower has insufficient pledgeable assets, which creates inefficiencies in the bank's credit decisions: projects with a low but positive net present value (NPV) are sometimes rejected. The bank's likelihood of extending credit is increasing in the borrower's pool of pledgeable assets. In other words, the borrower faces collateral constraints. Open banking can help mitigate these constraints by enabling the borrower to pledge assets like accounts receivable or inventory, which were previously considered too risky or difficult to monitor. This expansion of the borrower's pool of pledgeable assets can therefore improve access to credit (see, e.g., Aretz, Campello and Marchica (2020)).

Open banking affects lending through two mechanisms. First, it enhances loan screening, mitigating ex ante asymmetric information. Access to the borrower's bank financial data provides lenders with a more comprehensive understanding of the borrower's creditworthiness, thereby reducing perceived risk. Second, it facilitates ongoing monitoring, addressing ex post asymmetric information. By accessing to the borrower's real-time bank data even after loan approval, lenders can monitor the borrower's cash flows more effectively. With these enhanced screening and monitoring capabilities, lenders may be more willing to accept current assets, such as accounts receivable and inventory, as floating lien collateral. Open banking transaction data thus serves as a form of "digital collateral" because it can mitigate both ex ante and ex post asymmetric information problem in credit markets.[6]

I begin my empirical analysis by testing the model's predictions. First, I find that firms just below the CCDS threshold (i.e., those eligible to share data) are more likely to use floating liens as collateral compared to those just above the threshold. In contrast, I do not find any significant difference for the use of non-floating liens at the threshold. These results confirm that open banking enables lenders to gain better insight into the borrower's operations and transactions, allowing borrowers to pledge current assets like accounts receivable and inventory as floating liens. While non-floating liens, typically associated with fixed assets, are less information-intensive and thus less dependent on the financial data shared through open banking. Second, I find that firms just below the threshold are more likely to get new secured loans compared to those just above the threshold. This confirms that open banking increases the likelihood of borrowers securing new loans by reducing collateral constraints and expanding the pool of pledgeable assets.

To rule out the possibility that my main findings are driven by confounding preexisting

[6] Theoretical studies motivate collateral as a screening device to mitigate adverse selection (Bester, 1985; Besanko and Thakor, 1987) and as a way of reducing moral hazard (Boot and Thakor, 1994; Holmstrom and Tirole, 1997).

differences in loan collateralization, I study the dynamic effects of open banking. I find that the discontinuity in the use of floating liens is minimal and not statistically significant before the CCDS scheme (2015–2016) but becomes significantly larger after the CCDS scheme (2017–2018). In contrast, the results for non-floating liens show no significant changes over the same period. In addition, motivated by the difference-in-difference (DiD) framework, I conduct a difference-in-discontinuity (DiDC) analysis and find that the DiDC estimates are very similar to the baseline results. These results suggest that the treatment effects of open banking are robust and not driven by confounding preexisting differences.

In the second part of the paper, I explore why open banking leads to an increase in the use of floating liens. First, I investigate whether this effect operates through an information channel. Open banking provides lenders with more comprehensive and real-time financial information about small businesses, who are often perceived as high-risk due to limited verifiable information. To explore this mechanism, I conduct three tests to examine the heterogeneous effects of open banking based on: (i) *firm age* (proxy for information frictions): I find that younger firms experience a more significant effect from open banking than older firms, suggesting that open banking mitigates information frictions for newer businesses that typically have shorter track records. (ii) *asset information-intensity* (proxy for information asymmetry)[7]: I find that open banking has a more pronounced impact on firms with high level of asset information-intensity, suggesting that open banking helps reduce information asymmetries for firms with assets that are difficult to assess and monitor. (iii) *prior lending relationships* (proxy for preexisting information availability): I find that firms with prior CCDS lending relationship experience a stronger effect from open banking compared to those with prior non-CCDS lending relationships. This suggests that prior CCDS relationships provide more comprehensive information to share with lenders, and that firms with greater information availability derive greater benefits from open banking. This is consistent with Babina et al. (2024) who show that firms with prior lending relationships are more likely to build new lending relationships. Moreover, I find that firms without prior lending relationships also benefit from open banking, suggesting that open banking helps reduce information asymmetry for firms with thin credit files. This highlights the role of open banking in promoting financial inclusion by improving access to credit for previously underserved businesses.

Second, I explore whether there is a competition channel by examining the impact of open banking on loan renewals from existing lenders. On the one hand, open banking could introduce a *substitution effect*, where firms switch from existing lenders to new lenders due to the reduced information asymmetry. By making borrower data more accessible to alternative lenders, open banking may encourage firms to seek better terms elsewhere, leading

[7] A high level of asset information-intensity indicates that a firm holds more information-intensive assets, which requires lenders to rely on up-to-date and detailed financial data to accurately assess and monitor their value, highlighting the presence of information asymmetry.

to a negative effect on loan renewals. On the other hand, there could be a *competition effect*. With the increased competition introduced by open banking, existing lenders may be more likely to renew loans to retain their customers, which suggests a positive effect of open banking on loan renewals. I find that firms just below the threshold are more likely to renew loans–regardless of the type of collateral–compared to those just above the threshold. This suggests that open banking encourages existing lenders to renew loans, thereby retaining borrowers who might otherwise seek new lenders. In other words, the competition effect dominates the substitution effect.

Third, an alternative interpretation for the increase in the use of floating liens is that different lenders have different lending technology to process the information shared through open banking. For instance, non-banks, such as FinTech lenders, may process information more efficiently due to advanced data-processing technologies. Meanwhile, banks, as current account and payment service providers, may have a better understanding of bank financial data, which allows them to utilize open banking data more efficiently. However, I find that for new loans from banks and non-banks, the RD estimates are of similar magnitude, suggesting that the effect of open banking on the use of floating liens is unlikely to be explained by differences in lending technology across lender types.

In the last part of the paper, I seek to understand the real effects of open banking on small businesses. Using the employment and establishment data from the Inter-Departmental Business Register, I find that firms just below the CCDS threshold experience a statistically significant increase in employment growth compared to those just above the threshold. This result provides evidence that open banking has a positive impact on job creation. However, I do not observe a similar pattern for establishment growth, suggesting that the effect of open banking manifests through employment growth in existing establishments rather than through the expansion of establishments.

Related Literature. This paper speaks to several strands of literature. First, it contributes to a rapidly growing literature on the economics of data and the data economy. Jones and Tonetti (2020) shows that firms may hoard their data, which leads to the inefficient use of the nonrival data, suggesting that shifting the data ownership to consumers could improve allocations. Relative to this literature, this paper shows that transferring the ownership of bank financial data from banks to small business customers can mitigate collateral constraints and thus improve access to credit. Farboodi and Veldkamp (2020), Veldkamp (2023), and Farboodi et al. (2024) examine the role of data and develop methods for measuring its value in financial markets. Abis and Veldkamp (2024) quantifies the impact of data as a key input in production functions on labor market outcomes. This paper supplements this literature by exploring the role of data in credit markets. Additionally, Acemoglu et al. (2022) document the externalities in the data economy, showing that excessive data sharing among consumers can reduce welfare. Cong, Xie and Zhang (2021) develop an endogenous growth model that illustrates how, as a data economy matures, consumers' data provision initially accelerates but eventually declines, potentially leading to growth traps. This paper connects with this literature by examining the impact of data sharing within the context of the data economy.

Second, this paper contributes to the broader discussion on the impact of information sharing through credit bureaus and credit registries. Pagano and Jappelli (1993) theoretically show that information sharing through credit bureaus can increase bank lending. Using contract-level data from a US credit bureau, Doblas-Madrid and Minetti (2013) find that information sharing reduces contract defaults but does not reduce the use of guarantees. Karapetyan and Stacescu (2014a) argue that information sharing increases the likelihood of collateral requirements, particularly for borrowers with poor credit histories. Karapetyan and Stacescu (2014b) show that information sharing increases banks' incentives to gather information about their borrowers. Brown, Jappelli and Pagano (2009), using cross-country data, demonstrate that information sharing is associated with improved credit availability and lower borrowing costs. Dierkes et al. (2013) empirically show that information sharing enhances the quality of default predictions. Liberti, Sturgess and Sutherland (2022) examine the introduction of a US credit bureau and its effects on competition and credit access, finding that lenders initially reluctant to adopt the system eventually do so to retain borrowers, leading to the formation of an information-sharing system. Darmouni and Sutherland (2021) exploit the staggered entry of lenders into an information sharing platform and find that lenders align their contract terms with those offered by competitors. Gehrig and Stenbacka (2007) analyze information sharing with repeated banking competition and argue that information sharing relaxes competition. While conceptually related to this literature, this paper focuses on data sharing through open banking. Unlike the credit file data shared through credit bureaus, open banking provides more granular, high-frequency data, including real-time transaction information. This detailed information can provide lenders with new insights into borrower behavior and improve assessments of creditworthiness. Furthermore, this paper contributes to the literature by empirically examining the causal effect of data sharing on loan collateralization, providing new evidence on how data sharing impacts credit markets.

Third, this paper expands the recent work on open banking. In traditional relationship lending (Petersen and Rajan, 1994, 1995; Berger and Udell, 1995; Bolton et al., 2016), banks acquire "soft information" about borrowers through ongoing interactions over time. Open banking, however, gives lenders access to borrowers' real-time bank transaction data, providing information comparable to that obtained through traditional relationship lending (Rishabh, 2024), which emphasizes the importance of studying open banking. Theoretically, He, Huang and Zhou (2023) explore how open banking affects credit market competition between traditional banks and FinTech lenders. Goldstein, Huang and Yang (2022) show that open banking can improve borrower welfare and financial inclusion, though it may also lead to inefficient resource allocation under certain conditions. Empirically, studies such as Babina et al. (2024) using UK data show that open banking facilitates consumer access to financial advice and credit products and helps small businesses build new lending relationships. Nam (2023) analyze the German consumer credit market and find that open banking improves credit allocation efficiency. Alok et al. (2024) find that open banking expands credit access and improves financial inclusion in India. Build on this literature, this paper provides the first empirical evidence that how open banking impacts collateralization in small business lending.

Fourth, this paper complements the large and growing literature on FinTech lending. Allen, Gu and Jagtiani (2021) and Berg, Fuster and Puri (2022) provide comprehensive literature surveys on FinTech lending, highlighting its important role in reshaping the financial landscape. Iyer et al. (2016) find that screening through soft or nonstandard information is more important when evaluating lower-quality borrowers in online lending markets. Berg et al. (2020) document that a consumer's digital footprints can predict defaults and complement traditional credit files data from credit bureaus. Gopal and Schnabl (2022), using the US secured business loan data, show that FinTech lenders increased small business lending after the global financial crisis. Fuster et al. (2022) examine the distributional impacts of new technologies, such as machine learning, in predicting mortgage defaults. Ghosh, Vallee and Zeng (2023) study FinTech lending to small businesses in India and find that the use of cashless payments improves access to credit and reduces loan defaults. Gambacorta et al. (2023) document that the wider use of big tech credit makes lending more reactive to changes in firms' business activity. Di Maggio, Ratnadiwakara and Carmichael (2022) and Ouyang (2023) explore the financial inclusion implications of alternative data in FinTech lending. This paper contributes to this literature by identifying the channel for FinTech lenders to level the playing field with banks through the use of open banking data.

Finally, this paper is connected to the literature that examines the importance of collateral. Chaney, Sraer and Thesmar (2012) document a collateral channel, showing that collateral value can impact firm investments. Adelino, Schoar and Severino (2015) examine the employment effects of collateral lending channel for small businesses. Schmalz, Sraer and Thesmar (2017) demonstrate that collateral constraints restrict firm entry and hinder growth. Calomiris et al. (2017) find that the collateral laws affect lending and sectoral activity. Using a credit reform in France that expanded the set of assets eligible for collateral, Aretz, Campello and Marchica (2020) show that expanding pledgeable assets improves credit access. Ioannidou, Pavanini and Peng (2022) study the benefits and costs of collateral by estimating a structural model. Jensen, Leth-Petersen and Nanda (2022) highlight the role of housing collateral in enabling entrepreneurship, shedding light on which types of founders are most likely to benefit. Catherine et al. (2022) quantifies the aggregate costs of collateral constraints. Consistent with this literature, this paper finds that reducing collateral constraints positively affects firm outcomes. Additionally, Manove, Padilla and Pagano (2001), Asriyan, Laeven and Martin (2022) and Biswas (2023) theoretically explore the roles of ex ante screening and collateralization in credit markets, considering them as potential substitutes or complements. Cerqueiro, Ongena and Roszbach (2016) empirically show that collateralization and ex post monitoring act as complements. This paper adds to this literature by examining the impact of open banking on collateralization, showing that open banking may complement collateral by increasing the use of floating liens.

The rest of the paper is organized as follows. Section 2 provides the institutional background. Section 3 presents the motivating framework. Section 4 describes empirical design and data. Section 5 reports the main empirical findings. Section 6 discusses the mechanism, and Section 7 discusses the real effects. Section 8 presents the falsification and robustness test of the research design, and Section 9 concludes.

2 Institutional Background

This section provides an overview of the institutional background related to open banking and the UK small business lending market. This is based on relevant laws and precedents, existing literature, and conversations with open banking policymakers, banking supervisors, and researchers at the UK Financial Conduct Authority.

2.1 Open Banking

This section provides an overview of open banking. Open banking enables bank customers to share their bank financial data with third-party financial services providers, allowing for a wide range of applications. Key use cases include credit applications, financial advice, account aggregation, identity verification, and insurance. Additionally, open banking supports the development of innovative payment services, which could potentially compete with traditional card payments and serve as alternatives to direct debits.

Global Open Banking Adoption. Open banking has become a worldwide trend. As of October 2024, over 87 countries, accounting for 95% of the world's GDP in 2023, have implemented or are actively developing open banking frameworks through government initiatives, market-driven efforts, or a combination of both (see Figure A.1). Many countries, including the US, the UK, and EU member states, have introduced regulations to foster open banking. These regulations typically mandate banks to share their customer data (with customer's consent) through application programming interfaces (APIs), enabling third-party providers to offer innovative financial services.[8] In regions without formal regulations, such as China and Switzerland, market forces play a significant role in driving open banking. In these cases, the financial institutions, particularly FinTechs, develop and manage API standards collaboratively, with limited or no government involvement. For further details on the adoption of open banking across different regions, see Internet Appendix Section A.1.

UK Open Banking Regulation. The UK has introduced two open banking policies aimed at enhancing financial data sharing and fostering competition. The UK's general open banking policy, launched by the UK Competition and Markets Authority (CMA) in 2017, mandated that by 2018, the nine largest UK banks (known as the "CMA9 Banks") enable their customers to share their bank transaction data with third parties via APIs. This initiative applies to both individual and business customers. According to Open Banking Limited (OBL)–the UK's open banking implementation entity–around 10 million UK consumers and businesses were regularly using open banking services as of July 2024, highlighting the positive momentum and growing adoption of open banking.

Prior to the general open banking policy, the UK government introduced a small business-focused open banking policy–Commercial Credit Data Sharing (CCDS) scheme–in 2017.[9] Unlike the general open banking policy which applies to all individual and business customers, the CCDS scheme specifically targets small businesses with annual revenue up to £25 million. Under the CCDS scheme, nine leading banks (known as the "CCDS9 Banks")[10] are required to share small business customer data–including business current account data, corporate credit card data, and business loan performance data–with four major Credit Reference Agencies (CRAs): Experian, Equifax, CreditSafe, and Dun & Bradstreet. The CRAs are responsible for managing, cleaning, and integrating this data with their existing credit records. Lenders can then use this completely new package of information to build comprehensive business profile, assess financial performance and risk, and make informed lending decisions.

The CCDS scheme introduces a significant innovation by mandating the sharing of business current account data (i.e., cash flows data). While corporate credit card data and business loan performance data (i.e., credit files data) were largely accessible to lenders through the CRAs, current account data had previously been exclusively held by the banks providing transaction and payment services. By allowing this data to be shared, the CCDS scheme lowers entry barriers for alternative credit providers, fostering greater competition

[8] See Babina et al. (2024) for a detailed discussion on government-led open banking adoption.

[9] See Internet Appendix Section A.1.3 for details on the development of the UK open banking regulation.

[10] The CCDS9 Banks include HSBC, Barclays, Lloyds, NatWest, Santander, Danske, Bank of Ireland, Allied Irish Bank, and *Clydesdale*, whereas the CMA9 Banks consist of HSBC, Barclays, Lloyds, NatWest, Santander, Danske, Bank of Ireland, Allied Irish Bank, and *Nationwide*. The only difference is the inclusion of Clydesdale in the CCDS9 banks and Nationwide in the CMA9 banks. This discrepancy arises because the CCDS scheme focuses on the small business lending market, while Nationwide, despite being one of the largest banks and building societies in the UK, does not provide corporate banking services.

in the small business lending market. As a result, challenger banks and alternative lenders saw a substantial increase in their market share, rising from 53% of small business lending in 2015 to 64% in 2018. This regulatory environment makes the UK an ideal setting to study the impact of open banking on small business lending.

2.2 Small Business Lending in the UK

This section provides an overview of UK small business lending market, with more details available in Internet Appendix Section A.2. Small businesses are the backbone of the UK economy. Nearly 5.7 million small businesses operate in the UK, representing 99.9% of all private sector businesses. They employ three-fifths of the private sector workforce, and produce about half of the UK's private sector revenue.[11] Despite their significant contribution to the economy, small businesses often face challenges in accessing the finance they need for growth.[12] The lending market plays a crucial role in small business financing; for example, Nanda and Phillips (2023) show that 45% of UK small businesses rely on the core financial products, such as bank overdraft, credit card, bank loan, and commercial mortgage.

As small businesses grow, they are more likely to seek external finance for expansion or capital improvements. However, because small businesses tend to be young, opaque, and perceived as risky, there is an information asymmetry between them and potential lenders: small businesses, as borrowers, have more detailed knowledge about their business's financial health, risks, and potential than the lenders do. To mitigate this asymmetric information problem, lenders often require small businesses to pledge collateral.[13] Indeed, the vast majority of UK small business lending is collateralized: the Bank of England's 2015 survey of UK SME and Mid-Corporate Lending shows that 97% of loans to firms were secured, with 72% of firms using property as security.[14]

There are two types of security interests or liens in the UK: *fixed liens* and *floating liens*. A fixed lien is a general security interest attached to a specific asset that the borrower cannot sell or dispose of without the lender's permission, or until the loan is fully repaid. Common examples of fixed lien assets include land, property, and heavy machinery. In contrast, a floating lien typically covers general categories of assets that are not individually identified

[11] See UK Department for Business, Innovation and Skills: Business Population Estimates (2018). Small businesses are defined as private sector businesses with fewer than 250 employees.

[12] The UK Government estimated that the gap between the funding required by small businesses and the funding available was approximately £22 billion by 2017.

[13] Empirical studies show that collateral can mitigate adverse selection and moral hazard (Ioannidou, Pavanini and Peng, 2022; Berger, Frame and Ioannidou, 2011).

[14] This survey covered loans from the five major banks to firms with annual revenue up to £500 million and borrowing at least £250,000. The data is based on responses to the question: "*Does your bank hold any of the following as collateral?*" with multiple options available: (a) property; (b) debenture including charges over plant, equipment and vehicles; (c) cash or cash equivalent; (d) other tangible collateral/security; (e) personal guarantee. See Bahaj, Foulis and Pinter (2020) for more details on the survey.

and may fluctuate in quantity and value over time. This flexibility allows the borrower to use, sell, or replace these assets in the normal course of business without needing the lender's approval. Examples of floating lien assets include accounts receivable and inventory. To illustrate, one might think of the borrower as a farmer, where the fixed lien asset is his land and the floating lien asset is the fruit that grows on that land (Kiyotaki and Moore, 1997).

Lenders often prefer fixed liens over floating liens for several reasons. First, in the event of a company's bankruptcy, fixed liens take precedence over floating liens, giving lenders a better chance of recovering their funds. Second, fixed liens create a strong incentive for borrowers to repay their loans, as losing access to essential assets like properties, vehicles, or equipment can significantly disrupt business operations. Third, the fluctuating nature of floating lien assets makes their values difficult to predict. Firms with lower-quality or highly volatile assets may be more likely to apply for credit, potentially leading to an adverse selection problem. Lastly, assets under fixed liens cannot be sold or disposed of without the lender's approval, whereas assets under floating liens can be sold or diverted by the borrower without permission, posing a moral hazard risk.

In the UK, secured loans backed solely by floating liens is very rare in practice–especially among small businesses–with less than 10% of secured business loans relying exclusively on floating liens. Around 90% of secured loans are backed by fixed liens, while a significant portion (about 50%) are backed by a combination of fixed liens and floating liens.

3 Motivating Framework

In this section, I present a model of small business lending in which a bank makes its accept or reject credit decision. This model builds upon Inderst and Mueller (2007) and is simplified to include only the key elements necessary for studying the role of open banking. This section also serves as motivation for the empirical analysis.

Model Setup. Consider a risk-neutral economy in which a firm (the borrower) has a risky project with fixed investment cost $I > 0$. The project yields $X > 0$ if successful and zero otherwise. The borrower has assets W, of which a proportion α is pledgeable assets and the remaining $1-\alpha$ is non-pledgeable assets. I assume that $W < I$, so the project cannot be financed by issuing a safe claim. The risk-free interest rate is normalized to zero.

In this credit market, a monopolistic bank makes a credit decision based on its private information about the borrower. I assume that the bank's assessment of the borrower's project can be represented by (i) its internal rating of the borrower, $z \in [0,1]$, and (ii) the associated probability of success based on the rating, $p_z \in [0,1]$.[15] The probability

[15]Because banks typically do not disclose the internal ratings, the estimated probabilities of default, or the methods used to calculate them, I assume that z and p_z are private information known only to the bank.

of success p_z is strictly increasing in the rating z, implying that the conditional expected project cash flow $\pi_z = p_z X$ is also strictly increasing in z. I assume that $\pi_0 - I < 0$ and $\pi_1 - I > 0$, that is, the project's NPV is negative for low z and positive for high z. This allows the bank to differentiate between positive- and negative-NPV projects based on its private information about the borrower.

There are three stages. In stage 1, the bank offers a loan contract (R, C), where $R \geq 0$ represents the repayment the bank obtains when the project succeeds, and $C \geq 0$ is the collateral the bank can liquidate when the project fails. In stage 2, if the borrower chooses the offer, the lender evaluates the borrower's project. If the borrower is accepted, he obtains financing based on the terms of the contract (R, C). If the borrower is rejected, he can still seek non-bank financing. In stage 3, the project cash flow is realized, and the bank either receives the agreed repayment or liquidates the collateral.

Credit Decision. I begin with the first-best optimal credit decision. Because the project's NPV is positive for high z and negative for low z, and the conditional expected project cash flow π_z is strictly increasing in z, there exists a unique first-best cutoff $z^{FB} \in (0, 1)$ such that the project's NPV is zero, that is, $\pi_{z^{FB}} = I$. The first-best credit decision is therefore to accept the loan application if, and only if, $z \geq z^{FB}$, or equivalently,

$$p_z \geq p_{z^{FB}} = \frac{I}{X} \tag{1}$$

I next derive the bank's optimal credit decision. The bank accepts the loan application if, and only if, its conditional expected payoff

$$U_z(R, C) = p_z R + (1 - p_z)C \tag{2}$$

is equal to or greater than I. Because p_z is strictly increasing in z, $U_z(R, C)$ is also strictly increasing in z. As a result, there exists a unique cutoff z^* such that the bank's conditional expected payoff is equal to I, that is, $U_{z^*}(R, C) = I$. The bank's credit decision is therefore to accept the loan application if, and only if, $z \geq z^*$, or equivalently,

$$p_z \geq p_{z^*} = \frac{I - C}{R - C} \tag{3}$$

Lender's Problem. The bank chooses the loan repayment (R) and the collateral (C) for the borrower to maximize its profit:

$$\max_{R,C} U(R, C) = \int_{z^*}^{1} [U_z(R, C) - I] f(z) dz \tag{4}$$

subject to the condition $U_{z^*}(R, C) = I$ characterizing the lender's optimal credit decision and the borrower's participation constraint:

$$V(R,C) = \int_{z^*}^{1} V_z(R,C) f(z) dz \geq V_0 \tag{5}$$

where $V_z(R,C) = p_z(X-R)-(1-p_z)C$ is the borrower's conditional expected payoff. $f(z)$ is the density associated with z. V_0 represents the reservation utility of the borrower, as he has access to non-bank financing, so $V_0 > 0$.[16] Since the bank is monopolistic, it receives all the borrower's surplus in excess of V_0, therefore, the borrower's participation constraint must bind.

Because $V_0 > 0$, the bank cannot extract the full project cash flow, so its conditional expected payoff $U_z(R,C)$ is less than the conditional expected project cash flow π_z for all z. This implies that $U_{z^{FB}}(R,C) < \pi_{z^{FB}} = I$, that is, the bank does not break even at $z = z^{FB}$. Given that $U_z(R,C)$ is strictly increasing in z, the bank's optimal cutoff exceeds the first-best cutoff, that is, $z^* > z^{FB}$, implying that the bank rejects projects with a low but positive NPV.

Collateral can reduce this inefficiency. To see why optimally increasing collateral can achieve the first-best optimum. Consider a loan contract (R,C), increasing collateral C lowers z^*, pushing it closer to the first-best cutoff z^{FB}. However, to satisfy the borrower's binding participation constraint, the repayment R must be reduced at the same time. This adjustment raises z^*, pulling it further from z^{FB}. Although the two adjustments have opposite impacts on the bank's optimal cutoff z^*, Internet Appendix Section B shows that the overall impact is that z^* is pushed down. When collateral is sufficiently large, the bank's optimal cutoff z^* equals the first-best cutoff z^{FB}.

To solve the lender's profit maximization problem, I consider two cases:

(i) *Sufficient pledgeable assets*: If the borrower has sufficient pledgeable assets, then the first-best optimum can be achieved with a unique loan contract (R^{FB}, C^{FB}), which is jointly determined by the borrower's binding participation constraint:

$$V(R^{FB}, C^{FB}) = \int_{z^{FB}}^{1} V_z(R^{FB}, C^{FB}) f(z) dz = V_0 \tag{6}$$

and the condition characterizing the bank's optimal credit decision:

$$p_{z^{FB}} R^{FB} + (1-p_{z^{FB}}) C^{FB} = I \tag{7}$$

where $p_{z^{FB}}$ is defined in Equation (1). Solving these two equations yields the optimal loan

[16] If $V_0 = 0$, the bank will offer a loan contract $(X, 0)$ to extract the full project cash flow, which is first-best optimal, implying that there is no need for collateral.

contract:

$$R^{FB} = X - \frac{V_0(X-I)}{\int_{z^{FB}}^1 (p_z X - I) f(z) dz}, \quad C^{FB} = \frac{V_0 I}{\int_{z^{FB}}^1 (p_z X - I) f(z) dz}, \tag{8}$$

Because the borrower has sufficient pledgeable assets $\alpha W \geq C^{FB}$, the bank's optimal cutoff z^* can be pushed down towards the first-best cutoff z^{FB}. Projects with a low but positive NPV will be accepted, thereby increasing the total surplus. Because the borrower's participation constraint must bind, this additional surplus entirely accrues to the bank.

(ii) *Insufficient pledgeable assets*: If the borrower has insufficient pledgeable assets $\alpha W < C^{FB}$, then the bank's credit decision is inefficient, it rejects projects with a low but positive NPV. Given that z^* is decreasing in C for all $C < C^{FB}$, the optimal loan contract requires the borrower to use all pledgeable assets as collateral, that is, $C^* = \alpha W$. The optimal repayment R^* is determined by the borrower's binding participation constraint:

$$V(R^*, \alpha W) = \int_{z^*}^1 V_z(R^*, \alpha W) f(z) dz = V_0 \tag{9}$$

and the condition characterizing the bank's optimal credit decision:

$$p_{z^*} R^* + (1 - p_{z^*}) \alpha W = I \tag{10}$$

The interesting case is when the borrower has insufficient pledgeable assets, as this situation is more common in small business lending.[17] Therefore, I focus on this case in the following discussion.

The Role of Open Banking. Small businesses often face credit constraints because the collateral constraints bind. Open banking can help mitigate these constraints by enabling borrowers to pledge assets like accounts receivable or inventory, which were previously considered too risky or difficult to monitor. This effectively expands the pool of pledgeable assets, thereby increasing the proportion of pledgeable assets α. There are two key mechanisms:

(i) *Loan screening*: By accessing the borrower's bank current account data through open banking, lenders can obtain a more detailed understanding of the borrower's creditworthiness. This reduces the perceived risk, since current account data substantially improves default predictions (Norden and Weber, 2010). As a result, lenders may be more willing to accept a wider range of assets–such as accounts receivable and inventory–as collateral, which they might have previously declined due to the lack of reliable information. This implies that open banking can address ex ante asymmetric information such as adverse selection.

[17] The presence of collateral constraints for small businesses is widely recognized, see, e.g., Degryse and Van Cayseele (2000); Jimenez, Salas and Saurina (2006) for Europe, Berger and Udell (1995); Benmelech, Kumar and Rajan (2024) for US, and Hanedar, Broccardo and Bazzana (2014) for less-developed countries.

(ii) *Ongoing monitoring*: Open banking gives lenders access to the borrower's real-time bank data even after the loan has been approved, allowing them to monitor the borrower's cash flows more effectively (Mester, Nakamura and Renault, 2007), especially in the event of the borrower's financial distress or bankruptcy. With enhanced monitoring capability, lenders can reduce the potential costs associated with moral hazard, increasing their willingness to accept more current assets like accounts receivable and inventory as collateral. This suggests that open banking can mitigate ex post asymmetric information such as moral hazard.

Open banking transaction data thus serve as a form of "digital collateral" because it can mitigate both ex ante and ex post asymmetric information problem in credit markets.[18] This is similar to Ghosh, Vallee and Zeng (2023) who demonstrate that cashless payment records can also function as a form of "digital collateral."

Prediction 1. *Open banking leads to an increase in the use of floating liens.*

Lending Outcomes. With open banking, the borrower's pledgeable assets increase to $\hat{\alpha}W$, where $\hat{\alpha} > \alpha$. Given that z^* is decreasing in C for all $C < C^{FB}$ and the optimal contract requires the borrower to pledge all the pledgeable assets, the bank's optimal cutoff is lower under open banking, that is, $\hat{z}^* < z^*$. This implies that $F(\hat{z}^*) < F(z^*)$, suggesting that the bank is more likely to grant credit.

Prediction 2. *Open banking increases the likelihood of borrowers obtaining credit.*

4 Research Design and Data

4.1 Empirical Strategy

Empirically identifying how open banking affects collateral outcomes is challenging for three reasons. First, it requires a setting where the econometrician can observe whether a firm is treated by open banking.[19] Second, beyond the information shared through open banking, the information gathered through lending relationships can also affect collateral outcomes. For example, relationship lending allows lenders to gather private information over time, which can lead to relaxed collateral requirements, as suggested by Jimenez, Salas and Saurina (2006); Bharath et al. (2011). Third, credit demand also plays a role in collateral requirements. For example, Dell'Ariccia and Marquez (2006) demonstrate a negative relationship between credit demand and collateral requirements.

[18] This notion of "digital collateral" is different from that of Gertler, Green and Wolfram (2024), who show that the lender can digitally disable the functional value of collateral assets, such as solar panels, without physically repossessing them.

[19] In most cases of open banking, all bank customers are generally treated and eligible to share data, with no specific restrictions on who can participate. To mitigate the risk of selection bias, I do not simply compare the firms that share data with those that do not. Instead, I compare the firms that are eligible to share data with those that are not. Therefore, the treatment is defined as eligibility to share data through open banking.

To address these challenges, I first exploit the institutional features of the UK's SME-focused open banking policy–the CCDS scheme introduced in Section 2.1. Although the UK's general open banking policy covers all individual consumers and businesses, only firms with annual revenue up to £25 million are subject to the CCDS scheme and can share their bank data via the CRAs. This creates a quasi-random assignment of whether a firm is eligible to share data. One might be concerned that the eligible firms may not consent to sharing their data. Indeed, data sharing in principle requires the firm's consent (a key component of open banking), however, this consent typically occurs automatically as it is often embedded in the standard contract terms with lenders in the pre-screening stage.[20]

Second, my analysis focuses on the collateralization of loans from new lenders, allowing me to exclude the impacts of relationship lending. Following Ioannidou and Ongena (2010), I define a *new lender* as one with which a firm did not have a lending relationship in the prior 12 months. Conservatively, I assume that key inside information can become stale within 12 months, motivated by empirical evidence suggesting that bank account activity, which provides a real-time windows into the borrower's cash flows, exhibits abnormal patterns approximately 12 months before default events (Norden and Weber, 2010). In Table D.3, I show that my main results remain robust when using 24- and 36-month cutoffs.

Third, I exploit the sharp discontinuity in eligibility to share data through open banking to estimate its casual effect on collateralization. The firm's *normalized revenue*–the amount by which a firm's revenue falls below the CCDS threshold of £25 million–serves as the running variable. Specifically, I examine whether firms just below the CCDS threshold exhibit significant differences in the collateralization of loans from new lenders compared to those just above the threshold. Since firms near the threshold are similar in terms of fundamental characteristics and credit demand, this RD design allows me to isolate the effect of open banking on collateral outcomes, while holding the fundamental characteristics and credit demand constant.

The discontinuity can be estimated using both parametric and nonparametric methods. Gelman and Imbens (2019) argue that the parametric RD approach, which includes higher-order polynomial functions of the running variable in the regression, often produces RD estimates that are noisy, highly sensitive to the degree of the polynomial, and lead to poor statistical inference. Therefore, I employ the recommended non-parametric local linear approach and estimate the following equation:

$$Y_{i,t} = \alpha + \tau \mathbb{D}_{i,t} + \beta R_{i,t} + \gamma R_{i,t} \cdot \mathbb{D}_{i,t} + \mu_{j,t} + \nu_{r,t} + \varepsilon_{i,t} \tag{11}$$

[20]This is not always true for individual consumers, who are generally more concerned about their personal privacy and therefore more reluctant to consent to sharing data. For example, Babina et al. (2024) find that consumers with concerns about data sharing are less likely to use open banking services. They may be wary of how their financial information will be used and who will have access to it.

where $Y_{i,t}$ is an indicator variable that equals one if firm i uses floating liens to secure loans from new lenders in year $t = 2017$ or 2018, and zero otherwise. $R_{i,t}$ represents firm i's normalized revenue over the 12 months preceding year t, defined as the amount by which the firm's revenue falls short of the CCDS threshold £25 million. $\mathbb{D}_{i,t}$ is an indicator variable that equals one if firm i's normalized revenue is greater than or equal to zero in year t, and zero otherwise. The term $R_{i,t}$ captures the baseline relationship between the running variable and the outcome variable; and the interaction term $R_{i,t} \cdot \mathbb{D}_{i,t}$ allows this relationship to differ on either side of the threshold, providing a more flexible model to account for different slopes in the treatment and control groups.

To capture the industry- and location-specific effects and temporal effects in the non-parametric estimations, I include industry-by-year and geographic region-by-year fixed effects $\mu_{j,t}$ and $\nu_{r,t}$ in the baseline model. The estimation of this nonparametric RD model with fixed effects is implemented using the two-step approach suggested by Lee and Lemieux (2010): first running an OLS regression of outcome variable on a set of industry-by-year and region-by-year dummy variables, and then applying the nonparametric estimation to the residualized outcome variable. This method can improve the precision of the estimates without introducing bias.

Standard errors, $\varepsilon_{i,t}$, are clustered at the industry-by-region level to deal with the potential correlation of the error term within each industry and geographic region, as suggested by Cameron and Miller (2015). To deal with the bias-variance trade-off problem[21], I choose the MSE-optimal bandwidth following Calonico, Cattaneo and Titiunik (2014) for triangular kernel weighting method. The magnitude of the discontinuity, measured by τ, is estimated by the difference in the predicted values of the outcome variable at the cutoff. This coefficient should be interpreted locally in the immediate vicinity of the cutoff.

The key identifying assumption is that, in the absence of treatment, all relevant factors vary smoothly at the cutoff. Formally, let Y^1 and Y^0 denote potential outcomes under treatment and control, identification for the RD model requires that $\mathbb{E}[Y^1|R]$ and $\mathbb{E}[Y^0|R]$ are continuous (smooth) at the cutoff. This assumption is needed for firms just above the cutoff to serve as an appropriate counterfactual for those just below the cutoff. Although this assumption cannot be directly tested, indirect evidence can be persuasive. In Sections 4.4 and 8.1, I conduct five tests to support the credibility of this underlying assumption as well as the validity of the RD design.

I augment the baseline econometric specification in several different ways: (i) adding covariates; (ii) using quadratic RD polynomial of the running variable; (iii) using different kernel functions: Triangular, Epanechnikov, Uniform; (iv) selecting optimal bandwidth

[21] On one hand, selecting a larger bandwidth includes more data points for estimation, reducing the variance of the estimated coefficients but introducing more smoothing bias (also known as "misspecification error"). On the other hand, a smaller bandwidth decreases smoothing bias but increases variance, as it includes fewer data points.

based on two sides; (v) clustering the standard errors by the industry-by-region-by-year fixed effects; (vi) using bias correction methods. As will be elaborated in more detail in Section 8.2 Specification Curve Analysis, my main findings hold across all of these alternative models.

4.2 Data

I use a novel and comprehensive dataset of UK businesses by linking two proprietary administrative datasets: loan-level data from the Company Charge Register and firm-level profile data from the Inter-Departmental Business Register. This dataset is further enriched with firm-level financial accounting data from Bureau van Dijk. To the best of my knowledge, my paper is the first to put together a comprehensive dataset that covers nearly the entire universe of UK firms and captures the entire secured business loan market in the UK.

4.2.1 Loan Data

Company Charge Register (CCR). The CCR is a comprehensive credit register database maintained by Companies House, the UK government's executive agency responsible for incorporating all forms of companies in the UK. Companies House holds records for 11.3 million UK companies and the charges (or "claims") placed against their assets. According to the UK Companies Act 2006, all claims must be registered with Companies House using the MR01 form within 21 days, starting from the day after the claim is created. If this deadline is missed, the lender will need a court order to extend the registration period. The cost of registration is small: £15 for online filings and £24 for paper filings. Failure to register the claim within the prescribed time frame leads to severe legal consequences, effectively making the registration of company claims mandatory.[22]

The CCR data has three advantages. First, it captures nearly the entire secured business loan market in the UK, which helps mitigate the selection bias.[23] Second, it provides detailed information on the type of collateral used–fixed liens, floating liens, or a combination–along with the underlying assets, allowing me to analyze the borrower's reliance on floating liens or non-floating liens when securing loans from new lenders. Third, it provides contract terms, such as the instrument date, identities of borrowers and lenders,

[22]Failure to register a claim at Companies House has several consequences. First, it is a criminal offense for an interested personal not to register a company claim. Second, if a claim is not registered within 21 days, it becomes void and unenforceable against a liquidator, administrator, and any creditor of the company. This means the lender is still owed money by the company but is now unsecured, sharing the same level of losses as other unsecured creditor in the event of the company's bankruptcy. Third, when a claim becomes void, the money secured by it becomes immediately repayable by the company, potentially leading to bankruptcy and losses.

[23]Notably, the CCR data covers real estate lending–a common practice in the UK (for example, Bahaj, Foulis and Pinter (2020) show that about 70% of UK small businesses use corporate and residential property as security for business loans); while a similar secured business loan dataset commonly used in the US–the Uniform Commercial Code (UCC) data–does not cover real estate business loans.

as well as loan performance information such as satisfaction status and date.[24] However, one limitation is that it does not cover unsecured lending market. Nevertheless, as discussed in Section 2.2, the overwhelming majority of small business lending in the UK is secured, making the CCR data a highly representative data source for analyzing the small business lending market.

4.2.2 Firm Data

Inter-Departmental Business Register (IDBR). The IDBR is a comprehensive statistical business register database maintained by the Office for National Statistics (ONS), the executive office of the UK Statistics Authority, which is responsible for collecting and publishing official statistics in the UK. It serves as the primary administrative data source for government departments analyzing UK business activity. The IDBR is primarily created and maintained using data from two main tax returns: Value Added Tax (VAT) and Pay As You Earn (PAYE) records provided by His Majesty's Revenue and Customs (HMRC), the UK tax authority. Additional administrative and commercial data from Companies House, Dun & Bradstreet, along with ONS business surveys such as the Annual Business Survey and the Business Register Employment Survey, are used to supplement the IDBR.

The IDBR data covers approximately 3 million businesses across all sectors, representing 99% of UK economic activity. A key advantage of the IDBR data is that it reports firm revenue, which serves as the running variable in the RD design. Additionally, the IDBR data provides essential business profile information for my analysis, including address, entry and exit dates, industry, employment, number of establishments, legal status, and business structure. However, due to the thresholds for VAT and PAYE tax returns, very small businesses (e.g., sole traders) operating below these thresholds are not included.[25]

4.2.3 Other Data

However, since the IDBR data does not include other financial accounting information, I supplement this firm-level data with Bureau van Dijk's Orbis database.

Bureau van Dijk (BvD). The primary source of firm-level financial accounting data is BvD, a Moody's Analytics company. For the UK, BvD provides firm-level data through a product called FAME, the UK-specific version of BvD's Orbis database.[26] The BvD data includes balance sheet and profit-and-loss account information, sourced from publicly available fil-

[24]Unfortunately, pricing information and loan amount are not always consistently reported in the CCR.

[25]As of July 2024, the VAT threshold is set at £90,000 in taxable revenue over any 12-month period. If a business's taxable revenue exceeds this threshold, it must register for VAT with HMRC and submit VAT returns. The PAYE threshold is £123 per week. If any employee within a business is paid £123 or more per week, the business must register for PAYE with HMRC and submit PAYE returns.

[26]For more details about the BvD data, see Bahaj, Foulis and Pinter (2020) for the UK version and Kalemli-Özcan et al. (2024) for the global version of BvD's Orbis database.

ings at Companies House. All private limited and public firms are required to file statutory accounts (e.g., financial records) with Companies House, although reporting requirements vary by firm size.[27] A key advantage of the BvD data is that it can be perfectly linked with the CCR data using the unique Company Registration Number (CRN). However, one limitation is that many small businesses have incomplete financial data due to being exempt from certain reporting requirements.

Postcode Lending Data (PLD). The PLD is an administrative dataset on small business lending at the postcode sector level, reported quarterly by nine leading UK banks (the CCDS9 Banks). It provides data on outstanding balances for small business borrowing through business loans and overdrafts (excluding business credit card), representing around 60% of the total national market for small business lending by banks and building societies. One limitation of the PLD is that, due to customer confidentiality and data privacy rules, some lending values are redacted and thus not reported at the postcode sector level. However, these redacted figures are aggregated and reported at the postcode area level. For greater reliability, I use the aggregate postcode area-level PLD data.

4.3 Sample Selection and Summary Statistics

This section briefly discusses the sample selection, with further details are provided in Internet Appendix Section C.2. Although the CCDS scheme targets 99% of the entire business population in the UK, I focus on a narrower sample of firms with revenue between £10.2 million and £36 million for more clean identification. This strategy ensures that the information structure of the business lending market remains consistent for both the treatment group (firms with revenue between £10.2 million and £25 million) and the control group (firms with revenue between £25 million and £36 million) in the absence of the CCDS scheme.[28]

My sample includes nearly the entire universe of private sector firms in the UK with revenue between £10.2 million and £36 million over a four-year period: 2015–2016 (pre-policy) and 2017–2018 (post-policy).[29] Following the literature, I exclude financial firms and organizations involved in certain service activities. This yields a baseline sample of 67,143 firm-year observations (31,739 observations in the pre-policy period and 35,404

[27] A small firm can prepare and submit accounts according to special provisions in the UK Companies Act 2006 and the relevant regulations. This allows them to disclose less information than medium and large firms. For example, small firms typically submit an abridged version of their statutory accounts, which contains less detail by omitting certain balance sheet items, and they may also be exempt from audit requirements.

[28] According to the UK Companies Act 2006, small firms are defined as having annual revenue of no more than £10.2 million, while medium firms have revenue up to £36 million. Firms within the £10.2 million to £36 million revenue range are thus subject to the same information disclosure requirements under the Companies Act 2006. Other definitions of company size exist in the UK (see Table C.1), but the Companies Act 2006 is more relevant to corporate reporting and information disclosure standards.

[29] Data from 2019 onwards is excluded to avoid the potential confounding effects from the COVID-19 pandemic and the UK's general open banking policy.

observations in the post-policy period) and 25,215 unique firms. The exact sample size for each specification is reported in the regression tables.

Table 1 presents summary statistics on firm collateralization, characteristics and financial metrics. Panel A shows the likelihood of different collateral types used in new loan agreements. On average, firms are 6.02% likely to use floating liens, whereas non-floating liens are employed in only 2.70% of cases. The overall probability of securing new loans using any type of lien is 8.46%, suggesting that a relatively small fraction of firms secure new loans with collateral within a given year.

Panel B presents key firm characteristics. A substantial proportion of firms (81.66%) report revenue below £25 million, this indicates that the majority fall within the eligibility criteria of the CCDS scheme. The average firm age is approximately 25 years, highlighting a balanced representation of both younger and more mature firms. Furthermore, a high percentage (82.34%) of firms have established lending relationships, with 73.96% specifically having relationships with CCDS lenders, suggesting the prominent role of CCDS lenders in the small business lending market. Employment and establishment growth rates are modest, averaging 3.48% and 1.61%, respectively, which implies limited expansion in terms of workforce and physical footprint.

Panel C focuses on predetermined financial variables for the post-policy sample. The average leverage ratio is 60.39%, indicating a considerable dependence on debt financing among firms in the sample. Cash holdings represent 14.53% of total assets, while current assets account for 74.50% of total assets. This high proportion of current assets implies that firms possess a substantial amount of assets–such as accounts receivable and inventory–that could be pledged as collateral in the form of floating liens.

4.4 Validity of the RD Design

Covariate Balance Test. To rule out the possibility of discrete jumps in the observable characteristics of firms just below versus just above the cutoff, I test the smoothness of covariates across the cutoff. Specifically, I apply the standard RD methodology at the firm level for six predetermined firm characteristics: (i) log of total assets, (ii) leverage ratio, (iii) cash, (iv) current assets, (v) bank overdraft, and (vi) trade credit. Figure 2 presents the RD plots for each of these characteristics, suggesting that the observable characteristics of firms are continuous at the cutoff. The RD estimates, reported in Table D.11, are all very small and not statistically significant. In other words, there is no empirical evidence indicating that these predetermined covariates are discontinuous at the cutoff.[30]

Manipulation Test. To rule out the possibility that firms could precisely manipulate their revenues in tax returns, I begin by presenting a histogram of the running variable (normalized revenue) in Figure 3 Panel A, which suggests that the number of firms above and below the threshold is very similar. I then conduct a test of the null hypothesis that the density of the running variable is continuous at the threshold. Specifically, I implement the density test outlined in Cattaneo, Jansson and Ma (2020) based on local polynomial density estimators. The value of the t-statistics is 0.2475 and the associated p-value is 0.8045. This means I fail to reject the null hypothesis of no manipulation of the density at the threshold. Figure 3 Panel B further confirms this, showing no discontinuity in the density of observations at the threshold. The density estimates for treated and control group at the threshold (the two intercepts in the figure) are very close to each other, with overlapping confidence intervals (shaded areas), suggesting that no revenue manipulation by firms and ruling out other sources of bunching in the distribution of firms.

5 Baseline Results

5.1 Effects of Open Banking on Loan Collateralization

I begin the empirical analysis by visualizing the main findings. Figure 4 plots the firm-level collateralization of new loans during post-policy period (2017–2018) against normalized revenue. Each dot represents the average residualized likelihood of using collateral for firms within a specific bin of normalized revenue. A fitted line with 95% confidence intervals is overlaid to highlight the discontinuity around the threshold.

In Figure 4, Panel A presents the RD plot for floating liens, showing a sharp increase in the likelihood of using floating liens at precisely the CCDS scheme threshold. Firms just below the threshold are significantly more likely to use floating liens compared to those just above the threshold. In contrast, Panel B shows no significant discontinuity in the likelihood of using non-floating liens between firms just above and below the threshold. This suggests that open banking leads to an increase in the use of floating liens, aligning with the model prediction 1 in Section 3.

Table 2 quantifies the graphic findings in Figure 4. The estimate in column (1) is 0.0447, significant at the 1% level, indicating a sharp discontinuity at the CCDS threshold. Firms just below the threshold are 4.47 percentage points more likely to use floating liens compared

[30]Note that I use a bandwidth of 4 million across tests for consistency and ease of interpretation. However, for falsification purposes it may be more appropriate to choose the bandwidth to obtain confidence intervals minimizing the coverage error probability (CER)–i.e., CER-optimal bandwidth (Calonico, Cattaneo and Farrell, 2018)–because the primary interest is in testing the null hypothesis of no effect, while the point estimates are of no particular interest. As shown in Figure D.2, using CER-optimal bandwidth does not significantly change my empirical results.

to those just above the threshold. In column (2), after controlling for the industry-by-year and region-by-year fixed effects, the RD estimate slightly decreases to 0.0440, but remains significant. This suggests that the open banking increases the likelihood of firms using floating liens by about 4.4 percentage points, which represents a 73% increase compared to the average.

In contrast, columns (3) and (4) show no significant discontinuity for non-floating liens at the threshold. The RD estimates are small and not statistically significant, suggesting that the effect of open banking is specific to floating liens rather than non-floating liens. Moreover, as shown in Figure D.1 and Table D.1, the RD estimates for new loans (secured by any type of lien) are of similar magnitude to those for floating liens and are statistically significant. This suggests that open banking increases the likelihood of borrowers securing new loans, consistent with the model prediction 2 in Section 3.

In sum, the baseline results suggest that open banking reduces collateral constraints for small businesses by increasing their likelihood of using floating liens. This expansion of pledgeable assets enhances their borrowing capacity and improves access to credit.

5.2 Dynamic Effects

As discussed in Section 2.1, the CCDS scheme was launched in 2017. I hypothesize that the difference in the use of floating liens between firms just above and below the CCDS threshold should be pronounced after 2017. To fully examine the dynamics of this baseline discontinuity, I run the following RD regression for each year from 2015 to 2018:

$$Y_{i,t} = \alpha_t + \tau_t \mathbb{D}_{i,t} + \beta_t R_{i,t} + \gamma_t R_{i,t} \cdot \mathbb{D}_{i,t} + \mu_{j,t} + \nu_{r,t} + \varepsilon_{i,t} \tag{12}$$

Figure 5 plots the RD estimates separately for each year. I find that the discontinuity in the use of floating liens is very small and statistically insignificant from 2015 to 2016 but becomes significantly larger in 2017 and 2018. This highlights how open banking facilitates a shift towards more flexible forms of collateral, particularly through the use of assets like accounts receivable and inventory as floating liens. In contrast, the results for non-floating liens show no significant changes over the same period, suggesting that open banking has no significant impact on the use of non-floating liens. The corresponding regression results are summarized in Table D.2.

The absence of a significant discontinuity prior to 2017 is consistent with the balance tests and further supports the validity of the identifying assumption: in the absence of open banking, firms just above and below the CCDS threshold exhibited similar patterns in their use of collateral. The dynamic pattern of the RD coefficients is also reassuring, as it helps rule out alternative explanations: one might be concerned that the baseline results can be driven by certain confounding factors, such as factors would have to be specific not only to firms above and below the CCDS threshold or the use of floating liens versus non-floating liens, but also coincide specifically with the timing of the CCDS scheme's introduction to explain the observed discontinuity.

5.3 Difference-in-Discontinuities Analysis

Motivated by the "break in trends" between firms above and below the CCDS threshold in 2017, I adopt a difference-in-discontinuities (DiDC) framework following Grembi, Nannicini and Troiano (2016) to examine changes in the use of collateral before and after the introduction of the CCDS scheme in 2017. Specifically, I estimate the following DiDC regression:

$$\begin{aligned} Y_{i,t} = \alpha_1 + \tau_1 \mathbb{D}_{i,t} + \beta_1 R_{i,t} + \gamma_1 R_{i,t} \cdot \mathbb{D}_{i,t} + \tau_2 \mathbb{D}_{i,t} \cdot \text{Post}_t + \beta_2 R_{i,t} \cdot \text{Post}_t \\ + \gamma_2 R_{i,t} \cdot \mathbb{D}_{i,t} \cdot \text{Post}_t + \mu_{j,t} + \nu_{r,t} + \varepsilon_{i,t} \end{aligned} \tag{13}$$

where Post_t is a dummy variable that equals one if $t = 2017$ or 2018, and zero if $t =$ 2015 or 2016. In this specification, the coefficient, τ_2, on the interaction between the time dummy and the indicator for being below the threshold, measures the changes in the use of collateral due to the firm moving from above to below the threshold before and after the CCDS scheme.

The intuition of the difference-in-discontinuities design is very similar to that of the difference-in-differences design. In this framework, a pre-treatment RD identifies the pre-existing or time-invariant effects of other treatments or sorting that occur near the threshold. A post-treatment RD captures those preexisting effects plus the effect of the treatment of interest. By taking the difference between the pre- and post-treatment RD effects, this method isolates the causal impact of the treatment of interest, filtering out any confounding effects unrelated to the treatment itself.

Table 3 reports the difference-in-discontinuities estimates. In columns (1)–(2), I find that firms just below the CCDS threshold are about 4 percentage points more likely to use floating liens after the introduction of the CCDS scheme, compared to firms just above the threshold. In contrast, as shown in columns (3)–(4), there is no significant discontinuity between firms above and below the threshold in the use of non-floating liens. The estimates are very similar to the baseline results presented in Table 2, suggesting that the treatment effects of open banking are robust and not driven by confounding preexisting differences.

6 Mechanisms

In this section, I test several mechanisms that could possibly explain why open banking affects the collateralization of new loans.

6.1 Information Channel

Open banking provides lenders with more comprehensive and real-time financial information about small businesses, a group considered high-risk due to the limited availability of verifiable information. To explore the information channel of open banking, I use three measures to assess the heterogeneous effects on loan collateralization. First, I examine the heterogeneous effect by firm age (as a proxy for information frictions). I find that younger firms are more significantly impacted by open banking than older firms. Second, I investigate the heterogeneous effect by asset information-intensity by comparing firms with high versus low levels of information-intensive assets. I find that open banking has a more pronounced impact on firms with high level of asset information-intensity. Lastly, I compare firms with and without prior (CCDS) lending relationships. I find that firms with prior CCDS lending relationship experience a stronger effect from open banking compared to those with prior non-CCDS lending relationships, suggesting that prior CCDS relationships provide more comprehensive information to share with lenders. Moreover, I find that firms without prior lending relationships also benefit from open banking, highlighting the role of open banking in promoting financial inclusion. Together, these results confirm the information channel of open banking.

6.1.1 Firm Age

Finance literature often interprets firm age as a proxy for information frictions in bank lending, as younger firms tend to have shorter track records and less publicly available information, making it more challenging for banks to assess their credit risk. Petersen and Rajan (1995) suggest that firm age can proxy for latent credit quality. Zarutskie (2006) argues that younger firms have significantly less outside debt due to the informational asymmetries. Hadlock and Pierce (2010) show that firm age is a useful predictor of financial constraint.

To examine whether the effects of open banking on collateralization can be explained by firm age, I run the RD regression separately for two subsamples: firms younger than the median age and firms older than the median age. Table 4 presents the results. Column (1) shows that young firms just below the threshold are 9.73 percentage points more likely to use floating liens compared to those young firms just above the threshold. In contrast, as shown in column (2), I find no significant discontinuity in the use of floating liens for old firms. Thus, open banking has a greater impact on the use of floating liens for young firms, i.e., those likely characterized by greater informational asymmetries. Table D.4 further confirms that no significant discontinuity existed before the CCDS scheme, ruling out the possibility that these results can be driven by preexisting differences due to firm age.

6.1.2 Asset Information-Intensity

Udell (2004) refers to accounts receivable and inventory (so-called *inside collateral*) as "information-intensive" for the lender, because these assets tend to fluctuate frequently and have more volatile and less easily observable values compared to fixed assets. As a result, the lender needs to rely on up-to-date and detailed financial data to accurately assess their value,[31] which underscores the presence of information asymmetry. To measure a firm's level of asset information-intensity (as a proxy for information asymmetry), I use the ratio of current assets to total assets.

To examine whether the effects of open banking on collateralization can be explained by asset information-intensity, I run the RD regression separately for two subsamples: firms with current assets ratio higher than the median and firms with current assets ratio lower than the median. Table 5 presents the results. Column (1) shows that firms with higher current assets ratio just below the threshold are 10.51 percentage points more likely to use floating liens compared to those just above the threshold. Column (2) shows no significant discontinuity in the use of floating liens for firms with lower current assets ratio. These results suggest that open banking significantly affect the use of floating liens for firms with high level of asset information-intensity, i.e., those likely characterized by greater informational asymmetries. Table D.5 further confirms that no significant discontinuity existed before the CCDS scheme, ruling out the possibility that these results can be driven by pre-existing differences due to asset information-intensity.

6.1.3 Prior Lending Relationship

Open banking is expected to affect collateralization for firms with prior lending relationships, because their bank transaction data is more likely to reveal them as low-risk borrowers (otherwise, they would not have secured credit in the first place). With open banking, these firms may find it easier to leverage floating liens and secure additional loans from new lenders, as their low-risk status becomes more visible to other potential lenders by sharing their bank financial data. However, it is ambiguous whether open banking has a significant impact on collateralization for firms without prior lending relationships.

On the one hand, firms without prior lending relationships may benefit from open banking, because they often have thin credit records and are more likely to be characterized by greater informational asymmetries. In addition, they likely have assets such as accounts receivable and inventory that have not previously been used as collateral. Open banking allows potential lenders to access more information about these assets, making it easier for lenders to accept them as collateral in the form of floating liens (i.e., the information channel of open banking). This mechanism supports the view among policymakers that open

[31] Mester, Nakamura and Renault (2007) note that monitoring a firm's current account can provide lenders with information about the firm's information-intensive assets.

banking promotes financial inclusion by facilitating access to credit for previously underserved firms. On the other hand, a selection mechanism may also be at play. Firms without prior lending relationships may be revealed as high-risk through their bank financial data, leading to credit denial even when they share their data with lenders through open banking.

To examine whether the effects of open banking on collateralization can be explained by firm's prior lending relationships, I run the RD regression separately for two subsamples: firms with prior lending relationships and firms without. Table 6 presents the results. In Panel A, columns (1) and (2) present the RD estimates for the use of floating liens among these two groups. In Panel A column (1), firms just below the CCDS threshold with prior relationships are 3.45 percentage points more likely to use floating liens compared to those just above the threshold. For firms without prior relationships, as reported in Panel A column (2), the RD estimate is 0.0479. This shows that firms without prior relationships also benefit from open banking, suggesting that the information channel outweighs the countervailing selection mechanism. This can be seen as evidence that open banking promotes financial inclusion.

In Panel B, I split prior lending relationships into prior CCDS lending relationships and prior non-CCDS lending relationships. This allows me to examine the impact of type of prior relationship. Since the CCDS scheme required the CCDS9 banks ("CCDS lenders") to enable their small business customers to share financial data, firms with prior CCDS lending relationships likely had more comprehensive information available to share with potential lenders, and therefore are more likely to experience a greater impact from open banking compared to firms with prior non-CCDS lending relationships, which is consistent with Babina et al. (2024).

Panel B columns (1) and (2) confirm this: the RD estimate for floating liens is larger and more significant for firms with prior CCDS lending relationships (3.64 percentage points) than firms with prior non-CCDS lending relationships (1.80 percentage points, which is not statistically significant). These results suggest that the heterogeneous effect by prior lending relationship is primarily driven by the effect of prior CCDS lending relationships. Table D.6 further confirms that no significant discontinuity existed before the CCDS scheme, ruling out the possibility that these results can be driven by preexisting differences due to prior lending relationships.

The results for non-floating liens in Panel A and Panel B columns (3) and (4) show no significant RD estimates for any group. These findings suggest that, regardless of a firm's prior lending relationship, open banking has no significant effect on the use of non-floating liens. This is because non-floating liens are typically associated with fixed assets, which are not information-intensive and thus less dependent on the financial data shared through open banking.

Overall, my findings confirm the information channel of open banking: firms facing greater informational asymmetries and those with more comprehensive information available to share are more significantly impacted by open banking.

6.2 Competition Channel

Open banking promotes competition in credit markets. However, the relationship between competition and the use of collateral remains mixed in both theoretical and empirical literature.[32] To explore whether there is a competition channel, I examine the impact of open banking on loan renewals from existing lenders. I define a new loan as a *loan renewal* when a firm obtains a new loan from a lender with which it had a lending relationship during the prior 12 months (I refer to such lenders as *existing lenders*). My definition of loan renewals does not differentiate between firms that renew their existing loans and those that obtain new credit from their existing lenders. I believe that examining the circumstances under which a firm continues its lending relationship with an existing lender–whether through renewal or new credit–is the most relevant consideration.[33]

It is ambiguous how open banking affects loan renewals from existing lenders. On the one hand, open banking could introduce a *substitution effect*, where firms switch from existing lenders to new lenders due to the reduction in information asymmetry between borrowers and potential lenders. By making borrower data more accessible to alternative lenders, open banking may encourage firms to seek better terms elsewhere, leading to a negative effect on loan renewals. On the other hand, there could be a *competition effect*. With the increased competition introduced by open banking, existing lenders may be more inclined to renew loans to retain their customers, which suggests a positive effect of open banking on loan renewals.

Table 7 presents the results of the RD estimation for loan renewals. Columns (1) and (2) show that firms just below the threshold are about 1.8 percentage points more likely to renew loans secured by floating liens compared to those just above the threshold. Similarly, columns (3) and (4) show that firms just below the threshold are about 2.1 percentage points more likely to renew loans secured by non-floating liens compared to those just

[32] Theoretical work by Villas-Boas and Schmidt-Mohr (1999) suggests that the impact of competition on collateral requirements can move in either direction—collateral may increase or decrease as bank competition increases. Empirical studies provide similarly mixed evidence. For example, Jimenez, Salas and Saurina (2006), using data from Spain, find a positive relationship between collateral requirements and bank competition. In contrast, Hainz, Weill and Godlewski (2013) find a negative relationship between the use of collateral and bank competition when analyzing global data.

[33] Differentiating between renewals and new credit when a firm obtains a new loan from its existing lender does not necessarily provide a meaningful distinction. A firm renewing a loan could be classified as obtaining new credit if, at the time of renewal, previous loans had expired but were not renewed immediately. Similarly, firms obtaining new credit could be seen as renewing if their previous loans happened to expire shortly after the new credit was granted. Therefore, it is difficult to establish a clear distinction between these categories without relying on future (and potentially endogenous) information, as firms might adjust their decisions based on future offers from their existing lenders.

above the threshold. These findings suggest that open banking encourages existing lenders to renew loans regardless of the type of collateral, thereby retaining borrowers who might otherwise seek alternative lenders. In other words, the competition effect dominates the substitution effect, confirming the competition channel of open banking. Table D.7 further confirms that no significant discontinuity existed before the CCDS scheme, ruling out the possibility that these results can be driven by preexisting differences in loan renewals.

As a supplementary test of the competition channel, I examine whether the effects of open banking can be explained by regional heterogeneity in lending market competition, measured by Herfindahl-Hirschman Index (HHI) in CCDS9 banks' SME lending at postcode area level. Specifically, I run the RD regression separately for two subsamples: firms located in regions with below-median 2016 HHI (highly competitive markets) and firms in regions with above-median 2016 HHI (less competitive markets). Table 8 reports the results. Column (1) and (2) show that the effect of open banking on the use of floating liens is more pronounced in highly competitive markets. This provides additional evidence in support of the competition channel. Table D.8 further confirms that no significant discontinuity existed before the CCDS scheme, ruling out the possibility that these results can be driven by preexisting differences in lending market competition.

6.3 Lending Technology Channel

While open banking standardizes the format in which data is shared, the efficiency with which this information is processed may vary across lenders, particularly between banks and non-banks. It is not clear whether banks or non-banks are more likely to be impacted by open banking. On the one hand, non-banks, especially FinTech lenders, may process information more efficiently (Berg et al., 2020; Liu, Lu and Xiong, 2022; Hau et al., 2024). On the other hand, banks, as current account and payment service providers, may have a better understanding of bank financial data. This allows them to process open banking data without losing crucial content and integrate it more efficiently into their risk models, compared to non-banks.

Table 9 presents the results of the RD estimation for banks and non-banks. Column (1) shows that for new loans from banks, firms just below the threshold are 2.92 percentage points more likely to use floating liens compared to those just above the threshold. Similarly, as shown in column (3), for new loans from non-banks, firms just below the threshold are 1.74 percentage points more likely to use floating liens compared to those just above the threshold. These two RD estimates are relatively close in magnitude, suggesting that the effect of open banking on loan collateralization cannot be explained by differences in lending technology. Table D.9 further confirms that no significant discontinuity existed before the CCDS scheme, ruling out the possibility that these results can be driven by preexisting differences due to lending technology.

7 Real Effects of Open Banking

The empirical analyses in the previous sections demonstrate that open banking increases firms' use of floating liens and discuss several mechanisms that could explain this effect. In this section, I investigate the real effects of open banking. The IDBR provides data on the number of employees and the number of establishments for each firm, allowing me to study the impact of open banking on firm outcomes.

Table 10 presents the RD estimation results for employment and establishment growth. Columns (1) and (2) report the estimates for employment growth. Firms just below the threshold experience a statistically significant increase in employment growth compared to firms just above the threshold, with RD estimates of 0.0266 in Column (1) and 0.0273 in Column (2) when controlling for industry-by-year and region-by-year fixed effects. This represents an increase of approximately 76% to 78% compared to the average growth. These results suggest that open banking has a significantly positive impact on job creation. Columns (3) and (4) report the RD estimates for establishment growth. The estimates are small and not statistically significant. This suggests that the effect of open banking manifests through employment growth in existing establishments rather than through the expansion of establishments. Table D.10 further confirms that no significant discontinuity existed before the CCDS scheme, ruling out the possibility that these results can be driven by preexisting differences in firm outcomes.

8 Falsification and Robustness

8.1 Falsification Test

Placebo Cutoffs Test. To test for the absence of discontinuities away from the actual cutoff, I run the RD regression Equation (11) using artificial cutoffs. To avoid "contamination" from real treatment effects, the analysis is conducted as follows: for artificial cutoffs above the actual cutoff, only treated observations are used, and for artificial cutoffs below the actual cutoff, only control observations are included. This restriction ensures that the test at each placebo cutoff uses only observations with the same treatment status. Therefore, by construction, the treatment effect at each artificial cutoff should be zero. Figure 6 shows the test results from this falsification test. Panel A suggests that the use of floating liens does not jump discontinuously at the artificial cutoffs. The corresponding regression results are summarized in Table D.12.

Donut Hole Test. To assess the sensitivity of the results to the observations located closest to the cutoff, I conduct a "donut hole" test, where I run the RD regression Equation (11) excluding the observations closest to the cutoff. The intuition behind this test is that if systematic manipulation has occurred, the observations nearest to the cutoff are the most likely to involve in such manipulation. Even in the absence of manipulation, this test is helpful for checking the robustness of the results, as these observations can be the most influential when fitting the local polynomials in RD estimations. Figure 7 presents the results of this falsification test. Panel A and B suggests that the conclusions regarding floating and non-floating liens remain unchanged, even when excluding observations near the cutoff. The corresponding regression results are summarized in Table D.13.

Bandwidth Sensitivity Test. To assess how sensitive the results are to observations near the boundaries of the neighborhood, I conduct a bandwidth sensitivity test. The implementation of this test is straightforward, I run the RD regression Equation (11) with different bandwidth choices. It is important to note that as the bandwidth increases, the bias of the RD estimation also increases, while the variance of the RD estimation decreases. Therefore, testing the sensitivity to bandwidth choices is only useful within a small range around the MSE-optimal bandwidth. If the bandwidth is much larger than the MSE-optimal choice, the estimated RD effects will have too much bias, while using much smaller bandwidths will lead to RD effects with too much variance. In both cases, point estimates become unreliable, as do the conclusions drawn from this falsification test. Figure 8 presents the results, suggesting that the results are broadly consistent with the findings using the MSE-optimal bandwidth. The corresponding regression results are summarized in Table D.14.

8.2 Robustness Test: Specification Curve Analysis

Following the framework of Simonsohn, Simmons and Nelson (2020), I conduct a specification curve analysis to assess the robustness of the RD estimates in Equation (11). Specifically, I run 96 RD regressions under different model specifications: (i) adding covariates, (ii) including a quadratic RD polynomial of the running variable, (iii) selecting different kernel functions (Triangular, Epanechnikov, Uniform), (iv) employing separate optimal bandwidths for each side of the cutoff, (v) clustering standard errors at industry-by-region-by-year level, and (vi) applying bias correction methods.

Adding Covariates. The simplest way to implement local polynomial RD estimation is to fit the outcome variable on the running variable alone. However, covariates are often included in the basic specification to increase the precision of the RD treatment effect estimator. I follow the covariate-adjustment approach proposed by Calonico et al. (2019) to estimate the RD regression. This adjustment only includes predetermined covariates, as including post-treatment or imbalanced covariates can bias the estimated parameter. As a robustness test for my main specification, I include the key predetermined covariates that are used in the covariate balance test: the log of total assets, leverage ratio, cash, current assets, bank overdraft, and trade credit.

Choice of Polynomial Order. The choice of the local polynomial order is critical in the

RD implementation. Higher-order polynomials tend to produce overfittinng and lead to unreliable estimates (Gelman and Imbens, 2019). In general, the local linear estimator can strike a balance between simplicity, precision, and stability, making it a preferred choice in RD settings. I also test whether my main results hold when using a local quadratic polynomial, allowing for more flexibility in the RD framework.

Choice of Kernel Function. In local linear non-parametric RD regression, the kernel function assigns non-negative weights to each observation. While the triangular kernel is generally recommended, I also test the robustness of my main results using alternative kernel functions: the uniform kernel, which assigns equal weights to all observations within the bandwidth, and the Epanechnikov kernel, which assigns weights with a quadratic decay to observations within the bandwidth.

Bandwidth Implementation. In my main specification, following Calonico, Cattaneo and Titiunik (2014), I choose a common MSE-optimal bandwidth for both sides of the cutoff. Given that the distribution of observations is not symmetric, with more firms below the cutoff than above, I also test the robustness of my results by allowing for different bandwidths on each side. This approach applies separate MSE-optimal bandwidths for the treatment and control groups, which is particularly useful if the two groups differ in bias or variance due to the different curvature of the unknown regression functions.

Clustering the Standard Errors. Since the MSE-optimal bandwidth depends on the variance estimators, clustering the standard errors can affect the optimal bandwidth and the resulting point estimate. In my main specification, I cluster the standard errors at the industry-by-region level to deal with the potential correlation of the error term within each industry and geographic region, as suggested by Cameron and Miller (2015). For a robustness check, I also cluster the standard errors at the industry-by-region-by-year level for possible correlations within each industry, geographic region, and time period.

Bias Correction Method. The local polynomial approach is a non-parametric approximation that uses the polynomials to approximate the unknown regression functions, which can introduce a bias term. To address this, I apply the bias correction method proposed by Calonico, Cattaneo and Titiunik (2014), estimating and adjusting for this bias to obtain a bias-corrected point estimate for robustness checks.

By systematically altering these specifications, I provide a comprehensive view of the sensitivity of the RD estimates across plausible model choices. This reduces the room for selective reporting. Figure 9 presents the specification curves. Panel A shows the specification curve for floating liens. Each dot represents an RD estimate under a different specification, with the shaded area denoting the 95% confidence interval. The clustering of RD estimates around a positive effect, along with the consistent confidence intervals, suggests that the positive impact of open banking on the use of floating liens remains robust across various specifications. In other words, the observed effect is not driven by any specific model choices. Panel B shows the specification curve for non-floating liens. The RD estimates are very close to zero, with no significant shifts across different specifications. This suggests that the conclusion that open banking does not significantly affect the use of non-floating liens is robust to changes in model specifications. Overall, this analysis confirms that the main findings are robust.

9 Conclusion

In this paper, I examine the impact of open banking on credit market outcomes, with a focus on small business collateralization. Using a novel and comprehensive dataset that covers nearly the entire universe of UK firms and captures the entire secured business loan market in the UK, I find that open banking increases the use of floating liens, enabling firms to pledge assets like accounts receivable and inventory as collateral. This finding suggests that open banking helps reduce collateral constraints. Additionally, I document that open banking improves small businesses' access to credit, driven by the increased use of floating liens, as reduced collateral constraints make it easier for firms to secure loans.

My paper also sheds light on the mechanisms through which open banking affects collateralization. First, I find that firms facing greater informational asymmetries–such as young firms, firms with with high level of information-intensive assets, and firms without prior lending relationships–are more significantly impacted by open banking. This suggests that open banking helps lenders mitigate information asymmetries for riskier or more opaque borrowers. Second, I find that firms with more comprehensive information available to share gain greater benefits from open banking. Third, open banking can affect loan collateralization through the competition channel, as new lenders face increased competition from existing lenders. Finally, I find no significant differences in the effects of open banking across banks and non-banks, suggesting that the impact is unlikely to be driven by differences in lending technology.

By documenting how open banking expands the pool of pledgeable assets and alleviates collateral constraints, my paper provides new insights into the role of data as "digital collateral" in improving credit access for small businesses. As policymakers consider future directions for banking regulation, these findings underscore the potential of open banking to promote financial inclusion and spur innovation in lending markets.

My paper contributes to the growing discussion on the open data economy, where data sharing and enhanced data access can stimulate competition, foster innovation, reduce barriers to entry, and reshape the financial landscape. By examining the impact of open banking, this study lays a foundation for future research into the broader implications of data-sharing ecosystems and the open data economy, inviting further exploration into the costs and benefits of creating a more interconnected and data-driven economy.

References

Abis, Simona and Laura Veldkamp. 2024. "The changing economics of knowledge production." *Review of Financial Studies* 37(1):89–118.

Acemoglu, Daron, Ali Makhdoumi, Azarakhsh Malekian and Asu Ozdaglar. 2022. "Too much data: Prices and inefficiencies in data markets." *American Economic Journal: Microeconomics* 14(4):218–256.

Adelino, Manuel, Antoinette Schoar and Felipe Severino. 2015. "House prices, collateral, and self-employment." *Journal of Financial Economics* 117(2):288–306.

Allen, Franklin, Xian Gu and Julapa Jagtiani. 2021. "A survey of fintech research and policy discussion." *Review of Corporate Finance* 1:259–339.

Alok, Shashwat, Pulak Ghosh, Nirupama Kulkarni and Manju Puri. 2024. "Does open banking expand credit access." *Working Paper* .

Aretz, Kevin, Murillo Campello and Maria-Teresa Marchica. 2020. "Access to collateral and the democratization of credit: France's reform of the Napoleonic Security Code." *Journal of Finance* 75(1):45–90.

Asriyan, Vladimir, Luc Laeven and Alberto Martin. 2022. "Collateral booms and information depletion." *Review of Economic Studies* 89(2):517–555.

Babina, Tania, Saleem Bahaj, Greg Buchak, Filippo De Marco, Angus Foulis, Will Gornall, Francesco Mazzola and Tong Yu. 2024. "Customer data access and fintech entry: Early evidence from open banking." *Journal of Financial Economics* (forthcoming).

Bahaj, Saleem, Angus Foulis and Gabor Pinter. 2020. "Home values and firm behavior." *American Economic Review* 110(7):2225–2270.

Benmelech, Efraim, Nitish Kumar and Raghuram Rajan. 2024. "The decline of secured debt." *Journal of Finance* 79(1):35–93.

Berg, Tobias, Andreas Fuster and Manju Puri. 2022. "Fintech lending." *Annual Review of Financial Economics* 14(1):187–207.

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

Berger, Allen N. and Gregory F. Udell. 1995. "Relationship lending and lines of credit in small firm finance." *Journal of Business* 68(3):351–381.

Berger, Allen N., W. Scott Frame and Vasso Ioannidou. 2011. "Tests of ex ante versus ex post theories of collateral using private and public information." *Journal of Financial Economics* 100(1):85–97.

Besanko, David and Anjan V. Thakor. 1987. "Collateral and rationing: Sorting equilibria in monopolistic and competitive credit markets." *International Economic Review* 28(3):671–689.

Bester, Helmut. 1985. "Screening vs. rationing in credit markets with imperfect information." *American Economic Review* 75(4):850–855.

Bharath, Sreedhar T., Sandeep Dahiya, Anthony Saunders and Anand Srinivasan. 2011. "Lending relationships and loan contract terms." *Review of Financial Studies* 24(4):1141–1203.

Biswas, Sonny. 2023. "Collateral and bank screening as complements: A spillover effect." *Journal of Economic Theory* 212:105703.

Bolton, Patrick, Xavier Freixas, Leonardo Gambacorta and Paolo Emilio Mistrulli. 2016. "Relationship and transaction lending in a crisis." *Review of Financial Studies* 29(10):2643–2676.

Boot, Arnoud W. A. and Anjan V. Thakor. 1994. "Moral hazard and secured lending in an infinitely repeated credit market game." *International Economic Review* 35(4):899–920.

Brown, Marti, Tullio Jappelli and Marco Pagano. 2009. "Information sharing and credit: Firm-level evidence from transition countries." *Journal of Financial Intermediation* 18(2):151–172.

Calomiris, Charles W., Mauricio Larrain, José Liberti and Jason Sturgess. 2017. "How collateral laws shape lending and sectoral activity." *Journal of Financial Economics* 123(1):163–188.

Calonico, Sebastian, Matias D. Cattaneo and Max H. Farrell. 2018. "On the effect of bias estimation on coverage accuracy in nonparametric inference." *Journal of the American Statistical Association* 113(522):767–779.

Calonico, Sebastian, Matias D. Cattaneo, Max H. Farrell and Rocio Titiunik. 2019. "Regression discontinuity designs using covariates." *Review of Economics and Statistics* 101(3):442–451.

Calonico, Sebastian, Matias D. Cattaneo and Rocio Titiunik. 2014. "Robust nonparametric confidence intervals for regression-discontinuity designs." *Econometrica* 82(6):2295–2326.

Cameron, A. Colin and Douglas L. Miller. 2015. "A practitioner's guide to cluster-robust inference." *Journal of Human Resources* 50(2):317–372.

Catherine, Sylvain, Thomas Chaney, Zongbo Huang, David Sraer and David Thesmar. 2022. "Quantifying reduced-form evidence on collateral constraints." *Journal of Finance* 77(4):2143–2181.

Cattaneo, Matias D., Michael Jansson and Xinwei Ma. 2020. "Simple local polynomial density estimators." *Journal of the American Statistical Association* 115(531):1449–1455.

Cerqueiro, Geraldo, Steven Ongena and Kasper Roszbach. 2016. "Collateralization, bank loan rates, and monitoring." *Journal of Finance* 71(3):1295–1322.

Chaney, Thomas, David Sraer and David Thesmar. 2012. "The collateral channel: How real estate shocks affect corporate investment." *American Economic Review* 102(6):2381–2409.

Cong, Lin William, Danxia Xie and Longtian Zhang. 2021. "Knowledge accumulation, privacy, and growth in a data economy." *Management Science* 67(10):6480–6492.

Darmouni, Olivier and Andrew Sutherland. 2021. "Learning about competitors: Evidence from SME lending." *Review of Financial Studies* 34(5):2275–2317.

Degryse, Hans and Patrick Van Cayseele. 2000. "Relationship lending within a bank-based system: Evidence from European small business data." *Journal of Financial Intermediation* 9(1):90–109.

Dell'Ariccia, Giovanni and Robert Marquez. 2006. "Lending booms and lending standards." *Journal of Finance* 61(5):2511–2546.

Di Maggio, Marco, Dimuthu Ratnadiwakara and Don Carmichael. 2022. "Invisible primes: Fintech lending with alternative data." *NBER Working Paper* No. 29840.

Dierkes, Maik, Carsten Erner, Thomas Langer and Lars Norden. 2013. "Business credit information sharing and default risk of private firms." *Journal of Banking and Finance* 37(8):2867–2878.

Doblas-Madrid, Antonio and Raoul Minetti. 2013. "Sharing information in the credit market: Contract-level evidence from US firms." *Journal of Financial Economics* 109(1):198–223.

Farboodi, Maryam, Dhruv Singal, Laura Veldkamp and Venky Venkateswaran. 2024. "Valuing financial data." *Review of Financial Studies* (forthcoming).

Farboodi, Maryam and Laura Veldkamp. 2020. "Long-run growth of financial data technology." *American Economic Review* 110(8):2485–2523.

Fuster, Andreas, Paul Goldsmith-Pinkham, Tarun Ramadorai and Ansgar Walther. 2022. "Predictably unequal? The effects of machine learning on credit markets." *Journal of Finance* 77(1):5–47.

Gambacorta, Leonardo, Yiping Huang, Zhenhua Li, Han Qiu and Shu Chen. 2023. "Data versus collateral." *Review of Finance* 27(2):369–398.

Gehrig, Thomas and Rune Stenbacka. 2007. "Information sharing and lending market competition with switching costs and poaching." *European Economic Review* 51(1):77–99.

Gelman, Andrew and Guido Imbens. 2019. "Why high-order polynomials should not be used in regression discontinuity designs." *Journal of Business and Economic Statistics* 37(3):447–456.

Gertler, Paul, Brett Green and Catherine Wolfram. 2024. "Digital collateral." *Quarterly Journal of Economics* 139(3):1713–1766.

Ghosh, Pulak, Boris Vallee and Yao Zeng. 2023. "FinTech lending and cashless payments." *Journal of Finance* (forthcoming).

Goldstein, Itay, Chong Huang and Liyan Yang. 2022. "Open banking under maturity transformation." *Available at SSRN Electronic Journal* No. 4215016.

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

Grembi, Veronica, Tommaso Nannicini and Ugo Troiano. 2016. "Do fiscal rules matter?" *American Economic Journal: Applied Economics* 8(3):1–30.

Hadlock, Charles J. and Joshua R. Pierce. 2010. "New evidence on measuring financial constraints: Moving beyond the KZ index." *Review of Financial Studies* 23(5):1909–1940.

Hainz, Christa, Laurent Weill and Christophe J. Godlewski. 2013. "Bank competition and collateral: Theory and evidence." *Journal of Financial Services Research* 44(2):131–148.

Hanedar, Elmas Yaldız, Eleonora Broccardo and Flavio Bazzana. 2014. "Collateral requirements of SMEs: The evidence from less-developed countries." *Journal of Banking and Finance* 38:106–121.

Hau, Harald, Yi Huang, Chen Lin, Hongzhe Shan, Zixia Sheng and Lai Wei. 2024. "FinTech credit and entrepreneurial growth." *Journal of Finance* 79(5):3309–3359.

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

Holmstrom, Bengt and Jean Tirole. 1997. “Financial intermediation, loanable funds, and the real sector.” *Quarterly Journal of Economics* 112(3):663–691.

Inderst, Roman and Holger M. Mueller. 2007. “A lender-based theory of collateral.” *Journal of Financial Economics* 84(3):826–859.

Ioannidou, Vasso, Nicola Pavanini and Yushi Peng. 2022. “Collateral and asymmetric information in lending markets.” *Journal of Financial Economics* 144(1):93–121.

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

Iyer, Rajkamal, Asim Ijaz Khwaja, Erzo F.P. Luttmer and Kelly Shue. 2016. “Screening peers softly: Inferring the quality of small borrowers.” *Management Science* 62(6):1554–1557.

Jensen, Thais Laerkholm, Søren Leth-Petersen and Ramana Nanda. 2022. “Financing constraints, home equity and selection into entrepreneurship.” *Journal of Financial Economics* 145(2):318–337.

Jimenez, Gabriel, Vicente Salas and Jesus Saurina. 2006. “Determinants of collateral.” *Journal of Financial Economics* 81(2):255–281.

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

Kalemli-Özcan, Şebnem, Bent E. Sørensen, Carolina Villegas-Sanchez, Vadym Volosovych and Sevcan Yeşiltaş. 2024. “How to construct nationally representative firm-level data from the Orbis global database: New facts on SMEs and aggregate implications for industry concentration.” *American Economic Journal: Macroeconomics* 16(2):353–374.

Karapetyan, Artashes and Bogdan Stacescu. 2014a. “Does information sharing reduce the role of collateral as a screening device.” *Journal of Banking and Finance* 43:48–57.

Karapetyan, Artashes and Bogdan Stacescu. 2014b. “Information sharing and information acquisition in credit markets.” *Review of Finance* 18(4):1583–1615.

Kiyotaki, Nobuhiro and John Moore. 1997. “Credit cycles.” *Journal of Political Economy* 105(2):211–248.

Lee, David S. and Thomas Lemieux. 2010. “Regression discontinuity designs in economics.” *Journal of Economic Literature* 48(2):281–355.

Liberti, José, Jason Sturgess and Andrew Sutherland. 2022. "How voluntary information sharing systems form: Evidence from a US commercial credit bureau." *Journal of Financial Economics* 145(3):827–849.

Liu, Lei, Guangli Lu and Wei Xiong. 2022. "The big tech lending model." *NBER Working Paper* No. 30160.

Manove, Michael, A. Jorge Padilla and Marco Pagano. 2001. "Collateral versus project screening: A model of lazy banks." *RAND Journal of Economics* 32(4):726–744.

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

Nam, Rachel J. 2023. "Open banking and customer data sharing: Implications for fintech borrowers." *SAFE Working Paper* No. 364.

Nanda, Ramana and Gordon Phillips. 2023. "Small firm financing: Sources, frictions, and policy implications." *Handbook of the Economics of Corporate Finance* 1(1):107–135.

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

Ouyang, Shumiao. 2023. "Cashless payment and financial inclusion." *Available at SSRN Electronic Journal* No. 3948925.

Pagano, Marco and Tullio Jappelli. 1993. "Information sharing in credit markets." *Journal of Finance* 48(5):1693–1717.

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

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

Rishabh, Kumar. 2024. "Beyond the bureau: Interoperable payment data for loan screening and monitoring." *Available at SSRN Electronic Journal* No. 4782597.

Schmalz, Martin C., David A. Sraer and David Thesmar. 2017. "Housing collateral and entrepreneurship." *Journal of Finance* 72(1):99–132.

Simonsohn, Uri, Joseph P. Simmons and Leif D. Nelson. 2020. "Specification curve analysis." *Nature Human Behaviour* 4(11):1208–1214.

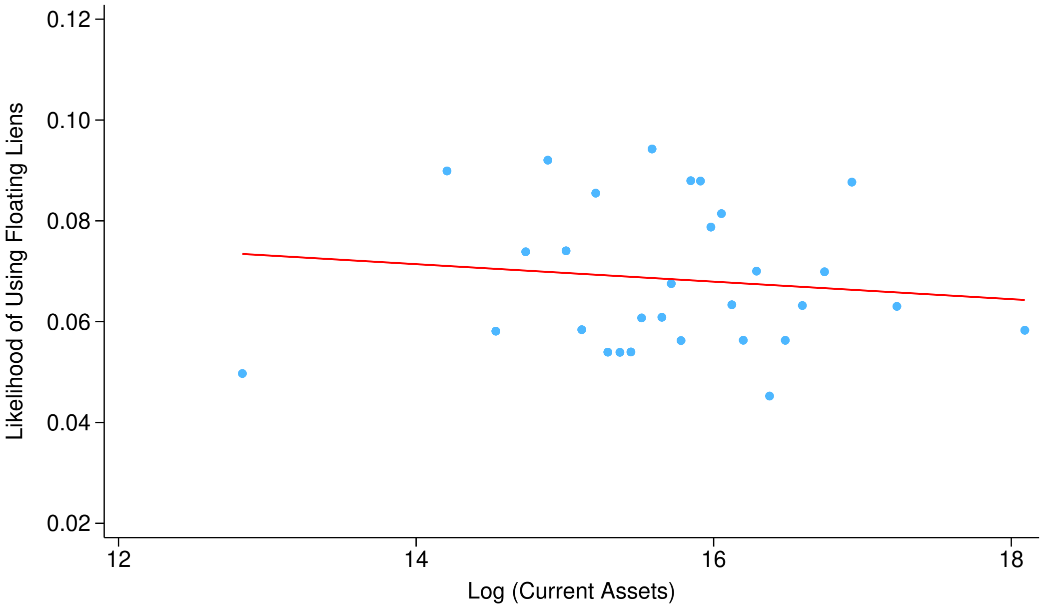
Udell, Grergory F. 2004. *Asset-based finance: Proven disciplines for prudent lending*. Commercial Finance Association, New York, NY.

Veldkamp, Laura. 2023. "Valuing data as an asset." *Review of Finance* 27(5):1545–1562.

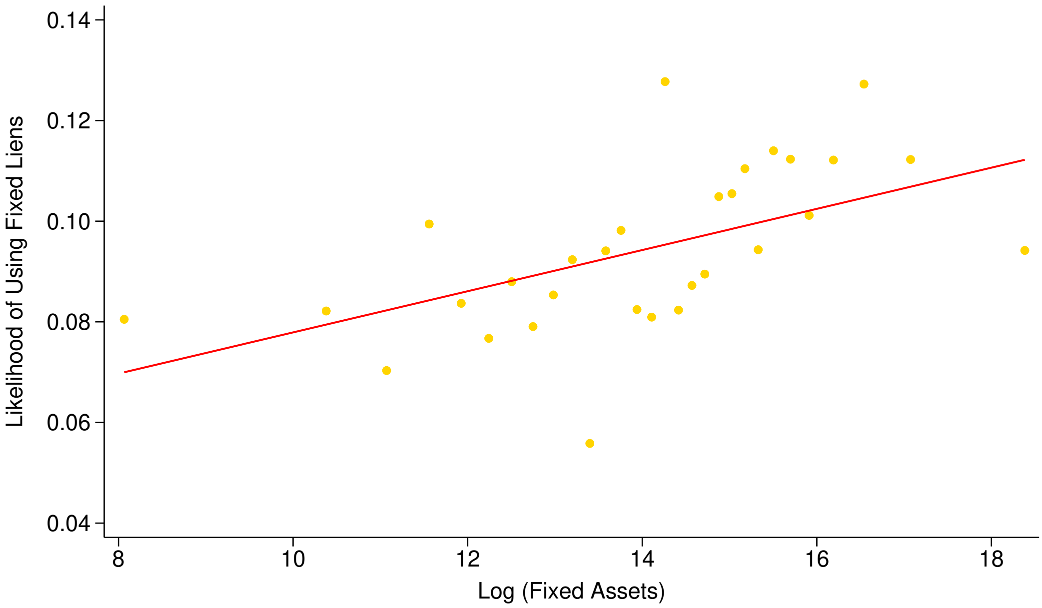
Villas-Boas, J. Miguel and Udo Schmidt-Mohr. 1999. "Oligopoly with asymmetric information: Differentiation in credit markets." *RAND Journal of Economics* 30(3):375–396.

Zarutskie, Rebecca. 2006. "Evidence on the effects of bank competition on firm borrowing and investment." *Journal of Financial Economics* 81(3):503–537.

Figure 1. Relationship Between Asset Size and Collateral Usage



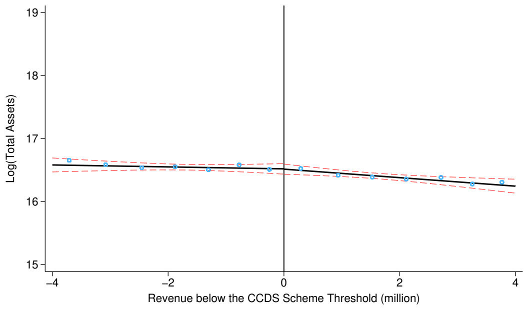
Panel A. Current Assets and Floating Liens



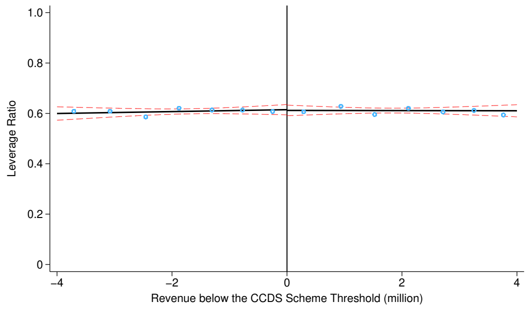
Panel B. Fixed Assets and Fixed Liens

Notes: This figure shows the binned scatter plots for the relationship between asset size and the use of collateral, based the 2016 BvD sample of small businesses. Panel A presents the binned scatter plot of the log of current assets versus the likelihood of using floating liens to secure loans from new lenders. The dependent variable is the likelihood of using floating liens in 2016, and the independent variable is the log of current assets in 2016. The log of fixed assets in 2016 is included as a control variable. Each blue dot represents the average value of the dependent variable for firms within a specific bin of the log of current assets and a red linear line is fitted to the scattered dots. Panel B presents the binned scatter plot of the log of fixed assets versus the likelihood of using fixed liens to secure loans from new lenders. The dependent variable is the likelihood of using fixed liens in 2016, and the independent variable is the log of fixed assets in 2016. The log of current assets in 2016 is included as a control variable. Each yellow dot represents the average value of the dependent variable for firms within a specific bin of the log of fixed assets, and a red linear line is fitted to the scattered dots.

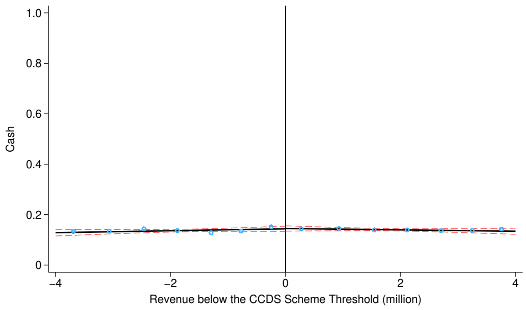
Figure 2. RD Plot: Covariate Balance Test



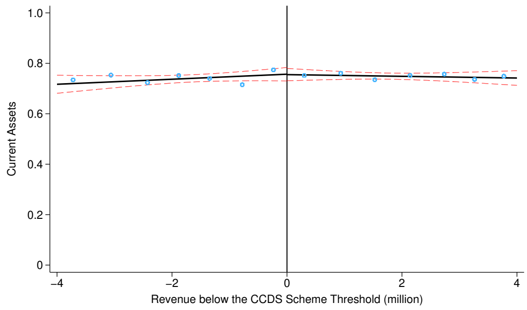
Panel A. Log (Total Assets)



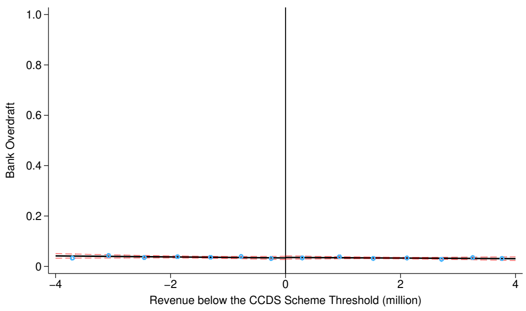
Panel B. Leverage Ratio



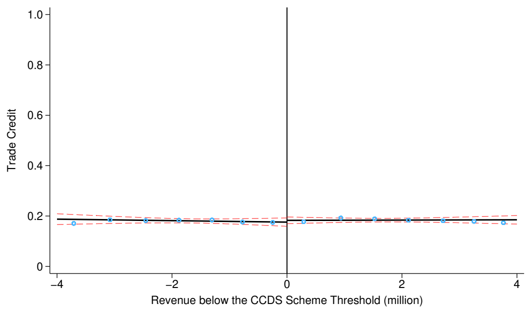
Panel C. Cash



Panel D. Current Assets



Panel E. Bank Overdraft



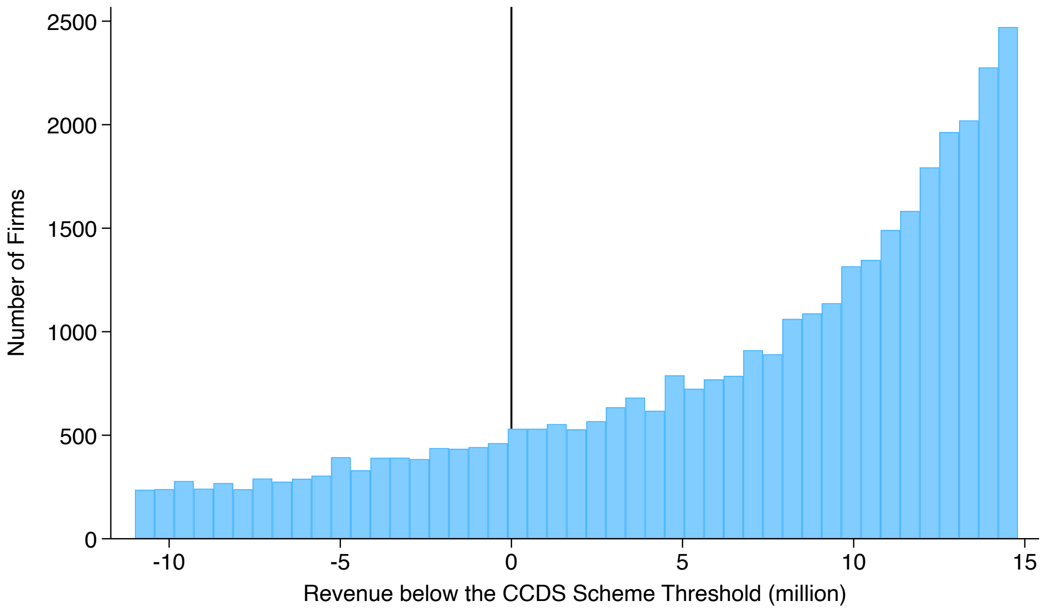
Panel F. Trade Credit

Notes: This figure shows the RD plots for the covariate balance test. Specifically, it reports results from the following RD regression:

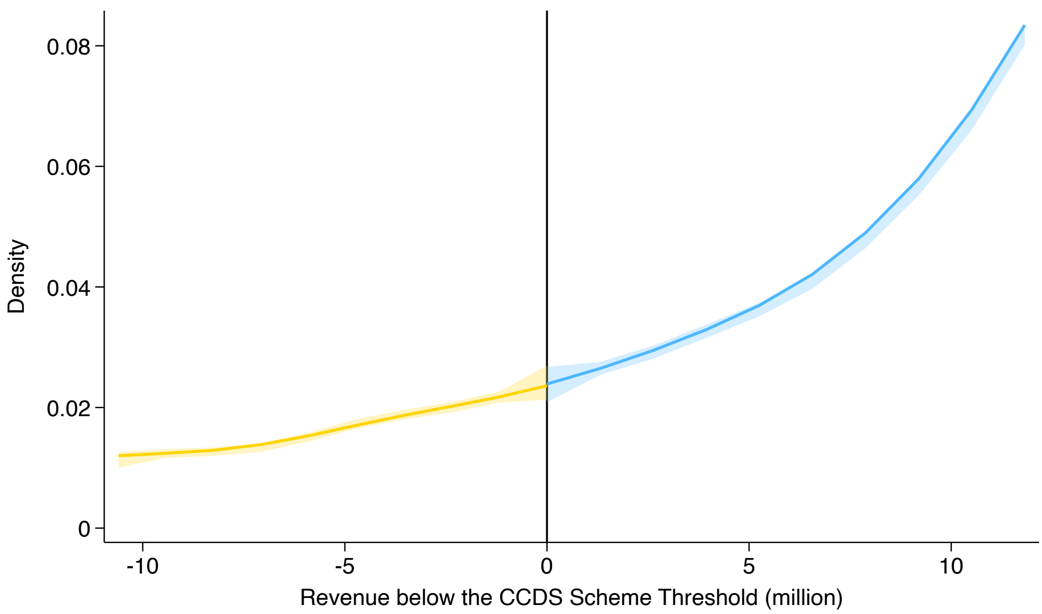
$$Y_{i,t} = \alpha + \tau\, \mathbb{D}_{i,t} + \beta\, R_{i,t} + \gamma\, R_{i,t} \cdot \mathbb{D}_{i,t} + \varepsilon_{i,t}$$

where $Y_{i,t}$ is one of the six different predetermined financial variables (from the 2016 BvD) for firm i in year $t = 2017$ or 2018. $R_{i,t}$ represents firm i's normalized revenue over the 12 months preceding year t, defined as the amount by which the firm's revenue falls short of the CCDS threshold £25 million. $\mathbb{D}_{i,t}$ is an indicator variable that equals one if firm i's normalized revenue is greater than or equal to zero in year t, and zero otherwise. I restrict the running variable to a bandwidth of £4 million across tests for consistency and ease of interpretation. Standard errors are clustered at the industry-by-region level. The corresponding regression results are summarized in Table D.11. Each blue dot represents the average value of the respective outcome variable for firms within a specific bin of normalized revenue. A linear line is fitted separately for observations above and below the threshold, with 95% confidence intervals in red dashed lines.

Figure 3. Manipulation Test



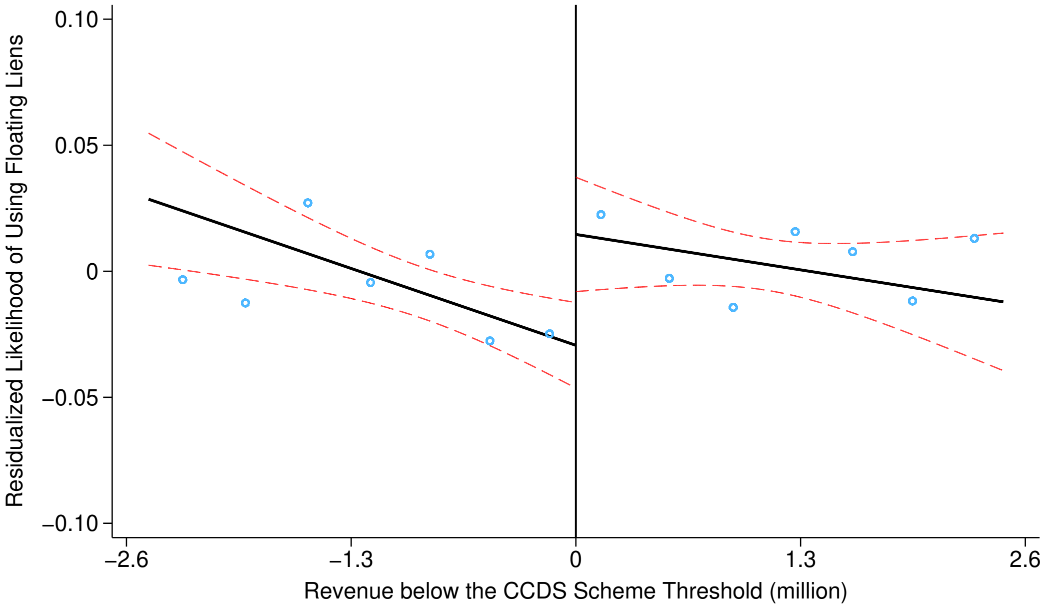
Panel A. Histogram



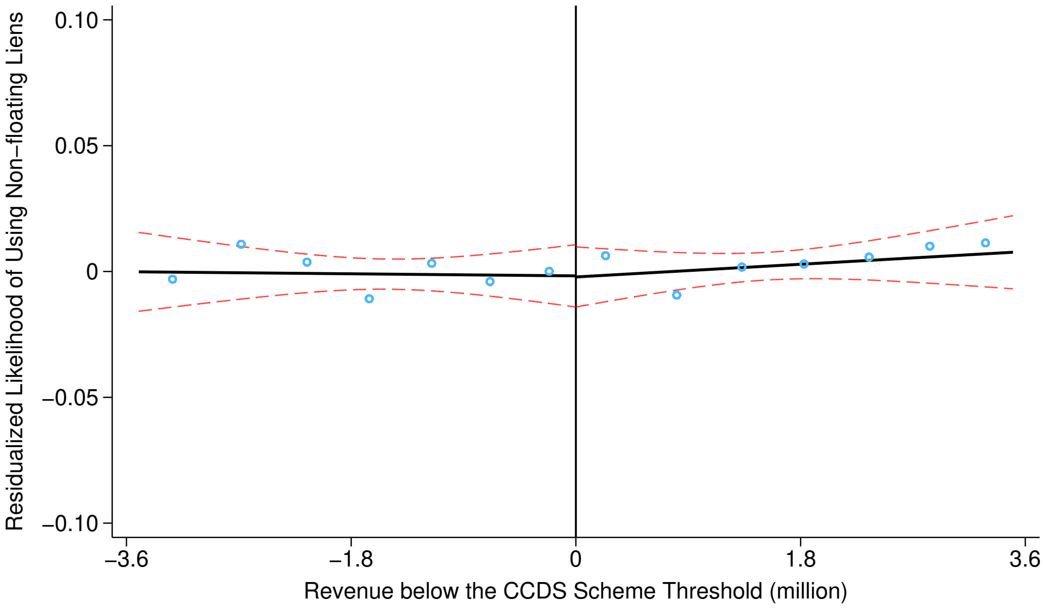
Panel B. Density Test

Notes: This figure shows the manipulation test of the running variable (normalized revenue). Panel A presents a histogram of the running variable. Panel B shows the plot of the density test proposed by Cattaneo, Jansson and Ma (2020). The solid line is a local polynomial estimate of the density of the running variable and the shaded regions represent 95% confidence intervals, both calculated separately for observations above and below the threshold.

Figure 4. RD Plot: Effects of Open Banking on Loan Collateralization



Panel A. Floating Liens



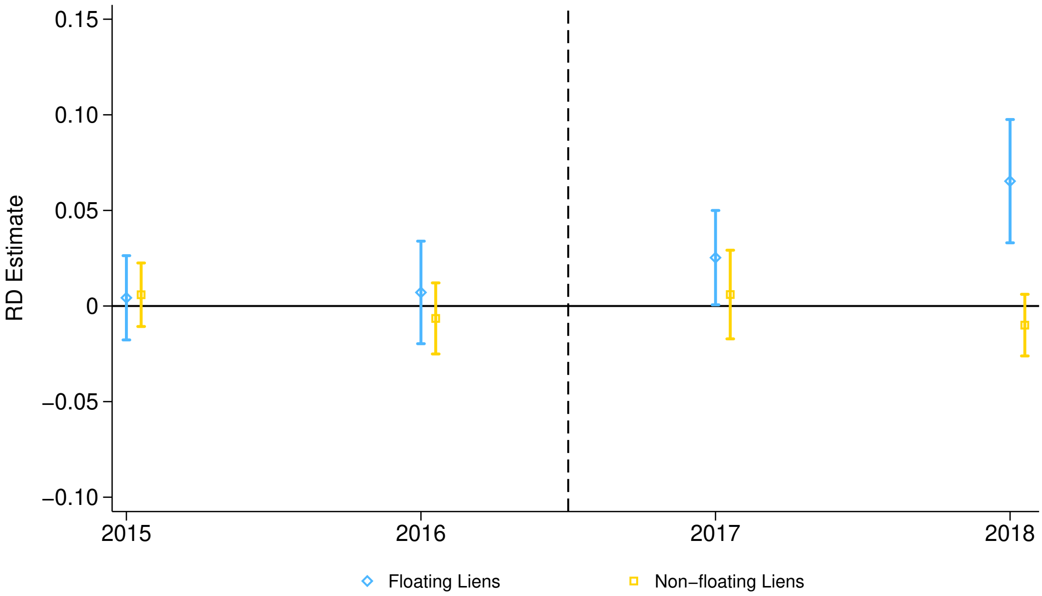
Panel B. Non-floating Liens

Notes: This figure shows the RD plots for the effects of open banking on loan collateralization. Specifically, it reports results from the following RD regression:

$$Y_{i,t} = \alpha + \tau \mathbb{D}_{i,t} + \beta R_{i,t} + \gamma R_{i,t} \cdot \mathbb{D}_{i,t} + \mu_{j,t} + \nu_{r,t} + \varepsilon_{i,t}$$

where $Y_{i,t}$ is an indicator variable that equals one if firm i uses floating liens (Panel A) or non-floating liens (Panel B) to secure loans from new lenders in year $t = 2017$ or 2018, and zero otherwise. $R_{i,t}$ represents firm i's normalized revenue over the 12 months preceding year t, defined as the amount by which the firm's revenue falls short of the CCDS threshold £25 million. $\mathbb{D}_{i,t}$ is an indicator variable that equals one if firm i's normalized revenue is greater than or equal to zero in year t, and zero otherwise. $\mu_{j,t}$ and $\nu_{r,t}$ represent the industry-by-year and geographic region-by-year fixed effects, respectively. Standard errors are clustered at the industry-by-region level. I use the MSE-optimal bandwidth following Calonico, Cattaneo and Titiunik (2014) for triangular kernel weighting method. The corresponding regression results are summarized in Table 2. Each blue dot represents the average value of the respective outcome variable for firms within a specific bin of normalized revenue. A linear line is fitted separately for observations above and below the threshold, with 95% confidence intervals in red dashed lines.

Figure 5. Dynamic Effects of Open Banking on Loan Collateralization

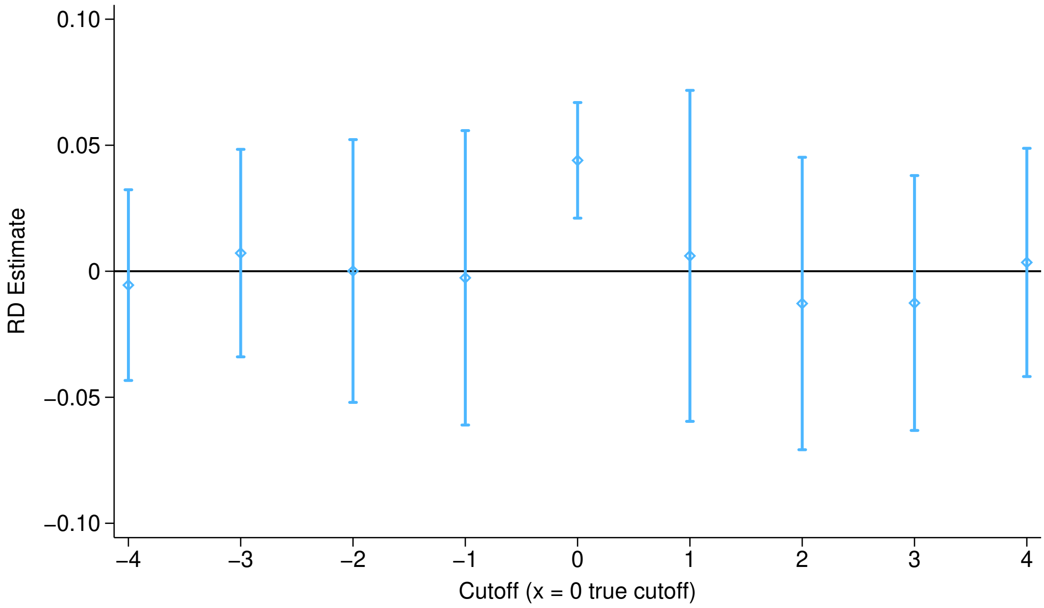


Notes: This figure shows the dynamic effects of open banking on loan collateralization. Specifically, it reports results from the following RD regression for each year from 2015 to 2018:

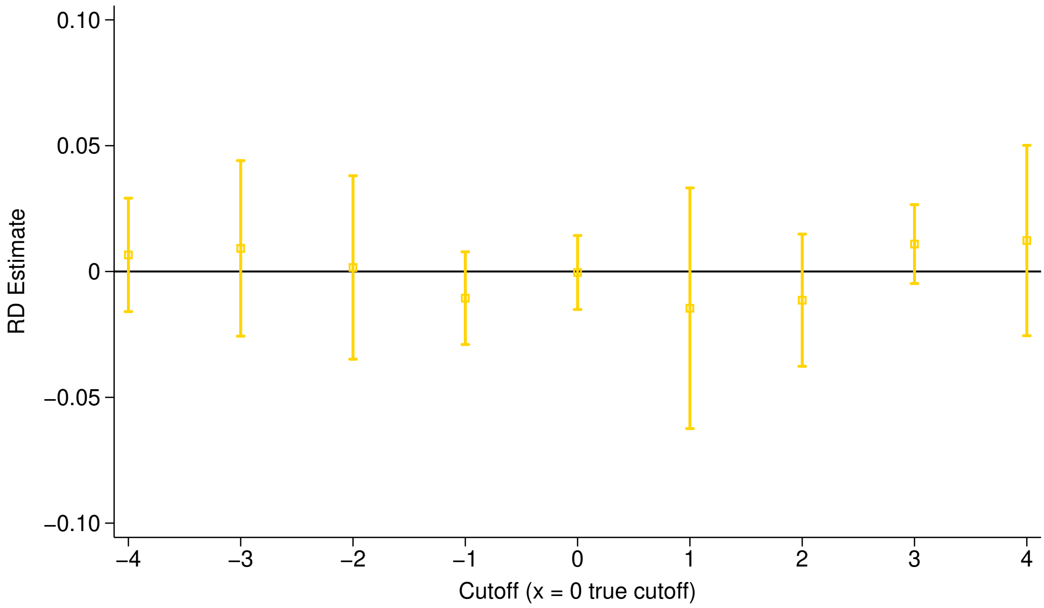
$$Y_{i,t} = \alpha_t + \tau_t \mathbb{D}_{i,t} + \beta_t R_{i,t} + \gamma_t R_{i,t} \cdot \mathbb{D}_{i,t} + \mu_{j,t} + \nu_{r,t} + \varepsilon_{i,t}$$

where $Y_{i,t}$ is an indicator variable that equals one if firm i uses floating liens (blue) or non-floating liens (yellow) to secure loans from new lenders in year t, and zero otherwise. $R_{i,t}$ represents firm i's normalized revenue over the 12 months preceding year t, defined as the amount by which the firm's revenue falls short of the CCDS threshold £25 million. $\mathbb{D}_{i,t}$ is an indicator variable that equals one if firm i's normalized revenue is greater than or equal to zero in year t, and zero otherwise. $\mu_{j,t}$ and $\nu_{r,t}$ represent the industry-by-year and geographic region-by-year fixed effects, respectively. Standard errors are clustered at the industry-by-region level. I use the MSE-optimal bandwidth following Calonico, Cattaneo and Titiunik (2014) for triangular kernel weighting method. The corresponding regression results are summarized in Table D.2. Each blue dot represents the RD estimates for floating liens in the respective year, with a 90% confidence interval shown by the solid blue lines. Similarly, each yellow dot represents the RD estimates for non-floating liens in the respective year, with a 90% confidence interval shown by the solid yellow lines.

Figure 6. Placebo Cutoffs Test



Panel A. Floating Liens



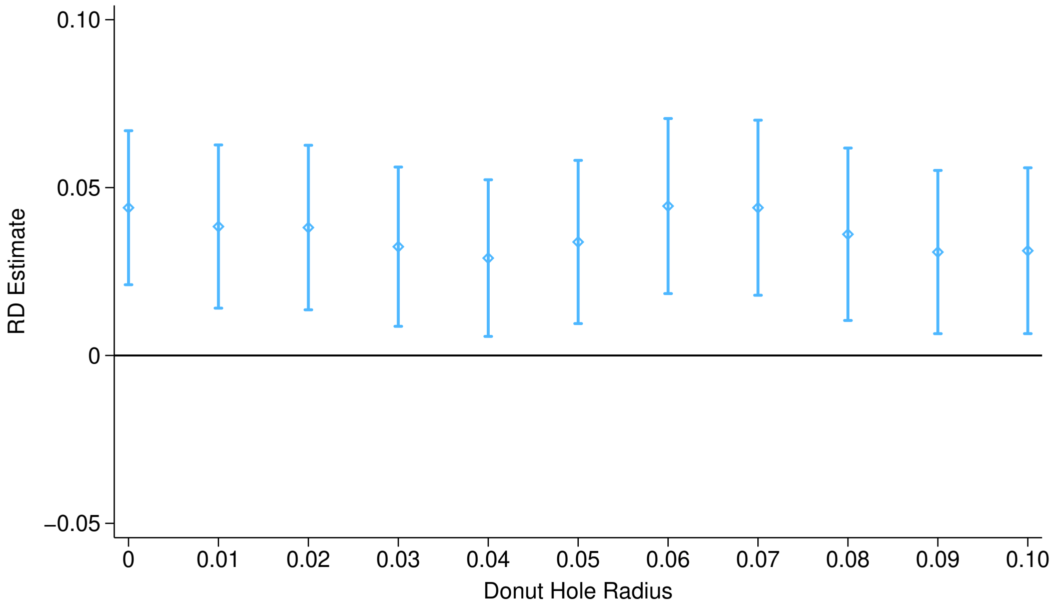
Panel B. Non-floating Liens

Notes: This figure shows the placebo cutoffs test. Specifically, it reports results from the following RD regression for each cutoff between £21 million to £29 million:

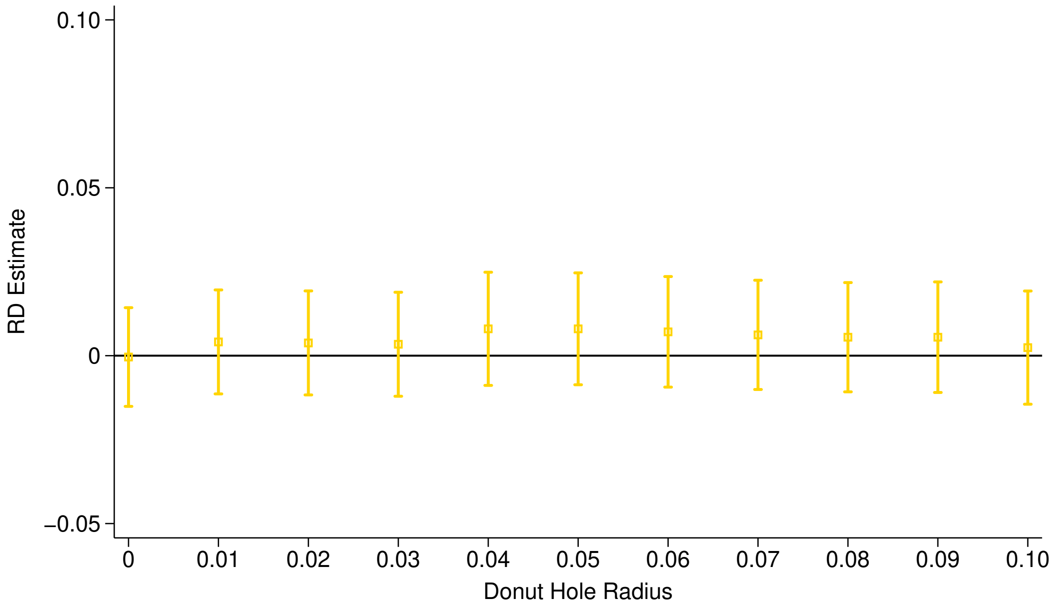
$$Y_{i,t} = \alpha + \tau \mathbb{D}_{i,t} + \beta R_{i,t} + \gamma R_{i,t} \cdot \mathbb{D}_{i,t} + \mu_{j,t} + \nu_{r,t} + \varepsilon_{i,t}$$

where $Y_{i,t}$ is an indicator variable that equals one if firm i uses floating liens (Panel A) or non-floating liens (Panel B) to secure loans from new lenders in year $t = 2017$ or 2018, and zero otherwise. $R_{i,t}$ represents firm i's normalized revenue over the 12 months preceding year t, defined as the amount by which the firm's revenue falls short of the tested threshold between £21 million and £29 million (with £25 million as the true cutoff). $\mathbb{D}_{i,t}$ is an indicator variable that equals one if firm i's normalized revenue is greater than or equal to zero in year t, and zero otherwise. $\mu_{j,t}$ and $\nu_{r,t}$ represent the industry-by-year and geographic region-by-year fixed effects, respectively. Standard errors are clustered at the industry-by-region level. I use the MSE-optimal bandwidth following Calonico, Cattaneo and Titiunik (2014) for triangular kernel weighting method. The corresponding regression results are summarized in Table D.12. Panel A presents the RD estimates by cutoff for floating liens, with 95% confidence intervals. Panel B presents the RD estimates by cutoff for non-floating liens, with 95% confidence intervals.

Figure 7. Donut Hole Test



Panel A. Floating Liens



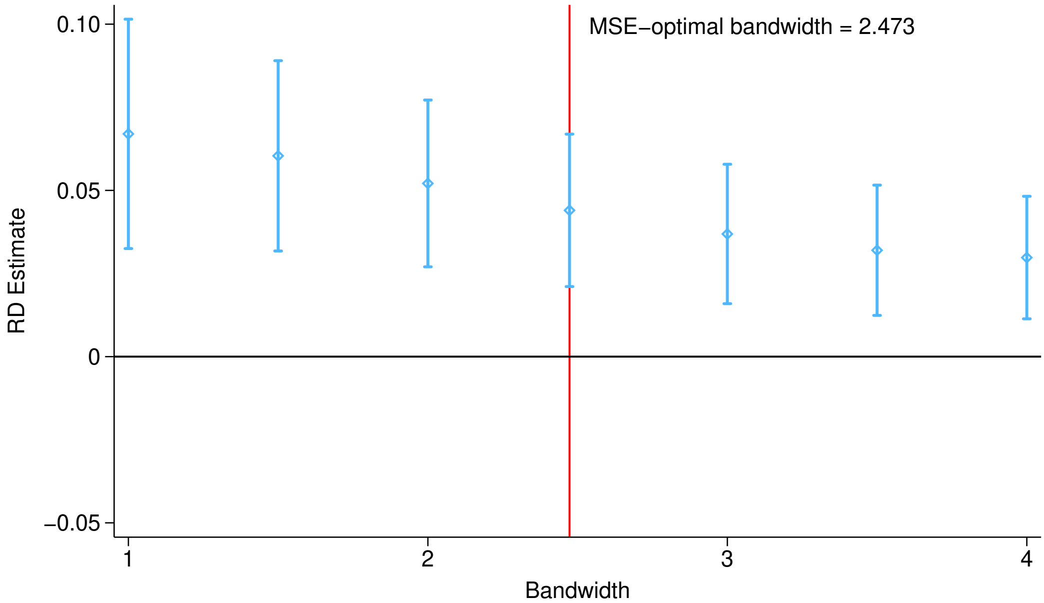
Panel B. Non-floating Liens

Notes: This figure shows the donut hole test. Specifically, it reports results from the following RD regression, excluding observations within a donut hole radius of 0 to 0.10 around the CCDS threshold:

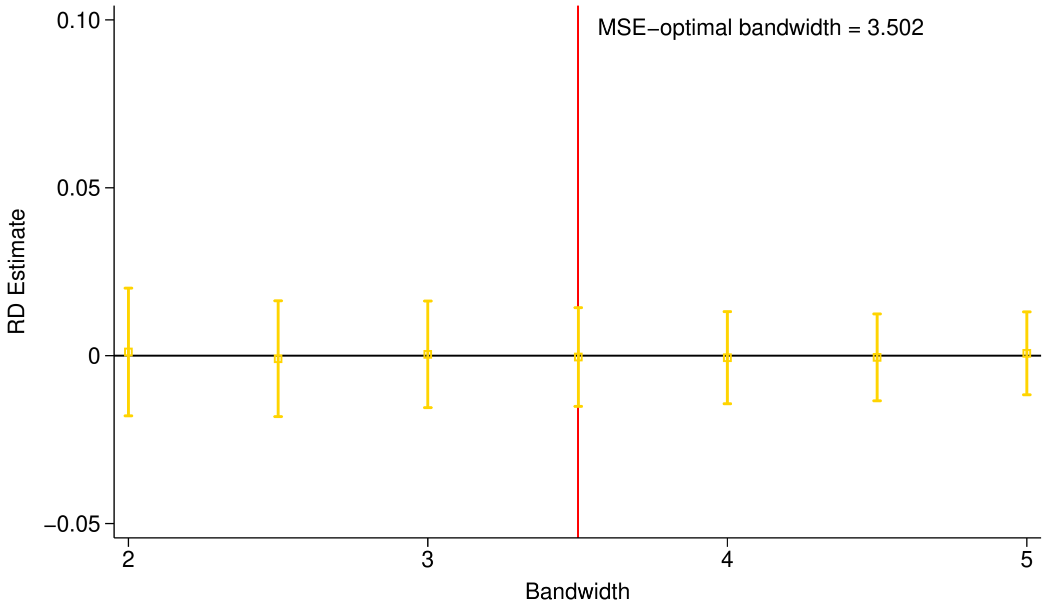
$$Y_{i,t} = \alpha + \tau\,\mathbb{D}_{i,t} + \beta\,R_{i,t} + \gamma\,R_{i,t}\cdot\mathbb{D}_{i,t} + \mu_{j,t} + \nu_{r,t} + \varepsilon_{i,t}$$

where $Y_{i,t}$ is an indicator variable that equals one if firm i uses floating liens (Panel A) or non-floating liens (Panel B) to secure loans from new lenders in year $t = 2017$ or 2018, and zero otherwise. $R_{i,t}$ represents firm i's normalized revenue over the 12 months preceding year t, defined as the amount by which the firm's revenue falls short of the CCDS threshold £25 million. $\mathbb{D}_{i,t}$ is an indicator variable that equals one if firm i's normalized revenue is greater than or equal to zero in year t, and zero otherwise. $\mu_{j,t}$ and $\nu_{r,t}$ represent the industry-by-year and geographic region-by-year fixed effects, respectively. Standard errors are clustered at the industry-by-region level. I use the MSE-optimal bandwidth following Calonico, Cattaneo and Titiunik (2014) for triangular kernel weighting method. The corresponding regression results are summarized in Table D.13. Panel A presents the RD estimates by donut hole radius for floating liens, with 95% confidence intervals. Panel B presents the RD estimates by donut hole radius for non-floating liens, with 95% confidence intervals.

Figure 8. Bandwidth Sensitivity Test



Panel A. Floating Liens



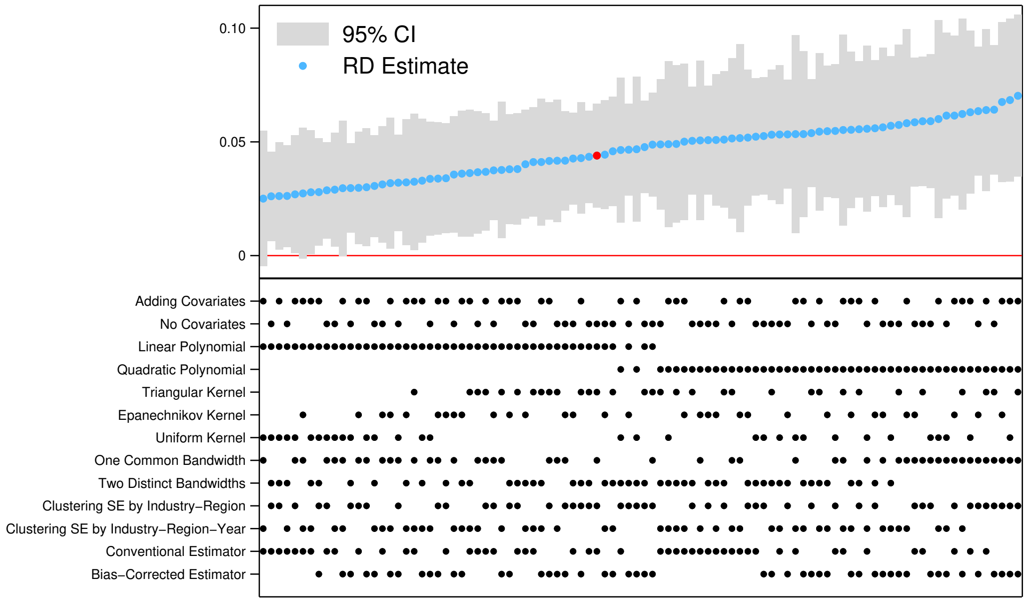
Panel B. Non-floating Liens

Notes: This figure shows the bandwidth sensitivity test. Specifically, it reports results from the following RD regression, using different bandwidths:

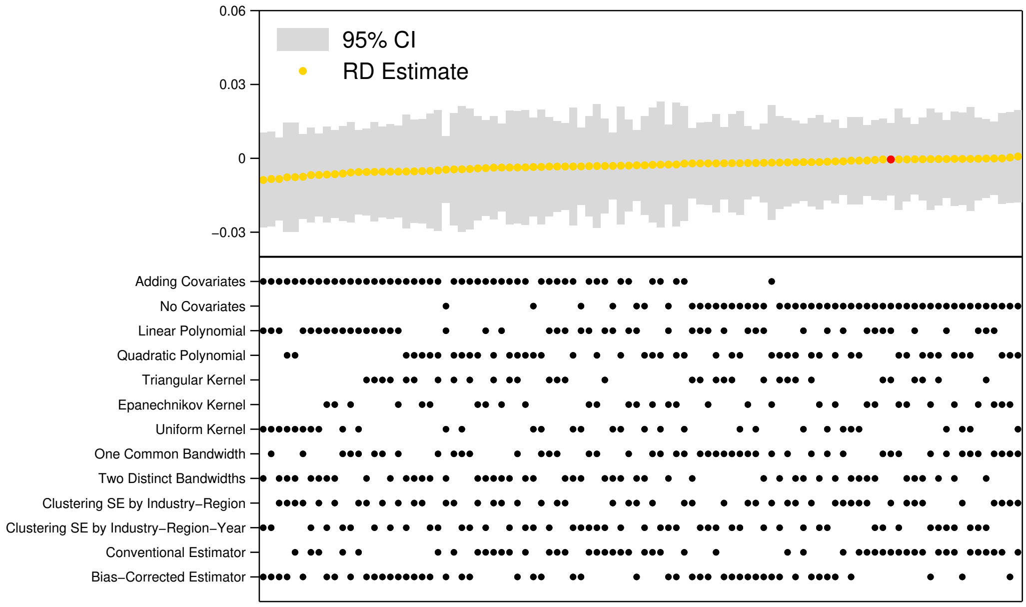
$$Y_{i,t} = \alpha + \tau\,\mathbb{D}_{i,t} + \beta\,R_{i,t} + \gamma\,R_{i,t} \cdot \mathbb{D}_{i,t} + \mu_{j,t} + \nu_{r,t} + \varepsilon_{i,t}$$

where $Y_{i,t}$ is an indicator variable that equals one if firm i uses floating liens (Panel A) or non-floating liens (Panel B) to secure loans from new lenders in year $t = 2017$ or 2018, and zero otherwise. $R_{i,t}$ represents firm i's normalized revenue over the 12 months preceding year t, defined as the amount by which the firm's revenue falls short of the CCDS threshold £25 million. $\mathbb{D}_{i,t}$ is an indicator variable that equals one if firm i's normalized revenue is greater than or equal to zero in year t, and zero otherwise. $\mu_{j,t}$ and $\nu_{r,t}$ represent the industry-by-year and geographic region-by-year fixed effects, respectively. Standard errors are clustered at the industry-by-region level. I use the triangular kernel weighting method for the RD estimation. The corresponding regression results are summarized in Table D.14. Panel A presents the RD estimates by bandwidth for floating liens, with 95% confidence intervals. Panel B presents the RD estimates by bandwidth for non-floating liens, with 95% confidence intervals.

Figure 9. Specification Curve Analysis



Panel A. Floating Liens



Panel B. Non-floating Liens

Notes: This figure shows the specification curve analysis. Specifically, it reports results from the augmented RD regression under different model specifications: (i) adding covariates, (ii) including a quadratic RD polynomial of the running variable, (iii) selecting different kernel functions (Triangular, Epanechnikov, Uniform), (iv) employing separate optimal bandwidths for each side of the cutoff, (v) clustering standard errors at industry-by-region-by-year level, and (vi) applying bias correction methods. Panel A presents the RD estimates by model choice for floating liens, with 95% confidence intervals. Panel B presents the RD estimates by model choice for non-floating liens, with 95% confidence intervals. The red dot represents the results for the main model specification.

Table 1. Summary Statistics

Notes: This table presents summary statistics on my sample of UK firms with revenue between £10.2 million and £36 million over a four-year period: 2015–2016 (pre-policy) and 2017–2018 (post-policy). Panel A reports statistics on new loan collateralization. *Floating liens* is an indicator variable that equals one if the firm uses floating liens to secure loans from new lenders in a given year, and zero otherwise. *Non-floating liens* is an indicator variable that equals one if the firm uses non-floating liens to secure loans from new lenders in a given year, and zero otherwise. *New loans* is an indicator variable that equals one if the firm obtains secured loans from new lenders in a given year, and zero otherwise. Panel B reports firm characteristics. *Revenue* is defined as the firm's revenue over the 12 months preceding a given year. *Revenue below £25 million* is an indicator variable that equals one if the firm's revenue is below the CCDS threshold of £25 million in a given year, and zero otherwise. *Firm age* is the firm's age in a given year. *Prior relationships* is an indicator variable that equals one if the firm has any secured lending relationship before a given year, and zero otherwise. *Prior relationships with CCDS lenders* is an indicator variable that equals one if the firm has any secured lending relationship with CCDS lenders (the "CCDS9 Banks") before a given year, and zero otherwise. *Prior relationships with non-CCDS lenders* is an indicator variable that equals one if the firm has any secured lending relationship exclusively with non-CCDS lenders before a given year, and zero otherwise. *Employment growth* is the firm's employment growth rate in a given year. *Establishment growth* is the firm's establishment growth rate in a given year. Panel C reports the predetermined financial variables for the post-policy sample. *Log(total assets)* is the log of total assets in 2016 for firms in the post-policy sample. *Leverage ratio* is the ratio of total liabilities to total assets in 2016 for firms in the post-policy sample. *Cash* is the ratio of cash holdings to total assets in 2016 for firms in the post-policy sample. *Current assets* is the ratio of current assets to total assets in 2016 for firms in the post-policy sample. *Bank overdraft* is the ratio of bank overdraft to total assets in 2016 for firms in the post-policy sample. *Trade credit* is the ratio of trade credit to total assets in 2016 for firms in the post-policy sample.

Variables	Observations	Mean	Standard	Percentiles				
			Deviation	10th	25th	50th	75th	90th
Panel A. New Loan Collateralization								
Floating liens	67,143	0.0602	0.2378	0	0	0	0	0
Non-floating liens	67,143	0.0270	0.1622	0	0	0	0	0
New loans	67,143	0.0846	0.2781	0	0	0	0	0
Panel B. Firm Characteristics								
Revenue (million)	67,143	18.177	6.787	11.028	12.528	16.125	22.496	29.045
Revenue below £25 million	67,143	0.8166	0.3870	0	1	1	1	1
Firm age	67,143	24.485	12.703	7	14	24	35	43
Prior relationships	67,143	0.8234	0.3813	0	1	1	1	1
Prior relationships with CCDS lenders	67,143	0.7396	0.4388	0	0	1	1	1
Prior relationships with non-CCDS lenders	67,143	0.0838	0.2771	0	0	0	0	0
Employment growth	66,948	0.0348	0.3063	–0.1280	0	0	0.0490	0.2353
Establishment growth	60,671	0.0161	0.2107	0	0	0	0	0
Panel C. Predetermined Financial Variables								
Log (total assets)	28,700	16.059	1.123	14.912	15.430	15.980	16.611	17.327
Leverage ratio	28,700	0.6039	0.2367	0.2673	0.4258	0.6186	0.7987	0.9466
Cash	28,700	0.1453	0.1622	0.0013	0.0195	0.0870	0.2163	0.3822
Current assets	28,700	0.7450	0.2496	0.3511	0.6002	0.8302	0.9523	0.9879
Bank overdraft	28,700	0.0316	0.0773	0	0	0	0.0173	0.1047
Trade credit	28,700	0.1790	0.1927	0.0017	0.0323	0.1158	0.2605	0.4463

Table 2. The Effect of Open Banking on Loan Collateralization

Notes: This table shows the effect of open banking on loan collateralization. Specifically, it reports results from the following RD regression:

$$Y_{i,t} = \alpha + \tau\,\mathbb{D}_{i,t} + \beta\,R_{i,t} + \gamma\,R_{i,t} \cdot \mathbb{D}_{i,t} + \mu_{j,t} + \nu_{r,t} + \varepsilon_{i,t}$$

where $Y_{i,t}$ is an indicator variable that equals one if firm i uses floating liens (Columns 1–2) or non-floating liens (Columns 3–4) to secure loans from new lenders in year $t = 2017$ or 2018, and zero otherwise. $R_{i,t}$ represents firm i's normalized revenue over the 12 months preceding year t, defined as the amount by which the firm's revenue falls short of the CCDS threshold £25 million. $\mathbb{D}_{i,t}$ is an indicator variable that equals one if firm i's normalized revenue is greater than or equal to zero in year t, and zero otherwise. $\mu_{j,t}$ and $\nu_{r,t}$ represent the industry-by-year and geographic region-by-year fixed effects, respectively. I use the MSE-optimal bandwidth following Calonico, Cattaneo and Titiunik (2014) for triangular kernel weighting method. Standard errors are clustered at the industry-by-region level and are in parentheses. *, **, *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

	Floating Liens		Non-floating Liens	
	(1)	(2)	(3)	(4)
RD Estimate	0.0447***	0.0440***	0.0004	–0.0004
	(0.0118)	(0.0117)	(0.0075)	(0.0075)
Industry-by-year FE	No	Yes	No	Yes
Region-by-year FE	No	Yes	No	Yes
Observations	35,404	35,402	35,404	35,402
Bandwidth (million)	2.440	2.473	3.494	3.502

Table 3. Difference-in-discontinuities Estimation

Notes: This table shows the difference-in-discontinuities estimation. Specifically, it reports results from the following difference-in-discontinuities (DiDC) regression:

$$Y_{i,t} = \alpha_1 + \tau_1 \mathbb{D}_{i,t} + \beta_1 R_{i,t} + \gamma_1 R_{i,t} \cdot \mathbb{D}_{i,t} + \tau_2 \mathbb{D}_{i,t} \cdot \text{Post}_t + \beta_2 R_{i,t} \cdot \text{Post}_t + \gamma_2 R_{i,t} \cdot \mathbb{D}_{i,t} \cdot \text{Post}_t + \mu_{j,t} + \nu_{r,t} + \varepsilon_{i,t}$$

where $Y_{i,t}$ is an indicator variable that equals one if firm i uses floating liens (Columns 1–2) or non-floating liens (Columns 3–4) to secure loans from new lenders in year t, and zero otherwise. $R_{i,t}$ represents firm i's normalized revenue over the 12 months preceding year t, defined as the amount by which the firm's revenue falls short of the CCDS threshold £25 million. $\mathbb{D}_{i,t}$ is an indicator variable that equals one if firm i's normalized revenue is greater than or equal to zero in year t, and zero otherwise. Post_t is a dummy variable that equals one if $t = 2017$ or 2018, and zero if $t = 2015$ or 2016. $\mu_{j,t}$ and $\nu_{r,t}$ represent the industry-by-year and geographic region-by-year fixed effects, respectively. I use the MSE-optimal bandwidth following Calonico, Cattaneo and Titiunik (2014) for triangular kernel weighting method. Standard errors are clustered at the industry-by-region level and are in parentheses. *, **, *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

	Floating Liens		Non-floating Liens	
	(1)	(2)	(3)	(4)
$\mathbb{D} \cdot \text{Post}$	0.0401**	0.0395**	0.0007	0.0006
	(0.0200)	(0.0197)	(0.0125)	(0.0123)
Industry-by-year FE	No	Yes	No	Yes
Region-by-year FE	No	Yes	No	Yes
Observations	67,143	67,137	67,143	67,137
Bandwidth (million)	3.522	3.623	3.500	3.607

Table 4. Information Channel: Heterogeneous Effect by Firm Age

Notes: This table shows the heterogeneous effect by firm age. Specifically, it reports results from the following RD regression:

$$Y_{i,t} = \alpha + \tau \mathbb{D}_{i,t} + \beta R_{i,t} + \gamma R_{i,t} \cdot \mathbb{D}_{i,t} + \mu_{j,t} + \nu_{r,t} + \varepsilon_{i,t}$$

where $Y_{i,t}$ is an indicator variable that equals one if firm i uses floating liens (Columns 1–2) or non-floating liens (Columns 3–4) to secure loans from new lenders in year $t = 2017$ or 2018, and zero otherwise. Columns (1) and (3) focus on the subsample of firms younger than the median age ("young firms"), and columns (2) and (4) focus on the subsample of firms older than the median age ("old firms"). $R_{i,t}$ represents firm i's normalized revenue over the 12 months preceding year t, defined as the amount by which the firm's revenue falls short of the CCDS threshold £25 million. $\mathbb{D}_{i,t}$ is an indicator variable that equals one if firm i's normalized revenue is greater than or equal to zero in year t, and zero otherwise. $\mu_{j,t}$ and $\nu_{r,t}$ represent the industry-by-year and geographic region-by-year fixed effects, respectively. I use the MSE-optimal bandwidth following Calonico, Cattaneo and Titiunik (2014) for triangular kernel weighting method. Standard errors are clustered at the industry-by-region level and are in parentheses. *, **, *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

	Floating Liens		Non-floating Liens	
	Young firms	Old firms	Young firms	Old firms
	(1)	(2)	(3)	(4)
RD Estimate	0.0973***	0.0098	0.0018	–0.0094
	(0.0221)	(0.0123)	(0.0089)	(0.0112)
Industry-by-year FE	Yes	Yes	Yes	Yes
Region-by-year FE	Yes	Yes	Yes	Yes
Observations	17,315	18,087	17,315	18,087
Bandwidth (million)	2.030	4.562	2.529	4.392

Table 5. Information Channel: Heterogeneous Effect by Asset Information-Intensity

Notes: This table shows the heterogeneous effect by asset information-intensity. Specifically, it reports results from the following RD regression:

$$Y_{i,t} = \alpha + \tau \mathbb{D}_{i,t} + \beta R_{i,t} + \gamma R_{i,t} \cdot \mathbb{D}_{i,t} + \mu_{j,t} + \nu_{r,t} + \varepsilon_{i,t}$$

where $Y_{i,t}$ is an indicator variable that equals one if firm i uses floating liens (Columns 1–2) or non-floating liens (Columns 3–4) to secure loans from new lenders in year $t = 2017$ or 2018, and zero otherwise. Columns (1) and (3) focus on the subsample of firms with current assets ratio higher than the median ("high level"), and columns (2) and (4) focus on the subsample of firms with current assets ratio lower than the median ("low level"). $R_{i,t}$ represents firm i's normalized revenue over the 12 months preceding year t, defined as the amount by which the firm's revenue falls short of the CCDS threshold £25 million. $\mathbb{D}_{i,t}$ is an indicator variable that equals one if firm i's normalized revenue is greater than or equal to zero in year t, and zero otherwise. $\mu_{j,t}$ and $\nu_{r,t}$ represent the industry-by-year and geographic region-by-year fixed effects, respectively. I use the MSE-optimal bandwidth following Calonico, Cattaneo and Titiunik (2014) for triangular kernel weighting method. Standard errors are clustered at the industry-by-region level and are in parentheses. *, **, *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

	Floating Liens		Non-floating Liens	
	High level	Low level	High level	Low level
	(1)	(2)	(3)	(4)
RD Estimate	0.1051***	0.0227	0.0040	–0.0088
	(0.0238)	(0.0197)	(0.0113)	(0.0157)
Industry-by-year FE	Yes	Yes	Yes	Yes
Region-by-year FE	Yes	Yes	Yes	Yes
Observations	14,350	14,350	14,350	14,350
Bandwidth (million)	1.556	3.448	4.217	3.393

Table 6. Information Channel: Heterogeneous Effect by Prior Lending Relationships

Notes: This table shows the heterogeneous effect by prior lending relationships. Specifically, it reports results from the following RD regression:

$$Y_{i,t} = \alpha + \tau \mathbb{D}_{i,t} + \beta R_{i,t} + \gamma R_{i,t} \cdot \mathbb{D}_{i,t} + \mu_{j,t} + \nu_{r,t} + \varepsilon_{i,t}$$

where $Y_{i,t}$ is an indicator variable that equals one if firm i uses floating liens (Columns 1–2) or non-floating liens (Columns 3–4) to secure loans from new lenders in year $t = 2017$ or 2018, and zero otherwise. In Panel A, columns (1) and (3) focus on the subsample of firms with prior lending relationships, and columns (2) and (4) focus on the subsample of firms with no prior lending relationships. In Panel B, columns (1) and (3) focus on the subsample of firms with prior CCDS lending relationships, and columns (2) and (4) focus on the subsample of firms with prior non-CCDS lending relationships. $R_{i,t}$ represents firm i's normalized revenue over the 12 months preceding year t, defined as the amount by which the firm's revenue falls short of the CCDS threshold £25 million. $\mathbb{D}_{i,t}$ is an indicator variable that equals one if firm i's normalized revenue is greater than or equal to zero in year t, and zero otherwise. $\mu_{j,t}$ and $\nu_{r,t}$ represent the industry-by-year and geographic region-by-year fixed effects, respectively. I use the MSE-optimal bandwidth following Calonico, Cattaneo and Titiunik (2014) for triangular kernel weighting method. Standard errors are clustered at the industry-by-region level and are in parentheses. *, **, *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Panel A. Prior lending relationships versus no prior lending relationships

	Floating Liens		Non-floating Liens	
	Prior relationships	No prior relationships	Prior relationships	No prior relationships
	(1)	(2)	(3)	(4)
RD Estimate	0.0345***	0.0479**	–0.0058	0.0129
	(0.0122)	(0.0244)	(0.0092)	(0.0138)
Industry-by-year FE	Yes	Yes	Yes	Yes
Region-by-year FE	Yes	Yes	Yes	Yes
Observations	29,168	6,234	29,168	6,234
Bandwidth (million)	3.056	2.543	3.449	4.139

Panel B. Prior lending relationships with CCDS versus non-CCDS lenders

	Floating Liens		Non-floating Liens	
	Prior relationships with CCDS lenders	Prior relationships with non-CCDS lenders	Prior relationships with CCDS lenders	Prior relationships with non-CCDS lenders
	(1)	(2)	(3)	(4)
RD Estimate	0.0364***	0.0180	–0.0049	–0.0101
	(0.0139)	(0.0174)	(0.0104)	(0.0239)
Industry-by-year FE	Yes	Yes	Yes	Yes
Region-by-year FE	Yes	Yes	Yes	Yes
Observations	26,207	2,961	26,207	2,961
Bandwidth (million)	3.071	3.168	3.258	2.864

Table 7. Competition Channel: The Effect of Open Banking on Loan Renewals

Notes: This table shows the effect of open banking on loan renewals. Specifically, it reports results from the following RD regression:

$$Y_{i,t} = \alpha + \tau\,\mathbb{D}_{i,t} + \beta\,R_{i,t} + \gamma\,R_{i,t}\cdot\mathbb{D}_{i,t} + \mu_{j,t} + \nu_{r,t} + \varepsilon_{i,t}$$

where $Y_{i,t}$ is an indicator variable that equals one if firm i uses floating liens (Columns 1–2) or non-floating liens (Columns 3–4) to renew loans from existing lenders in year $t = 2017$ or 2018, and zero otherwise. $R_{i,t}$ represents firm i's normalized revenue over the 12 months preceding year t, defined as the amount by which the firm's revenue falls short of the CCDS threshold £25 million. $\mathbb{D}_{i,t}$ is an indicator variable that equals one if firm i's normalized revenue is greater than or equal to zero in year t, and zero otherwise. $\mu_{j,t}$ and $\nu_{r,t}$ represent the industry-by-year and geographic region-by-year fixed effects, respectively. I use the MSE-optimal bandwidth following Calonico, Cattaneo and Titiunik (2014) for triangular kernel weighting method. Standard errors are clustered at the industry-by-region level and are in parentheses. *, **, *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

	Loan Renewals: Floating Liens		Loan Renewals: Non-Floating Liens	
	(1)	(2)	(3)	(4)
RD estimate	0.0181**	0.0176**	0.0223**	0.0214**
	(0.0090)	(0.0089)	(0.0087)	(0.0088)
Industry-by-year FE	No	Yes	No	Yes
Region-by-year FE	No	Yes	No	Yes
Observations	35,404	35,402	35,404	35,402
Bandwidth (million)	3.718	3.698	3.964	3.876

Table 8. Competition Channel: Heterogeneous Effect by Lending Market Competition

Notes: This table shows the heterogeneous effect by lending market competition. Specifically, it reports results from the following RD regression:

$$Y_{i,t} = \alpha + \tau\,\mathbb{D}_{i,t} + \beta\,R_{i,t} + \gamma\,R_{i,t} \cdot \mathbb{D}_{i,t} + \mu_{j,t} + \nu_{r,t} + \varepsilon_{i,t}$$

where $Y_{i,t}$ is an indicator variable that equals one if firm i uses floating liens (Columns 1–2) or non-floating liens (Columns 3–4) to secure loans from new lenders in year $t = 2017$ or 2018, and zero otherwise. Columns (1) and (3) focus on the subsample of firms located in regions with below-median 2016 HHI ("low HHI"), and columns (2) and (4) focus on the subsample of firms located in regions with above-median 2016 HHI ("high HHI"). $R_{i,t}$ represents firm i's normalized revenue over the 12 months preceding year t, defined as the amount by which the firm's revenue falls short of the CCDS threshold £25 million. $\mathbb{D}_{i,t}$ is an indicator variable that equals one if firm i's normalized revenue is greater than or equal to zero in year t, and zero otherwise. $\mu_{j,t}$ and $\nu_{r,t}$ represent the industry-by-year and geographic region-by-year fixed effects, respectively. I use the MSE-optimal bandwidth following Calonico, Cattaneo and Titiunik (2014) for triangular kernel weighting method. Standard errors are clustered at the industry-by-region level and are in parentheses. *, **, *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

	Floating Liens		Non-floating Liens	
	Low HHI	High HHI	Low HHI	High HHI
	(1)	(2)	(3)	(4)
RD Estimate	0.0588***	0.0077	0.0035	–0.0022
	(0.0150)	(0.0182)	(0.0101)	(0.0123)
Industry-by-year FE	Yes	Yes	Yes	Yes
Region-by-year FE	Yes	Yes	Yes	Yes
Observations	19,712	14,745	19,712	14,745
Bandwidth (million)	2.951	2.663	3.908	3.671

Table 9. Lending Technology Channel: Heterogeneous Effect by Lenders

Notes: This table shows the heterogeneous effect by lenders. Specifically, it reports results from the following RD regression:

$$Y_{i,t} = \alpha + \tau \mathbb{D}_{i,t} + \beta R_{i,t} + \gamma R_{i,t} \cdot \mathbb{D}_{i,t} + \mu_{j,t} + \nu_{r,t} + \varepsilon_{i,t}$$

where $Y_{i,t}$ is an indicator variable that equals one if firm i uses floating liens (Columns 1 and 3) or non-floating liens (Columns 2 and 4) to secure loans from new banks (Columns 1–2) or new non-banks (Columns 3–4) in year $t = 2017$ or 2018, and zero otherwise. $R_{i,t}$ represents firm i's normalized revenue over the 12 months preceding year t, defined as the amount by which the firm's revenue falls short of the CCDS threshold £25 million. $\mathbb{D}_{i,t}$ is an indicator variable that equals one if firm i's normalized revenue is greater than or equal to zero in year t, and zero otherwise. $\mu_{j,t}$ and $\nu_{r,t}$ represent the industry-by-year and geographic region-by-year fixed effects, respectively. I use the MSE-optimal bandwidth following Calonico, Cattaneo and Titiunik (2014) for triangular kernel weighting method. Standard errors are clustered at the industry-by-region level and are in parentheses. *, **, *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

	Banks		Non-banks	
	Floating liens	Non-floating liens	Floating liens	Non-floating liens
	(1)	(2)	(3)	(4)
RD Estimate	0.0292***	–0.0028	0.0174***	0.0050
	(0.0094)	(0.0063)	(0.0065)	(0.0041)
Industry-by-year FE	Yes	Yes	Yes	Yes
Region-by-year FE	Yes	Yes	Yes	Yes
Observations	35,402	35,402	35,402	35,402
Bandwidth (million)	2.865	3.172	2.784	4.446

Table 10. The Real Effects of Open Banking

Notes: This table shows the real effects of open banking. Specifically, it reports results from the following RD regression:

$$Y_{i,t} = \alpha + \tau\,\mathbb{D}_{i,t} + \beta\,R_{i,t} + \gamma\,R_{i,t}\cdot\mathbb{D}_{i,t} + \mu_{j,t} + \nu_{r,t} + \varepsilon_{i,t}$$

where $Y_{i,t}$ is firm i's employment growth (Columns 1–2) or establishment growth (Columns 3–4) in year $t =$ 2017 or 2018. $R_{i,t}$ represents firm i's normalized revenue over the 12 months preceding year t, defined as the amount by which the firm's revenue falls short of the CCDS threshold £25 million. $\mathbb{D}_{i,t}$ is an indicator variable that equals one if firm i's normalized revenue is greater than or equal to zero in year t, and zero otherwise. $\mu_{j,t}$ and $\nu_{r,t}$ represent the industry-by-year and geographic region-by-year fixed effects, respectively. I use the MSE-optimal bandwidth following Calonico, Cattaneo and Titiunik (2014) for triangular kernel weighting method. Standard errors are clustered at the industry-by-region level and are in parentheses. *, **, *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

	Employment Growth		Establishment Growth	
	(1)	(2)	(3)	(4)
RD estimate	0.0266**	0.0273**	0.0025	0.0008
	(0.0128)	(0.0130)	(0.0104)	(0.0101)
Industry-by-year FE	No	Yes	No	Yes
Region-by-year FE	No	Yes	No	Yes
Observations	35,298	35,296	31,986	31,985
Bandwidth (million)	4.310	4.220	4.036	4.222

Internet Appendix for "Data as Collateral: Open Banking for Small Business Lending"

Tong Yu
Imperial College London

December 2, 2024

Contents

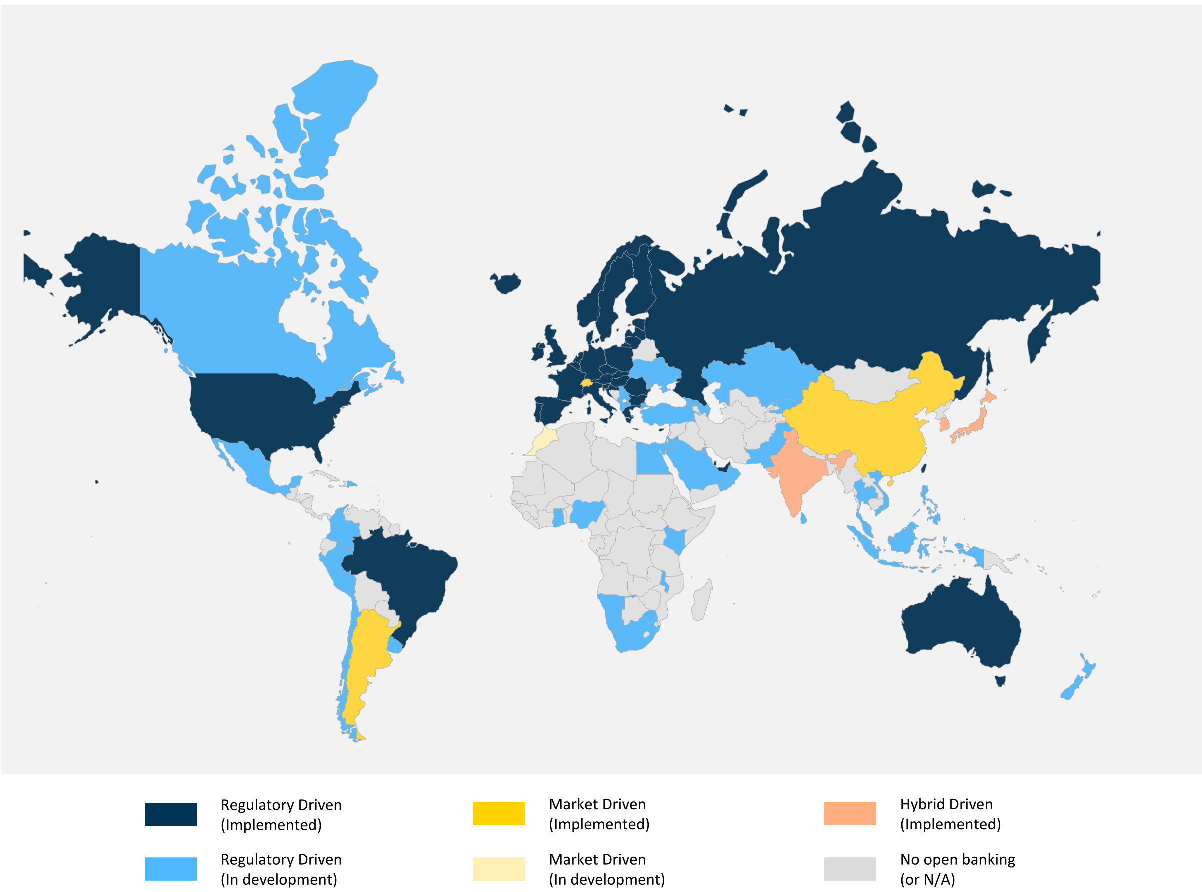
A Background Appendix

A.1 Open Banking

A.1.1 Global Open Banking Adoption

While the UK pioneered the global open banking adoption, an increasing number of countries are now joining this effort, with open banking initiatives proliferating worldwide. As shown in Figure A.1, as of October 2024, over 87 countries, accounting for 95% of the world's GDP in 2023, have implemented or are actively developing open banking frameworks through government initiatives, market-driven efforts, or a combination of both. There is no one-size-fits-all approach to open banking implementation; instead, countries are developing unique strategies tailored to their specific market conditions and policy objectives. Supported by regulatory frameworks, industry coordination, and technological advancements, open banking is poised to transform the financial landscape globally.

Figure A.1. The World Map of Open Banking (2024)



Notes: This figure shows the world map of open banking implementation status as of October 2024.

A.1.2 A Case Study in Market-Driven Open Banking: Switzerland

Switzerland, with a population of 8.7 million, operates within a market-driven open banking ecosystem. The country has 239 banks, and its GDP per capita stands at $92,101. In the absence of formal regulatory frameworks, market forces play a significant role in driving open banking. The Swiss Bankers Association (SBA) and its member financial institutions recognize substantial potential in the opportunities presented by open banking and fostering collaboration between financial institutions and third-party service providers. This case study aims to explore Switzerland's market-led approach to open banking, the role of the SBA and Swiss Fintech Innovations (SFTI) in coordinating industry efforts, and the implications of these initiatives for the financial sector and its participants.

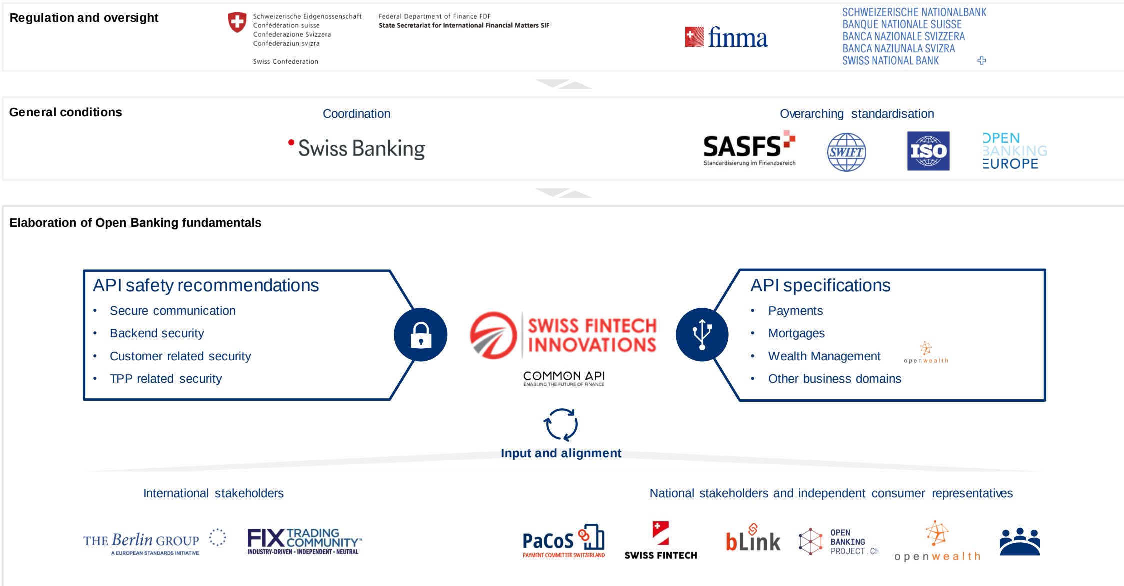
Challenges and Remedies for Traditional Banking. Traditional banks in Switzerland are facing significant challenges because of evolving customer needs, the rise of new stakeholders, and the rapid advancement of innovative technologies. These developments are fragmenting the financial services value chain, with customers increasingly turning to a diverse range of providers, including fintech companies and non-bank entities, to meet their financial needs. In response to these challenges, the SBA recognizes the significant potential of open banking for all market participants. Acknowledging the fragmented value chain and the need for innovation, the SBA is actively contributing to the creation of framework conditions that support these emerging business models, thereby enhancing the competitiveness of Switzerland's financial center.

Simultaneously, the SBA emphasizes the importance of maintaining trust in the financial center, particularly when opening interfaces to third parties. Rather than endorsing regulatory measures like the forced opening of interfaces, the SBA advocates for a market-driven approach where free competition and customer needs dictate the implementation of open banking in Switzerland. This approach ensures that individual banks retain the autonomy to decide whether and with which third-party providers they wish to collaborate.

Market-Driven Approach to Open Banking. Building on this foundation, the SBA, along with Swiss financial institutions, strongly supports a market-driven approach to the evolution of open banking. This approach relies on the financial sector itself to drive innovation and development, without depending on regulatory mandates. The belief is that a self-regulated, market-oriented strategy will naturally foster innovation, efficiency, and competition within the financial ecosystem. By allowing market forces to guide the process, both consumers and providers stand to benefit from more tailored and competitive financial services.

To successfully implement a market-driven approach to open banking, the SBA and SFTI have established a framework for collaboration among all relevant stakeholders (see Figure A.2). This collaboration focuses on creating a shared understanding of the roles each party will play in the future of open banking, particularly concerning API standardization.

Figure A.2. Switzerland's Market-Driven Open Banking Framework



Source: Swiss Bankers Association (2021)

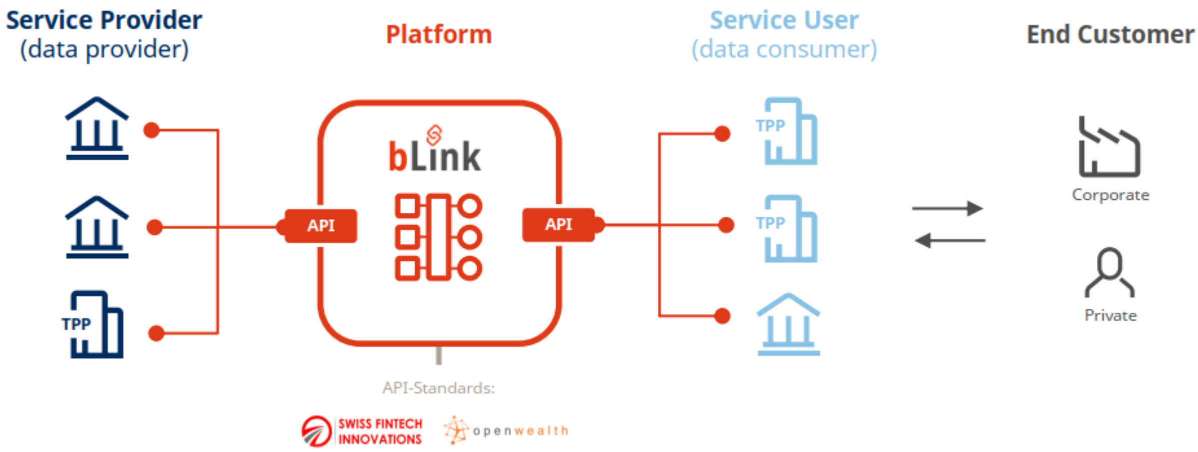
Within this framework, SFTI serves as a central forum, developing essential business and technical principles and providing recommendations for open banking in partnership with leading national and international stakeholders. This broad-based working group includes representatives from banks, insurers, FinTechs, and technology firms. Meanwhile, the SBA plays a vital coordinating role, effectively communicating the industry's concerns to politicians, authorities, and the general public. This clear allocation of roles strengthens the SBA's ability to advocate for the sector.

An important aspect of this collaborative effort is the development of multibanking solutions. Multibanking enables customers to manage accounts across multiple banks through a single platform, using application programming interfaces (APIs) to integrate data from different institutions. Functions may include viewing account balances, submitting payments for approval, and initiating payments from third-party accounts, streamlining financial management for users. Although multibanking solutions have long existed for corporate customers via standards like EBICS and SWIFT, they have struggled to gain traction with private customers in Switzerland due to their complexity. However, recent standardization efforts, such as SFTI's Common API, have begun to make multibanking more accessible to private customers.

A key challenge, however, remains the need for coordinated adoption among banks and companies, as multibanking only becomes viable if multiple players simultaneously offer open APIs. To address this, many Swiss banks, facilitated by the SBA, signed a Memorandum of Understanding (MoU) in May 2023. This MoU outlines: (i) the intention to enable and implement initial multibanking offerings for natural persons; (ii) the minimum content requirements for these initial multibanking offerings; and (iii) the key points regarding the timeline and procedure for implementation. Through this coordinated effort, the SBA aims to ensure the successful adoption of multibanking in Switzerland, supported by consistent technical standards and collaboration across the financial sector.

Innovation, Collaboration, and Future Prospects. The standardization of APIs and the promotion of collaboration among financial institutions are crucial for maintaining Switzerland's leadership in global financial services. A key initiative in this context is the bLink platform, launched by SIX in 2020 in collaboration with leading Swiss financial institutions. Designed as an open banking hub, as shown in Figure A.3, bLink connects data providers and users, such as financial institutions and third parties, via Common API-compatible interfaces.

Figure A.3. The bLink Platform: An Example of Open Banking Infrastructure



Source: SIX Group (2022)

bLink supports the development of services across various use cases, including accounts, payments, and wealth management, for both corporate and private customers. By managing the administrative aspects of participant relationships and ensuring compliance with security standards, bLink enhances efficiency and reliability within the financial ecosystem. Key features of the platform include a uniform platform contract, a standardized admission process for security, efficient digital consent management, and modern, standardized interfaces. As a result, bLink not only facilitates the market adoption of open banking but also drives innovation by enabling both established institutions and new entrants to develop customer-centric services. This dynamic environment fosters ongoing innovation and further strengthens the competitiveness of Switzerland's financial sector.

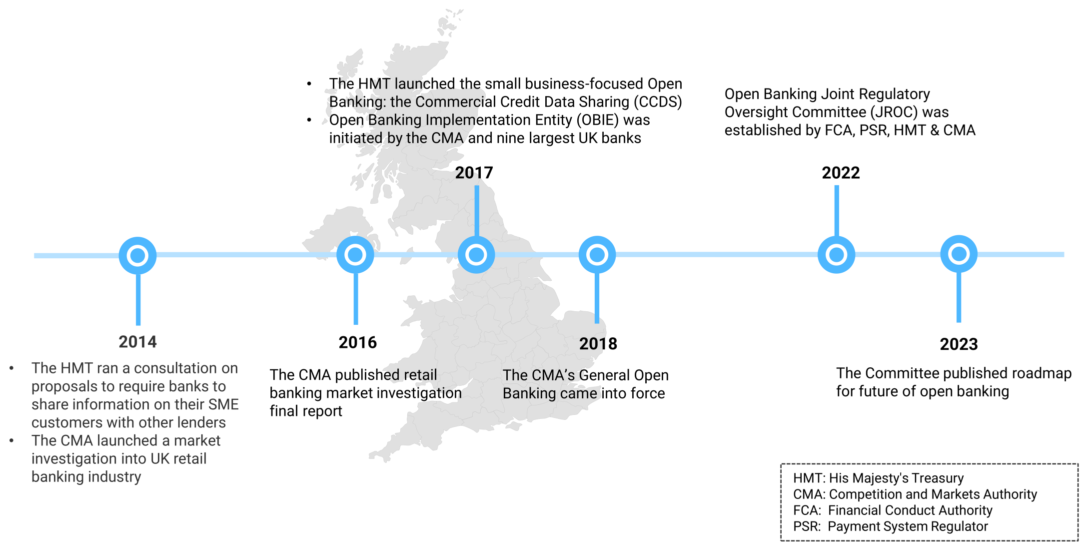
Switzerland is well-positioned to remain at the forefront of financial innovation as it advances open banking through industry-led initiatives. The continued success of this approach will hinge on sustained collaboration among financial institutions, fintechs, and other stakeholders, along with a balanced approach between market-driven efforts and potential regulatory oversight. The Swiss experience demonstrates the effectiveness of a market-driven strategy, offering valuable lessons for other countries developing their own financial frameworks.

In conclusion, Switzerland's commitment to a market-driven approach, reinforced by strong principles and collaboration, is setting a global benchmark for open banking. The country's ability to innovate while upholding high standards of security and customer focus will be crucial in maintaining its leadership in the evolving financial landscape.

A.1.3 UK Open Banking Regulation

The UK has introduced two open banking policies aimed at enhancing financial data sharing and fostering competition. Figure A.4 shows the timeline of UK open banking development.

Figure A.4. UK Open Banking Timeline



Notes: This figure shows the timeline of UK Open Banking regulation.

The development of open banking in the UK has been shaped by a series of regulatory initiatives aimed at fostering competition, transparency, and innovation in the financial services sector. This journey began in 2014, when His Majesty's Treasury (HMT) initiated a consultation on requiring banks to share SME customer data with other lenders, aiming to improve access to credit for small and medium-sized enterprises (SMEs). Concurrently, the Competition and Markets Authority (CMA) launched an in-depth investigation into the UK retail banking sector to address concerns about competition and customer choice.

By 2016, the CMA had published its final report, which highlighted the need for standardized, secure data sharing through open Application Programming Interfaces (APIs). This led to the creation of the Open Banking Implementation Entity (OBIE), tasked with establishing technical standards, data security protocols, and ensuring compliance among major banks. The CMA's recommendations emphasized empowering customers—both individuals and SMEs—by enabling seamless data portability and fostering innovation through third-party providers (TPPs).

Building on these efforts, the UK government introduced the Commercial Credit Data Sharing (CCDS) scheme in 2017 as part of the "Small Business Enterprise & Employment Act 2015." The CCDS initiative aimed specifically at improving credit access for small businesses. Under this scheme, nine major UK banks ("CCDS9 Banks") were mandated to share small business customer data with other lenders, thereby broadening financing opportunities for SMEs. That same year, the OBIE was officially established in collaboration with the CMA and nine largest UK banks ("CMA9 Banks"), setting the stage for a general open banking framework by developing APIs, security standards, and governance systems.

The CMA's general open banking framework was officially launched in 2018, allowing consumers and businesses to securely share their financial data with authorized TPPs. This regulatory milestone aimed to increase competition, expand consumer choice, and stimulate innovation in financial services. By improving information transparency and data accessibility, open banking empowered individuals and businesses to make more informed decisions about financial products.

Recognizing the importance of ongoing oversight, HMT, CMA, Financial Conduct Authority (FCA), and Payment Systems Regulator (PSR) established the Joint Regulatory Oversight Committee (JROC) in 2022. Co-chaired by the FCA and PSR, the JROC was tasked with setting strategic priorities to advance open banking and ensure its continued growth. By this time, open banking adoption had grown significantly, with over 7 million active users in the UK.

In 2023, the JROC published a detailed roadmap for the future of open banking, outlining its priorities for the next two years. These priorities include enhancing the availability and performance of open banking services, mitigating financial crime risks, improving consumer protection, ensuring efficient information flows to TPPs, and promoting innovative services such as variable recurring payments (VRPs). Additionally, the roadmap highlights plans to transition from the OBIE to a new governance structure, ensuring a sustainable and secure future for open banking in the UK.

CCDS versus General Open Banking. The CCDS and the general open banking are both UK government initiatives aimed at enhancing competition and innovation within the financial sector by promoting data sharing. While these initiatives share common goals, they differ in their origins, scope, target audiences, and implementation details.[1]

- Origins and Purpose: The CCDS was launched by the HMT as part of the "Small Business Enterprise & Employment Act 2015" to address challenges in small business lending. The general open banking is introduced under the CMA's Retail Banking Market Investigation Order 2017 to fulfill one of the remedies mandated by the CMA following a market investigation into UK retail banking.

- Participating Banks: The two initiatives involve overlapping but distinct sets of banks. The CCDS mandates participation from nine leading UK banks—HSBC, Barclays, Lloyds, NatWest, Santander, Danske, Bank of Ireland, Allied Irish Bank, and *Clydesdale* (the "CCDS9 Banks"). Meanwhile, the general Open Banking policy applies to the nine largest banks: HSBC, Barclays, Lloyds, NatWest, Santander, Danske, Bank of Ireland, Allied Irish Bank, and *Nationwide* (the "CMA9 Banks").

- Target Audiences: The CCDS targets small businesses with annual revenue up to £25 million, while the general open banking applies to all bank customers–both individuals and businesses–without specific restrictions.

- Data Sharing Channels: The CCDS requires participating banks to share data through the CRAs (who then share the data with other lenders), while the general open banking requires banks to share the data directly with third-parties via standardized APIs.

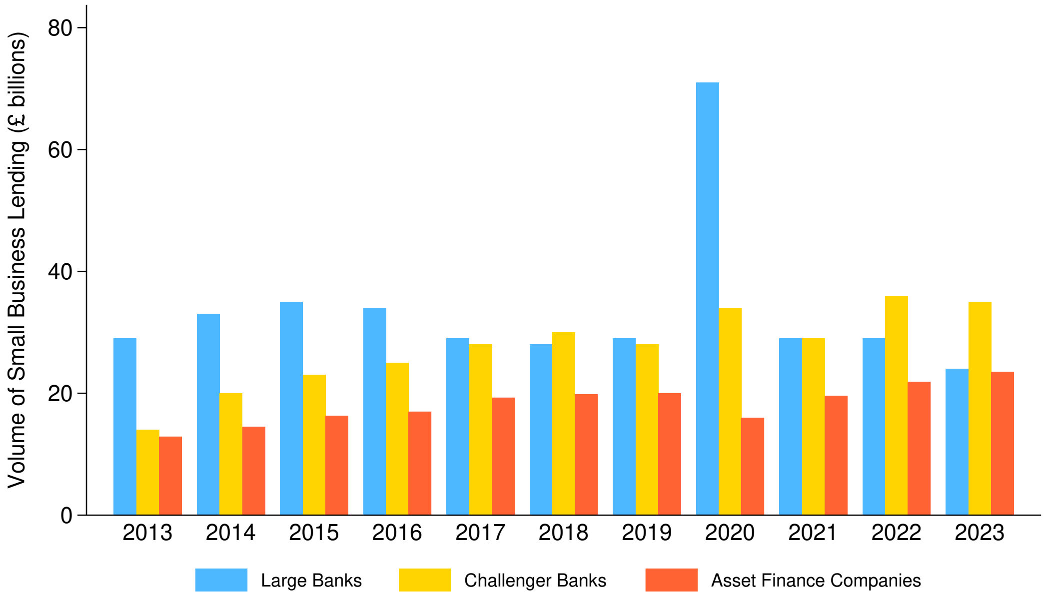
- Types of Data Shared: The CCDS shares business current account data, corporate credit card data, and business loan performance data, while the general open banking focuses on current account data.

A.2 UK Small Business Lending Market

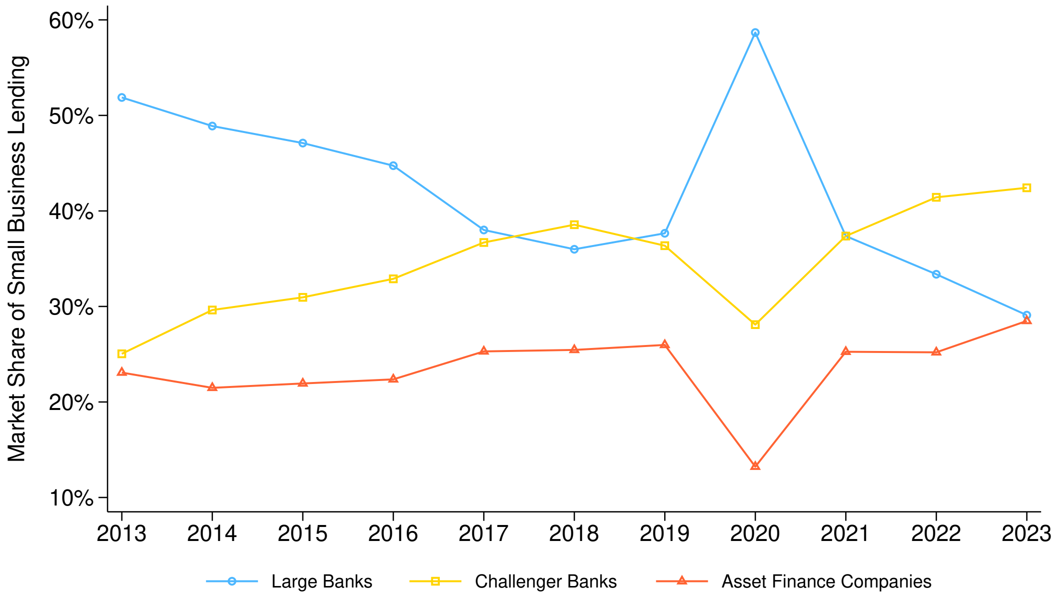
Figure A.5 illustrates some key trends in the UK small business lending market over the last decade (2013–2023): (i) banks continue to be the dominant players, accounting for over 75% of small business lending; (ii) the market share of large banks has been steadily declining since 2013, reaching a record low of 29% in 2023, with the exception of 2020 when large banks played a significant role in sustaining the market through government-backed lending programs during the COVID-19 pandemic; and (iii) challenger banks and alternative lenders (e.g., asset finance companies) have consistently increased their share of the small business lending market, collectively achieving a record high of 71% in 2023, despite a temporary setback in 2020 due to the market disruptions.

[1] See Babina et al. (2024) for a detailed discussion on the CCDS versus general open banking.

Figure A.5. Small Business Lending in the UK (2013–2023)



Panel A. Lending Volume



Panel B. Market Share

Notes: This figure presents the lending volume and market share of small business lending in the UK from 2013 to 2023. The data and definitions are sourced from multiple institutions and are not directly comparable across sources, so the information should be viewed as indicative. The primary data sources include the Bank of England, the British Business Bank, and the Finance & Leasing Association. *Large Banks* refer to the five largest banks in the UK. *Challenger Banks* refer to the challenger and specialist banks in the UK. *Asset Finance Companies* refer to firms categorized as asset finance providers by the Finance & Leasing Association.

B Theoretical Appendix

Consider a loan contract (R, C), increasing collateral C lowers z^*, pushing it closer to the first-best cutoff z^{FB}. However, to satisfy the borrower's binding participation constraint, the repayment R must be reduced at the same time. This adjustment raises z^*, pulling it further from z^{FB}. Although the two adjustments have opposite impacts on the bank's optimal cutoff z^*, the overall impact is that z^* is pushed down.

To see why z^* is pushed down, suppose that the bank's optimal cutoff is currently $z^* = z'$ and suppose that the bank increases C to C' and simultaneously decreases R to R' such that, conditional on $z \geq z'$, the borrower's expected payoff is unchanged:

$$V(R,C) = \int_{z'}^{1} V_z(R,C) f(z) dz = \int_{z'}^{1} V_z(R',C') f(z) dz = V(R',C') \tag{B.1}$$

Although on average, when $z \in [z', 1]$, the borrower remains equally well off, his conditional expected payoff

$$V_z(R',C') > V_z(R,C) \quad \text{for high values of } z \in [z',1] \tag{B.2}$$

and

$$V_z(R',C') < V_z(R,C) \quad \text{for low values of } z \in [z',1] \tag{B.3}$$

This is because

$$\begin{aligned} V_z(R',C') - V_z(R,C) &= p_z(X - R') - (1-p_z)C' - [p_z(X-R) - (1-p_z)C] \\ &= p_z(R - R') - (1-p_z)(C' - C) \end{aligned} \tag{B.4}$$

which is increasing in $z \in [z', 1]$

Similarly, the bank's conditional expected payoff

$$U_z(R',C') > U_z(R,C) \quad \text{for low values of } z \in [z',1] \tag{B.5}$$

and

$$U_z(R',C') < U_z(R,C) \quad \text{for high values of } z \in [z',1] \tag{B.6}$$

That is, when $z = z'$, $U_{z'}(R',C') > U_{z'}(R,C) = U_{z^*}(R,C) = I$, which implies that z' is no longer the optimal cutoff. Because $U_z(R,C)$ is strictly increasing in z, the new optimal cutoff must be lower than z', that is, z^* is pushed down.

C Data Appendix

C.1 Data Access and Matching

I obtain access to the data through my affiliation with the FCA under research agreements. The CCR data is matched to the IDBR data by the ONS data team using vintage versions of the IDBR database. It is important to note that firms in the IDBR dataset do not necessarily have a one-to-one correspondence with firms in the Companies House CRN system. This is because firms in the IDBR data are defined as enterprises, which are described as *the smallest combination of legal units with some degree of autonomy*. Consequently, an enterprise in the IDBR data can consist of more than one legal unit (i.e., company in the Companies House CRN system) if the companies share a common VAT group, PAYE system or report their statistics in the same tax return.

C.2 Sample Selection

My key sample selection criteria are outlined in the main text. For completeness, I describe here the specific conditions under which firms and observations are included in my sample:

1. IDBR Sample

    - I restrict my sample to firms classified as "Company" to which the UK Companies Act 2006 applies. Specifically, I exclude firms with a legal status of "Sole proprietor", "Partnership", "Public corporation", "Central government body", "Local authority", or "Non-profit making body".
    - I exclude firms in certain industries based on the 2007 UK Standard Industrial Classification (SIC) codes. Specifically, I exclude firms operating in financial and insurance activities (2007 SIC Section K: 64110-66300), public administration and defence (2007 SIC Section O: 84110-84300), other service activities (2007 SIC Section S: 94110-96090), activities of households as employers (2007 SIC Section T: 97000-98200), and activities of extraterritorial organisations and bodies (2007 SIC Section U: 99000-99999).
    - I also exclude firms whose death date is specified.
    - I include only firms with annual revenue between £10.2 million and £36 million. (see Table C.1 for UK company size definition)

2. BvD Sample

    - After matching with the IDBR sample, I drop observations with missing data on key balance sheet variables: current assets and fixed assets.

Table C.1. UK Company Size Definition

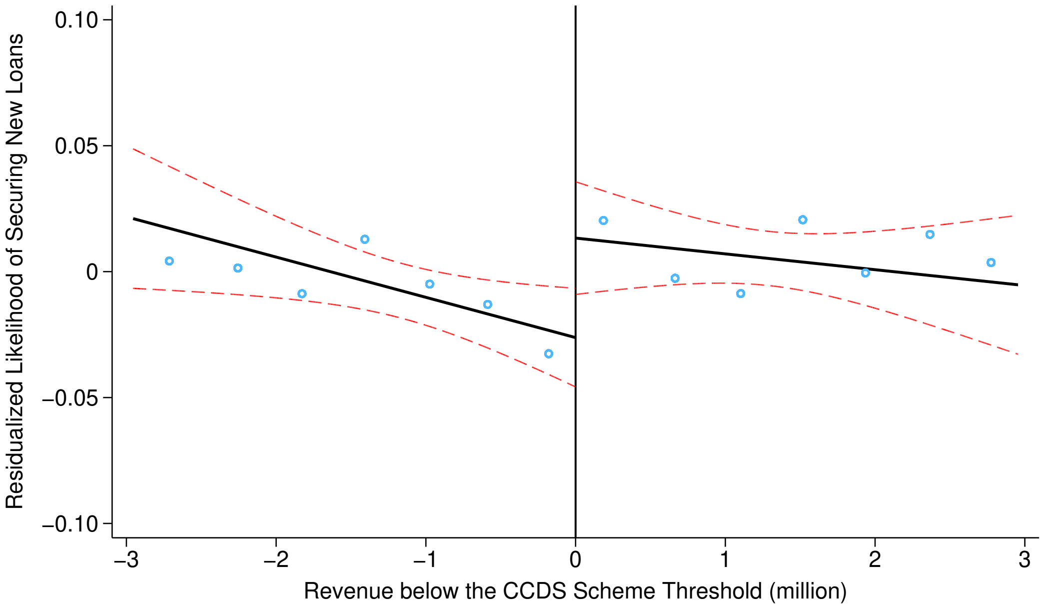
Notes: This table shows the commonly accepted definitions of UK company sizes based on revenue thresholds, as specified by various frameworks, including the UK Government, the Companies Act 2006, and HMRC's Corporation Tax Relief guidelines.

Company Size	**UK Government**	**Companies Act 2006**	**HMRC: Corporation Tax Relief**
Micro	≤ €2,000,000	≤ £632,000	≤ €100,000,000
Small	≤ €10,000,000	≤ £10,200,000	≤ €100,000,000
Medium	≤ €50,000,000	≤ £36,000,000	≤ €100,000,000
Large	> €50,000,000	> £36,000,000	> €100,000,000

D Empirical Appendix

D.1 Additional Results

Figure D.1. RD Plot: Effect of Open Banking on New Loans



Notes: This figure shows the RD plot for the effect of open banking on securing new loans. Specifically, it reports results from the following RD regression:

$$Y_{i,t} = \alpha + \tau\,\mathbb{D}_{i,t} + \beta\,R_{i,t} + \gamma\,R_{i,t}\cdot\mathbb{D}_{i,t} + \mu_{j,t} + \nu_{r,t} + \varepsilon_{i,t}$$

where $Y_{i,t}$ is an indicator variable that equals one if firm i secures loans from new lenders in year $t = 2017$ or 2018, and zero otherwise. $R_{i,t}$ represents firm i's normalized revenue over the 12 months preceding year t, defined as the amount by which the firm's revenue falls short of the CCDS threshold £25 million. $\mathbb{D}_{i,t}$ is an indicator variable that equals one if firm i's normalized revenue is greater than or equal to zero in year t, and zero otherwise. $\mu_{j,t}$ and $\nu_{r,t}$ represent the industry-by-year and geographic region-by-year fixed effects, respectively. Standard errors are clustered at the industry-by-region level. I use the MSE-optimal bandwidth following Calonico, Cattaneo and Titiunik (2014) for triangular kernel weighting method. The corresponding regression results are summarized in Table D.1. Each blue dot represents the average value of the respective outcome variable for firms within a specific bin of normalized revenue. A linear line is fitted separately for observations above and below the threshold, with 95% confidence intervals in red dashed lines.

Table D.1. The Effect of Open Banking on New Loans

Notes: This table shows the effect of open banking on securing new loans. Specifically, it reports results from the following RD regression:

$$Y_{i,t} = \alpha + \tau\,\mathbb{D}_{i,t} + \beta\,R_{i,t} + \gamma\,R_{i,t} \cdot \mathbb{D}_{i,t} + \mu_{j,t} + \nu_{r,t} + \varepsilon_{i,t}$$

where $Y_{i,t}$ is an indicator variable that equals one if firm i secures loans from new lenders in year $t = 2017$ or 2018, and zero otherwise. $R_{i,t}$ represents firm i's normalized revenue over the 12 months preceding year t, defined as the amount by which the firm's revenue falls short of the CCDS threshold £25 million. $\mathbb{D}_{i,t}$ is an indicator variable that equals one if firm i's normalized revenue is greater than or equal to zero in year t, and zero otherwise. $\mu_{j,t}$ and $\nu_{r,t}$ represent the industry-by-year and geographic region-by-year fixed effects, respectively. I use the MSE-optimal bandwidth following Calonico, Cattaneo and Titiunik (2014) for triangular kernel weighting method. Standard errors are clustered at the industry-by-region level and are in parentheses. *, **, *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

	New Loans	
	(1)	(2)
RD Estimate	0.0415***	0.0395***
	(0.0136)	(0.0133)
Industry-by-year FE	No	Yes
Region-by-year FE	No	Yes
Observations	35,404	35,402
Bandwidth (million)	2.812	2.956

Table D.2. The Dynamic Effect of Open Banking on Loan Collateralization

Notes: This table shows the dynamic effect of open banking on loan collateralization. Specifically, it reports results from the following RD regression for each year from 2015 to 2018:

$$Y_{i,t} = \alpha_t + \tau_t \mathbb{D}_{i,t} + \beta_t R_{i,t} + \gamma_t R_{i,t} \cdot \mathbb{D}_{i,t} + \mu_{j,t} + \nu_{r,t} + \varepsilon_{i,t}$$

where $Y_{i,t}$ is an indicator variable that equals one if firm i uses floating liens (Columns 1–2) or non-floating liens (Columns 3–4) to secure loans from new lenders in year t, and zero otherwise. $R_{i,t}$ represents firm i's normalized revenue over the 12 months preceding year t, defined as the amount by which the firm's revenue falls short of the CCDS threshold £25 million. $\mathbb{D}_{i,t}$ is an indicator variable that equals one if firm i's normalized revenue is greater than or equal to zero in year t, and zero otherwise. $\mu_{j,t}$ and $\nu_{r,t}$ represent the industry-by-year and geographic region-by-year fixed effects, respectively. I use the MSE-optimal bandwidth following Calonico, Cattaneo and Titiunik (2014) for triangular kernel weighting method. Standard errors are clustered at the industry-by-region level and are in parentheses. *, **, *** denote statistical significance at the 10%, 5%, and 1% level, respectively. Each cell in the table represents a separate RD regression, where observations and bandwidth vary by specification.

	Floating Liens		Non-floating Liens	
	(1)	(2)	(3)	(4)
2015	0.0049	0.0043	0.0084	0.0059
	(0.0138)	(0.0134)	(0.0102)	(0.0101)
2016	0.0073	0.0071	–0.0065	–0.0065
	(0.0164)	(0.0163)	(0.0110)	(0.0113)
2017	0.0255*	0.0253*	0.0082	0.0060
	(0.0153)	(0.0150)	(0.0142)	(0.0141)
2018	0.0665***	0.0653***	–0.0100	–0.0100
	(0.0198)	(0.0196)	(0.0098)	(0.0098)
Industry-by-year FE	No	Yes	No	Yes
Region-by-year FE	No	Yes	No	Yes

Table D.3. The Effect of Open Banking on Loan Collateralization

Notes: This table shows the effect of open banking on loan collateralization using different definitions of new lender. Specifically, it reports results from the following RD regression:

$$Y_{i,t} = \alpha + \tau\,\mathbb{D}_{i,t} + \beta\,R_{i,t} + \gamma\,R_{i,t}\cdot\mathbb{D}_{i,t} + \mu_{j,t} + \nu_{r,t} + \varepsilon_{i,t}$$

where $Y_{i,t}$ is an indicator variable that equals one if firm i uses floating liens (Columns 1–2) or non-floating liens (Columns 3–4) to secure loans from new lenders in year $t = 2017$ or 2018, and zero otherwise. In Panel A, I define a *new lender* as one with which a firm did not have a lending relationship in the prior 24 months. In Panel B, I define a *new lender* as one with which a firm did not have a lending relationship in the prior 36 months. $R_{i,t}$ represents firm i's normalized revenue over the 12 months preceding year t, defined as the amount by which the firm's revenue falls short of the CCDS threshold £25 million. $\mathbb{D}_{i,t}$ is an indicator variable that equals one if firm i's normalized revenue is greater than or equal to zero in year t, and zero otherwise. $\mu_{j,t}$ and $\nu_{r,t}$ represent the industry-by-year and geographic region-by-year fixed effects, respectively. I use the MSE-optimal bandwidth following Calonico, Cattaneo and Titiunik (2014) for triangular kernel weighting method. Standard errors are clustered at the industry-by-region level and are in parentheses. *, **, *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Panel A. New lender defined as no lending relationship in the prior 24 months

	Floating Liens		Non-floating Liens	
	(1)	(2)	(3)	(4)
RD Estimate	0.0416***	0.0407***	–0.0015	–0.0020
	(0.0123)	(0.0111)	(0.0073)	(0.0073)
Industry-by-year FE	No	Yes	No	Yes
Region-by-year FE	No	Yes	No	Yes
Observations	35,404	35,402	35,404	35,402
Bandwidth (million)	2.514	2.548	3.537	3.555

Panel B. New lender defined as no lending relationship in the prior 36 months

	Floating Liens		Non-floating Liens	
	(1)	(2)	(3)	(4)
RD Estimate	0.0341***	0.0336***	0.0051	0.0043
	(0.0112)	(0.0110)	(0.0069)	(0.0069)
Industry-by-year FE	No	Yes	No	Yes
Region-by-year FE	No	Yes	No	Yes
Observations	35,404	35,402	35,404	35,402
Bandwidth (million)	2.725	2.773	3.548	3.579

Table D.4. Pre-Policy Placebo Test: Heterogeneous Effect by Firm Age

Notes: This table shows the pre-policy placebo test of the heterogeneous effect by firm age. Specifically, it reports results from the following RD regression:

$$Y_{i,t} = \alpha + \tau\,\mathbb{D}_{i,t} + \beta\,R_{i,t} + \gamma\,R_{i,t} \cdot \mathbb{D}_{i,t} + \mu_{j,t} + \nu_{r,t} + \varepsilon_{i,t}$$

where $Y_{i,t}$ is an indicator variable that equals one if firm i uses floating liens (Columns 1–2) or non-floating liens (Columns 3–4) to secure loans from new lenders in year $t = 2015$ or 2016, and zero otherwise. Columns (1) and (3) focus on the subsample of firms younger than the median age ("young firms"), and columns (2) and (4) focus on the subsample of firms older than the median age ("old firms"). $R_{i,t}$ represents firm i's normalized revenue over the 12 months preceding year t, defined as the amount by which the firm's revenue falls short of the CCDS threshold £25 million. $\mathbb{D}_{i,t}$ is an indicator variable that equals one if firm i's normalized revenue is greater than or equal to zero in year t, and zero otherwise. $\mu_{j,t}$ and $\nu_{r,t}$ represent the industry-by-year and geographic region-by-year fixed effects, respectively. I use the MSE-optimal bandwidth following Calonico, Cattaneo and Titiunik (2014) for triangular kernel weighting method. Standard errors are clustered at the industry-by-region level and are in parentheses. *, **, *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

	Floating Liens		Non-floating Liens	
	Young firms	Old firms	Young firms	Old firms
	(1)	(2)	(3)	(4)
RD Estimate	0.0029	0.0085	0.0042	–0.0069
	(0.0133)	(0.0180)	(0.0108)	(0.0115)
Industry-by-year FE	Yes	Yes	Yes	Yes
Region-by-year FE	Yes	Yes	Yes	Yes
Observations	15,806	15,923	15,806	15,923
Bandwidth (million)	3.642	3.558	2.838	3.851

Table D.5. Pre-Policy Placebo Test: Heterogeneous Effect by Asset Information-Intensity

Notes: This table shows the pre-policy placebo test of the heterogeneous effect by asset information-intensity. Specifically, it reports results from the following RD regression:

$$Y_{i,t} = \alpha + \tau\,\mathbb{D}_{i,t} + \beta\,R_{i,t} + \gamma\,R_{i,t} \cdot \mathbb{D}_{i,t} + \mu_{j,t} + \nu_{r,t} + \varepsilon_{i,t}$$

where $Y_{i,t}$ is an indicator variable that equals one if firm i uses floating liens (Columns 1–2) or non-floating liens (Columns 3–4) to secure loans from new lenders in year $t = 2015$ or 2016, and zero otherwise. Columns (1) and (3) focus on the subsample of firms with current assets ratio higher than the median ("high level"), and columns (2) and (4) focus on the subsample of firms with current assets ratio lower than the median ("low level"). $R_{i,t}$ represents firm i's normalized revenue over the 12 months preceding year t, defined as the amount by which the firm's revenue falls short of the CCDS threshold £25 million. $\mathbb{D}_{i,t}$ is an indicator variable that equals one if firm i's normalized revenue is greater than or equal to zero in year t, and zero otherwise. $\mu_{j,t}$ and $\nu_{r,t}$ represent the industry-by-year and geographic region-by-year fixed effects, respectively. I use the MSE-optimal bandwidth following Calonico, Cattaneo and Titiunik (2014) for triangular kernel weighting method. Standard errors are clustered at the industry-by-region level and are in parentheses. *, **, *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

	Floating Liens		Non-floating Liens	
	High level	Low level	High level	Low level
	(1)	(2)	(3)	(4)
RD Estimate	0.0059	0.0032	–0.0140	0.0148
	(0.0187)	(0.0118)	(0.0131)	(0.0158)
Industry-by-year FE	Yes	Yes	Yes	Yes
Region-by-year FE	Yes	Yes	Yes	Yes
Observations	12,884	12,881	12,884	12,881
Bandwidth (million)	3.954	2.923	4.365	3.195

Table D.6. Pre-Policy Placebo Test: Heterogeneous Effect by Prior Lending Relationships

Notes: This table shows the pre-policy placebo test of the heterogeneous effect by prior lending relationships. Specifically, it reports results from the following RD regression:

$$Y_{i,t} = \alpha + \tau \mathbb{D}_{i,t} + \beta R_{i,t} + \gamma R_{i,t} \cdot \mathbb{D}_{i,t} + \mu_{j,t} + \nu_{r,t} + \varepsilon_{i,t}$$

where $Y_{i,t}$ is an indicator variable that equals one if firm i uses floating liens (Columns 1–2) or non-floating liens (Columns 3–4) to secure loans from new lenders in year $t = 2015$ or 2016, and zero otherwise. In Panel A, columns (1) and (3) focus on the subsample of firms with prior lending relationships, and columns (2) and (4) focus on the subsample of firms with no prior lending relationships. In Panel B, columns (1) and (3) focus on the subsample of firms with prior CCDS lending relationships, and columns (2) and (4) focus on the subsample of firms with prior non-CCDS lending relationships. $R_{i,t}$ represents firm i's normalized revenue over the 12 months preceding year t, defined as the amount by which the firm's revenue falls short of the CCDS threshold £25 million. $\mathbb{D}_{i,t}$ is an indicator variable that equals one if firm i's normalized revenue is greater than or equal to zero in year t, and zero otherwise. $\mu_{j,t}$ and $\nu_{r,t}$ represent the industry-by-year and geographic region-by-year fixed effects, respectively. I use the MSE-optimal bandwidth following Calonico, Cattaneo and Titiunik (2014) for triangular kernel weighting method. Standard errors are clustered at the industry-by-region level and are in parentheses. *, **, *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Panel A. Prior lending relationships versus no prior lending relationships

	Floating Liens		Non-floating Liens	
	Prior relationships	No prior relationships	Prior relationships	No prior relationships
	(1)	(2)	(3)	(4)
RD Estimate	0.0046	0.0060	–0.0006	0.0064
	(0.0162)	(0.0189)	(0.0100)	(0.0146)
Industry-by-year FE	Yes	Yes	Yes	Yes
Region-by-year FE	Yes	Yes	Yes	Yes
Observations	26,117	5,618	26,117	5,618
Bandwidth (million)	2.982	4.608	3.043	3.286

Panel B. Prior lending relationships with CCDS versus non-CCDS lenders

	Floating Liens		Non-floating Liens	
	Prior relationships with CCDS lenders	Prior relationships with non-CCDS lenders	Prior relationships with CCDS lenders	Prior relationships with non-CCDS lenders
	(1)	(2)	(3)	(4)
RD Estimate	0.0051	0.0041	–0.0033	0.0107
	(0.0178)	(0.0201)	(0.0113)	(0.0178)
Industry-by-year FE	Yes	Yes	Yes	Yes
Region-by-year FE	Yes	Yes	Yes	Yes
Observations	23,452	2,665	23,452	2,665
Bandwidth (million)	3.049	3.439	2.960	3.345

Table D.7. Pre-Policy Placebo Test: The Effect of Open Banking on Loan Renewals

Notes: This table shows the pre-policy placebo test of the effect of open banking on loan renewals. Specifically, it reports results from the following RD regression:

$$Y_{i,t} = \alpha + \tau\, \mathbb{D}_{i,t} + \beta\, R_{i,t} + \gamma\, R_{i,t} \cdot \mathbb{D}_{i,t} + \mu_{j,t} + \nu_{r,t} + \varepsilon_{i,t}$$

where $Y_{i,t}$ is an indicator variable that equals one if firm i uses floating liens (Columns 1–2) or non-floating liens (Columns 3–4) to renew loans from existing lenders in year $t = 2015$ or 2016, and zero otherwise. $R_{i,t}$ represents firm i's normalized revenue over the 12 months preceding year t, defined as the amount by which the firm's revenue falls short of the CCDS threshold £25 million. $\mathbb{D}_{i,t}$ is an indicator variable that equals one if firm i's normalized revenue is greater than or equal to zero in year t, and zero otherwise. $\mu_{j,t}$ and $\nu_{r,t}$ represent the industry-by-year and geographic region-by-year fixed effects, respectively. I use the MSE-optimal bandwidth following Calonico, Cattaneo and Titiunik (2014) for triangular kernel weighting method. Standard errors are clustered at the industry-by-region level and are in parentheses. *, **, *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

	Loan Renewals: Floating Liens		Loan Renewals: Non-Floating Liens	
	(1)	(2)	(3)	(4)
RD estimate	0.0082	0.0088	0.0042	0.0047
	(0.0113)	(0.0113)	(0.0120)	(0.0118)
Industry-by-year FE	No	Yes	No	Yes
Region-by-year FE	No	Yes	No	Yes
Observations	31,739	31,735	31,739	31,735
Bandwidth (million)	4.663	4.586	3.654	3.673

Table D.8. Pre-Policy Placebo Test: Heterogeneous Effect by Lending Market Competition

Notes: This table shows the pre-policy placebo test of the heterogeneous effect by lending market competition. Specifically, it reports results from the following RD regression:

$$Y_{i,t} = \alpha + \tau \mathbb{D}_{i,t} + \beta R_{i,t} + \gamma R_{i,t} \cdot \mathbb{D}_{i,t} + \mu_{j,t} + \nu_{r,t} + \varepsilon_{i,t}$$

where $Y_{i,t}$ is an indicator variable that equals one if firm i uses floating liens (Columns 1–2) or non-floating liens (Columns 3–4) to secure loans from new lenders in year $t = 2015$ or 2016, and zero otherwise. Columns (1) and (3) focus on the subsample of firms located in regions with below-median 2014 HHI ("low HHI"), and columns (2) and (4) focus on the subsample of firms located in regions with above-median 2014 HHI ("high HHI"). $R_{i,t}$ represents firm i's normalized revenue over the 12 months preceding year t, defined as the amount by which the firm's revenue falls short of the CCDS threshold £25 million. $\mathbb{D}_{i,t}$ is an indicator variable that equals one if firm i's normalized revenue is greater than or equal to zero in year t, and zero otherwise. $\mu_{j,t}$ and $\nu_{r,t}$ represent the industry-by-year and geographic region-by-year fixed effects, respectively. I use the MSE-optimal bandwidth following Calonico, Cattaneo and Titiunik (2014) for triangular kernel weighting method. Standard errors are clustered at the industry-by-region level and are in parentheses. *, **, *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

	Floating Liens		Non-floating Liens	
	Low HHI	High HHI	Low HHI	High HHI
	(1)	(2)	(3)	(4)
RD Estimate	0.0057	0.0051	−0.0089	0.0090
	(0.0114)	(0.0095)	(0.0120)	(0.0114)
Industry-by-year FE	Yes	Yes	Yes	Yes
Region-by-year FE	Yes	Yes	Yes	Yes
Observations	16,828	14,042	16,828	14,042
Bandwidth (million)	2.698	3.340	3.530	2.707

Table D.9. Pre-Policy Placebo Test: Heterogeneous Effect by Lenders

Notes: This table shows the pre-policy placebo test of the heterogeneous effect by lenders. Specifically, it reports results from the following RD regression:

$$Y_{i,t} = \alpha + \tau\,\mathbb{D}_{i,t} + \beta\,R_{i,t} + \gamma\,R_{i,t}\cdot\mathbb{D}_{i,t} + \mu_{j,t} + \nu_{r,t} + \varepsilon_{i,t}$$

where $Y_{i,t}$ is an indicator variable that equals one if firm i uses floating liens (Columns 1 and 3) or non-floating liens (Columns 2 and 4) to secure loans from new banks (Columns 1–2) or new non-banks (Columns 3–4) in year $t = 2015$ or 2016, and zero otherwise. $R_{i,t}$ represents firm i's normalized revenue over the 12 months preceding year t, defined as the amount by which the firm's revenue falls short of the CCDS threshold £25 million. $\mathbb{D}_{i,t}$ is an indicator variable that equals one if firm i's normalized revenue is greater than or equal to zero in year t, and zero otherwise. $\mu_{j,t}$ and $\nu_{r,t}$ represent the industry-by-year and geographic region-by-year fixed effects, respectively. I use the MSE-optimal bandwidth following Calonico, Cattaneo and Titiunik (2014) for triangular kernel weighting method. Standard errors are clustered at the industry-by-region level and are in parentheses. *, **, *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

	Banks		Non-banks	
	Floating liens	Non-floating liens	Floating liens	Non-floating liens
	(1)	(2)	(3)	(4)
RD Estimate	0.0093	–0.0026	–0.0040	0.0027
	(0.0128)	(0.0082)	(0.0074)	(0.0039)
Industry-by-year FE	Yes	Yes	Yes	Yes
Region-by-year FE	Yes	Yes	Yes	Yes
Observations	31,735	31,735	31,735	31,735
Bandwidth (million)	3.777	3.772	4.788	3.367

Table D.10. Pre-Policy Placebo Test: The Real Effects of Open Banking

Notes: This table shows the pre-policy placebo test of the real effects of open banking. Specifically, it reports results from the following RD regression:

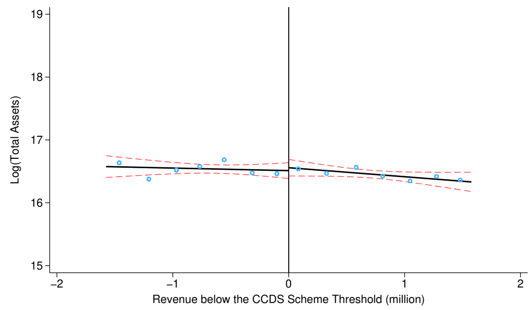
$$Y_{i,t} = \alpha + \tau\,\mathbb{D}_{i,t} + \beta\,R_{i,t} + \gamma\,R_{i,t} \cdot \mathbb{D}_{i,t} + \mu_{j,t} + \nu_{r,t} + \varepsilon_{i,t}$$

where $Y_{i,t}$ is firm i's employment growth (Columns 1–2) or establishment growth (Columns 3–4) in year $t =$ 2015 or 2016. $R_{i,t}$ represents firm i's normalized revenue over the 12 months preceding year t, defined as the amount by which the firm's revenue falls short of the CCDS threshold £25 million. $\mathbb{D}_{i,t}$ is an indicator variable that equals one if firm i's normalized revenue is greater than or equal to zero in year t, and zero otherwise. $\mu_{j,t}$ and $\nu_{r,t}$ represent the industry-by-year and geographic region-by-year fixed effects, respectively. I use the MSE-optimal bandwidth following Calonico, Cattaneo and Titiunik (2014) for triangular kernel weighting method. Standard errors are clustered at the industry-by-region level and are in parentheses. *, **, *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

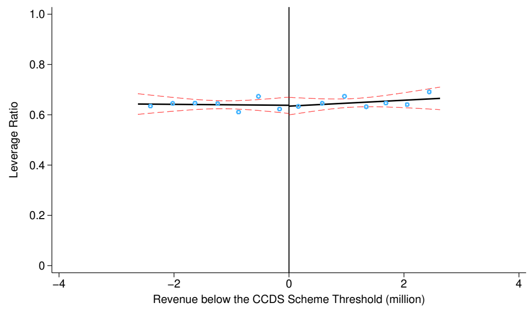
	Employment Growth		Establishment Growth	
	(1)	(2)	(3)	(4)
RD estimate	0.0081	0.0067	0.0236	0.0230
	(0.0165)	(0.0166)	(0.0158)	(0.0158)
Industry-by-year FE	No	Yes	No	Yes
Region-by-year FE	No	Yes	No	Yes
Observations	31,650	31,646	28,685	28,683
Bandwidth (million)	4.636	4.507	3.679	3.655

D.2 Robustness Check

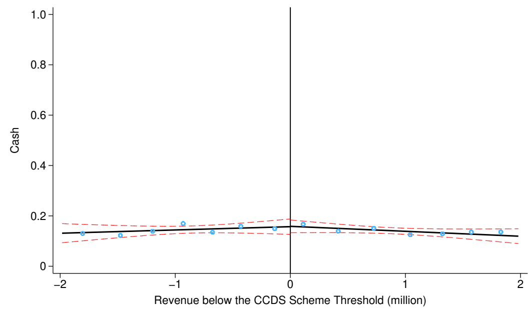
Figure D.2. RD Plot: Covariate Balance Test



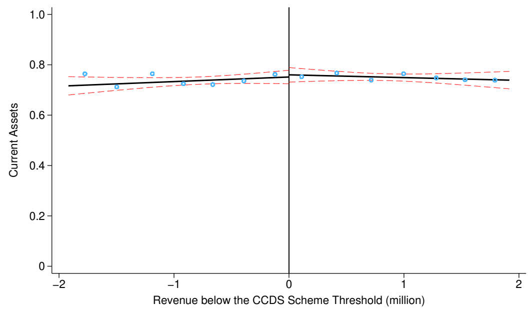
Panel A. Log (Total Assets)



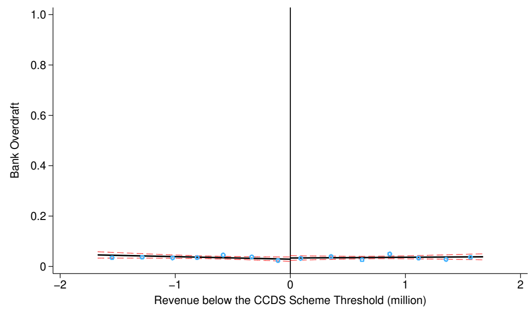
Panel B. Leverage Ratio



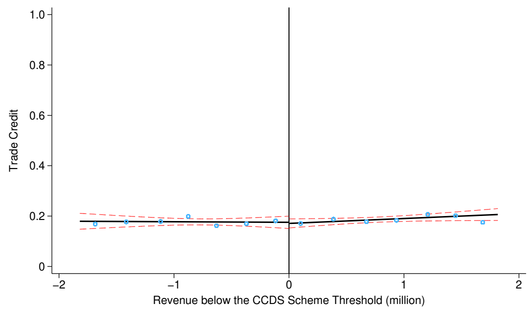
Panel C. Cash



Panel D. Current Assets



Panel E. Bank Overdraft



Panel F. Trade Credit

Notes: This figure shows the RD plots for the covariate balance test using the CER-optimal bandwidth. Specifically, it reports results from the following RD regression:

$$Y_{i,t} = \alpha + \tau\,\mathbb{D}_{i,t} + \beta\,R_{i,t} + \gamma\,R_{i,t} \cdot \mathbb{D}_{i,t} + \varepsilon_{i,t}$$

where $Y_{i,t}$ is one of the six different predetermined financial variables (from the 2016 BvD) for firm i in year $t = 2017$ or 2018. $R_{i,t}$ represents firm i's normalized revenue over the 12 months preceding year t, defined as the amount by which the firm's revenue falls short of the CCDS threshold £25 million. $\mathbb{D}_{i,t}$ is an indicator variable that equals one if firm i's normalized revenue is greater than or equal to zero in year t, and zero otherwise. Standard errors are clustered at the industry-by-region level. I use the CER-optimal bandwidth following Calonico, Cattaneo and Farrell (2018) for triangular kernel weighting method. Each blue dot represents the average value of the respective outcome variable for firms within a specific bin of normalized revenue. A linear line is fitted separately for observations above and below the threshold, with 95% confidence intervals in red dashed lines.

Table D.11. Covariate Balance Test

Notes: This table shows the covariate balance test. Specifically, it reports results from the following RD regression:

$$Y_{i,t} = \alpha + \tau \mathbb{D}_{i,t} + \beta R_{i,t} + \gamma R_{i,t} \cdot \mathbb{D}_{i,t} + \varepsilon_{i,t}$$

where $Y_{i,t}$ is one of the six different predetermined financial variables (from the 2016 BvD) for firm i in year $t = 2017$ or 2018. $R_{i,t}$ represents firm i's normalized revenue over the 12 months preceding year t, defined as the amount by which the firm's revenue falls short of the CCDS threshold £25 million. $\mathbb{D}_{i,t}$ is an indicator variable that equals one if firm i's normalized revenue is greater than or equal to zero in year t, and zero otherwise. I restrict the running variable to a bandwidth of £4 million across tests for consistency and ease of interpretation. As shown in Figure D.2, I also use the CER-optimal bandwidth, which does not significantly change the results. Standard errors are clustered at the industry-by-region level and are in parentheses. *, **, *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

	Log(Total Assets)	Leverage Ratio	Cash Ratio	Current Assets Ratio	Bank Overdraft Ratio	Trade Credit Ratio
RD Estimate	–0.0047	–0.0029	0.0013	–0.0016	0.0019	0.0070
	(0.0478)	(0.0123)	(0.0055)	(0.0170)	(0.0042)	(0.0093)
Observations	28,700	28,700	28,700	28,700	28,700	28,700
Bandwidth (million)	4.000	4.000	4.000	4.000	4.000	4.000

Table D.12. Placebo Cutoffs Test

Notes: This table shows the placebo cutoffs test. Specifically, it reports results from the following RD regression for each cutoff between £21 million to £29 million:

$$Y_{i,t} = \alpha + \tau\,\mathbb{D}_{i,t} + \beta\,R_{i,t} + \gamma\,R_{i,t}\cdot\mathbb{D}_{i,t} + \mu_{j,t} + \nu_{r,t} + \varepsilon_{i,t}$$

where $Y_{i,t}$ is an indicator variable that equals one if firm i uses floating liens (Column 1) or non-floating liens (Column 2) to secure loans from new lenders in year $t = 2017$ or 2018, and zero otherwise. $R_{i,t}$ represents firm i's normalized revenue over the 12 months preceding year t, defined as the amount by which the firm's revenue falls short of the tested threshold between £21 million and £29 million (with £25 million as the true cutoff). $\mathbb{D}_{i,t}$ is an indicator variable that equals one if firm i's normalized revenue is greater than or equal to zero in year t, and zero otherwise. $\mu_{j,t}$ and $\nu_{r,t}$ represent the industry-by-year and geographic region-by-year fixed effects, respectively. I use the MSE-optimal bandwidth following Calonico, Cattaneo and Titiunik (2014) for triangular kernel weighting method. Standard errors are clustered at the industry-by-region level and are in parentheses. *, **, *** denote statistical significance at the 10%, 5%, and 1% level, respectively. Each cell in the table represents a separate RD regression, where bandwidth varies by specification.

	Floating Liens	Non-floating Liens
	(1)	(2)
Cutoff = 21	–0.0055	0.0066
	(0.0193)	(0.0115)
Cutoff = 22	0.0072	0.0092
	(0.0210)	(0.0178)
Cutoff = 23	0.0001	0.0016
	(0.0266)	(0.0186)
Cutoff = 24	–0.0026	–0.0106
	(0.0298)	(0.0094)
Cutoff = 25 (true cutoff)	0.0440***	–0.0004
	(0.0117)	(0.0075)
Cutoff = 26	0.0061	–0.0146
	(0.0335)	(0.0244)
Cutoff = 27	–0.0128	–0.0114
	(0.0296)	(0.0134)
Cutoff = 28	–0.0126	0.0109
	(0.0258)	(0.0080)
Cutoff = 29	0.0035	0.0123
	(0.0231)	(0.0193)
Industry-by-year FE	Yes	Yes
Region-by-year FE	Yes	Yes
Observations	35,402	35,402

Table D.13. Donut Hole Test

Notes: This table shows the donut hole test. Specifically, it reports results from the following RD regression, excluding observations within a donut hole radius of 0 to 0.10 around the CCDS threshold:

$$Y_{i,t} = \alpha + \tau\,\mathbb{D}_{i,t} + \beta\,R_{i,t} + \gamma\,R_{i,t} \cdot \mathbb{D}_{i,t} + \mu_{j,t} + \nu_{r,t} + \varepsilon_{i,t}$$

where $Y_{i,t}$ is an indicator variable that equals one if firm i uses floating liens (Column 1) or non-floating liens (Column 2) to secure loans from new lenders in year $t = 2017$ or 2018, and zero otherwise. $R_{i,t}$ represents firm i's normalized revenue over the 12 months preceding year t, defined as the amount by which the firm's revenue falls short of the CCDS threshold £25 million. $\mathbb{D}_{i,t}$ is an indicator variable that equals one if firm i's normalized revenue is greater than or equal to zero in year t, and zero otherwise. $\mu_{j,t}$ and $\nu_{r,t}$ represent the industry-by-year and geographic region-by-year fixed effects, respectively. I use the MSE-optimal bandwidth following Calonico, Cattaneo and Titiunik (2014) for triangular kernel weighting method. Standard errors are clustered at the industry-by-region level and are in parentheses. *, **, *** denote statistical significance at the 10%, 5%, and 1% level, respectively. Each cell in the table represents a separate RD regression, where observations and bandwidth vary by specification.

	Floating Liens	Non-floating Liens
	(1)	(2)
Donut Hole Radius = 0	0.0440***	–0.0004
	(0.0117)	(0.0075)
Donut Hole Radius = 0.01	0.0384***	0.0041
	(0.0124)	(0.0079)
Donut Hole Radius = 0.02	0.0381***	0.0038
	(0.0125)	(0.0079)
Donut Hole Radius = 0.03	0.0324***	0.0034
	(0.0121)	(0.0079)
Donut Hole Radius = 0.04	0.0290**	0.0080
	(0.0119)	(0.0086)
Donut Hole Radius = 0.05	0.0338***	0.0080
	(0.0124)	(0.0085)
Donut Hole Radius = 0.06	0.0445***	0.0071
	(0.0133)	(0.0084)
Donut Hole Radius = 0.07	0.0440***	0.0062
	(0.0134)	(0.0083)
Donut Hole Radius = 0.08	0.0361***	0.0055
	(0.0131)	(0.0083)
Donut Hole Radius = 0.09	0.0308**	0.0055
	(0.0124)	(0.0084)
Donut Hole Radius = 0.10	0.0312**	0.0024
	(0.0126)	(0.0086)
Industry-by-year FE	Yes	Yes
Region-by-year FE	Yes	Yes

Table D.14. Bandwidth Sensitivity Test

Notes: This table shows the bandwidth sensitivity test. Specifically, it reports results from the following RD regression, using different bandwidths:

$$Y_{i,t} = \alpha + \tau\,\mathbb{D}_{i,t} + \beta\,R_{i,t} + \gamma\,R_{i,t}\cdot\mathbb{D}_{i,t} + \mu_{j,t} + \nu_{r,t} + \varepsilon_{i,t}$$

where $Y_{i,t}$ is an indicator variable that equals one if firm i uses floating liens (Column 1) or non-floating liens (Column 2) to secure loans from new lenders in year $t = 2017$ or 2018, and zero otherwise. $R_{i,t}$ represents firm i's normalized revenue over the 12 months preceding year t, defined as the amount by which the firm's revenue falls short of the CCDS threshold £25 million. $\mathbb{D}_{i,t}$ is an indicator variable that equals one if firm i's normalized revenue is greater than or equal to zero in year t, and zero otherwise. $\mu_{j,t}$ and $\nu_{r,t}$ represent the industry-by-year and geographic region-by-year fixed effects, respectively. I use the triangular kernel weighting method for the RD estimation. Standard errors are clustered at the industry-by-region level and are in parentheses. *, **, *** denote statistical significance at the 10%, 5%, and 1% level, respectively. Each cell in the table represents a separate RD regression, where bandwidth varies by specification.

	Floating Liens		Non-floating Liens
Bandwidth (million)	(1)	Bandwidth (million)	(2)
Bandwidth = 1	0.0670***	Bandwidth = 2	0.0011
	(0.0176)		(0.0097)
Bandwidth = 1.5	0.0604***	Bandwidth = 2.5	–0.0009
	(0.0146)		(0.0088)
Bandwidth = 2	0.0521***	Bandwidth = 3	0.0004
	(0.0128)		(0.0081)
Bandwidth = 2.473	0.0440***	Bandwidth = 3.502	–0.0004
(MSE-optimal)	(0.0117)	(MSE-optimal)	(0.0075)
Bandwidth = 3	0.0369***	Bandwidth = 4	–0.0006
	(0.0107)		(0.0070)
Bandwidth = 3.5	0.0320***	Bandwidth = 4.5	–0.0005
	(0.0100)		(0.0066)
Bandwidth = 4	0.0298***	Bandwidth = 5	0.0007
	(0.0094)		(0.0063)
Industry-by-year FE	Yes	Industry-by-year FE	Yes
Region-by-year FE	Yes	Region-by-year FE	Yes
Observations	35,402	Observations	35,402