Sharing SME Cash Flow Data

January 2023

Abstract

This paper shows that non-banks lend more to SMEs when they gain access to banks' private information. For identification, we exploit a government policy that forced nine banks in the UK to share their corporate clients' data, including monthly current account balances, with other lenders. We find that treated SMEs, i.e. those with turnover just below £25 million, are more likely to start new relationships with non-banks compared to firms above the cutoff. This result holds even if firms' credit histories are already observable by other lenders, suggesting a unique role of cash flow information. The policy has modest real effects, as affected firms increase their liquidity buffers rather than investment. This evidence suggests that open banking can alleviate the hold-up problem, and stimulate competition in the SME lending market.

Keywords: Open banking, small business lending, big data, bank competition;

JEL: G21; G23; G38;

1 Introduction

After the 2008 financial crisis, non-bank lenders have gained significant market shares in the small business lending market, partly substituting credit from traditional banks. Figure 1 shows that non-banks represent an increasingly large share of new secured loans made to Small and Medium Enterprises (SMEs) in the UK, going from 24% in 2007 to 37% in 2019 (Gopal and Schnabl (2022) reports similar patterns for the US). This growth may be explained by the increased availability of large amounts of data that are used to support lending decisions.[1] By accessing newly available big data, in fact, non-banks or FinTech companies can leverage their supposed competitive advantage in processing and utilising information compared to traditional banks.

[Figure 1 here]

But is the competitive advantage of non-bank lenders driven by information processing abilities (i.e. better internal credit models) or by convenience of the customer-lender relationship (i.e. app-based or online interactions) which reduce overall processing times (Buchak et al., 2018; Fuster et al., 2019)? Answering this question empirically is difficult because banks and non-banks may serve different customers and may have access to different types of information. For example, banks not only lend but also provide payment services to prospective borrowers, thus gathering information on their checking accounts.

To study the effect of information processing by banks and non-banks in the SME lending market, this paper empirically tests the effects of a data sharing policy the UK Government launched in the Spring 2016. The "Commercial Credit Data Sharing" (CCDS) policy mandated the nine largest UK banks[2] to share loan repayment, corporate credit card and current account information of their SME clients with any other UK lenders that requests the information, including both other banks and alternative finance providers (i.e. non-banks). Upon borrower consent, these new finance providers receive credit and business account information to check credit worthiness of potential customers. Importantly, the initiative is applicable only to SMEs with annual turnover below £25 million, providing a quasi-random treatment incidence that is useful for identification. Moreover, previous initiatives in the UK made the credit histories of borrowers already publicly available. The new policy mostly revealed information about cash flows that were previously solely the domain of the banks providing payment services.

[1] Part of the increase is due to banks' retreating from riskiest segments of the credit market due to the increased regulatory burden after the financial crisis (Cortés et al., 2020)

[2] The list of banks participating in the scheme is available here.

Using near universal data on UK SME's secured borrowing relationships combined with their financial statements, this paper shows that sharing data allows SMEs to form new borrowing relationships. The scheme increased the probability of a treated SME receiving a new secured loan from a different lender by almost 30%. When decomposing these new relationships by lender type, we find that this effect is only statistically significant for new relationships created with non-banks and the point estimate is 50% larger than the equivalent for new relationships formed with a bank-lender. This finding is consistent with the idea that non-bank lenders are better able to take advantage of new, proprietary information about SME borrowers. In turn, this contributes to the erosion of market power in banks.

We conduct several robustness tests to address any endogeneity concerns that may bias our estimates. In particular, although our identification strategy helps to alleviate concerns about the comparability between treated and control firms, a matching strategy on firm observables confirm the baseline results. Moreover, we show that our choices for the length of turnover windows, or the lagged period within which we consider existing relationships, do not qualitatively affect our results.

We continue the analysis by exploring heterogeneous effects of the data sharing scheme. In our baseline we focus on all firms with an exist lending relationship. However, the main effects come from companies that had a relationship with one of the nine banks the CCDS scheme targeted. Although the scheme requires reciprocal information sharing, this is with a lag.[3] Hence, consistent with our results it is the original 9 lenders the ones providing most of the new information.

Moreover, we find that the probability a SME switches to a new lender increases with the number of lenders the firm had at the time of the policy. This result suggests that sharing data from multiple sources (possibly about multiple products) can increase even further the the ability to switch lender. Finally, we find that the probability to take a loan from a new lender decreases in the liquidity (cash over assets) ratio of the firm. This result suggests that financially (cash) constrained firms benefit

[3] Non-designated lenders that decide to join the scheme are required to start providing equivalent data within 12 months of receiving the data from the CRA.

the most from data sharing, as the hold-up problem may be more severe for them (Sharpe, 1990; Rajan, 1992).

In the second part of the paper, we conduct a real effect analysis to investigate how the CCDS policy affected the SMEs that made use of the scheme. Focusing on firm-level outcomes, we first confirm that, in line with the idea that data sharing produces better outcomes for borrowers, the scheme affects loan pricing: switchers to non-bank lenders are able to decrease their interest rate expenses. We also find that firms that take CCDS loans from non-banks primarily increase their short-term debt. This may reflect the short sample of data being shared (i.e. a 12 month history). Consistent with the results on liquidity and new relationship formation, we find that the firms use the new credit to increase their liquid asset positions, arguably for precautionary reasons. We do not find that firms invest more (i.e. increase their fixed assets).

The contribution of this paper is threefold. First, academic scholars have started investigating the effects that government-led open banking initiatives may have on competition and welfare. Sharing bank clients's transaction and repayment data with non-banks may improve financial innovation and financial inclusion in the venture capital industry (Babina et al., 2022). In this paper, we show that opening bank SME clients' data can increase competition from non-bank lenders. In the long-run, however, open banking may disrupt information spillover between bank lending and payment service (Parlour et al., 2022), exacerbate winner's curse in terms of depositor funding and monitoring (Goldstein et al., 2022), and over-empower FinTech lenders (He et al., 2020).

Second, this paper contributes to the growing literature on lending from non-bank or "FinTech" lenders. The share of non-bank lending has been increasing in recent years and many papers investigate whether non-bank lending is a substitute or a complement to bank lending, especially in the consumer or mortgage credit market (Buchak et al., 2018; Tang, 2019; Fuster et al., 2019; Di Maggio and Yao, 2020; de Roure et al., 2021). Fewer papers have focused on lending to SMEs (Gopal and Schnabl, 2022; Beaumont et al., 2022; Eça et al., 2021). The paper most similar to ours is Ghosh et al. (2021), which makes a link between information advantage in screening and monitoring and the ability to observe payments and deposit flows by a FinTech lender in India. Our novel contribution is to exploit a policy shock to the information set of all lenders (including non-banks), which allows them to get new information about existing banks' clients. The fact that non-bank lenders are the ones taking advantage of it indicates their comparative advantage in data processing.

Third, our findings are related to the vast literature studying the effects of sharing credit information among lenders through credit bureaus or registers (Giannetti et al., 2017; Liberti et al., 2022). These studies generally find that the formation of information sharing technology have positive effects on credit allocation. Our paper contributes to this literature by showing that forcing a bank to disclose information with non-bank lenders can help them gaining market share. An additional feature of our setting, is that the shared information relates to cash flows (not credit histories). Our findings suggest that firms' transactions contain additional information that lenders view as informative. Transactions data has traditionally been available to the bank providing payment services alone suggesting an entrenched informational advantage for these banks that forced information sharing can unwind.

2 Institutional Background

Bank credit is the most prominent source of external financing for SMEs because of asymmetric information problems (Diamond, 1984; Ramakrishnan and Thakor, 1984). Banks engage in relationship lending to overcome these problems and "know" their borrowers (Boot, 2000). If this information remains private, banks can extract rents and their older SME clients end up paying higher premiums, ending up in the so-called *hold-up* problem (Sharpe, 1990; Rajan, 1992). As small businesses represent a large share of overall employment,[4] there are potentially large economic costs of the hold-up problems and bank market power.

Data-driven innovations of the twenty-first century have reduced the cost of pro-

[4]In the UK, SMEs represent 99% of all businesses, and account for 60% of total employment (see U.K., Department for Business, p. 3). In the EU, the figures are similar.

cessing, storing and sharing information (Byrne et al., 2015), and have encouraged policy-makers all around the world to start thinking about ways to tackle the *hold-up* problem pervading SME credit markets.[5] One of the first countries that enabled consumers and businesses to share their data – possibly through their relationship bank – was the United Kingdom. In November 2015, concerned with the high concentration of its national banking market, the UK Government enacted the "Small Business Enterprise & Employment Act 2015" (the Act). The intention of the Act was to lower entry barriers for alternative credit providers in the SME credit segment, thereby actively stimulating competition. The Act comprised of several initiatives, including the Commercial Credit Data Sharing (CCDS) scheme.[6] We describe this initiative in detail as it provides the key identifying variation we use in the empirical analysis.

The CCDS regulation required nine (large) UK banks to share current account data, as well as up-to-date performance of loans and corporate credit cards, of all their SME customers (defined as those annual turnover below £25 million) with alternative lenders (including both banks and non-banks) via four designated Credit Reference Agencies (CRAs).[7] As of April 2016, each CRA receives the raw transactions on a monthly basis, and consolidates data in a common format such that can be easily delivered to any credit providers that want to make a lending decision to a particular new borrower. The designated CRA must provide the raw data to any alternative finance providers which, in turn, commit to sharing their own SME portfolio data with the designed CRA within one year (*reciprocity rule*). While data sharing with the CRA is in principle dependent upon SME approval (as it is the case for open banking), consent typically happens automatically because it may be buried in general (historic) contract terms with the main lender. Active consent from the SME is instead needed when the CRA shares data with alternative lenders, which happens when the SME applies for credit with the alternative lender. Essentially, the CCDS policy creates

[5]g., Joint Resolution CMN-BCB No. 1/20 in Brazil, Dodd-Frank Section 1033 in US. See Babina et al. (2022) for a cross-country evidence of open banking policies.

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

[7]These are: Experian, Equifax, Dun & Bradstreet, and Creditsafe

what amounts to a credit register (with additional non-credit information).

It is important to note that this is not the first credit data sharing agreement among UK lenders: for example, the 2014 Credit Account Information Sharing (CAIS) from Experian provides lenders with past credit history and default status. However, CCDS offers greater coverage in terms of breadth and depth of detailed data. Under CCDS, lenders are required to share each individual facility of eligible SMEs, including all loan types (e.g., secured vs. unsecured, mortgage, primary vs. secondary lease, etc.), and asset finance products (e.g., dealer buy-back, contract hire, credit sale fixed term, etc.) as well as a twelve-month history snapshots of their monthly current account balance. The latter item is one of the key innovations of the CCDS policy compared to previous data sharing agreements. Accessing a firm's current account data allows alternative lenders to precisely estimate monthly cash flows and better assess firm performance and risk.

In principle, there are considerable benefits for finance providers. SME business account data can be easily integrated into lenders' risk models to better measure viability and affordability of the customer. Also, underwriter time can be released to improve productivity. Removal of process friction reduces costs and increases credit decision speed, creating competitive advantage that drives higher proposal volumes. Moreover, harmonized financial data reduces fraud risk possibly coming from different bank statements. Finally, lenders can benefit from easier monitoring of ongoing liquidity and resilience of customers.

3 Data and Empirical Strategy

Firms in the UK are required to report all claims ("charges") against their assets by lenders, including the name of the charge holder (inclusive of non-banks), the date the claim commenced, and when the charge ceases, to Companies House (the UK firm Registrar).[8] The charge can be against a specific asset or it can be floating a charge

[8] These reports are similar to Uniform Commercial Code (UCC) data on SME lending in the US where lenders make filings on all secured loans to preserve priority in bankruptcy (Gopal and Schnabl, 2022).

covering entirety of the firm's balance sheet or its outstanding invoices in the case of invoice financing. There are strong incentives to ensure this data are accurately reported. Lenders have 21 days to formally register their claim (or face legal barriers to repossessing the assets). Borrowers have an incentive to declare when a charge is satisfied to unencumber their assets.[9]

Observing the charges amounts to observing the set of secured lending relationships between firm and it lenders of all different types. We do not observed unsecured claims. However, the overwhelming majority of loans to UK SMEs are collateralized and hence this data provides a highly representative and timely view of a firm's lending relationships. [10]

The raw information on charges is collected from Companies House by Bureau Van Dijk (BVD) and released in a usable, electronic form as part of their FAME database. BvD data also provides annual firm-level financial information matched to chargeholders information. However, BvD data is well known for suffering from survivorship bias and various issues with constructing consistent historical panels (Kalemli-Ozcan et al., 2015). To alleviate this concern and maximize coverage with historical observations, we use annually sampled archived vintages of the FAME database, as in Bahaj et al. (2020), to compile our final panel dataset.

3.1 Data

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

[9]As we observe the start date of outstanding claims prior to the start of our firm panel, our data is not susceptible to the problems of truncation that often afflict the computation of relationship lengths (see e.g., (Ongena and Smith, 2001)).

[10]The Bank of England's 2015 survey of UK SME and Mid-Corporate Lending covered loans from the five major UK banks to businesses with annual revenue of at most £500 million and which were borrowing at least £250k. In this survey, 97% of lending to limited liability companies was secured.

Even though CCDS targets 99.9% of the entire business population, we focus on a relatively narrow sample window of firms with a 2016 turnover between £10 million and £40 million to cleanly identify the effect of data sharing on relationship formations (alternative bandwidths yield similar results). This strategy ensures that we compare firms around the threshold that are similar with each other except for the policy incidence. The CCDS effective date was April 1st 2016, but it went live in 2017 only for technical complications in supplying data. We assign treatment based on turnover reported in 2016, and track treated and control firms over a six-year time period. In turn, $I< (25m)_i$ is the treatment indicator taking a value of one if SME i is below the treatment threshold in 2016.

One limitation of our study is that Turnover is frequently not reported in BVD. our treatment assignment is based on turnover in 2016, which may not always be reported and may not capture all firms that in reality used the scheme.

Following Ioannidou and Ongena (2010), we consider a firm as forming a "new" relationship if in a particular year it adds one (or more) new lender(s) that is not part of the set of lenders from whom the firm had a charge outstanding in the previous three years.[11] This choice is consistent with empirical evidence suggesting that most bank information is produced during the first years of a lending relationship (Cole, 1998). We define $NewLender_{i,t}$ a dummy equal to one when SME i takes out forms a new relationship in year t. We construct alternative dummy variables that take a value of 1 if the firm forms a new relationship with a bank, $NewBank_{i,t}$, or a non-bank, $NewNonBank_{i,t}$. Additionally, we define $NoCredit_{i,t}$, $SingleRelati, t$ and $MultiRelati, t$ as dummy variables that take a value of one if the firm has, respectively, no outstanding relationships (i.e. no charges), a relationship with a single lender or a relationship with multiple lenders.

The stage of existing relationships is important for the investigation of new relationship formations and the strength of potential hold up problems. Hence, we compute the length of a firm's outstanding relationships at the time of the policy change. We define $RelLength_i = log(1 + \text{number of months since start})$. In case of a relationship

[11]Our results are robust to different (either shorter or longer) time windows.

with multiple lenders, we average the length of all the outstanding relationships the firm has at a given point in time. For a firm that takes more than one charge from the same lender over time, we consider the oldest one and calculate the relationship length from there.

To construct our final sample, we restrict our attention to limited liability firms that are not in the following industries: mining, utilities, finance, public administration, education and health. This allows focus on non-financial firms with their own balance sheet in sectors with limited public intervention. Also, we condition the sample on firms reporting turnover as well as information to compute our baseline control variables: total assets, non-equity liabilities, cash holdings, and QuiScore (a measure of credit risk) for at least one year in each of the pre- and post-treatment time windows. This leaves us with a sample of 39,089 observations on 6,886 unique firms.

Table 1 presents the summary statistics of the variables. All ratios are winsorized at 1% level.

[Table 1 here]

The probability that an SME takes a loan from a new lender is 5.3%. Over the entire sample period, SMEs switch more often to banks than to non-banks. Note that banks were responsible for 72.3% of observed relationships in 2016 in our sample so this may simply reflect that there are more banks than non-banks in the SME lending market. The 57.2% of our firms have a relationship with a CCDS-designated bank in 2016, and this increases to 72.8% when conditioning on the firm having a charge outstanding.

A significant number of firms (21%) have no secured credit relationship (i.e. there is no charge against their assets in the data), and 56% of those that do have only one lender at the time of the policy introduction. The average firm in our sample is 24 years old and has a leverage ratio (defined as the ratio of total non-equity liabilities to total assets) of 58.7%.

Next, we can see that firms are in general considered safe, as the average quiscore is 90.3 (where larger values mean lower risk). This is also reflected in the high mean value for $LowRisk$ indicator which takes value one if the quiscore is above 80, and zero otherwise. The final rows of 1 consider the industrial composition of our sample by considering dummy variables for whether the firm belongs to a particular sector. We can see that the largest proportion of firms comes from the manufacturing, services, and retail sector; this mix is broadly in line with the aggregate economy.

3.2 Empirical Strategy

Figure 2 presents binned scattered plots of rates of new relationship formation against firm turnover before and after the reform. The figure suggests that, before the policy is implemented, the probability to create a new lending relationship is the same around the £25 million threshold, but there is a discontinuity following the policy implementation in 2016. This is due to a fall in the new relationship formation rate among the control firms, which could either be explained by the negative correlation between relationship formation and age or a general declining trend in switching behavior; but the relatively smaller drop in the switching probability for treated firms is already suggestive evidence for the treatment effect.

[Figure 2 here]

To estimate the effect of the policy on new relationship formation more formally, we use a linear probability model in a difference-in-differences (DiD) design.

The DiD regression equation takes the following form:

$$NewLender_{i,t} = \beta I\!<(25m)_i \times Post_t + \eta X_{i,t} + \alpha_i + \gamma_{s,t} + \eta_{g,t} + \nu_{r,t} + \varepsilon_{i,t} \qquad (1)$$

The CCDS treatment indicator is $I\!<(25m)_i$, which is equal to one for firms with turnover below $25m$ in 2016, and zero otherwise. To form the DiD variable, $I\!<(25m)_i$ is interacted with $Post_t$, an indicator variable taking value one in years after the scheme went live, and zero before and including 2016 (policy implementation year).

Although this strategy alleviates omitted omitted variable bias concerns, it may not be perfect. Therefore, we also control for firm annual financials in the $X_{i,t}$ vector to address any remaining potential issues. Specifically, we control for log total assets, a low credit risk dummy (i.e. indicator taking value 1 if the firm's QuiScore is below 80, and zero otherwise), cash to total assets and the firms leverage ratio all lagged one period.

Finally, we saturate the model with a rich set of fixed effects. In addition to firm (α_i) fixed effects, we include sector-times-year ($\gamma_{s,t}$), and region-times-year ($\eta_{g,t}$) and relationship stage-times-year ($\nu_{r,t}$) fixed effects, which captures credit demand by firm i at a given point in time. Regions correspond to the 124 UK post code areas. Relationship stages are calculated as the deciles of the relationship duration (in months) an SME has with its lenders up to year t.

We next assess the comparability of treated and control firms in 2016, the year in which the policy was implemented. Table 2 tests for any differences between the sample mean of the variables for the treated group and control group.

[Table 2 here]

In the first three rows, we can see that the both groups had a similar rate of switching to a new lender, both in total and by bank and non-bank lenders. When it comes to financial characteristics, treated firms are bigger in size than control firms, but this is naturally driven by the definition of treatment, which is based on turnover amounts. Control firms (above £25m in turnover) hold slightly more liquid (cash) assets, and have leverage ratios 3 percentage points higher. These differences are statistically significant, but rather small in economic magnitudes. Our baseline specification control for these potentially important differences. Importantly, the two groups are almost identical in terms of profitability, credit risk, age and lending relationship characteristics (number and length).

Finally, the distribution of the two groups in terms of industry slightly differs. The incidence of CCDS may appear falling to a lesser extent on manufacturing, and transportation firms, but it affects more the services and wholesale industries. In equation 1 industrial differences are absorbed by sector (interacted with time) fixed effects.

4 Results

4.1 New Relationships

Table 3 shows the estimation results of the model in equation (1). In the first 5 columns we study the effect the CCDS policy has on $NewLender_{i,t}$, the probability to switch to a lender with which the borrower had no previous relationships.

[Table 3 here]

In column 1, where we control for year fixed effects only, the DiD interaction coefficient is positive and statistically significant. It indicates that the CCDS reform increases the probability to form a new borrowing relationship for a firm by 1.38 percentage points (i.e. 26% compared to the mean). Importantly, the coefficient on the treatment itself is close to zero and not statistically significant, which reflects the absence of pre-trends in the treated group.

When we add firm fixed effects in column 2, the standalone indicator coefficient $I_{<(25m)_i}$ is absorbed but, most importantly, the DiD coefficient remains positive, statistically significant, and the point estimate increases slightly.

In the next two columns we add relationship stage-times-year fixed effects first, and sector-times-year fixed effects together with geographic area-times-year fixed effects to the specifications. In all cases, the estimated coefficient of interest is barely affected. Finally, in column 5, we maintain the most stringent fixed effect structure and include lagged time-varying firm controls: $LogTA_{i,t-1}$, $LowRisk_{i,t-1}$, $CashTA_{i,t-1}$, and $Leverage_{i,t-1}$. Once again, the DiD coefficient is similar to previous specifications, indicating that any differences between the two groups are unlikely to play a confounding role. In terms of economic significance, the point estimate is non-negligible. The information sharing reform increases the probability that an SME switches to a new lender by 29.25% (=.0155/.0530).

In the last two columns, we explore the nature of these target lenders. Column 6 replaces the previous dependent variable with $Bank_{i,t}$, which is a dummy indicator taking value one if the lender that firm i takes a loan from is a bank, and zero otherwise.

We include the same firm controls and fixed effect structure as in the baseline specification (column 5). The DiD coefficient is positive, but not statistically significant. Instead, column 7 replaces $Bank_{i,t}$ with $NonBank_{i,t}$, and we can see that the coefficient becomes statistically significant and large, indicating that the probability to add to a non-bank lender increases by 45% (=.009/.021) because of the reform. This result may point to a processing advantage that non-banks leverage with the new set of data received under the CCDS. Importantly, this implies that non-banks can enter the market and gain market share through the removal of proprietary information in the banking sector.

To obtain the estimate of the dynamic effect of the policy, and test for parallel trends more formally, we plot the interaction coefficient between $I_{<(25m),i}$ and year indicators over time, with the policy approval year (2016) being the omitted category. We do this for the most stringent specification using all fixed-effects and firm controls (column 5 of Table 3). The two panels in Figure 3 present the event-study plots separately for new lenders (panel a) and new non-banks (panel b).

[Figure 3 here]

The point estimates in both panels are positive but not significant in the pre-reform period. This evidence confirms the parallel trend assumption is satisfied. In 2017, the coefficient turns positive and significant, especially for non-banks and it remains positive for the whole duration of the post-reform period. The decrease in the effect in 2019 can be interpreted as the result of the reciprocity rule, that forces new lender that joins the CCDS to also share information on their existing SME lenders: it is then natural to obtain diminishing returns of sharing information over time. Although we cannot study long-term effect of the policy in 2020 due to to other confounding events (such as the Government guaranteed credit programs enacted during Covid), the event-study results show that the effects of the policy are not short-lived.

4.1.1 Robustness

Before studying the main channels and consequences of the new relationship formations, we provide a battery of robustness tests for the baseline result in Table 4.[12]

[Table 4 here]

First of all, the CCDS policy does not specify which year the turnover-eligibility refers to. Although firm turnover is persistent over time and we include firm fixed-effects in the regression, the cutoff may be measured with some error in 2016. Thus, in the first column of Table 4 we assign treatment based on firm turnover in 2017 instead of 2016. We can see that the DiD coefficient remains positive and statistically significant, and is comparable in magnitude to the baseline result.

As shown in the balance test in Table 2, the treated and control group differ slightly from one another in terms of some observable characteristics (liquidity and leverage, other than obvious difference in size). Therefore, in column 2 we estimate the DiD coefficient using a matching procedure. We match each firm in the control group to at most four firms in the treated group based on 2016 values of lagged total assets, leverage, cash-to-asset ratio, and risk indicator, sector and area. As we can see from the estimate, the DiD coefficient is positive and statistically significant.

In the next two columns, we change the length of the turnover window (£10-£40m) used to identify the total firm sample. We shrink the window to £15-£35m or £20-£30m (columns 3-4). Standard errors increase likely due to the reduced sample size and larger measurement error in the cutoff definition. Nonetheless, the point estimates remain very similar to the baseline coefficient, confirming the positive effect of the policy on the probability of adding borrowing relationships.

Finally, the last two columns change the window used to identify existing lending relationships a firm has in previous years. We move from the baseline window (3

[12]These are estimated on the specification in column 5 of Table 3, keeping constant fixed effects and controls. Table A2 in the Appendix presents the same robustness tests for $NonBank_{i,t}$.

years) to 1 year (column 5) or 5 years (column 6). In both cases the DiD coefficient is positive and statistically significant. Also, it is very similar in magnitude with respect to our baseline result. Therefore, it seems that information production is indeed spread in early years of a relationship.

4.2 Channels

To test any heterogenous effect of the policy across firm, we multiply the DiD interaction term with one channel at a time and present the results in table 5.

[Table 5 here]

The first channel is to study quantity of information shared per firm. For the first test we create an indicator dummy variable, $OldCCDS_i$, which takes value 1 for a firm i that at the time of the policy (2016) had a relationship with at least one of the nine designated banks, and zero otherwise. The triple interaction coefficient is positive and statistically significant. Therefore, new lenders can gain more customers when these large banks share information. Interestingly, the DiD interaction coefficient is insignificant, suggesting that non-CCDS corporate clients do not manage to switch, arguably because of the asymmetric information still present with new lenders.

In the next two columns, we study whether the policy exhibits increasing or decreasing returns to scale. Sharing more and more data about potential clients may help new lenders to better integrate their risk models, thereby making more accurate decisions. Specification in column 2 interacts the DiD coefficient with the number of lenders the firm had in 2016. We create three dummy variables: $NoCredit_i$, $SingleRel_i$, and $MultiRel_i$. The results show that sharing data has no effect on firms without credit outstanding, as the coefficient on $I_{<(25m)i} \times Post_t$ is statistically insignificant ($NoCredit_i$ is left out to avoid multi-collinearity). On the contrary, sharing data from one lender, increases the probability to get a loan from a new lender (coefficient is 0.0130). This effect is even more pronounced when the firm has multiple accounts at different lenders, as shown by the positive and statistically significant coefficient on the triple interaction with $MultiRel_i$. Therefore, these results suggest that there are increasing returns to scale in data sharing.

In column 3 instead, we investigate returns to scale of information over a time dimension. We first split $RelLength_i$ into quartiles, and then interact the each quartile dummy with $I_{<(25m)i} \times Post_t$. The DiD coefficient represents the category excluded for multi-collinearity reasons, namely the first quartile of $RelLength_i$. This coefficient is statistically insignificant, suggesting that data sharing when existing credit relationships are short has little effect on forming new relationships. Then, the coefficient on the interaction term between $I < (25m)_i \times Post_t$ and Len_iQtl2 and the coefficient on $I < (25m)_i \times Post_t \times Len_iQtl3$ are positive and statistically significant (0.0300, and 0.0319). Finally, the last interaction with the fourth quartile is insignificant, suggesting that very long relationships do not add anything to the probability to switch. This is intuitive, since the policy mandated the nine designated banks to share only the last few months of data for each customer rather than the entire history.

Finally, in column 4 we explore what role firm financial constraints play in the formation of new relationships. We interact the DiD term with $Cash_i$, the ratio between firm i current assets and its total assets (in 2016), which is a proxy for financial constraint. We can see that firms with more cash do benefit less from the sharing policy (triple interaction coefficient equals -0.0515).Instead, cash-constrained firms now can start leverage on the their past data being shared and can form new relationships more easily.

4.3 Real effects

After determining the effect of the policy on new relationships, we analyze whether and how firm outcomes are affected. Table 6 shows estimates for the CCDS effects on firm annual outcomes using BvD data.

[Table 6 here]

In the first two columns of Table 6, we investigate the effects on the maturity breakdown of firm liabilities. We show that current or short-term liabilities increase for treated firms after the policy, especially when switching to non-banks because of the CCDS scheme. We focus on the triple interaction between $I < (25m)_i$, $Post_t$, and $NonBank_{i,t}$ as it is where the policy has stronger effects (i.e., see table 3). This coefficient in column 1 is positive and statistically significant. However, the triple interaction coefficient under column 2, whose dependent variable is long-term liabilities, is insignificant. These results suggest that new non-banks use new data to gain customers and to provide short-term credit. This could be consistent with the policy opening access to non-traditional (shorter) sources of finance. Another reason may be that the data shared under CCDS covers only the last few months of SME performance. New lenders may be able to learn just a customer's recent creditworthiness, rather than their long-term risk profile.

Column 3 to 6 investigate the asset side of the balance sheet. New credit through data sharing leads to an increase in total assets (column 3), primarily driven by an increase in non-fixed assets (column 5) or cash (column 6). This evidence may suggest that new credit SMEs receive because of the data sharing scheme helps them in building up a precautionary buffer against uncertainty. Finally, in column 7 we explore loan pricing effect of the policy. Reporting quality of interest paid variable is not always good in BvD. For each firm, we address this data issue by replacing missing values with the average interest expenses over its pre- or post-reform period. We find that loans from new non-bank lenders lower interest expenses for the firm. Therefore, removing proprietary information to banks allow firms to get new loans and at better terms, consistent with the existence of hold-up problem in SME lending (Ioannidou and Ongena, 2010).

5 Conclusion

This paper studies how lending relationship formation of Small and Medium Enterprises (SMEs) in UK develop once their current account and past performance data are shared with other lenders. To identify the effect of removing proprietary information we take advantage of a reform implemented in 2016 which mandated nine banks to share their corporate clients' data with other finance providers. Importantly, the policy targeted firms with turnover below £25 million only. By means of a difference-in-differences design, we find that affected firms start building new relationships with new lenders more often after the reform. The results are robust to multiple definitions of treatment, existing relationship lenders, and firm samples. Moreover, we find that non-banks are the types of lenders that gain more customers, arguably because of their advantage in incorporating a large influx of data in their risk modeling.

The effects are stronger the more data is shared. Firms with relationships with multiple lenders, and those with longer relationships, can benefit more from the policy. Also, cash-constrained firms switch lenders more than unconstrained ones, consistent with the existence of the hold-up problem in SME lending pre-reform. Finally, we find that the new credit affected firms receive through the scheme has short-term maturity, lower interest-rates, and contributes to increase cash buffers of the SME.

The results of this paper have important policy implications. Removing proprietary information from banks reduce their market power, as non-banks can enter and gain new customers. Increased lending competition can benefit SMEs that start choosing from a larger variety of finance products, possibly at better rates. Further research is needed to understand the repayment behavior of SMEs when borrowing from non-banks that use more data.

References

Babina, T., Buchak, G., and Gornall, W. (2022). Customer data access and fintech entry: Early evidence from open banking. *Available at SSRN.*

Bahaj, S., Foulis, A., and Pinter, G. (2020). Home values and firm behavior. *American Economic Review*, 110(7):2225–70.

Beaumont, P., Tang, H., and Vansteenberghe, E. (2022). The Role of FinTech in Small Business Lending.

Boot, A. (2000). Relationship lending: What do we know? *Journal of Financial Intermediation*, 9:7–25.

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

Byrne, D. M. et al. (2015). Prices for data storage equipment and the state of it innovation. Technical report, Board of Governors of the Federal Reserve System (US).

Cole, R. A. (1998). The importance of relationships to the availability of credit. *Journal of Banking & Finance*, 22(6-8):959–977.

Cortés, K. R., Demyanyk, Y., Li, L., Loutskina, E., and Strahan, P. E. (2020). Stress tests and small business lending. *Journal of Financial Economics*, 136(1):260–279.

de Roure, C., Pelizzon, L., and Thakor, A. (2021). P2P Lenders versus Banks: Cream Skimming or Bottom Fishing? *The Review of Corporate Finance Studies*, 11(2):213–262.

Di Maggio, M. and Yao, V. (2020). Fintech Borrowers: Lax Screening or Cream-Skimming? *The Review of Financial Studies*, 34(10):4565–4618.

Diamond, D. W. (1984). Financial intermediation and delegated monitoring. *The review of economic studies*, 51(3):393–414.

Eça, A., Ferreira, M. A., Prado, M. P., and Rizzo, A. E. (2021). The real effects of fintech lending on smes: Evidence from loan applications.

Fuster, A., Plosser, M., Schnabl, P., and Vickery, J. (2019). The Role of Technology in Mortgage Lending. *The Review of Financial Studies*, 32(5):1854–1899.

Ghosh, P., Vallee, B., and Zeng, Y. (2021). Fintech lending and cashless payments. In *Proceedings of Paris December 2021 Finance Meeting EUROFIDAI-ESSEC.*

Giannetti, M., Liberti, J. M., and Sturgess, J. (2017). Information sharing and rating manipulation. *The Review of Financial Studies*, 30(9):3269–3304.

Goldstein, I., Huang, C., and Yang, L. (2022). Open banking with depositor monitoring. Technical report, Working paper.

Gopal, M. and Schnabl, P. (2022). The Rise of Finance Companies and FinTech Lenders in Small Business Lending. *The Review of Financial Studies*. hhac034.

He, Z., Huang, J., and Zhou, J. (2020). Open banking: credit market competition when borrowers own the data. Technical report, National Bureau of Economic Research.

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

Kalemli-Ozcan, S., Sorensen, B., Villegas-Sanchez, C., Volosovych, V., and Yesiltas, S. (2015). How to construct nationally representative firm level data from the orbis global database: New facts and aggregate implications. Technical report, National Bureau of Economic Research.

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

Ongena, S. and Smith, D. C. (2001). The duration of bank relationships. *Journal of financial economics*, 61(3):449–475.

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

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

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

Sharpe, S. A. (1990). Asymmetric information, bank lending, and implicit contracts: A stylized model of customer relationships. *The journal of finance*, 45(4):1069–1087.

Tang, H. (2019). Peer-to-Peer Lenders Versus Banks: Substitutes or Complements? *The Review of Financial Studies*, 32(5):1900–1938.

Figures

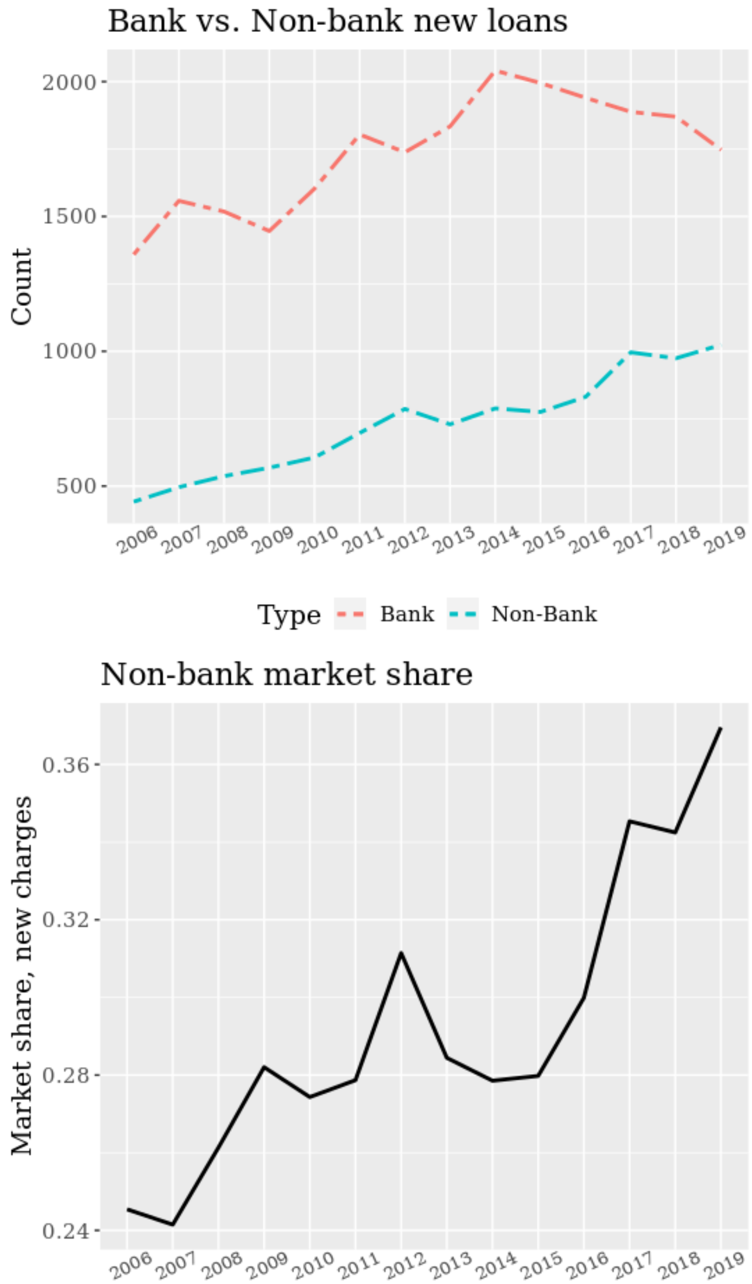
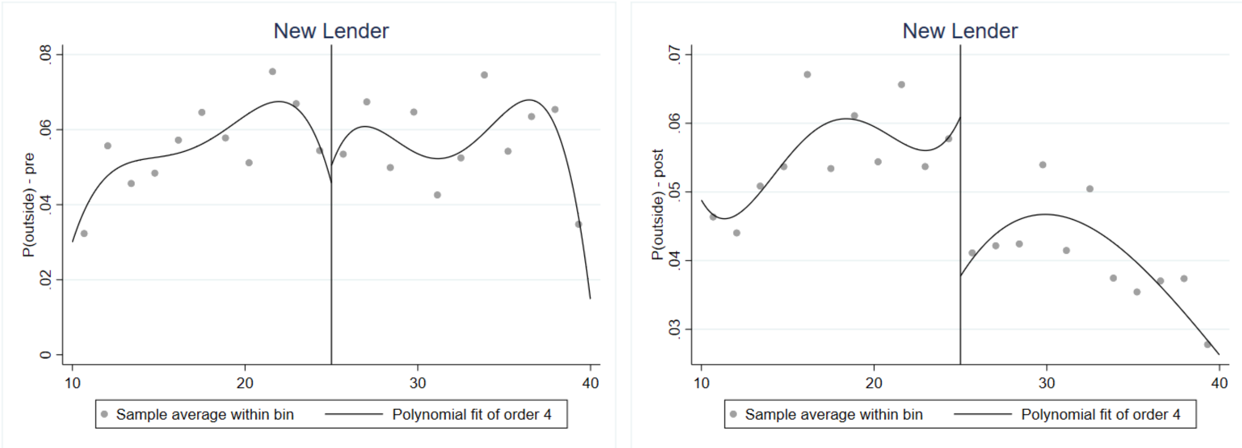


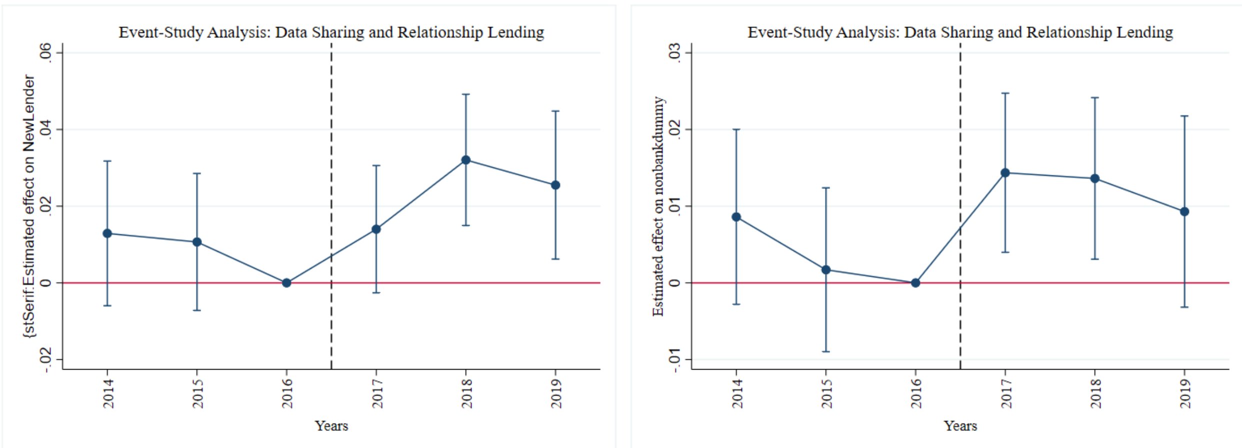
Figure 1: Bank vs. Non-bank, number of new secured loans in the UK corporate (£10m-£40m in turnover) sector. Authors' calculation. Data source: Companies House



(a) Pre-reform (before 2016)

(b) Post-reform (after 2016)

Figure 2: Threshold Discontinuity: these graphs plot a polynomial of order 4 on the average (by turnover-bins) of $NewLender_{i,t}$ variable against firm turnover in £ million (running variable), and split the samples across time periods: before the CCDS policy (2014-2016), and after the policy (2017-2019).



(a) NewLender

(b) Non-Bank

Figure 3: DiD estimates (β_t) of the probability to add a new lender according to the following equation: $NewLender_{i,t} = \beta_t I_{<(25m)_i} \times Year_t + \eta X_{i,t} + \mu_i + \mu_{f,rt,st,gt} + \varepsilon_{i,t}$. Panel (a) reports the estimates for new loans from both banks and non-banks and panel (b) restricts the dependent variable to be equal to one for loans from non-banks.

Tables

Table 1: Summary statistics

	Observ.	Mean	Std.Dev.	5^{th}	95^{th}
$\text{NewLender}_{i,t}$	39,089	.053	.224	0	1
$\text{NewBank}_{i,t}$	39,089	.035	.184	0	0
$\text{NewNonBank}_{i,t}$	39,089	.021	.144	0	0
NoCredit_{i}	39,089	.214	.410	0	1
SingleRelat_{i}	39,089	.437	.496	0	1
MultiRelat_{i}	39,089	.349	.477	0	1
OldCCDS_{i}	39,089	.572	.495	0	1
RelLength_{i}	39,089	1.758	1.135	0	3.376
Turnover16_{i}	39,089	19.50	7.34	10.79	34.42
FirmAge_{i}	39,089	24.33	20.75	4	67
$\text{I}< (25m)_{i}$	39,089	.771	.420	0	1
$\text{LogTA}_{i,t-1}$	39,089	9.259	.817	8.174	10.725
$\text{cashTA}_{i,t-1}$	39,089	.685	.258	.172	.992
$\text{Leverage}_{i,t-1}$	39,089	.587	.272	.164	.989
$\text{Quiscore}_{i,t-1}$	39,089	90.322	13.461	59	99
$\text{Lowrisk}_{i,t-1}$	39,089	.915	.279	0	1
Manufact_{i}	39,089	.198	.399	0	1
Services_{i}	39,089	.305	.461	0	1
RealEstate_{i}	39,089	.141	.348	0	1
Retail_{i}	39,089	.049	.216	0	1
Wholesale_{i}	39,089	.205	.404	0	1
Transport_{i}	39,089	.053	.225	0	1
OtherSec_{i}	39,089	.047	.212	0	1

Note: This table shows summary statistics (mean, standard deviation, main percentiles) of the variables used.

Table 2: Balance Test

	Control			Treated			
	N	Mean	SD	N	Mean	SD	Difference
NewLender_i	1,656	.06	.24	5,230	.05	.23	.008
NewBank_i	1,656	.04	.20	5,230	.04	.19	.005
NewNonBank_i	1,656	.02	.14	5,230	.02	.14	.001
Turnover_i	1,656	30.80	4.20	5,230	16.29	4.02	14.515***
logTA_i	1,656	9.70	.81	5,230	9.16	.77	.547***
cash2TA_i	1,656	.70	.26	5,230	.68	.26	.017**
Leverage_i	1,656	.61	.27	5,230	.58	.27	.028***
LowRisk_i	1,652	.92	.27	5,219	.92	.26	-.002
Profit2TA_i	1,655	.10	.12	5,226	.10	.12	-.001
FirmAge_i	1,656	23.37	21.15	5,230	23.11	20.30	.254
NoCredit_i	1,656	.22	.42	5,230	.21	.41	.014
SingleRelat_i	1,656	.43	.50	5,230	.44	.50	-.010
AvgLengthRel_i	1,656	1.71	1.13	5,230	1.76	1.13	-.051
Manufacturing_i	1,656	.17	.38	5,230	.20	.40	-.032***
Services_i	1,656	.33	.47	5,230	.31	.46	.026**
RealEstate_i	1,656	.15	.35	5,230	.14	.35	.006
Retail	1,656	.05	.22	5,230	.05	.22	-.00
Wholesale_i	1,656	.22	.41	5,230	.20	.40	.019*
Transport_i	1,656	.04	.20	5,230	.05	.23	-.010*
OtherSec_i	1,656	.04	.19	5,230	.05	.21	-.008

Note: This table presents the results of a balance test exercise, which compares the average value of the main variables used in the analysis across treated and control groups, taking 2016 as the reference period.

Table 3: Information Sharing and New Lenders

	$Newlender_{i,t}$					$Bank_{i,t}$	$NonBank_{i,t}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$I < (25m)_i \times Post_t$	.0138***	.0157***	.0154***	.0158***	.0155***	.0062	.0098***
	(2.67)	(3.01)	(2.96)	(3.02)	(2.94)	(1.41)	(2.94)
$I < (25m)_i$	-.0022						
	(-.54)						
$\text{LogTA}_{i,t-1}$					-.0037	-.0062	.0033
					(-.56)	(-1.17)	(.72)
$\text{LowRisk}_{i,t-1}$					.0033	-.0010	.0013
					(.54)	(-.20)	(.32)
$\text{CashTA}_{i,t-1}$					.0377*	.0296*	.0093
					(1.79)	(1.80)	(.65)
$\text{Leverage}_{i,t-1}$					-.0239	-.0233*	-.0055
					(-1.59)	(-1.91)	(-.52)
Year FE	✓	✓	X	X	X	X	X
Firm FE	X	✓	✓	✓	✓	✓	✓
Rel × Year FE	X	X	✓	✓	✓	✓	✓
Sec × Year FE	X	X	X	✓	✓	✓	✓
Reg × Year FE	X	X	X	✓	✓	✓	✓
# of Observations	39,089	39,089	39,089	39,089	39,089	39,089	39,089
R^2	0.00	.253	.261	.276	.277	.242	.288
adj. R^2	.00	.059	.067	.064	.064	.020	.079

Note: This table presents OLS coefficient estimates of regression equation 1. The dependent variable $NewLender_{i,t}$ takes value 1 if firm i takes a loan in year t from a "new" bank, that is one that had no prior relationship with firm i in the previous three years, and zero otherwise. The dependent variable in the last two columns are $Bank_{i,t}$ or $NonBank_{i,t}$, which are dummy variables taking value 1 if the new lender is a bank (column 6), or a non-bank (column 7), and zero otherwise. The dummy variable $I_{<(25m)i}$ takes value 1 if firm i has turnover below £25m in 2016 and, therefore, is affected by the policy, and zero otherwise. $Post_t$ is a time dummy taking value 1 after CCDS scheme introduction (2017 onwards), and zero otherwise. Lagged control firm variables are: $ln(TA_{i,t-1})$, $lowRisk_{i,t-1})$, $cash2TA_{i,t-1}$ and $Leverage_{i,t-1}$. Set of fixed effects includes firm, and time (year) interacted with relationship stage (decile indicators of firm-lender(s) relationship duration), sector s, and geographical area g fixed effects. Standard errors clustered at firm-level. *t-statistics* are in parentheses. *, **, *** represent p-values below 0.05, 0.01 and 0.001, respectively.

Table 4: Robustness Tests

Dep. var.: $NewLender_{i,t}$	Treated$_{17}$ (1)	Matched (2)	Turnover Window £15-35m (3)	£20-30m (4)	Relationship Window -1year (5)	-5years (6)
$I<(25m)_i \times Post_t$	.0201***	.0140**	.0131**	.0128	.0138***	.0152***
	(4.14)	(2.33)	(2.07)	(1.37)	(2.66)	(2.88)
Firm FE	✓	✓	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓	✓	✓
Rel × Year FE	✓	✓	✓	✓	✓	✓
Sec × Year FE	✓	✓	✓	✓	✓	✓
Reg × Year FE	✓	✓	✓	✓	✓	✓
# of Observations	39,089	22,644	23,587	11,064	39,089	39,089
R^2	.277	.280	.280	.318	.279	.276
adj. R^2	.064	.068	.068	.077	.067	.063

Note: This table presents OLS coefficient estimates of regression equation 1. The dependent variable $NewLender_{i,t}$ takes value 1 if firm i takes a loan in year t from a bank with no prior relationship in the previous three years, and zero otherwise. The dummy variable $I_{<(25m)i}$ takes value 1 if firm i has turnover below £25m and, therefore, is targeted by the policy, and zero if above. $Post_t$ is a time dummy taking value 1 after scheme introduction, and zero otherwise. Column 1 assigns treatment based on 2017 turnover. Column 2 excludes treated firms that under a matching exercises are not matched with a control firm. Columns 3 and 4 change the firm sample based on a turnover window (15m-35 in column 3 and 20-30 in column 4). Finally, columns 5 and 6 shrink and enlarge, respectively, the time window to identify existing relationship lenders. Lagged control firm variables are $ln(TA_{i,t-1})$, $lowRisk_{i,t-1})$, $cash2TA_{i,t}$ and $Leverage_{i,t}$. Set of fixed effects includes firm, and time (year) interacted with relationship stage, sector s, and geographical area g fixed effects. Standard errors clustered at firm-level. *t-statistics* are in parentheses. *, **, *** represent p-values below 0.05, 0.01 and 0.001, respectively.

Table 5: Channels

Dep. var.: $NewLender_{i,t}$	(1)	(2)	(3)	(4)
$I<(25m)_i \times Post_t$	.0004	.0036	-.0024	.0512***
	(.09)	(.58)	(-.32)	(3.01)
$I<(25m)_i \times Post_t \times OldCCDS_i$	.0200**			
	(2.01)			
$I<(25m)_i \times Post_t \times SingleRel_i$		.0130*		
		(1.67)		
$I<(25m)_i \times Post_t \times MultiRel_i$		.0280**		
		(2.28)		
$I<(25m)_i \times Post_t \times Len_iQtl2$			.0300**	
			(2.04)	
$I<(25m)_i \times Post_t \times Len_iQtl3$			.0319**	
			(2.41)	
$I<(25m)_i \times Post_t \times Len_iQtl4$			.0077	
			(.65)	
$I<(25m)_i \times Post_t \times Cash_i$				-.0515**
				(2.34)
Firm FE	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓
Rel × Year FE	✓	✓	✓	✓
Sec × Year FE	✓	✓	✓	✓
Reg × Year FE	✓	✓	✓	✓
# of Observations	39,089	39,089	39,089	39,089
R^2	.277	.283	.284	.277
adj. R^2	.064	.072	.064	.064

Note: This table presents OLS coefficient estimates of regression equation 1. The dependent variable $NewLender_{i,t}$ takes value 1 if firm i in sector s and in geographical area g takes a loan in year t from a bank with no prior relationship in the previous three years, and zero otherwise. The dummy variable $I_{<(25m)i}$ takes value 1 if firm i has turnover below £25m and, therefore, is targeted by the policy, and zero if above. $Post_t$ is a time dummy taking value 1 after scheme introduction, and zero otherwise. Column 1 interacts the DiD coefficient with $OldCCDS_i$, a dummy variable taking value 1 in case at least one of the relationship lenders of firm i was one of the nine designated banks by the policy, and zero otherwise. Column 2 includes the interaction with $SingleRel_i$ and $MultiRel_i$, indicators equal to 1 in case the firm up to 2016 had relationships with one or more lenders, respectively. Column 3 interacts $I_{<(25m)i} \times Post_t$ with quartile dummy variables for the relationship length the firm had in 2016 with its lenders. Column 4 includes the triple interaction between$I_{<(25m)i}$, $Post_t$, and $Cash_i$ that is the ratio between current assets and total assets of firm i in 2016. Lagged control firm variables are $ln(TA_{i,t-1})$, $lowRisk_{i,t-1})$, $cash2TA_{i,t}$ and $Leverage_{i,t}$. Set of fixed effects includes firm, and time (year) interacted with relationship stage, sector s, and geographical area g fixed effects. Standard errors clustered at firm-level. *t-statistics* are in parentheses. *, **, *** represent p-values below 0.05, 0.01 and 0.001, respectively.

Table 6: Real Effects

Dep. var. (log):	STLiab (1)	LTLiab (2)	TotAssets (3)	FixAsst (4)	NFixAsst (5)	Cash (6)	IntPaid (7)
$I{<}\ (25m)_i \times Post_t \times NonBank_{i,t}$	.119*	.173	.154***	.137	.0973*	.129**	-.0118*
	(1.76)	(0.62)	(3.18)	(1.03)	(1.84)	(2.27)	(-1.68)
$I{<}\ (25m)_i \times Post_t \times Bank_{i,t}$	.0878*	.301	.0195	-.0607	.0251	.0222	.0044
	(1.77)	(1.43)	(.63)	(-.74)	(.67)	(.59)	(1.25)
$I < (25m)_i \times Post_t$	.0251**	.0385	.0211***	-.0135	.0253***	.0268***	.0005
	(2.48)	(.83)	(3.84)	(-.74)	(3.39)	(3.59)	(1.04)
Firm FE	✓	✓	✓	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓	✓	✓	✓
Pairwise interactions	✓	✓	✓	✓	✓	✓	✓
Rel × Year FE	✓	✓	✓	✓	✓	✓	✓
Sec × Year FE	✓	✓	✓	✓	✓	✓	✓
Reg × Year FE	✓	✓	✓	✓	✓	✓	✓
# of Obs.	35,101	34,710	35,115	34,611	34,645	35,112	30,530
R^2	.876	.872	.966	.959	.927	.927	.645
adj. R^2	.837	.831	.955	.947	.904	.905	.528

Note: This table presents OLS coefficient estimates of regression equation 1, replacing the dependent variables with firm-level outcomes. The dummy variables $Bank_{i,t}$ and $NonBank_{i,t}$ take value 1 if the new lender is a bank or a non-bank, respectively. The dummy variable $I_{<(25m)i}$ takes value 1 if firm i has turnover below £25m and, therefore, is targeted by the policy, and zero if above. $Post_t$ is a time dummy taking value 1 after scheme introduction, and zero otherwise. The dependent variables across columns are the logarithm of: current liabilities (1), long-term liabilities (2), total assets (3), fixed assets (4), non-fixed assets (5), and current assets (6). Interest paid in column 7 is the ratio of interest expenses to total assets. Lagged control firm variables are $ln(TA_{i,t-1})$, $lowRisk_{i,t-1})$, $cash2TA_{i,t}$ and $Leverage_{i,t}$. Set of fixed effects includes firm, and time (year) interacted with relationship stage, sector s, and geographical area g fixed effects. Standard errors clustered at firm-level. *t-statistics* are in parentheses. *, **, *** represent p-values below 0.05, 0.01 and 0.001, respectively.

Online Appendix

Table A1: Collapsed version

Dep. var.:	$NewLender_{i,t}$ (1)	(2)	$Bank_{i,t}$ (3)	$NonBank_{i,t}$ (4)
$I{<}\,(25m)_i \times Post_t$	.0501***	.0527**	-.0035	.0132***
	(4.23)	(4.43)	(-.62)	(3.00)
LogTA$_{i,t-1}$		.0563***	.0110	.0301**
		(3.06)	(.76)	(2.34)
LowRisk$_{i,t-1}$		-.0006	-.0010	-.0009
		(-.03)	(-.06)	(-.07)
CashTA$_{i,t-1}$		.0208	.0530	-.0515
		(.34)	(1.00)	(-1.20)
Leverage$_{i,t-1}$		.0446	.0307	.0226
		(1.01)	(.84)	(.74)
Firm FE	✓	✓	✓	✓
TimeDummy	✓	✓	✓	✓
# of Observations	14,374	14,374	14,374	14,374
R^2	.555	.556	.521	.567
adj. R^2	.111	.112	.042	.134

Note: This table presents OLS coefficient estimates of regression equation 1, with the sample period collapsed into pre- and post-reform. Therefore, for each firm, all variables are averaged in the pre-policy, and within the post-policy period. The dependent variable $NewLender_{i,t}$ takes value 1 if firm i takes a loan in year t from a bank with no prior relationship in the previous three years, and zero otherwise. The dummy variable $I_{<(25m)i}$ takes value 1 if firm i has turnover below £25m and, therefore, is targeted by the policy, and zero if above. $Post_t$ is a time dummy taking value 1 after scheme introduction, and zero otherwise. Lagged control firm variables are $ln(TA_{i,t-1})$, $lowRisk_{i,t-1})$, $cash2TA_{i,t}$ and $Leverage_{i,t}$. Set of fixed effects includes firm, and time (year) interacted with relationship stage, sector s, and geographical area g fixed effects. Standard errors clustered at firm-level. *t-statistics* are in parentheses. *, **, *** represent p-values below 0.05, 0.01 and 0.001, respectively.

Table A2: Robustness Tests - NonBank loans

			Turnover Window		Relationship Window	
	Treated$_{17}$	Matched	£15-35m	£20-30m	-1year	-5years
Dep. var.: $NonBank_{i,t}$	(1)	(2)	(3)	(4)	(5)	(6)
$I < (25m)_i \times Post_t$	.0113***	.0066*	.0099**	.0068	.0094***	.0101***
	(3.73)	(1.70)	(2.51)	(1.21)	(2.84)	(3.04)
Firm FE	✓	✓	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓	✓	✓
Rel × Year FE	✓	✓	✓	✓	✓	✓
Sec × Year FE	✓	✓	✓	✓	✓	✓
Reg × Year FE	✓	✓	✓	✓	✓	✓
# of Observations	39,089	22,638	23,587	11,064	39,089	39,089
R^2	.288	.296	.289	.321	.291	.286
adj. R^2	.079	.089	.079	.081	.082	.076

Note: This table presents OLS coefficient estimates of regression equation 1. The dependent variable $NonBank_{i,t}$ takes value 1 if firm i takes a loan in year t from a mon-bank with no prior relationship in the previous three years, and zero otherwise. The dummy variable $I_{<(25m)i}$ takes value 1 if firm i has turnover below £25m and, therefore, is targeted by the policy, and zero if above. $Post_t$ is a time dummy taking value 1 after scheme introduction, and zero otherwise. Column 1 assigns treatment based on 2017 turnover. Column 2 excludes treated firms that under a matching exercises are not matched with a control firm. Columns 3 and 4 change the firm sample based on a turnover window (15m-35 in column 3 and 20-30 in column 4). Finally, columns 5 and 6 shrink and enlarge, respectively, the time window to identify existing relationship lenders. Lagged control firm variables are $ln(TA_{i,t-1})$, $lowRisk_{i,t-1})$, $cash2TA_{i,t}$ and $Leverage_{i,t}$. Set of fixed effects includes firm, and time (year) interacted with relationship stage, sector s, and geographical area g fixed effects. Standard errors clustered at firm-level. *t-statistics* are in parentheses. *, **, *** represent p-values below 0.05, 0.01 and 0.001, respectively.