DISCUSSION PAPER SERIES

DP15829
(v. 2)

Move a Little Closer? Information Sharing and the Spatial Clustering of Bank Branches

Shusen Qi, Ralph De Haas, Steven Ongena and Stefan Straetmans

FINANCIAL ECONOMICS

ISSN 0265-8003

Move a Little Closer? Information Sharing and the Spatial Clustering of Bank Branches

Shusen Qi, Ralph De Haas, Steven Ongena and Stefan Straetmans

Discussion Paper DP15829
First Published 19 February 2021
This Revision 19 July 2022

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programmes:

- Financial Economics

Move a Little Closer? Information Sharing and the Spatial Clustering of Bank Branches

Abstract

We study how information sharing between banks influences the geographical clustering of branches. A spatial model of credit market competition first explains how information sharing impacts loan pricing and equilibrium branch clustering. With data on 56,555 branches of 614 banks in 19 countries between 1995 and 2012, we test key model hypotheses. We find that information sharing increases branch clustering as banks open branches in localities that are new to them but that are already served by other banks. This branch clustering is associated with less spatial credit rationing as information sharing allows firms to borrow from more distant banks.

JEL Classification: D43, G21, G28, L13, R51

Keywords: Branch clustering, information sharing, spatial oligopoly model

Shusen Qi - shusenqi@xmu.edu.cn
Xiamen University

Ralph De Haas - dehaasr@ebrd.com
EBRD, Tilburg University and CEPR

Steven Ongena - steven.ongena@bf.uzh.ch
University of Zurich and CEPR

Stefan Straetmans - s.straetmans@maastrichtuniversity.nl
Maastricht University

Acknowledgements
The authors thank Martin Brown, Stephanie Chan, Francesca Cornelli, Hans Degryse, Thomas Mathä, Marco Pagano, Francesc Rodriguez Tous, Vahid Saadi, Koen Schoors, Nathaniel Young, Yunqi Zhang, Kuncheng Zheng, participants at the Bocconi University–Sapienza University Conference on "New Frontiers in Banking: from Corporate Governance to Risk Management" (Milan), the American Economic Association ASSA Annual Meeting (Philadelphia), the 11th Swiss Winter Conference on Financial Intermediation (Lenzerheide), the Great China Area Finance Conference (Xiamen), the 7th MoFiR Workshop on Banking (Ancona), the China International Conference in Finance (Hangzhou), the Deutsche Bundesbank–IWH–CEPR Conference on the Future of Financial Intermediation (Eltville am Rhein), the 2nd EBC Network Workshop (Luxembourg), the 32nd European Economic Association Annual Meeting (Lisbon), and seminar participants at the University of Antwerp, Xiamen University, Shandong University, and Huazhong University of Science and Technology (Wuhan) for useful comments. The views expressed are those of the authors and not necessarily those of the EBRD. Shusen Qi acknowledges financial support from the National Natural Science Foundation of China (71903164, 71790601) and the Social Science Foundation of Fujian Province (FJ2019B140). Steven Ongena acknowledges financial support from ERC ADG 2016 - GA 740272 lending.

Move a Little Closer? Information Sharing and the Spatial Clustering of Bank Branches

Shusen Qi
Xiamen University

Ralph De Haas
EBRD, KU Leuven, and CEPR

Steven Ongena
University of Zurich, Swiss Finance Institute, KU Leuven,
NTNU Business School, and CEPR

Stefan Straetmans
Maastricht University and University of Antwerp

July 13, 2022

The authors thank Martin Brown, Stephanie Chan, Francesca Cornelli, Hans Degryse, Thomas Mathä, Marco Pagano, Francesc Rodriguez Tous, Vahid Saadi, Koen Schoors, Nathaniel Young, Yunqi Zhang, Kuncheng Zheng, participants at the Bocconi University–Sapienza University Conference on "New Frontiers in Banking: from Corporate Governance to Risk Management" (Milan), the American Economic Association ASSA Annual Meeting (Philadelphia), the 11th Swiss Winter Conference on Financial Intermediation (Lenzerheide), the Great China Area Finance Conference (Xiamen), the 7th MoFiR Workshop on Banking (Ancona), the China International Conference in Finance (Hangzhou), the Deutsche Bundesbank–IWH–CEPR Conference on the Future of Financial Intermediation (Eltville am Rhein), the 2nd EBC Network Workshop (Luxembourg), the 32nd European Economic Association Annual Meeting (Lisbon), and seminar participants at the University of Antwerp, Xiamen University, Shandong University, and Huazhong University of Science and Technology (Wuhan) for useful comments. The views expressed are those of the authors and not necessarily those of the EBRD. Shusen Qi acknowledges financial support from the *National Natural Science Foundation of China (71903164, 71790601)* and the Social Science Foundation of Fujian Province (FJ2019B140). Steven Ongena acknowledges financial support from *ERC ADG 2016 - GA 740272 lending.*

Move a Little Closer? Information Sharing and the Spatial Clustering of Bank Branches

Abstract

We study how information sharing between banks influences the geographical clustering of branches. A spatial model of credit market competition first explains how information sharing impacts loan pricing and equilibrium branch clustering. With data on 56,555 branches of 614 banks in 19 countries between 1995 and 2012, we test key model hypotheses. We find that information sharing increases branch clustering as banks open branches in localities that are new to them but that are already served by other banks. This branch clustering is associated with less spatial credit rationing as information sharing allows firms to borrow from more distant banks. (100 words)

Key words: Branch clustering, information sharing, spatial bank competition

JEL codes: D43, G21, G28, L13, R51

1. Introduction

Over the past three decades, banks across the world have adjusted their branch networks in response to regulatory changes, increased competition, and progress in information and communication technology. Many banks have also pruned their branch networks in the aftermath of the Great Recession. Importantly, these dynamics did not play out in a spatially uniform way within countries. Banks increasingly cluster their branches together while leaving other areas underserved. The resulting emergence of banking deserts – localities almost entirely devoid of bank branches (Morgan, Pinkovsky and Yang, 2016) – has raised concerns about adverse local effects on firms' funding costs (Bonfim, Noguiera and Ongena, 2021) and concomitant declines in small-business lending and employment growth (Nguyen, 2019).

Against this background, we study theoretically and empirically how the introduction of information sharing among banks shapes the spatial footprint of their branch networks. This remains a pertinent question because even in an increasingly digitalized world, stark geographical differences remain in firms' ability to borrow from banks (Guiso, Sapienza and Zingales, 2004; Lee and Luca, 2019; Granja, Leuz and Rajan, 2021). This ability to access bank credit also continues to depend on the physical presence of bank branches (Pollard, 2003; Zhao and Jones-Evans; Beck et al, 2018). Policy-induced changes in the geographical location of these branches, such as in response to the introduction of information sharing, may therefore influence firms' access to credit and have real-economic consequences. More pithily: financial agglomeration can shape real agglomeration.

A first contribution of this paper is to develop a spatial model, in the vein of Hauswald and Marquez (2003, 2006), to clarify how information sharing between banks can influence branch clustering. While the model is stylized, its main building blocks are representative and it yields testable hypotheses. In the model, information asymmetries increase with distance and geographical credit rationing makes it difficult for entrepreneurs to successfully apply for a loan at a branch that is further away (Petersen and Rajan, 2002; Hauswald and Marquez, 2006). The main intuition is that such distance constraints bind less when banks can credibly share information about loan applicants. More specifically, the introduction of information sharing has a two-sided effect on incumbent bank branches. On one hand, these branches can now expand their local market as they can screen firms that are more distant. We call this the market-size effect. On the other hand, information sharing hollows out the information advantage of incumbent bank branches. We call this the competition effect. In contrast, for entrant banks the impact of information sharing is one-sided. It diminishes the information advantage of incumbent branches so that banks can now successfully enter local credit markets that were previously off limits to them. This holds especially in localities with more incumbent branches, which have already built up extensive information about local borrowers.

Using this basic theoretical framework, we derive predictions about the impact on branch clustering of a formal mechanism (a public credit registry or a private credit bureau) through which banks share hard (that is, codified and transferable) borrower information. Such data can include 'negative' information about prior defaults and late payments of loan applicants as well as 'positive' information about whether they have outstanding debt elsewhere. We derive four testable hypotheses.

First, our model predicts that information sharing increases the likelihood that banks open new branches in localities with more pre-existing branches of other banks. In such localities, entrant branches can benefit the most from information collected by incumbent branches. This increases the spatial clustering of bank branches. Second, adding more branches of the same bank in a locality where a bank is already present does not facilitate easier and better information access. We therefore expect that after the introduction of information sharing banks are more likely to open new branches in localities with fewer pre-existing own branches. Third, our model predicts that information sharing especially influences the clustering behavior of relationship banks. To obtain proprietary information, relationship banks need to interact frequently and intensely with borrowers. In our model, this translates into a negative distance impact on the effectiveness of screening that is stronger for relationship banks than for transactional banks. Fourth, we expect that the impact of information sharing on bank clustering is more pronounced in countries with higher quality information sharing systems.

To test these predictions, we use detailed bank branch data – geographical coordinates and the dates of establishment (and/or closure) of each branch – from 19 Eastern European countries. Our sample covers 56,555 branches from 614 banks that were active during the period 1995-2012 across 8,536 villages, towns or cities (henceforth indicated as 'localities').[1] The data set further contains information on the ownership of these branches so that we can distinguish between branches of different types of banks.

Eastern Europe constitutes a natural testing ground for our model because information asymmetries are pervasive while creditor rights remain relatively weak (Brown, Jappelli and Pagano, 2009). Importantly, during our sample period many eastern European countries

[1] According to mainstay English Dictionary definitions, a village is a small community in a rural area, a town is a populated area with fixed boundaries and a local government, and a city is a large or important town. These definitions also apply in our sample.

institutionalized information sharing among lenders – either through a public credit registry or through a private credit bureau. We use the introduction of these information-sharing regimes as country-level shocks that push banks towards a new clustering equilibrium. This setting also provides insights into how bank clustering may respond to similar but slower improvements in borrower transparency in more developed banking markets.

In terms of methodology, we implement a difference-in-differences-in-differences framework with the treatment (presence of information sharing) varying across countries and years. We follow Cengiz, Dube, Lindner, and Zipperer (2019) to deal with the staggered treatment timing where different countries introduced information sharing at different points in time. We then compare how, within the same country, the introduction of information sharing differentially affects branch openings across localities with different numbers of pre-existing bank branches. This strategy enables us to mitigate selection bias and, by including granular fixed effects, alleviates concerns about omitted variables. In particular, we saturate our specifications with *locality*year*treatment event* fixed effects; *bank*year*treatment event* fixed effects; and *locality*bank*treatment event* fixed effects. This removes the possibility that dynamics in branch clustering merely reflect changes in local firms' demand for credit or by other potential confounds, such as the depth of local labor markets.

By way of preview, we find that information sharing has a strong positive effect on bank branch clustering. Banks are more likely to open new branches in localities where they did not yet operate but where other banks were already present. This clustering is more pronounced for relationship banks and in countries where information sharing is more effective. Moreover, due to information sharing banks become more likely to open a branch in a new locality when they simultaneously close a branch elsewhere in the country. This holds true especially when banks close branches in relatively underbanked localities. Auxiliary data analysis confirms that around the introduction of information sharing, banks increasingly start to cluster their branches within regions as defined at various levels in the *Nomenclature des Unités Territoriales Statistiques* (Nomenclature of Territorial Units for Statistics or NUTS). Lastly, data on bank-firm relationships show that, in line with a reduction in geographical credit rationing, information sharing allows firms to borrow from banks that are more distant. In sum, our results show how information sharing makes it more important for banks to move closer to each other than to their borrowers.

This paper contributes to three strands of the literature. First, we provide new insights into the causes and consequences of spatial agglomeration in the supply of financial services. In line with Robinson's (1952) adage that "where enterprise leads finance follows", banks tend to locate branches in areas with fast population and economic growth, and where the demand for their services is high (Carbo-Valverde et al., 2011; Deller and Sundaram-Stukel, 2012; Brown, Guin and Kirschenmann, 2015).[2] At the same time, an influential 'finance-and-growth' literature has convincingly shown that the supply of bank credit not just passively follows economic activity but also has a causal impact on subsequent growth (see Levine (2005) for a comprehensive overview). Relatedly, a recent literature exploits plausibly *exogenous* spatial variation in bank branches – reflecting historical 'quirks' or waves of financial deregulation. This work shows that the local presence of bank branches remains a first-order determinant of local economic dynamism and continues to have strong effects on socioeconomic outcomes including employment, inequality and firm innovation.[3] Our contribution is to develop a simple and intuitive framework to study how information sharing among banks reshapes the spatial footprint of branch networks.[4] We then test our model predictions in a rich international context, using the introduction of information sharing as country-level shocks that push banks towards a new clustering equilibrium (while carefully controlling for initial real agglomeration).

Second, we add to the literature on the economic impact of information sharing. Theoretical contributions explore how information sharing reduces moral hazard and adverse selection, improves loan quality, and lowers interest rates (Padilla and Pagano, 1997, 2000), and shapes inter-bank competition (Bouckaert and Degryse, 2004, 2006; Bennardo, Pagano, and Piccolo, 2015). On the empirical side, cross-country evidence indicates that information sharing is associated with more private-sector lending, fewer defaults, and lower interest rates (Jappelli and Pagano, 2002). Evidence suggests that (voluntary) private credit bureaus are more effective than (mandatory) public registries in this regard (Martinez-Peria and Singh, 2014). Yet, it remains unclear exactly *how* information sharing affects bank behavior. We uncover an

[2] Branches may also cluster due to herding behavior. Reputational concerns can lead bank managers to open a new branch in a neighborhood with many pre-existing branches rather than in uncharted territory, even if this hurts the profitability of the new branch (Chang, Chaudhuri and Jayaratne, 1997).
[3] See Jayaratne and Strahan (1996), Beck, Levine and Levkov (2010), Rice and Strahan (2010), Kroszner and Strahan (2014), Favara and Imbs (2015), Célérier and Matray (2019) for the US; Guiso, Sapienza and Zingales (2004), Herrera and Minetti (2007) and Benfratello, Schiantarelli and Sembenelli (2008) for Italy; and Berkowitz, Hoekstra and Schoors (2014) and Bircan and De Haas (2020) for Russia.
[4] A small strand of work studies the size of banks' branch networks rather than their geographical distribution. Cerasi, Chizzolini and Ivaldi (2002) and Cohen and Mazzeo (2010) investigate the impact of competition on the size of branch networks. Temesvary (2015) shows theoretically and empirically that locational market power allows banks with larger branch networks to charge an interest-rate premium, while Coccorese (2012) incorporates branch decisions in a price competition model. Lastly, Levine, Lin and Wang (2020) show that banks with a higher degree of geographic branch overlap are more likely to merge.

important mechanism: the central availability of hard borrower information leads to a different branch-clustering equilibrium that is associated with less spatial credit rationing.[5]

Third, our findings add to the industrial organization literature on firm location. This literature asserts that customers trade off the utility they derive from products and the geographic distance to the firms where they can buy these products. As a result, firms have greater market power when they are closer to their customers. This literature starts with the Hotelling (1929) model where firms compete and price their products in geographic locations along a line of fixed length. Salop (1979) introduced a circle model on which firms are located and compete. Much sophistication has been built into such competition models over the years. Syverson (2004), for example, extends the Salop model to allow for heterogeneous producer costs and adds asymmetric information among producers about their production costs (see also, e.g., Barros, 1999; Dell'Ariccia, 2001; Kim, Lozano-Vivas and Morales, 2007). Our assumptions are less stringent than those in the original Salop model. In our model, borrowers are uniformly distributed on a two-dimensional plane and banks can cluster in a locality (in contrast to the Salop model where banks are equidistant).

We proceed as follows. Section 2 introduces our theoretical background and hypotheses development after which Section 3 describes our data. Section 4 then sets out our methodology and Section 5 reports the empirical findings. Section 6 concludes.

2. Theory and hypotheses

2.1 General setup

To derive testable hypotheses about the impact of information sharing on bank clustering, we use a stylized model that builds on Hauswald and Marquez (2003, 2006).[6] The model considers two types of borrowers: high quality (h) and low quality (l) ones.[7] The probability that a

[5] Van Cayseele, Bouckaert, and Degryse (1994) analyze theoretically the effect of sharing 'negative' borrower information about past defaults and 'positive' information about indebtedness on the number of branches per bank. Unlike our paper, the authors do not analyze the spatial distribution of branches.

[6] In an earlier working paper (Qi et al, 2021) we show that similar predictions can be derived from a simple spatial oligopoly model of financial agglomeration. That model builds on Konishi (2005) – who models the spatial concentration of retail stores – and does not explicitly deal with information sharing between banks. The main intuition of that alternative model is that branch clustering increases the likelihood that an entrepreneur will visit a locality and obtain a loan. This boosts the size of the local banking market (market-size effect). At the same time, inter-bank proximity also implies more vigorous competition (price-cutting effect). If the first effect dominates, banks earn higher profits by locating closer to each other so that they attract more clients. If the second effect dominates, banks try to decrease competition by dispersing their branches geographically.

[7] To ensure tractability, we assume that depositors put all their savings in the nearest bank branch and that the introduction of information sharing has no impact on the deposit market, which is much less affected by information asymmetries. Our focus on lending as the key banking activity is consistent with much of the literature

borrower is of high quality is q and the distribution of borrower types is public knowledge. Each potential borrower has a project that requires an initial investment of $1 and that generates a final cash flow X. The size of cash flow X will be an amount R with probability p_θ and 0 with probability $1 - p_\theta$ where $\theta \in \{l, h\}$ indicates the borrower's type. Like Hauswald and Marquez (2003) we assume that the success probability for the borrower with the better investment opportunity is higher: $p_h > p_l$. Final cash flows are observable and contractible but neither the bank nor the borrower itself do initially know borrower type θ. We also assume that borrowers have no private resources, so that it is ex post efficient to lend to good borrowers but not to bad ones. Moreover, letting $\bar{p} = qp_h + (1 - q)\, p_l$ denote the average success probability, we assume that $\bar{p}R > 1$. This means that it is ex ante efficient for banks to lend.

Our spatial economy consists of a finite number of localities (villages, towns, and cities) on a two-dimensional plane. Each locality is associated with a segmented credit market that is represented by a circle with a specific radius r. The length of the radius, and hence the spatial boundary of the local credit market, is defined by the information advantage of the bank branches in the locality (see below).[8]

The economy contains an exogenous and fixed number of incorporated banks.[9] Each bank operates a network of branches across a strictly positive number of localities. These branch networks only overlap partially and branch location is assumed to be exogenous. The initial dispersion of branches may for example reflect the agglomeration of people and economic activity across space. Each local credit market hence contains an initial number of incumbent bank branches (n) that compete for local borrowers. The incumbent bank branches are informed or "inside" branches because they have access to a screening and monitoring technology that generates a meaningful signal about borrower quality. Without operating a branch in a locality,

(e.g., Stein, 2002; Berger and Udell, 2006; Hauswald and Marquez, 2006, among others). An interesting exception is Park and Pennacchi (2009) who concurrently model credit granting and deposit taking. Several recent contributions also highlight the importance of deposit taking for value creation in banking (Egan, Lewellen and Sunderam, 2017; Drechsler, Savov and Schnabl, 2017, 2021). We leave the spatial modelling of the information derived from observing checking account turnover, for example, for future research.

[8] According to the president of the Italian Bankers' Association, "*the banker's rule of thumb is to never lend to a client located more than three miles from his office*" (quoted in Guiso, Sapienza and Zingales, 2004). The median Belgian small business borrower in Degryse and Ongena (2005) is located 2.5 kilometers (1.6 miles) from the lending branch. In U.S. data analyzed in Petersen and Rajan (2002) and Agarwal and Hauswald (2010) this median distance is 3.7 km (2.3 miles) and 4.2 km (2.6 miles), respectively. More recently, Granja, Leuz and Rajan (2021) have shown that the distance between U.S. banks and their small business borrowers is not constant over time but instead behaves procyclically as banks take on more (less) risk during boom periods (downturns).

[9] This aligns well with the fact that during our sample period only few greenfield banks entered Central and Eastern Europe by establishing a new branch network from scratch. Our empirical analysis is therefore effectively investigates how incumbent banks reshape and adjust, at the margin, the spatial distribution of their existing branch networks in the immediate aftermath of the introduction of new information sharing regimes.

banks cannot lend to local borrowers. That is, banks can only create a meaningful signal about borrower quality if they operate a local brick-and-mortar branch.

The screening of borrowers in the first period results in credit assessments that provide incumbent bank branches with an informative signal about the type of a local borrower. In particular, screening yields a signal $\eta \in \{l, h\}$ about a borrower's quality. The probability of successful and erroneous credit assessments is given by (where ϕ is the signal quality):

$$\begin{aligned} \Pr(\eta = h \mid \theta = h) = \phi = \Pr(\eta = l \mid \theta = l) \\ \Pr(\eta = h \mid \theta = l) = 1 - \phi = \Pr(\eta = l \mid \theta = h) \end{aligned} \tag{1}$$

Screening loan applicants in the first period is not perfect and depends on the effort I that incumbent bank branches put into generating information and interpreting the results of the credit assessment. Here as well, we follow Hauswald and Marquez (2006) and assume that signal quality ϕ decreases in the borrower's distance x to the bank branch:

$$\phi = (1 + I)/2 - x, \quad I \in (0, 1), \quad x \in [0, I/2) \tag{2}$$

This ensures that the screening in the first period is informative: $\phi \geq 1/2$. Yet, effort I is costly and this cost is given by a function $c(I)$, where $c' > 0$ and $c'' > 0$ for $I > 0$, and $c'(0) = 0$, $c'(1) = \infty$. This captures that branches enjoy an informational advantage in the local credit market in which they operate but that this expertise declines once they try to transact with customers located further away. In this setup, bank branches can use discriminatory pricing through their interest rate offers as a function of the informational content of their screening. The size of the segmented local credit market (in which branches have an information advantage so that $\phi > 1/2$) comprises a circle with a radius r equal to $I/2$.

The timing of our model is as follows. In the first period, incumbent bank branches decide how much costly effort I they will exert to screen local borrowers. In the second period, we then consider an uninformed outside bank that wants to expand its branch network by opening a new branch in a locality in which it does not yet operate one. The outside bank needs to decide in which locality to locate this new branch. In all local credit markets where the outside bank does not yet operate a branch, it is initially uninformed: it cannot generate meaningful signals about local borrower quality. The outside bank then decides which locality to enter by evaluating the equilibrium expected profits of the existing incumbent branches in each local credit market. In period 2, both incumbent branches (informed) and the new outside branch (uninformed) compete by simultaneously making interest rate offers. Local borrowers then accept a loan from the cheapest bank branch. The equilibrium outcome depends on whether the information collected by incumbent bank branches in the first period must be shared with the entrant bank branch in period 2. We first consider the case without information sharing.

2.2 No information sharing

We first derive a potential borrower's success probability in light of the incumbent bank branches' credit assessment. By Bayes' rule, the probability of a project being of high or low quality given a credit assessment of $\eta = h$ or $\eta = l$ is:

$$\begin{aligned}\Pr(\theta = h \mid \eta = h) &= \phi q \,/\, [\phi q + (1-\phi)(1-q)] = H \\ \Pr(\theta = l \mid \eta = l) &= \phi(1-q) \,/\, [\phi(1-q) + (1-\phi)\, q] = L\end{aligned} \tag{3}$$

We obtain the project's success probability conditional on screening result $p(\eta)$ as:

$$\begin{aligned}p(h) &= \Pr(X = R \mid \eta = h) = Hp_h + (1-H)\, p_l \\ p(l) &= \Pr(X = R \mid \eta = l) = (1-L)\, p_h + Lp_l\end{aligned} \tag{4}$$

The incumbent branches' strategy has two components: the effort they put into screening loan applicants in the first period and their interest rate offer in the second period. The latter strategic decision can be made conditional on the screening results. The uninformed entrant bank branch U, by contrast, has no information and can only decide on its interest rate offer.

We index the incumbent bank branches (informed) by i and the entrant bank (uninformed) by u. Let $\pi_i(\eta)$ represent the expected profits, gross of the cost of effort $c(I)$, of an incumbent branch after observing a signal $\eta \in \{h, l\}$. We can then characterize the equilibrium in the lending game once the incumbent bank branches have decided how much effort to invest in the first period. Incumbent and entrant branches compete in interest rate offers in the second period.

While this game does not have a pure strategy equilibrium, there exists a unique mixed strategy equilibrium.[10] It is also a standard result in models of competition under asymmetric information that a bidder, all of whose information is known by some other competitor, cannot make positive expected profits (see, for example, Engelbrecht-Wiggans, Milgrom and Weber, 1983). We therefore conclude that the entrant bank makes zero expected profits in equilibrium.

[10] Von Thadden (2004) and Hauswald and Marquez (2006) provide proofs for the nonexistence of a pure-strategy equilibrium in a similar framework.

To calculate profits for the informed bank, we follow Hauswald and Marquez (2003) and proceed as follows. Define π_u as the expected profits of the uninformed bank. Since the uninformed bank must make zero profit for every one of its possible bids, it must make zero profits also at the lowest possible bid, $\underline{r}$. Offering that rate must guarantee the uninformed bank of having the lowest rate and winning the interest rate auction. Hence $\pi_u(\underline{r}) = 0$, which leads to $\bar{p}\,\underline{r} - 1 = 0$ and $\underline{r} = 1 / \bar{p}$. When observing a low signal ($\eta = l$) the informed branch does not make a loan offer and its profits are $\pi_i(l) = 0$. However, upon observing a high signal ($\eta = h$) the informed branch bids and realizes expected profits of $\pi_i(h) = \underline{r}\, p(h) - 1 = p(h) / \bar{p} - 1 > 0$. In equilibrium, all its loan offers conditional on a high-quality credit assessment yield the same profit. The calculation of ex ante expected profits then follows directly by substituting for $\pi_i(l) = 0$ and $\pi_i(h)$ from the above expression and simplifying. Therefore, the equilibrium expected profits for the informed bank branch are:

$$E\,[\pi_i(\eta)] - c(I) = \Pr(\eta = h)\,\pi_i(h) + [1 - \Pr(\eta = h)]\,\pi_i(l) - c(I) = (p_h - p_l)\,q\,(1 - q)\,(2\phi - 1) / \bar{p} - c(I) \tag{5}$$

By substituting the expression for ϕ from Equation (2) into Equation (5), the expected profits of the incumbent branches are:

$$E\,[\pi_i(\eta)] - c(I) = (p_h - p_l)\,q\,(1 - q)\,(I - 2x) / \bar{p} - c(I) \tag{6}$$

We see that $\partial E\,[\pi_i(\eta)] / \partial I > 0$. This indicates that when the incumbent branches screen applicants more thoroughly, their profits (gross of cost of effort) increase (all else equal). This reflects the increased relative information advantage of these branches. The entrant branch is, in contrast, at a competitive disadvantage as it faces a larger adverse selection problem.[11]

2.3 Information sharing

We now introduce information sharing, which enables incumbent bank branches to screen distant borrowers more effectively, so that the distance parameter x decreases by t. The signal quality derived from the screening technology becomes:

$$\phi = (1 + I)/2 - x/t, \quad I \in (0, 1), \quad t \in [1, \infty) \tag{7}$$

[11] Banks use information sets that only partly overlap (Qian, Strahan and Yang, 2015; Liberti and Petersen, 2019) and that depend on each bank's expertise (Hauswald and Marquez, 2006). Banks also use different screening techniques to process information (Beck, Ioannidou and Schäfer, 2018) so that even when they assess exactly the same information, lending outcomes may differ.

As t increases, the efficiency of screening distant borrowers improves. For $t = 1$, there is no enhancement in screening technology but as $t \to \infty$, the introduction of information sharing completely eliminates the negative impact of distance. For $t > 1$, with information sharing in place, the size of the segmented local credit market (within which local branches have an information advantage $\phi > 1/2$) is now defined by a circle with a radius r equal to $tI/2 > I/2$. This means that the marginally viable borrower is located further away: informed branches only start to ration credit at a larger distance. Information sharing therefore benefits incumbent branches because the local credit market is expanded as more applicants become 'within reach' as potentially viable (or at least: verifiable) borrowers.

Information sharing also affects the new bank branch. This entrant branch can now acquire some information about a local applicant in the new credit registry or bureau, even without having screened or monitored any local borrowers in the first period. That is, a portion of any collected information in the first stage by the incumbent branches disseminates quickly and becomes freely available to all other branches including potential new entrants. We capture this effect by assuming that information sharing implies that an entrant branch costlessly observes a public signal about borrower type: $\eta_p \in \{l, h\}$. The probability of a successful credit assessment based on this public signal is:

$$\phi_p = (2 - 1/t)\ /2,\ t \in [1, \infty) \tag{8}$$

where $\phi_p = \Pr(\eta_p = h \mid \theta = h) = \Pr(\eta_p = l \mid \theta = 1)$. As t increases, the quality of the publicly available signal improves. For $t = 1$, the public signal has no information content. As $t \to \infty$, the signal becomes perfectly informative. The introduction of information sharing hence corresponds to an increase in t. Assuming that any incumbent bank branch also observes the public signal η_p, in addition to the result of its own private appraisal η, one can solve the model of competition with asymmetric information as in the previous section. The key difference is that the entrant branch now also has some information about borrowers instead of being completely uninformed. Importantly, we assume that banks can only lend to borrowers if they operate a physical branch in the local credit market. So even with information sharing in place, an entrant bank will still need to open a local physical branch in period 2 in order to truly benefit from the public information.[12]

[12] In the extreme, in which all banks only use hard information to make lending decisions, the introduction of (hard) information sharing could in principle make bank branches obsolete. However, in most real-world settings, including in the countries in our sample, lenders use a combination of hard and soft information when making lending decisions (Iyer, Khwaja, Luttmer and Shue, 2016; Liberti and Petersen, 2019). The central availability of

Furthermore, we assume that t is a function of the number of incumbent bank branches n and the quality of the information sharing system s, as $t(n, s)$, where $t' > 0$ for both $n > 0$ and $s > 0$. The intuition is that, with more incumbent branches gathering and processing information in the first period, the quality of the public signal in the second period increases. This is in line with Pagano and Jappelli (1993) who assume that the net benefit of information sharing per bank increases with the number of banks joining the system. With a better information sharing system in place, entrant bank branches get better public signals.

Because information sharing erodes the competitive advantage of incumbent branches, it reduces their return to screening in the first period.[13] After all, entrant branches can now obtain some of this information cost-free in period 2. An equivalent approach is to treat information sharing as decreasing the quality of an incumbent branch's credit assessment. Combining the aforementioned impact of information sharing that enables branches to screen distant borrowers more efficiently (so that distance parameter x decreases with information sharing t) we therefore redefine the screening technology from the previous section to incorporate these effects. We again assume that only the incumbent branch obtains private information:

$$\phi = (1 + I/t)\,/2 - x/t,\; I \in (0, 1),\; t \in [1, \infty) \tag{9}$$

As demonstrated by Hauswald and Marquez (2003), this more tractable specification is equivalent to assuming that other potential banks have access to a public signal such as the one in Equation (8). As t increases, the relative informational advantage of an incumbent branch diminishes, as indicated by I/t, precisely because other branches can gain access to at least some of its privately generated information.

Using this framework, we now consider the impact of information sharing on the equilibrium expected profits of incumbent bank branches. In particular, we analyze the effect of an increase in t on the equilibrium that obtains when incumbent and entrant branches compete for borrowers in period 2. Specifically, we substitute the expression for ϕ from Equation (9) into Equation (5), and get the expected profits of incumbent bank branches as:

$$E\,[\pi_i(\eta)] - c(I) = (p_h - p_l)\, q\, (1 - q)\, (I - 2x)\, /\, \bar{p}t - c(I) \tag{10}$$

hard information then reduces information asymmetries – and eliminates the absolute distance threshold beyond which borrowers cannot apply successfully – but still requires borrowers to visit a branch to apply in person.

[13] See also Karapetyan and Stacescu (2014) and von Thadden (2004).

We see that $\partial E\,[\pi_i(\eta)] / \partial t < 0$. The introduction of information sharing reduces the information asymmetries across branches and levels the playing field for entrants. Easier access to information allows entrant branches to better sort borrowers and therefore reduces some of the problems associated with competition under asymmetric information. All else equal, information sharing also reduces the profits earned by incumbent branches (gross of their cost of effort) because of their decreased information advantage. That is, a new entrant branch is no longer completely uninformed and hence becomes more competitive and more willing to enter a market with incumbent branches. We incorporate $t(n, s)$ in Equation (10) to obtain:

$$E\,[\pi_i(\eta)] - c(I) = (p_h - p_l)\, q\, (1 - q)\, (I - 2x) \;/\; \bar{p}t(n, s) - c(I) \tag{11}$$

Crucially, $\partial E\,[\pi_i(\eta)] / \partial t \partial n < 0$. This means that information sharing gives an entrant branch access to more private information (collected by incumbent branches in period 1) in localities with more incumbents. In such localities, entrant banks' relative competitive advantage increases the most. As a result, after the introduction of information sharing, new branches are more likely to open in localities with more incumbent bank branches.

Taken together, the introduction of information sharing has a two-sided effect on incumbent branches. On one hand, these branches benefit from information sharing. They can now expand their local credit market by screening more distant loan applicants that they previously had to disregard. This increases credit availability to the extent that some safe borrowers are no longer priced out of the market by adverse selection (see also Pagano and Jappelli, 1993). We call this the market-size effect. On the other hand, however, information sharing hollows out the information advantage of incumbent bank branches. We call this the competition effect.[14]

In contrast, for the entrant bank, the impact of information sharing is one-sided. It diminishes the information advantage of incumbent branches and thus increases the competitive advantage of entrants (competition effect). Entrant branches can now enter local credit markets that were previously off limits to them. This will especially be the case in localities with more incumbent branches, which in period 1 already collected relatively extensive private information about local borrowers. This yields our first testable hypothesis:

[14] Local loan applicants will not experience an increase in their choice set if the entrant bank buys an existing branch from a competitor rather than establishes a new branch of its own. Wang (2019) shows for the U.S. how entrant banks are more likely to buy existing branches rather than open new ones if the enforcement of non-compete contracts makes it difficult to poach local employees (who harbor soft information).

Hypothesis 1: *Information sharing increases the likelihood that banks open new branches in localities with more pre-existing branches of other banks (all else equal). This increases the spatial clustering of bank branches.*

Our model predicts that banks exploit the opportunities of sharing borrower information by extending their branch network to localities where new branches can benefit from the information gathered by incumbent branches of other banks. In contrast, adding more branches of the same bank in a locality where a bank is already present does not facilitate easier and better information access. Our second hypothesis is therefore:

Hypothesis 2: *After the introduction of information sharing, banks are more likely to open new branches in localities with no (or fewer) pre-existing own branches (all else equal).*

Our model also speaks to how information sharing differentially affects relationship lenders and transactional banks. Whereas relationship banks depend on in-depth screening and long-term lending relationships during which they obtain and exploit proprietary borrower information, transactional banks rely more on public information (Boot, 2000; Mian, 2006; Kysucky and Norden, 2016; Beck, Ioannidou, and Schäfer, 2018). To obtain proprietary information, relationship banks need to interact frequently and intensely with borrowers. In our model, this translates into a negative distance impact on the effectiveness of screening that is stronger for relationship banks than for transactional banks, $x^{relation} > x^{transaction}$. As a result, $E\,[\pi_i(\eta)]^{relation} < E\,[\pi_i(\eta)]^{transaction}$. The impact of information sharing on relationship banks is hence stronger than on transactional banks. This leads to our third hypothesis:

Hypothesis 3: *The impact of information sharing on bank clustering is stronger for relationship banks.*

Lastly, our model predicts that the extent to which information sharing eliminates information asymmetries, and thus fosters branch clustering, directly reflects how comprehensive and trustworthy the shared borrower information is. We can see from Equation (11) that $\partial E\,[\pi_i(\eta)] \,/\, \partial t \partial s < 0$. This indicates that the competitive advantage of incumbent bank branches is decreasing with the quality of information sharing. In countries with more comprehensive and trustworthy information sharing systems, entrant banks' relative competitive advantage is larger. Our fourth and final hypothesis is therefore:

Hypothesis 4: *The impact of information sharing on bank clustering is stronger in countries with higher quality information sharing systems.*

3. Empirical setting and data

3.1 Empirical setting

After the fall of the Berlin Wall in 1989, the countries of Central and Eastern Europe started a process of profound economic transition from socialist, centrally planned economies to capitalist market economies. The first decade of this transition process was highly disruptive. Large-scale privatization and widespread liberalization brought about a massive reallocation of labor and capital across and within economic sectors, rectifying the distortions inherited from central planning (EBRD, 1999; Roland, 2000; Svejnar, 2002).

An important element of the transition process was the development of well-functioning banking sectors. After having recapitalized state banks during the first half of the 1990s, countries started to privatize these banks. Privatization often involved selling majority stakes to foreign strategic investors, mostly multinational banking groups from Western Europe. These foreign-owned banks currently regard Emerging Europe as a second home market where they compete with the remaining domestic banks. In many countries, foreign banks presently own between 20 and 90 percent of all banking assets (De Haas, et al., 2015).

After the deep and disruptive reforms of the 1990s, policy makers realized that liberalization and privatization needed to be complemented with new institutions that actively supported the functioning of markets and private enterprise. It was during this later wave of institutional reforms that countries introduced credit registries and bureaus. Appendix Table A3 shows that the average credit registry (credit bureau) was only introduced in 2003 (2005).

The above background to our empirical setting also helps to clarify to what extent our findings carry over to other countries that introduced, or will introduce, information sharing. It can be argued that the introduction of information sharing relatively soon after a period of economic upheaval, as was the case in Emerging Europe, was not a one-off anomaly. In fact, historically, credit registries and bureaus have often been introduced (and continue to be introduced) in countries that recently experienced structural transformation. For example, Germany established its public credit registry in 1934, in between the two World Wars, while Ukraine only introduced a registry in 2019 after an economic crisis in 2014-16 (it had been operating a credit bureau since 2007). While this suggests that our setting may be quite representative, it also underlines the importance of absorbing local time-varying developments in population and economic activity and of robustness checks that control for concomitant policy changes and reforms. We will discuss both issues in the sections to come.

3.2 Data

To test our hypotheses, we use the introduction of information sharing regimes as country-level shocks that push banks towards a new clustering equilibrium. This approach requires time-varying data on branch locations for countries that introduce information sharing – either through a public credit registry or through a private credit bureau – at different points in time. We have access to information on the geographical coordinates of 56,555 branches owned by 614 banks in 8,536 villages, towns or cities (i.e., 'localities') across 19 emerging European countries. These branches represent over 95 percent of all bank assets in our sample countries. A team of consultants with extensive banking experience collected these data by contacting banks or downloading data from bank websites. We double-checked all information with the banks as well as with the SNL Financial database. This data collection exercise was part of the second Banking Environment and Performance Survey (BEPS II).[15] The data paint a precise, complete and gradually changing picture – reflecting branch openings and closures – of the banking landscape during the years 1995 to 2012. Figure 1 depicts the spatial branch distribution in these countries at the start and the end of our sample.

[Insert Figure 1 here]

Appendix Table A2 summarizes the number of branches that opened or closed by year and country: 31,927 (1,065) branches opened (closed) during our sample period. Many branches were established during 2001-07, a period of rapid credit growth. The expansion of branch networks slowed down after the global financial crisis when fewer branches opened while branch closures (rare before the crisis) accelerated. Approximately half of all branch openings took place when a country had a credit registry or bureau in place.

The unit of observation in our main analysis is the bank-locality-year (see Section 4). This means that for each bank in our data set, we track the number of existing branches (if any), the number of newly opened branches, and the number of closed branches. For each branch, we

[15] For more details, see Beck, Degryse, De Haas and Van Horen (2018) and www.ebrd.com/what-we-do/economics/data/banking-environment-and-performance-survey.html.

know its geo-coordinates and its address (street; name of village, town or city (the 'locality') and postal code). We use this information to aggregate all branches up to the locality level. We have information on 8,536 such localities for every year in the period 1995-2012. The resulting dependent variables capture the opening of new bank branches across localities and over time. Table 1 contains summary statistics while Appendix Table A1 provides all definitions.

New branch opening is a dummy variable that captures whether a particular bank opens a new branch in a locality in a given year. *Net branch opening* is a dummy that also takes branch closures into account: it equals one if in a particular year and locality a bank adds at least one branch in net terms (that is, the number of branch openings minus closures is strictly positive), and equals zero otherwise. Table 1 shows that on average 4 percent of all bank-locality-year observations see a new branch opening. Given the small number of branch closures, this percentage is virtually the same for the variable *Net branch opening*. We also count the log of (one plus the) number of pre-existing branches in a locality that are owned by other banks (*No. branches other banks*) and by individual banks (*No. branches own bank*). Variation is substantial, with some localities not being served by any bank whereas some of the largest localities contain many bank branches.

[Insert Table 1 here]

To contrast the impact of information sharing on relationship lenders versus transactional lenders, we use three empirical proxies of banks' lending techniques: size, ownership, and a direct measure of a bank's main lending technology. We first classify a bank as small if the number of branches it operates is strictly below the country median. The existing literature suggests that small banks are more likely to apply relationship-lending techniques and hence have a comparative advantage in lending to small and informationally opaque firms. In contrast, large banks tend to be better at lending to larger and more transparent firms (Cole, Goldberg and White, 2004; Berger et al., 2005). We therefore expect that the introduction of information sharing affects smaller banks more. Table 1 shows that 32 percent of all banks in our data set are small and that these banks own 10 percent of all bank branches in our sample.

Second, we merge our data with detailed, hand-collected bank ownership information to distinguish between branches of foreign and domestic banks. A bank is classified as foreign if at least half of its equity is in foreign hands. Domestic banks can possess a comparative advantage in reducing information asymmetries vis-à-vis local firms (Mian, 2006; Beck, Ioannidou, and Schäfer, 2016). In this view, domestic banks tend to have a deeper understanding of local businesses and typically base their lending decisions on 'soft' qualitative information on these firms (Berger and Udell, 1995, 2002; Petersen and Rajan, 2002). In contrast, foreign banks may have difficulties in processing soft information and therefore tend to grant loans on a transaction-by-transaction basis using standardized decision methodologies (Berger, Klapper and Udell, 2001). Table 1 shows that only 43 percent of the banks in our country sample are still in domestic hands, reflecting the high levels of foreign direct investment in these banking systems (see also Section 3.1). Domestic banks tend to be relatively large and on average account for 51 percent of all bank branches.

Third, we determine more directly whether a bank is a relationship lender or a transactional lender when providing credit to small businesses. Recent contributions argue that foreign banks, just like their domestic competitors, can successfully lend to small businesses (Berger and Udell, 2006). Indeed, Beck, Degryse, De Haas and Van Horen (2018) show that among both domestic *and* foreign banks in emerging Europe, large proportions operate as relationship lenders. Banks' ownership and their lending techniques may thus be more orthogonal than previously thought.

To characterize banks' lending technologies, we follow Beck et al. (2018) and use question Q6 of the 2nd Banking Environment and Performance Survey (BEPS II). As part of this unique survey, the CEOs of banks participated in in-depth, face-to-face interviews in 2012. Question Q6 asked CEOs to rate on a five-point scale the importance (frequency of use) of the following techniques when dealing with small businesses: relationship lending; fundamental and cash-flow analysis; business collateral; and personal collateral (personal assets pledged by the entrepreneur). Although, as expected, almost all banks find building a relationship (knowledge of the client) of some importance, 59 percent of the banks find building relationships "very important", while the rest considers it only "important" or "neither important nor unimportant". We categorize banks that find client relationships to be "very important" as relationship lenders and all other banks as transactional lenders.[16] Table 1 also shows that while relationship banks make up 59 percent of all banks, they own only 45 percent of all branches (among the banks for which we have data on lending technologies). This confirms that relationship lenders are typically somewhat smaller than transactional lenders.

[16] We have this information for slightly over half of all the banks in our sample. Beck et al. (2018) use credit registry data to show that when CEOs consider relationship lending to be very important, according to BEPS II, this is indeed reflected in the lending practices of their bank.

Next, we collect data on the introduction of information sharing regimes from the World Bank Doing Business database. Appendix Table A3 shows that during 1995-2012, 13 out of the 19 countries in our data set introduced a public credit registry and 15 a private credit bureau. There exists substantial variation in the timing of these introductions, which is crucial for our empirical identification. We also measure the quality of these information-sharing regimes through the World Bank Doing Business credit information index. This index ranges from zero to six and reflects the rules and practices that affect the coverage, scope, and accessibility of credit information (higher values indicate information sharing that is more effective).[17] Unconditional (conditional) on either a credit registry or a credit bureau being in place, the average quality score across countries and years is 1.3 (2.4).

Lastly, to test whether firms can borrow from more distant bank branches after the introduction of information sharing, we merge our branch data with the Kompass database on firm-bank relationships. Kompass provides information on firms' address, industry and – critically for our purposes – the primary bank relationship.[18] We have these data for the years 2000 and 2005. We collect the geographical coordinates of Kompass firms based on their name and address and identify the name of their primary bank. We then match each Kompass firm to all the branches of their primary lender (using BEPS II information) and calculate the distance from the firm to each of these branches. We assume that firms borrow from the nearest branch of their primary bank and use this nearest distance as the *Firm-branch distance* in kilometers. The median distance between a firm and its primary bank is 1.8 km (Table 1).

The Kompass data also allow us to create proxies for firms' relative opaqueness. We construct the following three dummy variables: whether the firm has a publicly available email address (*Has email address*); whether the firm has a tax number (*Has tax number*); and whether the firm has formal opening/working hours (*Has formal opening hours*). Table 1 shows that 60 percent of all firms have a publicly available email address, almost 74 percent of them have an official tax number, and 74 percent of the firms work based on formal openings hours.

[17] A score of one is assigned for each of six features: both positive credit information (outstanding loans and on-time repayments) and negative information (late payments and defaults) are distributed; data on both firms and individuals are distributed; data from retailers, utility companies, and financial institutions are distributed; more than two years of historical data are distributed; data on loan amounts below one percent of income per capita are distributed; and by law borrowers have the right to access their data in the largest credit bureau or registry.
[18] Other papers that employ Kompass include Ongena and Şendeniz-Yüncü (2011), Giannetti and Ongena (2012), Kalemli-Özcan, Laeven and Moreno (2018), and Beck, Ongena and Şendeniz-Yüncü (2019).

4. Identification

To test our hypotheses, we apply a difference-in-differences-in-differences (DDD) framework in which (i) the treatment (the presence of information sharing) varies across countries and years and (ii) localities within countries are affected differentially depending on the pre-existing bank branch structure. Because our treatment is introduced in a staggered fashion over time, and because treatment effects may be heterogeneous across countries, a standard two-way fixed effects framework can yield biased estimates. This will be the case if some already-treated countries (with information sharing in place) incorrectly act as controls for later events (Goodman-Bacon, 2018; Abraham and Sun, 2021; Callaway and Sant'Anna, 2021).

To address this issue, we follow Cengiz, Dube, Lindner, and Zipperer (2019) and create event-specific data sets. Each event includes all observations from the countries in which information sharing (the treatment) is introduced in the same calendar year as well as the observations from all clean control countries for a six-year panel by event time ($t=-3, \ldots, 2$) with information sharing introduced at $t=0$.[19] Clean control countries are those without any information sharing system in place during the full six-year event window. We stack these event-specific data sets to estimate a single average DDD result. Aligning events by event time instead of calendar time is equivalent to a setting where all events happen simultaneously. This approach avoids biases due to the negative weighting of some events (which can occur in a staggered design) or due to heterogeneous treatment effects (Goodman-Bacon, 2018).

We then compare how, within the same country, the introduction of information sharing differentially affects branch openings across localities with different numbers of pre-existing branches from other banks. To test Hypothesis 1, we estimate the linear probability model:

$$\text{New branch opening}_{ijct} = \beta_1 * \text{Information sharing}_{ct} * \text{No. branches other banks}_{ijct} + \beta_2 * \text{No. branches other banks}_{ijct} + \Phi_{jt} + \Phi_{it} + \Phi_{ij} + \varepsilon_{ijct} \quad (1)$$

where the dependent variable $\text{New branch opening}_{ijct}$ is a dummy that equals one if bank i opens a new branch in locality j of country c in year t, and equals zero otherwise. $\text{Information sharing}_{ct}$ is a dummy that equals one if banks in country c share borrower information in year t, and equals zero otherwise (the level effect of this variable is absorbed by the fixed effects). $\text{No. branches other banks}_{ijct}$ measures the number of pre-existing branches by banks other than

[19] In addition to the nineteen countries in our dataset, we also have bank branch data for Hungary, Lithuania, and Slovenia. However, because these countries already introduced information sharing in 1994 or 1995, we cannot include them in our analysis, as no pre-treatment data are available.

bank i in locality j. Based on our model, we expect β_1 to be positive as the introduction of information sharing induces banks to cluster to attract more borrowers. That is, after the introduction of a credit registry or bureau, banks are more likely to open new branches in localities that already had more branches of other banks to begin with.

The most saturated version of this model includes three types of interactive fixed effects: *locality*year* (Φ_{jt}); *bank*year* (Φ_{it}); and *bank*locality* (Φ_{ij}). We also allow these effects to vary by treatment event. *Locality*year* fixed effects absorb all time-varying and time-invariant historical, social, economic and cultural differences across villages, towns or cities. Importantly, this includes local trends in credit demand that may affect the location choice of banks. These fixed effects also wipe out local variation in labor markets or in the available IT infrastructure. *Bank*year* fixed effects flexibly account for time variation in individual banks' operational strategies and financial health that affect their branch network as a whole. *Bank*locality* fixed effects absorb time-invariant variation across banks in each specific locality. Lastly, ε_{ijct} is the error term and we cluster standard errors at the *country*treatment event* level.[20]

To test Hypothesis 2, we measure the number of pre-existing branches of the *same* bank in each locality (*No. branches own bank*) and run the following linear probability model:

$$\textit{New branch opening}_{ijct} = \beta_1 * \textit{Information sharing}_{ct} * \textit{No. branches own bank}_{ijct} + \beta_2 * \textit{No. branches own bank}_{ijct} + \Phi_{jt} + \Phi_{it} + \Phi_{ij} + \varepsilon_{ijct} \quad (2)$$

Our theoretical model predicts that information sharing reduces the probability that banks open a new branch in localities where they themselves already operate one or several branches. We thus expect β_1 to be negative.

To test Hypothesis 3, we also examine whether information sharing differentially affects relationship versus transactional lenders. We do so by further interacting our treatment with *Bank type* and running the following model:

$$\textit{New branch opening}_{ijct} = \beta_1 * \textit{Information sharing}_{ct} * \textit{No. branches other banks}_{ijct} + \beta_2 * \textit{Information sharing}_{ct} * \textit{No. branches other banks}_{ijct} * \textit{Bank type}_{ic} + \beta_3 * \textit{No. branches other banks}_{ijct} * \textit{Bank type}_{ic} + \beta_4 * \textit{No. branches other banks}_{ijct} + \Phi_{jt} + \Phi_{it} + \Phi_{ij} + \varepsilon_{ijct} \quad (3)$$

[20] We use linear models so we can fully saturate them with the aforementioned fixed effects. Including these fixed effects in a non-linear (e.g., probit) model would introduce an incidental parameters problem: parameter estimates would not converge to their true value as the number of parameters grows with the number of observations.

Bank type is one out of three time-invariant proxies for whether a bank is a relationship lender: a small bank dummy; a domestic bank dummy; or a dummy for whether relationship lending is the main technique when lending to small businesses. Based on our theoretical model, we expect information sharing to have a bigger impact on relationship than on transactional lenders so that especially relationship lenders start to open new branches in localities with more pre-existing branches of other banks. That is, we expect both β_1 and β_2 to be positive.

Lastly, we investigate whether the relationship between information sharing and branch clustering is more pronounced when the quality of information sharing is higher (Hypothesis 4). The time-varying variable *Quality information sharing*$_{ct}$ measures the rules and practices affecting the accessibility, coverage, scope, and quality of the borrower information that is publicly available. Augmenting the base regression (1) with this variable renders:

$$New\ branch\ opening_{ijct} = \beta_1 * Information\ sharing_{ct} * No.\ branches\ other\ banks_{ijct} + \beta_2 * Quality\ information\ sharing_{ct} * No.\ branches\ other\ banks_{ijct} + \beta_3 * No.\ branches\ other\ banks_{ijct} + \Phi_{jt} + \Phi_{it} + \Phi_{ij} + \varepsilon_{ijct} \quad (4)$$

Quality information sharing$_{ct}$ is by definition only available for country-years in which banks exchange borrower information (that is, when *Information sharing*$_{ct}$ equals one). It equals zero if there is no information sharing in a specific country and year. Based on our theoretical model, we expect β_2 (and β_1) to be positive.

5. Empirical results

5.1. Baseline results

Table 2 presents regression results based on the linear probability models (1) and (2). The dependent variable is *New branch opening*, which indicates whether a bank opens a branch in a particular locality in a particular year. We investigate hypotheses 1 and 2 in columns 1 to 4 and 5 to 8, respectively, while increasingly saturating the models with interactive fixed effects. Columns 2 and 6 include *Locality*Year*Treatment event* fixed effects while in columns 3 and 7 we add *Bank*Year*Treatment event* fixed effects. We further saturate the specifications with *Locality*Bank*Treatment event* fixed effects in columns 4 and 8. These granular fixed effects together capture unobserved variation at various levels, including changes in local credit demand, which might otherwise bias our results.

In line with our first hypothesis, columns 1 to 4 show that when a country introduces information sharing, banks become more likely to open new branches in localities with more pre-existing branches of competitor banks. This effect of establishing a credit registry or credit bureau is also economically significant.[21] Our preferred (most complete) specification in column 4 indicates that with information sharing in place, a one standard deviation higher number of pre-existing bank branches in a locality increases the probability that an additional new branch is opened in that locality by 62 percent. Especially the inclusion of *Locality*Year*Treatment event* fixed effects – which absorb local and time-varying conditions such as changes in credit demand – is quantitatively important. These fixed effects allow us to compare how different banks – with different numbers of pre-existing branches in the same locality in the same year – react differentially to the introduction of information sharing. The second row of coefficients shows that also in the absence of information sharing, a higher presence of other bank branches increases the chances of additional branches opening. Yet, the introduction of information sharing significantly increases this tendency of banks to cluster together, as predicted by our model.

Next, columns 5 to 8 show that in line with Hypothesis 2 information sharing induces banks to open new branches in localities where they operate fewer existing branches of their own. The effect is again sizable. Column 8 shows that after the introduction of information sharing, a one standard deviation increase in the number of pre-existing own branches in a locality reduces the likelihood that a bank opens another branch in that locality by 7 percentage points. In sum, Table 2 shows that the introduction of information sharing induces banks to open branches in localities where they did not yet operate themselves but where relatively many other banks were already present. As a result, the spatial clustering of bank branches intensifies once a country starts operating a credit registry or a credit bureau.

[Insert Table 2 here]

To gain more insights into the dynamics at play, we conduct an event-study analysis where we define an event as the year in which a country introduces an information-sharing regime. We present results for a six-year window around these events (the year of introduction is $t=0$). Figure 2 shows the pre- and post-trends for the probability of opening a new branch in localities

[21] When we assess the separate impact of the introduction of credit registries versus credit bureaus, we find that both influence bank clustering in a similar way.

with more pre-existing branches owned by other banks (Panel A) and by the bank itself (Panel B). All estimates are expressed as changes relative to event date $t=-1$ (the estimates for which we normalize to zero) and based on the most saturated specifications in column 4 (Panel A) and column 8 (Panel B) of Table 2.

Figure 2 reveals sharp changes in where banks open new branches at the time of the introduction of information sharing. Banks become more likely to open new branches in localities with more pre-existing branches from other banks (Panel A) but fewer branches of their own (Panel B). While the magnitude of the estimated effects gets somewhat smaller over time, the impact continues to be substantial and statistically significant three years out. This suggests that the introduction of information sharing pushes banks towards a durable new clustering equilibrium.

We observe some slight trends prior to the introduction of information sharing but these leads appear small relative to the post-treatment effect estimates. To investigate more formally how robust our results are to deviations from the parallel trends assumption, we follow Rambachan and Roth (2021). We assume that the observed small pre-trend differences persist and can be extrapolated into the treatment period. Extrapolation can either be linear or we can allow the slope of the differential trend to change between consecutive periods by a parameter M. Figure 3 reports the results. It shows the confidence set for the treatment parameters of interest: the number of pre-existing branches of other banks in a locality (Panel A) and the number of pre-existing branches of the bank itself (Panel B). The baseline DiD estimate is shown in blue and the estimates that allow for extrapolated pre-trends are shown in red. The first red estimate corresponds to $M=0$, which indicates a linear extrapolation of the pre-existing trend. Our results are robust to this extrapolation. Moving further to the right in each panel, we let M increase and thus allow for deviations from the pre-trend during the treatment period. The estimates remain very stable in Panel A and only become insignificant in Panel B once we reach an M of about 0.02.

Taken together, the sharp changes that we observe at $t=0$ in terms of where banks open new branches; the lack of substantial pre-treatment trends; the robustness of our results to extrapolating the limited pre-treatment effects into the treatment period; and the persistent post-treatment effects all support our research design.

[Insert Figures 2 and 3 here]

5.2. Information sharing, relationship lending, and branch clustering

The introduction of information sharing may not affect all banks equally. In particular, our third hypothesis states that the impact of information sharing will be stronger among relationship banks as compared with transactional banks. In Table 3, we test this hypothesis by further interacting our main interaction term – *Information sharing*No. branches other banks* – with the variable *Bank type*. *Bank type* is one of three proxies for a bank's reliance on relationship lending: whether the bank is relatively small (columns 1-2); whether it is domestically owned (columns 3-4); and whether its CEO finds relationship lending a very important technique to provide credit to small businesses (columns 5-6). From hereon we focus on the two regression specifications that are most saturated with interactive fixed effects.

The first two columns of Table 3 show that while the introduction of information sharing increases the tendency of large banks to cluster their branches, this impact is somewhat larger for small banks. To the extent that smaller banks rely more on relationship lending, this finding is therefore in line with our third hypothesis. Economically, when information sharing is introduced, a one standard deviation higher number of pre-existing branches of competitor banks in a locality increases the probability of a new branch opening by 64 and 68 percent for large and small banks, respectively. This difference does not simply reflect that small banks are more likely to open new branches. Instead, it shows that conditional on a new branch being opened, small banks are particularly likely to do so in a locality with more pre-existing branches once information sharing is introduced.

Next, in columns 3 and 4 of Table 3, we assess heterogeneity by bank ownership. As discussed in Section 2, some prior studies have proxied lending technologies by comparing domestic versus foreign banks. The traditional dichotomy is then that domestic banks are mostly relationship lenders while foreign banks rely more on transactional lending. Yet, we find no evidence for heterogeneous effects of information sharing by bank ownership. The triple interaction terms in columns 3 and 4 are small and imprecisely estimated. This null result is in line with Beck, Degryse, De Haas and Van Horen (2018) who find that the often (implicitly) assumed 'bank stereotype' that domestic banks are relationship lenders while foreign banks are transactional lenders, does not necessarily hold in reality – at least not in the emerging markets in their sample and in ours.

We then proceed by using our most direct measure of a bank's main lending technique when dealing with small businesses. The results in columns 5 and 6 of Table 3 show that while information sharing leads to more branch clustering among transactional lenders (as shown by the coefficient for the interaction term in the first line) the impact on relationship banks is even larger – again in line with our third hypothesis. However, in absolute terms this difference is limited, at about a fourth of the difference between small and large banks.

[Insert Table 3 here]

5.3. The quality of information sharing regimes and branch clustering

Our model posits that the extent to which information sharing successfully eliminates the distance threshold due to information asymmetries, and thus fosters branch clustering, depends directly on how comprehensive and trustworthy the shared borrower information is (Hypothesis 4). In Table 4, we now test whether in countries that introduce a particularly effective information-sharing system, subsequent bank branch clustering is stronger. We can only measure the variable *Quality information sharing* in countries with information sharing in place; in countries without information sharing we set this variable to zero.

Columns 1 and 2 of Table 4 show that in line with our fourth hypothesis the introduction of information sharing boosts branch clustering particularly in countries where the system is more effective. The results in column 2 indicate that an improvement of the registry quality by one point (out of six) increases branch clustering due to information sharing by 18 percent.

In columns 3 and 4, we restrict the data to only those observations from countries and years in which some form of information sharing was in place. That is, we now focus on the intensive margin of information sharing to see whether, conditional on a credit registry and/or bureau being in place, more effective information sharing is associated with more bank clustering. In line with the first two columns, the results indicate that this is indeed the case.

[Insert Table 4 here]

5.4. Robustness, placebo tests, and extensions

<u>Instrumental variables regressions</u>

One may worry that the introduction of information sharing in a country is endogenous as it reflects unobservable national circumstances that also bear directly on branch clustering. However, such country and time specific confounds are at least partly controlled for by our *Locality*Year*Treatment event* fixed effects. A related issue concerns reverse causality whereby the structure of a country's banking sector influences the (timing of) the introduction of information sharing. To alleviate this concern, we instrument the introduction of information sharing in a country-year with the percentage of all neighboring countries that introduced information sharing in the past five years (Martinez Peria and Singh, 2014). This instrumentation strategy builds on the notion that financial reforms tend to converge regionally (Abiad and Mody, 2005). The exclusion restriction is that the introduction of information sharing in nearby countries only has an impact on domestic bank clustering via an increase in the probability that information sharing is introduced domestically as well.

Because the country*year-level variable *Information sharing* gets absorbed by our interactive fixed effects, the endogenous variables in our most saturated baseline specifications are in fact the interaction terms *Information sharing*Number branches other banks* and *Information sharing*Number branches own bank*. We make use of the fact that interactions of instruments with exogenous variables are valid instruments for endogenous variables interacted with exogenous variables (Wooldridge, 2002, p. 122). As first-stage instruments we therefore use interaction terms between the percentage of neighboring countries that introduced information sharing in the previous five years and a locality-level measure of the number of pre-existing branches of other banks (column 1) or the bank itself (column 2).

Table 5 reports our IV results. The first stages (columns 1 and 2) show a strong and positive correlation between the introduction of information sharing in neighboring countries in the recent past (interacted with the local pre-existing branch structure) and the introduction of a credit registry or bureau in the country of observation (similarly interacted). The second-stage estimates are comparable to our baseline results though larger by a factor of three. There are two reasons why the IV estimates may be larger. First, as discussed above, information sharing may have emerged later in countries with relatively strong branch clustering to begin with. Correcting for this endogenous treatment timing then increases the (IV) estimate. A second explanation is the Local Average Treatment Effect (LATE) when the impact of information sharing on branch clustering differs across countries. If information sharing has a larger impact on branch clustering in complier countries (that is, those countries where the introduction of a credit registry or a credit bureau had been delayed by a lack of ‘example’ information-sharing systems in neighboring countries) than in non-complier countries, then the IV estimates will be larger than their OLS counterparts.

[Insert Table 5 here]

Net branch openings and per capita branch openings

Next, we replace our dependent variable, *New branch opening*, by *Net branch opening*. This dummy variable also takes the closure of local bank branches into account by measuring whether there is a *net* increase in the number of branches of a bank in a specific locality and year. We present the results in columns 1-4 of Appendix Table A4. They are very similar to our baseline results in Table 2, both statistically and economically.

In columns 5 to 8 of Table A4, we normalize the number of branches by the population of the locality, using data from the World Cities Database. We construct the variables *No. branches other banks per 1,000 population* and *No. branches own bank per 1,000 population*. Normalizing the presence of bank branches by the local population does not affect our results. This reflects that our most saturated specifications already include *Locality*Year*Treatment event* fixed effects, so that we effectively compare how different banks – with different numbers of pre-existing branches in the *same* locality (of a given population size) in the same year – react differentially to the introduction of information sharing at the country level.[22] Note also that we were able to collect precise population data for about a quarter of all observations. It is reassuring that our results hold up well in this (non-random) sub-sample.

Next, we explore in Table A6 the role of simultaneous branch closures and openings. This analysis yields some interesting additional results. First, the last two coefficient rows in this table show consistently that *before* the introduction of information sharing, banks are less likely to open a new branch in a locality when they close another branch elsewhere in the same year. This indicates that in the absence of information sharing, banks use both margins (closures and openings) in parallel to expand (or shrink) their network.

Interestingly, this pattern changes after the introduction of information sharing. Columns 1 and 2 show that with information sharing in place, banks are more likely to open a branch in a new locality when they close a branch elsewhere in the same year. This holds true especially when banks close a branch in a relatively underbanked locality (that is, a locality in the bottom quartile in terms of bank density) – cf. columns 3 and 4. The coefficient estimates in the most saturated regressions (columns 2 and 4) are borderline statistically significant (*p*-values of 0.11

[22] In Appendix Table A5, we replace the black box of our fixed effects with locality-level measures of the number of bank branches per 1,000 inhabitants; annual population growth; and local economic growth as proxied by the change in nighttime light intensity. When we explicitly control for these locality characteristics (instead of absorbing this variation through fixed effects) our results hold up well. In the main tables, we keep our fully saturated baseline regressions in which fixed effects absorb all observable and unobservable variation. This also allows us to work with the largest possible sample size.

and 0.06, respectively).[23]

In columns 5 to 8, we test whether after the introduction of information sharing, banks are especially more likely to open a branch in an already heavily banked locality when they close a branch elsewhere. We find strong evidence for this, both when we look at branch closures in general (columns 5-6) and when we focus on branch closures in sparse banking markets (columns 7-8). This indicates that the increased clustering due to information sharing works as a double-edged sword: new bank openings in densely banked areas go hand-in-hand with closures in less densely populated areas.

Bank branch clustering in rural versus urban areas

A somewhat separate issue is that our findings could mostly reflect branch clustering in specific parts of countries. A secular urbanization process can induce a disproportionate increase in the opening of new bank branches in urban areas. We may then pick up clustering forces in urban areas that are largely unrelated to (but coincide with) the introduction of information-sharing regimes. Figure 2, which shows *sharp* changes in clustering behavior right after the introduction of information sharing regimes, should already partly dispel concerns about gradual trends driving our estimates.

To look into this in more detail, we superimpose a grid consisting of 30x30 km cells on the 19 countries in our dataset. We use gridded population data from NASA (CIESIN, 2021) and calculate for each cell the local population density (number of inhabitants to the size of the cell) and the local bank branch density (number of bank branches to the size of the cell). Figure 4 provides a binned scatterplot that groups all grid cells into 20 equal-sized bins, based on their population density. We plot the mean population density (horizontal axis) and branch density (vertical axis) for each bin. We do this for the average in the (country-specific) years before (blue) and after (red) the introduction of information sharing. The lines reflect a linear OLS fit with 95 percent confidence intervals.[24]

Figure 4 shows that the introduction of information sharing is associated with an increase in bank branch density across the whole population density distribution, including in rural (less densely populated) areas. Yet, branch density increases most in densely populated areas, suggesting that in relative terms banks open more branches in larger agglomerations than in

[23] We also test whether information sharing is associated with an increased likelihood that singleton branches close down (branches that are the only branch in a locality). We do not find that such solo branches are more likely to close down in general nor are they more likely to close down after the introduction of information sharing.
[24] This binned scatterplot looks very similar when we partial out country fixed effects or use more granular grid cells at the 5x5 km level.

sparsely populated areas. Information sharing has enabled this relative shift as banks no longer need to be present "everywhere" but can now instead count on (potential) borrowers to travel to larger agglomerations with a deeper credit market.

To investigate this issue more formally, we split our sample into localities with less than 50,000 inhabitants; localities with between 50,000 and 250,000 inhabitants; and localities with over 250,000 inhabitants.[25] We then rerun our baseline (fully saturated) regression specifications on all three samples. Appendix Table A7 confirms that our estimates point to a somewhat stronger impact of information sharing in larger localities.[26] Yet, the impacts in more rural areas are highly significant and economically sizable too. We therefore conclude that our baseline findings do not mainly reflect secular urbanization trends.

[Insert Figure 4 here]

Lastly and relatedly, Figure 5 explores the geo-spatial dimension of our data in more descriptive detail. We use Eurostat's Nomenclature of Territorial Units for Statistics (NUTS) hierarchy to divide each country into NUTS 1 'major socio-economic regions' and split these further into smaller NUTS 3 regions. We then take geographic boundary shape files for all NUTS 1 and the related NUTS 3 sub-regions and use the geo-coordinates of individual branches to match them to these (sub-) regions. We do this for every year, so that we get an annual spatial picture of where bank branches are located within the NUTS hierarchy.

For each NUTS 1 region, we then calculate a Herfindahl-Hirschman Index (HHI) as a measure of spatial concentration of bank branches across the NUTS 3 sub-regions in that NUTS 1 region. We calculate this HHI three years before and three years after the introduction of information sharing in the respective country. We also calculate, again at the NUTS 1 level and over the same six-year window, the percentage change in the share of branches located in the most densely banked NUTS 3 sub-region (that is, the NUTS 3 region with the largest share of all bank branches in that NUTS 1 region).

The dashed kernel density plot in Figure 5 (Panel A) shows the distribution of the

[25] We use data from the World Bank-EBRD Business Environment and Performance Survey (BEEPS) to divide localities into these three broad size buckets. This allows us to retain more observations as compared with the approach in Table A4 where we collect the exact population size of localities using the World Cities Database.
[26] When we include in our baseline regressions both the variable *Information sharing * No. branches other banks* and its square, both coefficients are significant and statistically positive. This is suggestive of a convex relationship in which, after the introduction of information sharing, financial agglomeration forces are especially strong in deeper credit markets. This is in line with the fact that in Table A7 the estimated coefficient for *Information sharing * No. branches other banks* increases monotonically when considering increasingly large localities.

percentage change in the NUTS 1 level HHI indices. Likewise, the solid kernel density plot shows the distribution of the percentage change in the share of all branches clustered in the most densely banked sub-region. It is quite striking that both distributions show how in the six-year window around the introduction of information sharing, bank branches tend to agglomerate more within individual NUTS 1 regions.

The strengthening of financial agglomeration occurs soon after the introduction of information sharing (see also Figure 2) and is hence unlikely to reflect concomitant changes in the clustering of people or economic activity. Indeed, Panel B of Figure 5 shows clearly how there is no equivalent increase in the concentration of local population over the same period. This further underlines that the sharp increase in bank branch clustering around the time of the introduction of credit registries and bureaus indeed reflects these sudden information-regime changes rather than more gradual demographic shifts.

[Insert Figure 5 here]

<u>Controlling for other country-level reforms</u>

Our sample countries went through a process of economic and political transformation after the fall of the Berlin Wall in 1989. One may hence worry that our information-sharing treatment partially picks up other reforms as well. We note though that much of the structural reform agenda was heavily concentrated in the first decade of transition (EBRD, 2013) – that is, *before* countries introduced information sharing (see Section 3.1 and also Table A3).

To address this concern more formally, we include four additional interactions with the locality-level number of pre-existing bank branches. This allows us to control for key reform dimensions that might confound our estimates of the impact of new information-sharing regimes. The first variable we interact with is a dummy that equals one if a country is a member of the European Union in a particular year and equals zero otherwise. According to the principle of single authorization, a bank authorized to operate in any one EU country can provide its services throughout the whole Single Market. Acceding the EU may therefore expose a country to foreign-bank entry and an associated change in branch clustering dynamics. The countries in our sample were on average an EU member for 17 percent of the sample period (Table 1).

Second, we control for countries' progress with setting up effective competition policies. We take the EBRD Transition Indicator for competition policy, which can range between 1 ("no competition legislation and institutions") and 4^+ ("Standards and performance typical of advanced industrial economies: effective enforcement of competition policy; unrestricted entry to most markets"). The average score on this measure across the years and countries in our sample is 2.2 (Table 1). Enhanced competition policy may change the clustering of real economic activity and, via demand effects, eventually influence the clustering of the supply of financial services.

Third, we interact with the EBRD Transition Indicator for small-scale privatization. This indicator also ranges between 1 ("Little progress") and 4+ ("Standards and performance typical of advanced industrial economies: no state ownership of small enterprises; effective tradability of land"). This indicator averages 3.7 across the countries and years in our data set (Table 1). Progress with small-scale privatization makes lending to small businesses more attractive and can therefore have an independent impact on the branching decisions of commercial banks.

Fourth, we employ a measure of how pro-competitive bank regulation is. This measure, taken from the IMF Financial Reforms Database, can range between 0 and 3 and averages 2.8 in our data set (Table 1). It measures whether the government allows the entry of new domestic banks; whether there are restrictions on bank branching; and whether the government allows banks to engage in a wide range of activities. This variable therefore provides a direct gauge of whether governments constrain banks' branching decisions in a top-down manner. Where and when such constraints bind less, it is easier for banks to optimize their branching decisions, including in response to the introduction of new information-sharing regimes.

In Appendix Table A8, we first add these additional interaction terms one-by-one (columns 1-4 and 6-9) and then all at the same time (columns 5 and 10). Throughout all specifications, the interaction between information sharing and the pre-existing locality-level number of branches remains statistically significant at the 1 percent level. Moreover, controlling for the impact of EU membership, competition policy, and progress with small-scale privatization hardly makes a dent in the magnitude of the coefficient. Interestingly, this differs in columns 4-5 and 9-10, where we (also) control for the state of bank regulation (of which restrictions on bank branching is a key component). When we add an additional interaction term with this variable, our coefficient of interest declines by about two-thirds. This indicates that changes in top-down restrictions on banks' branching decisions had an important impact on local clustering equilibria too. However, even when controlling for this, we find large effects of the introduction of information sharing. A one standard deviation higher number of pre-existing bank branches in a locality increases the probability that a bank opens an additional branch by 17 percent after the introduction of information sharing. Likewise, a one standard deviation higher number of own bank branches in a locality, reduces the likelihood of another branch opening by the same bank by more than 2 percent.

Placebo test

We conduct a placebo test in which, within each treatment event, we randomize the countries that introduced information sharing. For example, the 2001 treatment event consists of all countries that introduced information sharing in 2001 (the real treated) as well as all countries that did not have or introduce information sharing in the six-year window around 2001 (the clean controls). Suppose the number of real treated countries in 2001 is three. We then randomly pick three placebo treatment countries from the set of all real treated and clean controls in the 2001 event sample. We do this for each event, stack the resulting randomized event samples, and rerun our baseline regressions (columns 4 and 8 of Table 2) to estimate the coefficient of our interaction terms of interest. We repeat this process 500 times and plot the distribution of the point estimates for these placebo treatments in Figure 6. The top (bottom) panel shows the estimates related to column 4 (8) of Table 2. The vertical red lines indicate the 95^{th} percentile of this distribution. Reassuringly, we find that in both panels the real coefficient estimate from Table 2 (0.395 for the top figure and -0.125 for the bottom figure) lies outside the corresponding distribution of the placebo treatment coefficients.

[Insert Figure 6 here]

5.5. Information sharing and geographical credit rationing

An important model prediction that we have not yet been testing with our branch-level data is that the introduction of information sharing reduces spatial credit rationing: firms will be able to borrow from more distant bank branches. To test this prediction empirically, we merge our branch data with information from the Kompass database on firm-bank relationships. We then assume that firms borrow from the nearest branch of their primary bank and use this nearest distance as the *Firm-branch distance* in kilometers.

Of all countries in Kompass, there are four that introduced information sharing between 2000 and 2005 and that are also included in our BEPS data: the Czech Republic, Estonia, Latvia, and Poland. Because the bank information in Kompass and in BEPS can only be matched poorly for Estonia and Latvia, we focus on the Czech Republic and Poland. These countries introduced information sharing in 2002 and 2001, respectively. We also include two countries that did not introduce information sharing between 2000 and 2005. There are four such BEPS countries (Croatia, Hungary, Slovak Republic, and Ukraine) but because the matching of bank information is very poor for the Slovak Republic and Ukraine, we focus on the first two. We thus compare the change in firm-branch distance between 2000 and 2005 in two countries that introduced information sharing during this period (Czech Republic and Poland) with the change in firm-branch distance in two similar countries that did not (Croatia and Hungary). The final merged data set contains 9,348 and 4,960 firm records in 2000 and 2005, respectively, across these four countries.

The upper panel of Table 6 shows summary statistics and a two-sample t-test with unequal variances. In the countries that introduced information sharing between 2000 and 2005, firms on average borrow from more distant bank branches in 2005 than in 2000 (2 km and 8 km further for the Czech Republic and Poland, respectively). In contrast, firms do not borrow from more distant branches in the comparator countries that did not introduce information sharing during this period (Croatia and Hungary). We analyze this more formally in a difference-in-differences regression framework (lower panel of Table 6). We cluster standard errors by country using the wild cluster bootstrap-t to account for the small number of clusters (Cameron, Gelbach, and Miller, 2008). Column 1 shows that after the introduction of information sharing, firms borrow from branches that are around 15 km further away as compared with firms in countries that did not introduce information sharing during the same period.[27]

If the sharing of hard information reduces geographical credit rationing, allowing firms to borrow from more distant bank branches, then we expect this to be particularly important for firms that are more opaque. For these firms, information asymmetries are more of an issue and the new publicly available information may therefore have more 'bite'. To test whether this is the case, we use the Kompass data to construct three dummy variables that proxy for a firm's opaqueness. We then use these opaqueness proxies to construct triple interaction terms with *Information sharing*. Each model is fully saturated with additional (unreported) interaction terms between the country and year fixed effects and the respective opaqueness proxy.

Columns 2, 3, and 4 of Table 6 present the results. We find that the effect of information sharing on the reduction in spatial credit rationing is about twice as large for relatively opaque firms than for more transparent firms. For instance, while the average effect of information

[27] We compare the average distance between firms and their lender for a cross-section of firms in 2000 with that of a cross-section of firms in 2005. This average distance can increase faster in countries that introduce information sharing because existing borrowers switch to a new, more distant lender or because previously credit rationed borrowers now have access to a larger variety of (more remote) lenders.

sharing is an increase in the firm-bank distance of 15.1 km (column 1), column 2 shows that this effect is 19.2 km for more opaque firms (here proxied as those without an email address) and only 11.3 km for less opaque firms (with an email address). Because of these differential impacts, opaque and less opaque firms partially converge in terms of the geographical radius within which they can successfully seek out attractive borrowing opportunities.

[Insert Table 6 here]

6. Concluding remarks

This paper provides a stylized spatial model that helps to clarify how information sharing between banks influences branch clustering. Using this theoretical framework, we derive key predictions and test these by exploiting dynamic information on the geographical locations of bank branches. We use the introduction of information sharing regimes across a large number of countries, at different points in time, as plausibly exogenous shocks that shift the relative advantages and disadvantages of branch clustering.

In line with our theoretical priors, we show that information sharing has a positive impact on bank clustering and that this impact is stronger for relationship banks. We also find that after the introduction of information sharing, banks are more likely to establish new branches in localities where they themselves did not yet have a branch presence. Importantly, new bank openings in already densely banked areas go hand-in-hand with closures in sparsely banked localities. Due to these spatial shifts, we observe that bank branches start to agglomerate more within individual NUTS 1 regions. Lastly, we provide evidence indicating that due to these changes the average firm is able to borrow from more distant bank branches.

In summary, information sharing makes it more important for banks to move closer to each other than to be closer to potential clients. We expect these findings to apply also in other emerging markets experiencing structural change (especially those with a Communist legacy such as China) as well as in richer countries that are restructuring in the wake of the Covid-19 pandemic or are transitioning towards a greener production structure. In all these cases, new information sharing regimes may play a role in reshaping local banking markets. Moreover, our findings may shed light on how bank branching will respond to the increasing availability of alternative data on loan applicants, such as information on rent and mobile phone payments and on social media usage. To the extent that such new types of hard information become centrally available to all lenders, banks may respond in a similar way as they did when credit registries and bureaus made borrower information more broadly available: their branches may increasingly cluster together and further away from their clients.

An important implication of our results is that banking markets become more homogeneous in terms of composition – as they are served by the same banks that now operate across the country – but less homogeneous in terms of size. While the public availability of hard information leads to further clustering of bank branches in well-served locations, other (smaller and more rural) locations may lose out as access to credit deteriorates. Assessing the real-economic impacts of such spatial variation in access to credit due to information sharing is a promising avenue for further research. In particular, it would be of interest to study empirically how the agglomeration of bank branches may contribute to the widening of regional disparities. While older rural firms with an established operational and borrowing track record may benefit from information sharing (as they can now be screened by more distant bank branches), younger firms (without a track record) may lose out when branches leave remote areas. To the extent that young firms create most jobs (Haltiwanger, Jarmin and Miranda, 2013) the emergence of banking deserts may hence further reduce economic dynamism in rural areas.

Another interesting area for further research relates to the possible longer-term implications of mandatory information sharing, such as changes in banks' incentives to invest in the production of new borrower information and the impact of borrower switching. These are phenomena we cannot observe in our current data but that we believe to be worthy of future empirical investigation.

References

Abiad, A., and Mody, A. (2005). Financial reform: What shakes it? What shapes it? *American Economic Review, 95*(1), 66-88.

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

Agarwal, S., and Hauswald, R. (2010). Distance and private information in lending. *Review of Financial Studies, 23*(7), 2757-2788.

Barros, P.P. (1999). Multimarket competition in banking, with an example from the Portuguese market. *International Journal of Industrial Organization, 17*(3), 335-352.

Beck, T., Degryse, H., De Haas, R., and Van Horen, N. (2018). When arm's length is too far: Relationship banking over the credit cycle. *Journal of Financial Economics, 127*(1), 174-196.

Beck, T., Ioannidou, V., and Schäfer, L. (2018). Foreigners vs. natives: Bank lending technologies and loan pricing. *Management Science, 64*(8), 3792-3820.

Beck, T., Levine, R., and Levkov, A. (2010). Big bad banks? The winners and losers from bank deregulation in the United States. *Journal of Finance, 65*(5), 1637-1667.

Beck, T., Ongena, S., and Şendeniz-Yüncü, İ. (2019). Keep walking? Geographical proximity, religion, and relationship banking. *Journal of Corporate Finance, 55*, 49-68.

Benfratello, L., Schiantarelli, F., and Sembenelli, A. (2008). Banks and innovation: Microeconometric evidence on Italian firms. *Journal of Financial Economics, 90*(2), 197-217.

Bennardo, A., Pagano, M., and Piccolo, S. (2015). Multiple bank lending, creditor rights, and information sharing. *Review of Finance, 19*(2), 519-570.

Berger, A.N., L.F. Klapper and G.F. Udell (2001). The ability of banks to lend to informationally opaque small businesses. *Journal of Banking and Finance, 25*(12), 2127-2167.

Berger, A.N., N.H Miller, M.A. Petersen, R.G. Rajan and J.C. Stein (2005). Does function follow organizational form? Evidence from the lending practices of large and small banks. *Journal of Financial Economics, 76*, 237-269.

Berger, A.N. and Udell, G.F. (1995). Relationship lending and lines of credit in small firm finance. *Journal of Business, 68*(3), 351-381.

Berger, A.N. and Udell, G.F. (2002). Small business credit availability and relationship lending: The importance of bank organisational structure. *Economic Journal*, 112(477), F32-F53.

Berger, A.N., and Udell, G.F. (2006). A more complete conceptual framework for SME finance. *Journal of Banking and Finance, 30*(11), 2945-2966.

Berkowitz, D., Hoekstra, M., and Schoors, K. (2014). Bank privatization, finance, and growth. *Journal of Development Economics, 110*, 93-106.

Bircan, Ç., and De Haas, R. (2020). The limits of lending? Banks and technology adoption across Russia. *Review of Financial Studies, 33*(2), 536-609.

Bonfim, D., Nogueira, G. and Ongena, S. (2021). "Sorry, we're closed". Bank branch closures, loan pricing, and information asymmetries. *Review of Finance, 25*(4), 1211-1259.

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

Bouckaert, J., and Degryse, H. (2004). Softening competition by inducing switching in credit markets. *Journal of Industrial Economics, 52*(1), 27-52.

Bouckaert, J., and Degryse, H. (2006). Entry and strategic information display in credit markets. *Economic Journal, 116*(513), 702-720.

Brown, M., Guin, B., and Kirschenmann, K. (2016). Microfinance banks and financial inclusion. *Review of Finance, 20*(3), 907-946.

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

Callaway, B., and Sant'Anna, P.H.C (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics, 225*(2), 200-230.

Cameron, A. Colin, J.B. Gelbach, and D.L. Miller (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics, 90*(3): 414-427.

Carbo-Valverde, S., Hannan, T.H., and Rodriguez-Fernandez, F. (2011). Exploiting old customers and attracting new ones: The case of bank deposit pricing. *European Economic Review, 55*(7), 903-915.

Célérier, C. and A. Matray (2019). Bank-branch supply, financial inclusion, and wealth accumulation. *Review of Financial Studies, 32*(12), 4767-4809.

Cengiz, D., Dube, A., Lindner, A., and Zipperer, B. (2019). The effect of minimum wages on low-wage jobs. *Quarterly Journal of Economics*, *134*(3), 1405-1454.

Cerasi, V., Chizzolini, B., and Ivaldi, M. (2002). Branching and competition in the European banking industry. *Applied Economics, 34*(17), 2213-2225.

Chang, A., Chaudhuri, S., and Jayaratne, J. (1997). Rational herding and the spatial clustering of bank branches: an empirical analysis. *Columbia University Working Paper.*

Center for International Earth Science Information Network (CIESIN). Columbia University. *Gridded Population of the World*, Version 4 (GPWv4): Population Density, Revision 11. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H49C6VHW. Accessed 09/04/2021.

Coccorese, P. (2012). Banks as 'fat cats': Branching and price decisions in a two-stage model of competition. *Journal of Economics and Business, 64*(5), 338-363.

Cohen, A., and Mazzeo, M. J. (2010). Investment strategies and market structure: An empirical analysis of bank branching decisions. *Journal of Financial Services Research, 38*(1), 1-21.

Cole, R.A., L.G. Goldberg and L.J. White (2004). Cookie cutter vs. character: The micro structure of small business lending by large and small banks. *Journal of Financial and Quantitative Analysis*, *39*, 227-251.

Degryse, H., and Ongena, S. (2005). Distance, lending relationships, and competition. *Journal of Finance*, *60*(1), 231-266.

De Haas, R., Korniyenko, Y., Pivovarsky, A. and Tsankova, T. (2015). Taming the herd? foreign banks, the Vienna Initiative, and crisis transmission. *Journal of Financial Intermediation*, *24*(3), 325-355.

Dell'Ariccia, G. (2001). Asymmetric information and the structure of the banking industry. *European Economic Review*, *45*(10), 1957-1980.

Deller, S., and Sundaram-Stukel, R. (2012). Spatial patterns in the location decisions of US credit unions. *Annals of Regional Science*, *49*(2), 417-445.

Drechsler, I., A. Savov, and Schnabl, P. (2017). The deposits channel of monetary policy. *Quarterly Journal of Economics*, *132*(4), 1819-1876.

Drechsler, I., A. Savov, and Schnabl, P. (2021). Banking on deposits: Maturity transformation without interest rate risk, *Journal of Finance, 76*(3), 1091-1143.

Egan, M., S. Lewellen, and Sunderam, A. (2017). The cross section of bank value, NBER Working Paper No. 23291, National Bureau of Economic Research.

EBRD (1999). *Transition Report 1999: Ten Years of Transition*, European Bank for Reconstruction and Development, London.

EBRD (2013). *Transition Report 2013: Stuck in Transition?*, European Bank for Reconstruction and Development, London.

Engelbrecht-Wiggans, R., P. Milgrom, and Weberz, R. (1983). Competitive bidding with proprietary information. *Journal of Mathematical Economics*, 11, 161-169.

Estrin, S. (1998). Privatization and restructuring in Central and Eastern Europe. In: Boone, P., Gomulka, S. and Layard, R. (eds.) *Emerging From Communism: Lessons From Russia, China and Eastern Europe.* MIT Press, Cambridge, MA, 73-98.

Favara, G., and Imbs, J. (2015). Credit supply and the price of housing. *American Economic Review*, *105*(3), 958-92.

Giannetti, M., and Ongena, S. (2012). "Lending by example": Direct and indirect effects of foreign banks in emerging markets. *Journal of International Economics, 86*(1), 167-180.

Goodman-Bacon, A. (2018). Difference-in-differences with variation in treatment timing. *NBER Working Paper No. 25018*.

Granja, J., C. Leuz, and Rajan, R.G. (2021). Going the extra mile: Distant lending and credit cycles. *Journal of Finance*, forthcoming.

Guiso, L., Sapienza, P., and Zingales, L. (2004). Does local financial development matter?. *Quarterly Journal of Economics*, *119*(3), 929-969.

Gupta, R. C., and Tao, H. (2010). A generalized correlated binomial distribution with application in multiple testing problems. *Metrika*, *71*(1), 59.

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

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

Hauswald, R., and Marquez, R. (2006). Competition and strategic information acquisition in credit markets. *Review of Financial Studies, 19*(3), 967-1000.

Herrera, A. M., and Minetti, R. (2007). Informed finance and technological change: Evidence from credit relationships. *Journal of Financial Economics, 83*(1), 223-269.

Hotelling, H. (1929). Stability in competition. *Economic Journal, 39*(153), 41-57.

Iyer, R., Khwaja, A.I., Luttmer, E.F., and Shue, K. (2016). Screening peers softly: Inferring the quality of small borrowers. *Management Science*, 62(6), 1554-1577.

Jappelli, T., and Pagano, M. (2002). Information sharing, lending and defaults: Cross-country evidence. *Journal of Banking and Finance*, *26*(10), 2017-2045.

Jayaratne, J., and Strahan, P.E. (1996). The finance-growth nexus: Evidence from bank branch deregulation. *Quarterly Journal of Economics, 111*(3), 639-670.

Kalemli-Özcan, S., Laeven, L., and Moreno, D. (2018). Debt overhang, rollover risk, and corporate investment: Evidence from the European crisis. CEPR Discussion Paper No. 12881, Centre for Economic Policy Research, London.

Karapetyan, A. and Stacescu, B. (2014). Information sharing and information acquisition in credit markets. Review of Finance, *18*(4), 1583-1615.

Kim, M., Lozano-Vivas, A., and Morales, A.J. (2007). Multistrategic spatial competition in banking. *International Journal of the Economics of Business, 14*(3), 351-366.

Konishi, H. (2005). Concentration of competing retail stores. *Journal of Urban Economics, 58*(3), 488-512.

Kroszner, R. S., and Strahan, P.E. (2014). Regulation and deregulation of the US banking industry: causes, consequences, and implications for the future. In: *Economic regulation and its reform: what have we learned?* (pp. 485-543). University of Chicago Press.

Kysucky, V. and Norden, L. (2016). The benefits of relationship lending in a cross-country contyext: A meta-analysis. *Management Science 61*(2), i-vii, 1-301.

Lee, N. and Luca, D. (2019). The big-city bias in access to finance: Evidence from firm perceptions in almost 100 countries. *Journal of Economic Geography*, *19*(1), 199-224

Levine, R. (2005). Finance and growth: Theory and evidence, In: *Handbook of Economic Growth*, P. Aghion and S.N. Durlauf (Eds.), Volume 1 (Part A), Chapter 12, 865-934.

Levine, R., C. Lin, and Wang, Z. (2020). Bank networks and acquisitions. *Management Science, 66*(11), 4921-5484.

Liberti, J.M. and Petersen, M.A. (2019). Information: Hard and soft. *Review of Corporate Finance Studies*, 8(1), 1-41.

Mian, A. (2006). Distance constraints: The limits of foreign lending in poor economies. *Journal of Finance*, *61*(3), 1465-1505.

Morgan, D. P., Pinkovsky, M. L., and Yang, B. (2016). Banking deserts, branch closings, and soft information. *Federal Reserve Bank of New York Working Paper*.

Nguyen, H. L. Q. (2019). Are credit markets still local? Evidence from bank branch closings. *American Economic Journal: Applied Economics*, *11*(1), 1-32.

Ongena, S., and Şendeniz-Yüncü, İ. (2011). Which firms engage small, foreign, or state banks? And who goes Islamic? Evidence from Turkey. *Journal of Banking and Finance, 35*(12), 3213-3224.

Padilla, A. J., and Pagano, M. (1997). Endogenous communication among lenders and entrepreneurial incentives. *Review of Financial Studies, 10*(1), 205-236.

Padilla, A. J., and Pagano, M. (2000). Sharing default information as a borrower discipline device. *European Economic Review, 44*(10), 1951-1980.

Pagano, M., and Jappelli, T. (1993). Information sharing in credit markets. *Journal of Finance, 48*(5), 1693-1718.

Park, K., and Pennacchi, G. (2008). Harming depositors and helping borrowers: The disparate impact of bank consolidation. *Review of Financial Studies, 22*(1), 1-40.

Peria, M. S. M., and Singh, S. (2014). The impact of credit information sharing reforms on firm financing. *World Bank Working Paper*.

Petersen, M. A., and Rajan, R.G. (2002). Does distance still matter? The information revolution in small business lending. *Journal of Finance, 57*(6), 2533-2570.

Pollard, J.S. (2003). Small firm finance and economic geography. *Journal of Economic Geography, 3*(4), 429-452.

Qi, S., De Haas, R., Ongena, S. and Straetmans, S. (2021). Move a little closer? Information sharing and the spatial clustering of bank branches. CEPR Discussion Paper 15829, Centre for Economic Policy Research, London.

Qian, J., Strahan, P.E., and Yang, Z. (2015). The impact of incentives and communication costs on information production and use: Evidence from bank lending, *Journal of Finance*, *70*(4), 1457-1493.

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

Rambachan, A. and Roth, J. (2021). An honest approach to parallel trends. Mimeo.

Rice, T., and Strahan, P.E. (2010). Does credit competition affect small-firm finance? *Journal of Finance*, *65*(3), 861-889.

Robinson, J. (1952). *The Rate of Interest and Other Essays*, MacMillan, London.

Roland, G. (2000). *Transition and Economics: Politics, Markets and Firms*, MIT Press.

Salop, S.C. (1979). Monopolistic competition with outside goods. *Bell Journal of Economics, 10*(1), 141-156.

Stein, J.C. (2002). Information production and capital allocation: Decentralized versus hierarchical firms. *Journal of Finance, 57*(5), 1891-1921.

Svejnar, J. (2002). Transition economies: Performance and challenges. *Journal of Economic Perspectives*, 16(1), 3-28.

Syverson, C. (2004). Market structure and productivity: A concrete example. *Journal of Political Economy, 112*(6), 1181-1222.

Temesvary, J. (2015). Dynamic branching and interest rate competition of commercial banks: Evidence from Hungary. *International Journal of Industrial Organization, 43*, 98-110.

Van Cayseele, P., Bouckaert, J., and Degryse, H. (1995). Credit market structure and information sharing mechanisms. In: *Market evolution* (pp. 129-143). Springer, Dordrecht.

von Thadden, E.L. (1995). Long-term contracts, short-term investment and monitoring, *Review of Economic Studies*, *62*(4), 557-575.

von Thadden, E-L. (2004). Asymmetric information, bank lending and implicit contracts: The winner's curse, *Finance Research Letters, 1*(1), 11-23.

Wang, T. (2019). To build or to buy? The role of local information in credit market development. *Management Science*, *65*(12), 5449-5956.

Wooldridge, J.M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

Zhao, T. and Jones-Evans, D. (2017). SMEs, banks and the spatial differentiation of access to finance. *Journal of Economic Geography*, *17*(4), 791-824

Table 1

Summary Statistics

This table provides the number of observations, mean, median, standard deviation, minimum and maximum for all variables used in the analysis.

Variable	**Obs.**	**Mean**	**Median**	**St. Dev.**	**Min.**	**Max.**
*Dependent variables (bank*locality*year level)*						
New branch opening	833,916	0.040	0	0.195	0	1
Net branch opening	833,916	0.039	0	0.195	0	1
No. branches other banks (log)	833,916	1.611	1.386	1.561	0	7.290
No. branches own bank (log)	833,916	0.428	0.693	0.537	0	5.347
New branch opening in top quartile	833,916	0.025	0	0.157	0	1
New branch closure in bottom quartile	833,916	0.003	0	0.056	0	1
No. branches other banks per 1,000 population (log)	200,274	0.136	0.118	0.097	0	0.615
No. branches own bank per 1,000 population (log)	200,274	0.007	0	0.017	0	0.239
No. bank branches per 1,000 population (log)	200,274	0.142	0.121	0.100	0	0.625
Population growth rate	165,948	-0.232	0	1.306	-34	15.158
Night-time light growth rate	830,636	0.060	-0.006	0.639	-1	108.500
*Country characteristics (country*year level)*						
Information sharing	342	0.532	1	0.500	0	1
Quality information sharing	342	1.289	0	2.132	0	6
EU membership	342	0.167	0	0.373	0	1
Competition policy	342	2.226	2.33	0.699	1	3.67
Small-scale privatisation	342	3.701	4	0.683	1	4.33
Pro-competition bank regulation	121	2.802	3	0.641	0	3
Bank characteristics (bank level)						
Small bank	614	0.318	0	0.466	0	1
Domestic bank	614	0.430	0	0.495	0	1
Relationship bank	316	0.592	1	0.492	0	1
Bank characteristics (branch level)						
Small bank	56,555	0.104	0	0.262	0	1
Domestic bank	56,555	0.505	1	0.500	0	1
Relationship bank	38,439	0.446	0	0.497	0	1
Firm characteristics (firm level)						
Firm-branch distance	14,308	15.447	1.809	45.266	0.010	443.515
Has email address	14,308	0.602	1	0.489	0	1
Has tax number	14,308	0.736	1	0.441	0	1
Has formal opening hours	14,308	0.743	1	0.437	0	1

Table 2

Information Sharing and the Clustering of Bank Branches

This table reports linear probability regressions to estimate the impact of the introduction of information sharing on bank branch clustering. The dependent variable measures whether a bank opens a new branch in a locality in a year. Observations from all countries that introduce information sharing in the same calendar year are grouped together and combined with the observations from not (yet) treated (control) countries within a six-year window around the introduction year. Following Cengiz et al. (2019) these event-specific data sets are then stacked to estimate a single coefficient. Table A1 contains the definitions and Table 1 the summary statistics for all variables. *Country * Treatment Event* clustered robust *p* -values are reported in parentheses. ***, **, * correspond to the 1%, 5%, and 10% level of statistical significance, respectively.

Dependent variable →	New branch opening							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Information sharing * No. branches other banks	0.017***	0.334***	0.165**	0.395***				
	(0.000)	(0.003)	(0.014)	(0.002)				
No. branches other banks	0.002	0.030***	0.039***	0.662***				
	(0.221)	(0.002)	(0.002)	(0.000)				
Information sharing * No. branches own bank					-0.073**	-0.080***	-0.070***	-0.125***
					(0.043)	(0.000)	(0.000)	(0.002)
No. branches own bank					-0.012*	-0.002	0.007*	-0.389***
					(0.091)	(0.816)	(0.076)	(0.000)
Information sharing	-0.007				0.085*			
	(0.710)				(0.087)			
Locality * Year * Treatment Event Fixed Effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Bank * Year * Treatment Event Fixed Effects	No	No	Yes	Yes	No	No	Yes	Yes
Locality * Bank * Treatment Event Fixed Effects	No	No	No	Yes	No	No	No	Yes
R-squared	0.005	0.433	0.628	0.712	0.009	0.432	0.628	0.727
Observations	833,916	833,916	833,916	833,916	833,916	833,916	833,916	833,916

Table 3

Information Sharing, Relationship Lending, and the Clustering of Bank Branches

This table reports linear probability regressions to estimate the impact of the introduction of information sharing on bank branch clustering by relationship lenders as compared with transaction-based lenders. The dependent variable measures whether a bank opens a new branch in a locality in a year. Observations from all countries that introduce information sharing in the same calendar year are grouped together and combined with the observations from not (yet) treated (control) countries within a six-year window around the introduction year. Following Cengiz et al. (2019) these event-specific data sets are then stacked to estimate a single coefficient. Table A1 contains the definitions and Table 1 the summary statistics for all variables. *Country * Treatment Event* clustered robust *p* -values are reported in parentheses. ***, **, * correspond to the 1%, 5%, and 10% level of statistical significance, respectively.

Dependent variable →	New branch opening					
	(1)	(2)	(3)	(4)	(5)	(6)
Bank type →	Small banks		Domestic banks		Relationship banks	
Information sharing * No. branches other banks	0.167**	0.407***	0.164**	0.392***	0.128**	0.418***
	(0.012)	(0.001)	(0.014)	(0.002)	(0.013)	(0.000)
Information sharing * No. branches other banks * Bank type	0.012**	0.027***	0.002	0.005	0.005*	0.006**
	(0.015)	(0.002)	(0.607)	(0.262)	(0.096)	(0.048)
No. branches other banks * Bank type	-0.011***	0.005	-0.003***	-0.046***	-0.001	0.063***
	(0.002)	(0.442)	(0.002)	(0.000)	(0.342)	(0.005)
No. branches other banks	0.037***	0.663***	0.040***	0.696***	0.078***	0.697***
	(0.002)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
Locality * Year * Treatment Event Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Bank * Year * Treatment Event Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Locality * Bank * Treatment Event Fixed Effects	No	Yes	No	Yes	No	Yes
R-squared	0.628	0.712	0.628	0.712	0.703	0.767
Observations	833,916	833,916	833,916	833,916	592,383	592,383

Table 4

Quality of Information Sharing and the Clustering of Bank Branches

This table reports linear probability regressions to estimate the relationship between the quality of a country's information-sharing regime and bank branch clustering. The dependent variable measures whether a bank opens a new branch in a locality in a year. Columns 1-2 are based on all observations from all countries that introduce information sharing in the same calendar year. These are grouped together and then combined with the observations from not (yet) treated (control) countries within a six-year window around the introduction year. Following Cengiz et al. (2019) these event-specific data sets are then stacked to estimate a single coefficient. Columns 3-4 are based on only those countries and years in which information sharing is in place. Table A1 contains the definitions and Table 1 the summary statistics for all variables. *Country * Treatment Event* clustered robust *p*-values are reported in parentheses. ***, **, * correspond to the 1%, 5%, and 10% level of statistical significance, respectively.

Dependent variable →	New branch opening			
	(1)	(2)	(3)	(4)
Information sharing * No. branches other banks	0.157**	0.383***		
	(0.027)	(0.003)		
Quality information sharing * No. branches other banks	0.122***	0.180***	0.122***	0.273***
	(0.000)	(0.000)	(0.001)	(0.000)
No. branches other banks	0.039***	0.662***	0.196**	1.584***
	(0.002)	(0.000)	(0.022)	(0.000)
Locality * Year * Treatment Event Fixed Effects	Yes	Yes	Yes	Yes
Bank * Year * Treatment Event Fixed Effects	Yes	Yes	Yes	Yes
Locality * Bank * Treatment Event Fixed Effects	No	Yes	No	Yes
R-squared	0.628	0.712	0.652	0.791
Observations	833,916	833,916	81,240	81,240

Table 5

Information Sharing and the Clustering of Bank Branches: IV Results

This table reports IV regressions to estimate the impact of the introduction of information sharing on bank branch clustering. The dependent variable in the first stages (columns 1-2) is an interaction term between a dummy variable indicating whether in a given year and country an information-sharing system is in place and a locality-level measure of the number of pre-existing branches of other banks (column 1) or the bank itself (column 2). The instruments in these first stages are interaction terms between the percentage of neighboring countries that introduced information sharing in the previous five years and a locality-level measure of the number of pre-existing branches of other banks (column 1) or the bank itself (column 2). The dependent variable in the second stage (columns 2 and 3) measures whether a bank opens a new branch in a locality in a year. Observations from all countries that introduce information sharing in the same calendar year are grouped together and combined with the observations from not (yet) treated (control) countries within a six-year window around the introduction year. Following Cengiz et al. (2019) these event-specific data sets are then stacked to estimate a single coefficient. Table A1 contains the definitions and Table 1 the summary statistics for all variables. *Country * Treatment Event* clustered robust *p*-values are reported in parentheses. ***, **, * correspond to the 1%, 5%, and 10% level of statistical significance, respectively.

	First stage		**Second stage**	
Dependent variable →	Information sharing * No. branches other banks	Information sharing * No. branches own bank	New branch opening	
	(1)	(2)	(3)	(4)
% neighboring countries introduced information sharing	0.174***			
* No. branches other banks	(0.000)			
% neighboring countries introduced information sharing		0.202***		
* No. branches own bank		(0.000)		
Information sharing * No. branches other banks			1.174***	
			(0.000)	
No. branches other banks	0.008***		0.622***	
	(0.000)		(0.000)	
Information sharing * No. branches own bank				-0.422***
				(0.000)
No. branches own bank		0.102***		-0.343***
		(0.000)		(0.000)
F-Statistic	4,005	8,615		
Locality * Year * Treatment Event Fixed Effects	Yes	Yes	Yes	Yes
Bank * Year * Treatment Event Fixed Effects	Yes	Yes	Yes	Yes
Locality * Bank * Treatment Event Fixed Effects	Yes	Yes	Yes	Yes
Observations	833,916	833,916	833,916	833,916

Table 6

Information Sharing and Spatial Credit Rationing

This table reports, by country, summary statistics for the variable *Firm-branch distance* and regressions to estimate the impact of the introduction of information sharing on spatial credit rationing. All diff-in-diff-in-diff regressions in the lower panel are fully saturated with additional (unreported) interaction effects between the year and country dummies and the firm characteristics. Standard errors are clustered by country and p-values based on the wild cluster bootstrap-t, which accounts for the small number of clusters (Cameron, Gelbach and Miller, 2008) are reported in parentheses. ***, **, * correspond to the 1%, 5%, and 10% level of significance, respectively.

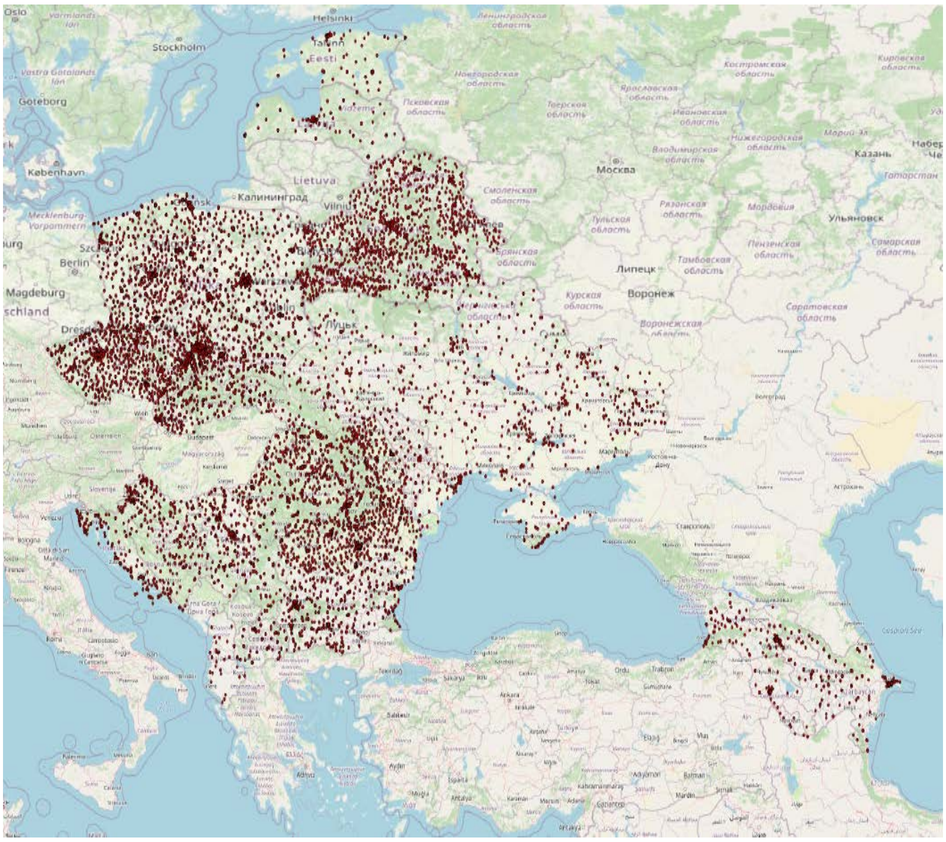
Dependent variable → Firm-branch distance (in km)											
Czech Republic (Introduced information sharing in 2002)						**Poland (Introduced information sharing in 2001)**					
	Obs.	Mean	St. Dev.	5%	95%		Obs.	Mean	St. Dev.	5%	95%
2000	1,650	3.01	5.16	2.76	3.26	2000	5,286	19.13	56.57	17.60	20.65
2005	1,892	5.01	14.02	4.38	5.64	2005	1,242	27.22	68.88	23.38	31.05
2005-2000	2.00***					2005-2000	8.09***				
Croatia (Introduced information sharing in 2007)						**Hungary (Introduced information sharing in 1995)**					
	Obs.	Mean	St. Dev.	5%	95%		Obs.	Mean	St. Dev.	5%	95%
2000	953	16.65	48.97	13.54	19.77	2000	1,459	24.08	34.51	22.31	25.85
2005	409	20.92	47.43	16.31	25.53	2005	1,417	8.54	13.65	7.83	9.25
2005-2000	4.26					2005-2000	-15.54***				

Difference-in-Differences (-in-Differences) regression				
	(1)	(2)	(3)	(4)
Information sharing	15.14***	19.15***	21.02***	19.48***
	(0.000)	(0.000)	(0.000)	(0.000)
Information sharing*Has email address		-7.89***		
		(0.000)		
Information sharing*Has tax number			-15.77***	
			(0.000)	
Information sharing*Has formal opening hours				-11.63***
				(0.000)
Year Fixed Effects	Yes	Yes	Yes	Yes
Country Fixed Effects	Yes	Yes	Yes	Yes
R-squared	0.027	0.030	0.030	0.029
Observations	14,308	14,308	14,308	14,308

Figure 1

Distribution of Localities with Bank Branches in 1995 and in 2012

Panel A. This map plots all localities in our data set with at least one bank branch in 1995



Panel B. This map plots all localities in our data set with at least one bank branch in 2012.

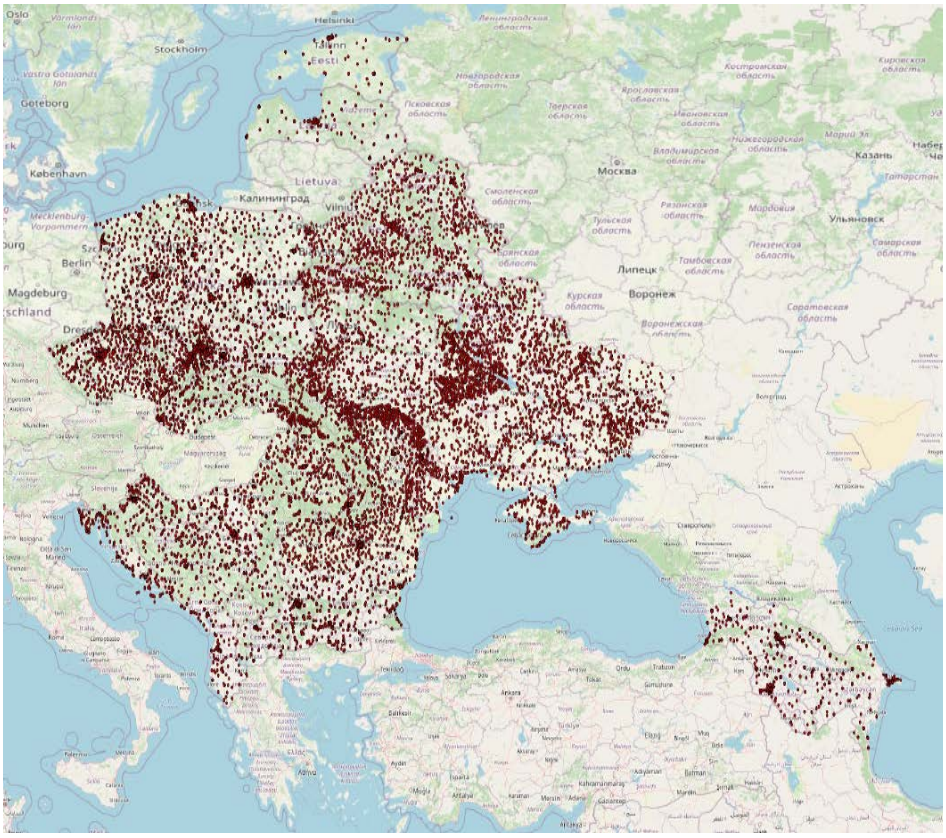
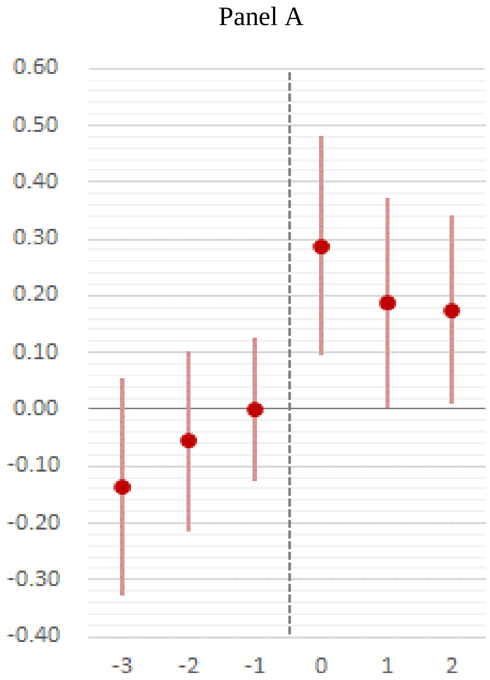


Figure 2

Information Sharing and the Clustering of Bank Branches: Event Study

This figure summarizes an event-study analysis in which a binary variable indicating whether a bank opens a new branch in a locality in a particular year is regressed on a set of year dummies around the introduction of information sharing in a country at $t=0$, each interacted with either the number of pre-existing branches of other banks in a locality (Panel A) or with the number of pre-existing branches of the bank itself (Panel B). All coefficients are based on specifications with the same interactive fixed effects and covariates as in column 4 (Panel A) and column 8 (Panel B) of Table 2.



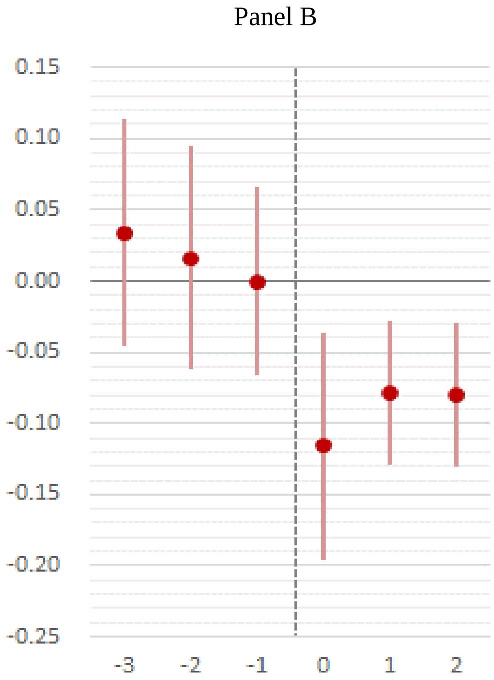


Figure 3

Sensitivity to Linear and Non-Linear Deviations from Parallel Trends

This figure shows confidence sets for the treatment parameters of interest – the number of pre-existing branches of other banks in a locality (Panel A) and the number of pre-existing branches of the bank itself (Panel B) – following Rambachan and Roth (2019). The original estimate is shown in blue and the estimates allowing for deviations from pre-trends – fixed length confidence intervals (FLCI) – in red.

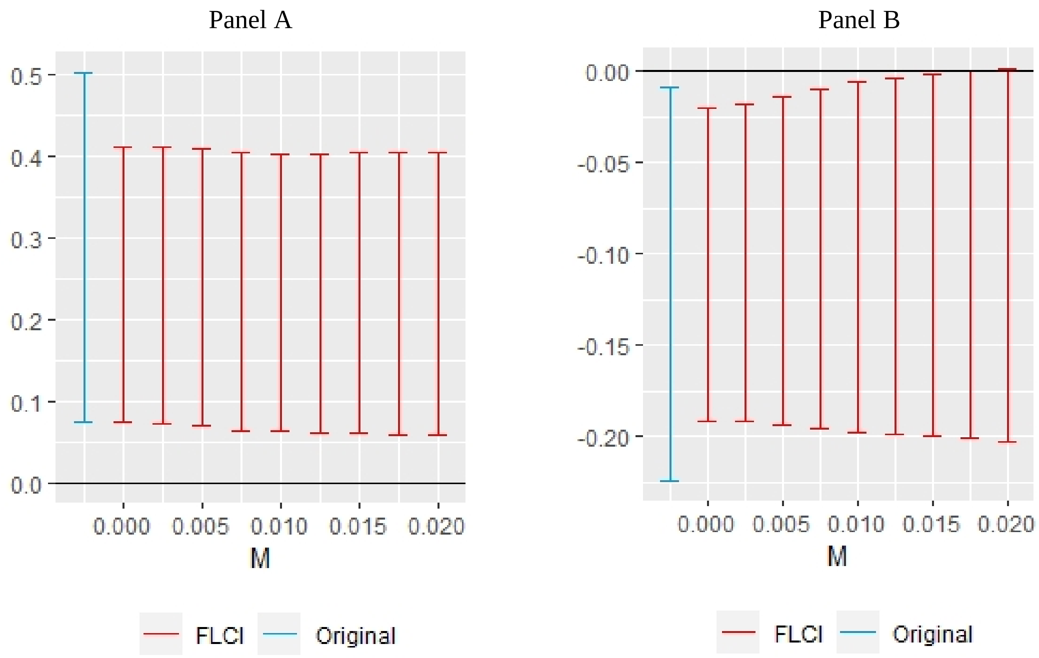


Figure 4

Population Density and Bank Branch Density Before and After the Introduction of Information Sharing

This binned scatter plot reflects spatial data at the 30x30 km grid level across 19 countries in Emerging Europe. For each grid cell, the local population density (number of inhabitants to the size of the cell) is calculated as well as the local bank branch density (number of bank branches to the size of the cell). The scatterplot groups all grid cells into 20 equal-sized bins, based on their population density, after which the mean population density (horizontal axis) and branch density (vertical axis) is plotted for each bin. This is done for the average in the years before (blue) and after (red) the introduction of information sharing. The lines reflect a linear OLS fit with 95% confidence intervals.

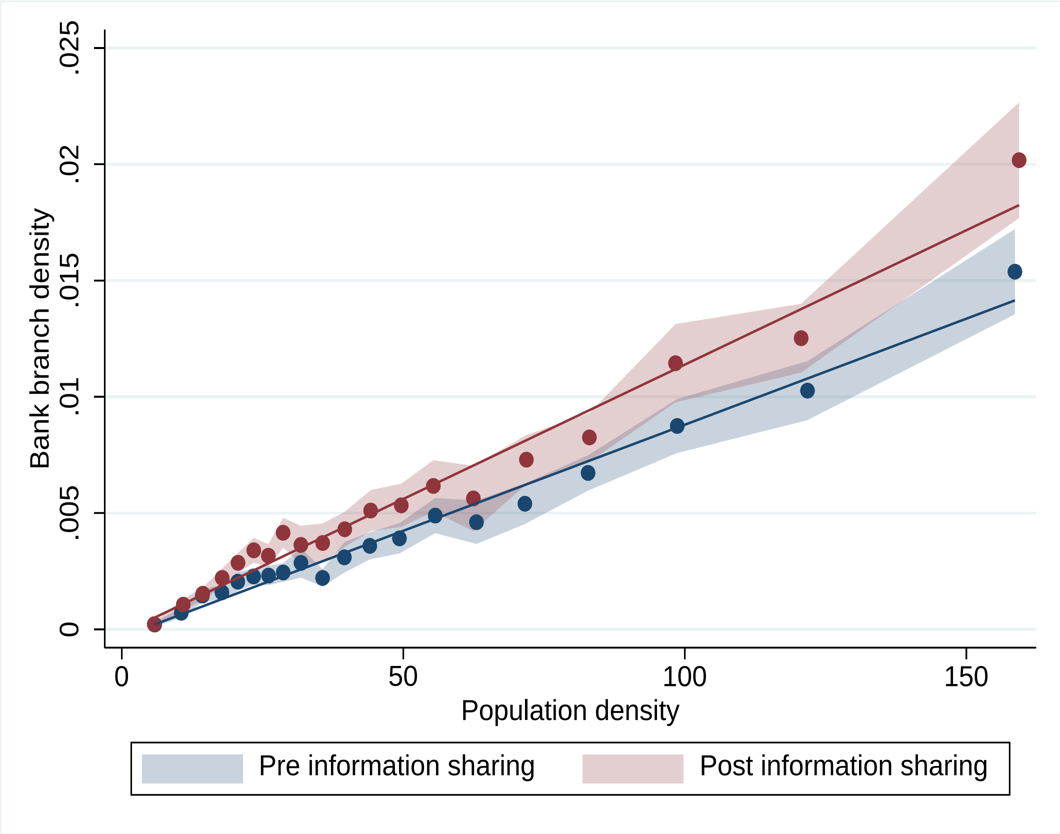
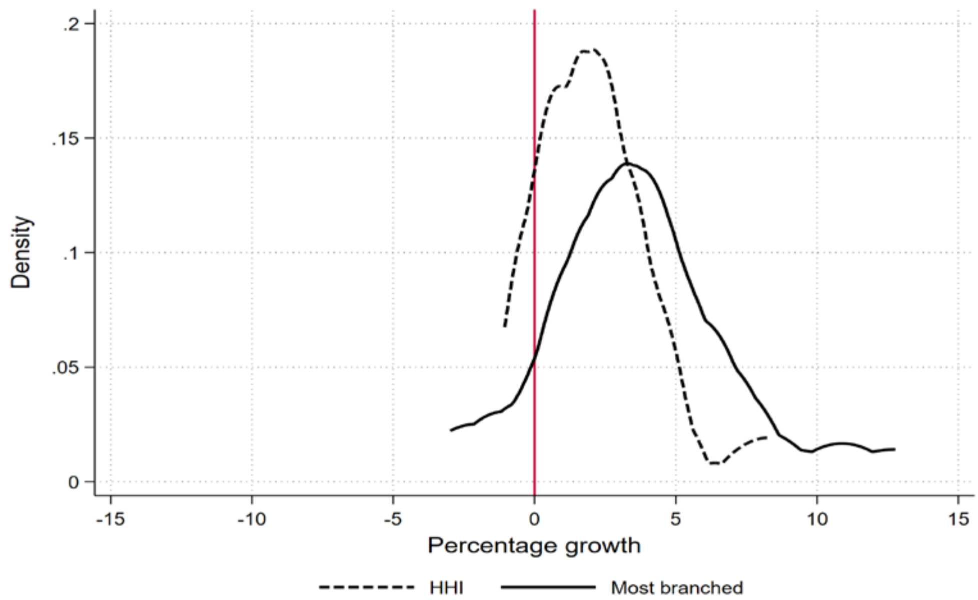


Figure 5

Information Sharing and the Clustering of Bank Branches at the NUTS 1 Level

Panel A summarizes the change in geo-spatial concentration of bank branches at the NUTS 1 level in the three years after as compared with the three years before the introduction of information sharing. Each country is divided into NUTS 1 'major socio-economic regions', each of which is further split into smaller NUTS 3 regions, using the official Eurostat Nomenclature of Territorial Units for Statistics (NUTS) hierarchy. For each NUTS 1 region, we calculate a Herfindahl-Hirschman Index (HHI) as a measure of spatial concentration of bank branches across the NUTS 3 sub-regions in that NUTS 1 region. We calculate this HHI three years before and three years after the introduction of information sharing in the respective country. The dashed kernel density plot shows the distribution of the percentage change in these NUTS 1 level HHI indices over this period. In a similar vein, the solid kernel density plot shows the distribution of the percentage change at the NUTS 1 level of the share of all branches clustered in the most densely banked NUTS 3 region. Panel B shows similarly the distribution of changes in the geo-spatial concentration of populations within NUTS 1 regions (data source: Eurostat).

Panel A



Panel B

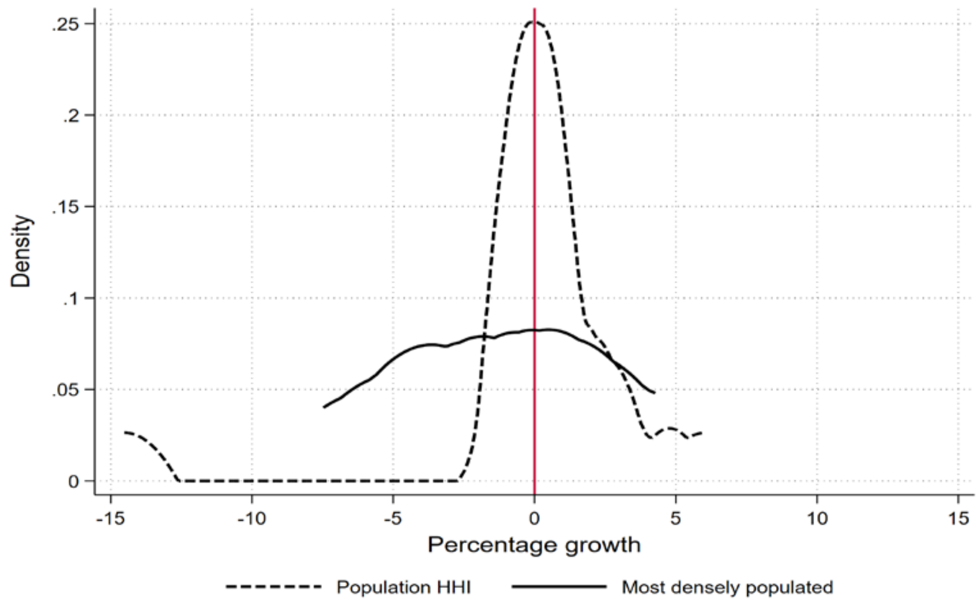


Figure 6

Information Sharing and the Clustering of Bank Branches: Placebo Treatments

These figures present the results of placebo tests in which, within each treatment event, the countries that introduced information sharing are randomized. Using that randomized sample, the baseline regressions (columns 4 and 8 of Table 2) are rerun to estimate the coefficient estimate of the interaction term of interest. We repeat this process 500 times and plot the distribution of the point estimates for these placebo treatments. The top (bottom) panel shows the estimates related to column 4 (8) of Table 2. The vertical red lines indicate the 95th percentile of this distribution. In both panels the real coefficient estimate from Table 2 (0.395 for the top figure and -0.125 for the bottom figure) is outside the corresponding distribution of the placebo treatment coefficients.

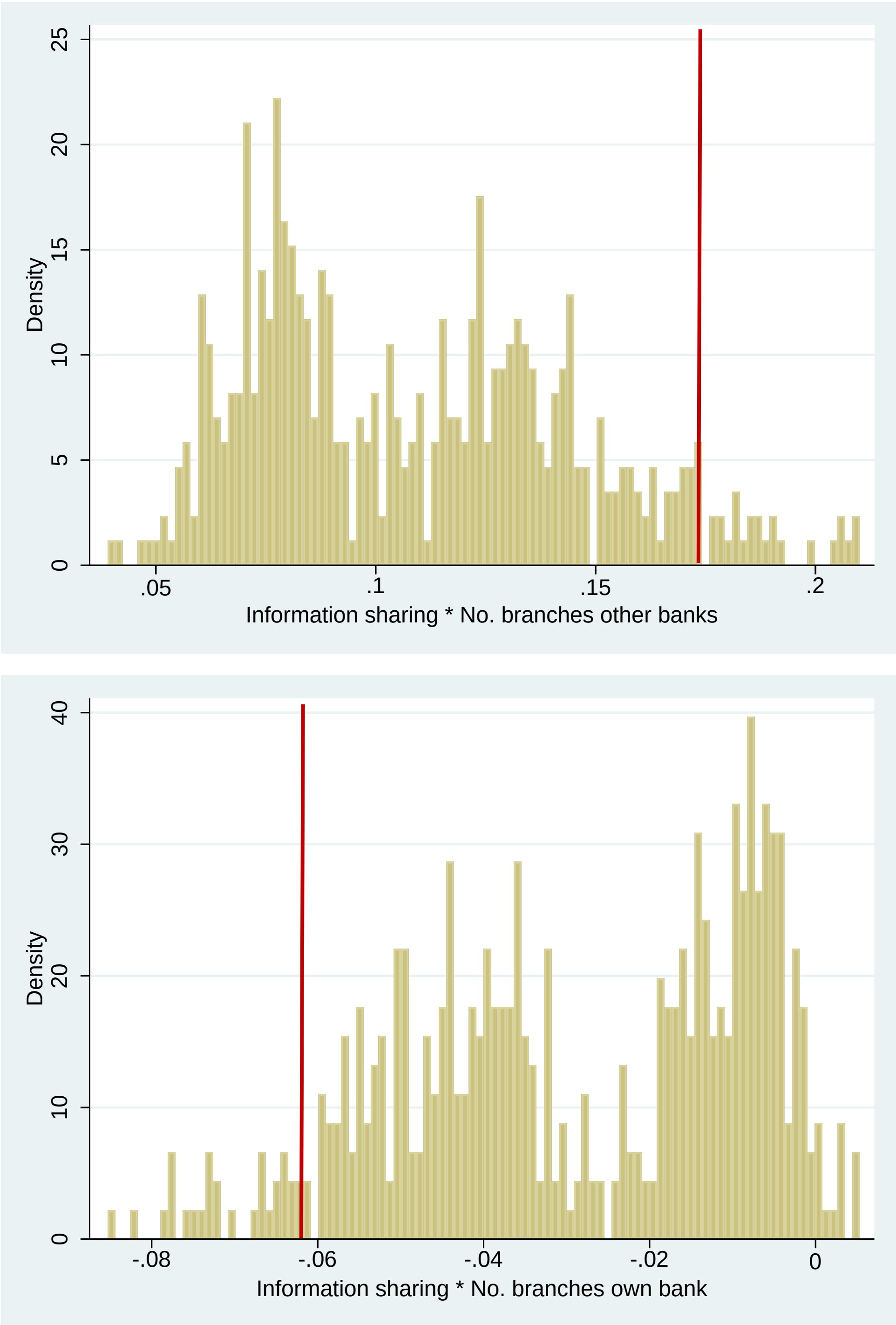


Table A1

Variable Definitions and Sources

This table provides the definition and data sources for all variables used in the analysis. BEPS II is the second round of the EBRD Banking Environment and Performance Survey (BEPS) which was conducted among 611 banks across 32 countries. "Doing Business" is the Doing Business Database by the World Bank. "Kompass" refers to the Kompass business directory. EBRD TI refers to the EBRD transition indicators. IMF FRD is the IMF Financial Reform Database. WCD is the World Cities Database.

	Definition	*Data Sources*
Dependent variables		
New branch opening	= 1 if there is any bank branch opening in a locality in a year, = 0 otherwise	BEPS II
Net branch opening	= 1 if the number of bank branch openings is larger than the number of bank branch closures in a locality in a year, = 0 otherwise	BEPS II
No. branches other banks (log)	Log number of existing branches of other banks within a locality in a year	BEPS II
No. branches own bank (log)	Log number of existing branches of the same bank within a locality in a year	BEPS II
No. branches other banks per 1,000 population (log)	Log number of existing branches of other banks within a locality per 1,000 inhabitants in a year	BEPS II, WCD
No. branches own bank per 1,000 population (log)	Log number of existing branches of the same bank within a locality per 1,000 inhabitants in a year	BEPS II, WCD
New branch opening in top quartile	= 1 if any branch opens in a locality belonging to the top 25% most densely banked localities in a country and year, = 0 otherwise	BEPS II
Independent variables		
Information sharing	= 1 if there is information sharing (credit registry and/or credit bureau) in the country in that year, = 0 otherwise	World Bank/EBRD
Quality information sharing	= 0 to 6, higher values indicate a higher quality of information sharing in the country in that year	Doing Business
Domestic bank	= 1 if more than 50% of a bank's shares are foreign-owned, = 0 otherwise	Bank websites
Relationship bank	= 1 if according to the bank CEO relationship lending is "very important" when providing credit to SMEs, = 0 otherwise	BEPS II
Small bank	= 1 if the no. branches of a bank is below the median no. branches operated by banks in a country and year, = 0 otherwise	BEPS II
Closed branch in bottom quartile	= 1 if any branch closes in a locality belonging to the bottom 25% most densely banked localities in a country and year, = 0 otherwise	BEPS II
No. bank branches per 1,000 population (log)	Log number of existing bank branches within a locality per 1,000 inhabitants in a year	BEPS II, WCD
Population growth rate	Annual percentage change in the size of the population in a locality	WCD
Night-time light growth rate	Annual percentage change in the locally emitted night-time light as measured by satellite imagery	NOAA/NASA
Firm-branch distance	Distance to the nearest branch of a firm's primary bank in km	Kompass
Has email address	= 1 if the firm has an email address, = 0 otherwise	Kompass
Has tax number	= 1 if the firm has a tax number, = 0 otherwise	Kompass
Has formal opening hours	= 1 if the firm has listed formal opening hours in Kompass, = 0 otherwise	Kompass
EU membership	= 1 if a country is part of the European Union in a particular year, = 0 otherwise	European Commission
Competition policy	= 1 to 4+, higher values indicate that a country has created more market-based competition policies and institutions	EBRD TI
Small-scale privatisation	= 1 to 4+, higher values indicate more progress of a country in terms of the privatisation of small- and medium-sized enterprises	EBRD TI
Pro-competition bank regulation	= 0 to 3, higher values indicate fewer entry barriers in the banking sector of a country in a given year	IMF FRD

Table A2

Overview of Branch Openings and Closures

This table provides an overview of the opening and closure of branches in our dataset by year (left) and by country (right).

Year	Opened branches	Closed branches	Country	Opened branches	Closed branches
1995	2,388	0	Albania	443	11
1996	489	0	Armenia	448	19
1997	546	0	Azerbaijan	335	13
1998	525	0	Belarus	2,481	9
1999	543	0	Bosnia & Herzegovina	617	10
2000	974	6	Bulgaria	1,405	100
2001	1,361	3	Croatia	608	48
2002	1,389	7	Czech Republic	382	19
2003	2,571	9	Estonia	60	56
2004	4,307	34	Georgia	703	108
2005	2,122	20	Latvia	195	9
2006	2,535	19	Moldova	1,300	180
2007	7,833	61	Montenegro	206	12
2008	1,753	92	North Macedonia	189	16
2009	548	199	Poland	3,192	51
2010	709	223	Romania	2,053	177
2011	1,060	201	Serbia	1,080	227
2012	274	191	Slovak Republic	153	0
			Ukraine	16,077	0
Total	*31,927*	*1,065*	*Total*	*31,927*	*1,065*

Table A3
Introduction of Information Sharing

This table provides an overview of the introduction years of public credit registries and private credit bureaus in our 19 sample countries. N.a.: No credit bureau or registry has as yet been introduced in this country. Source: World Bank Doing Business Database and EBRD.

Country	Public Credit Registry	Private Credit Bureau
Albania	2008	2009
Armenia	2003	2004
Azerbaijan	2005	n.a.
Belarus	2007	n.a.
Bosnia & Herzegovina	2006	2001
Bulgaria	1999	2005
Croatia	n.a.	2007
Czech Republic	2002	2002
Estonia	n.a.	2001
Georgia	n.a.	2005
Latvia	2003	n.a.
Moldova	n.a.	2011
Montenegro	2008	n.a.
North Macedonia	1998	2010
Poland	n.a.	2001
Romania	2000	2004
Serbia	2002	2004
Slovak Republic	1997	2004
Ukraine	n.a.	2007

Table A4

Information Sharing and the Clustering of Bank Branches: Net Branch Openings and Branch Openings per 1,000 Inhabitants

This table reports regressions to estimate the impact of the introduction of information sharing on bank branch clustering using the Cengiz et al. (2019) methodology to address the potential concern in staggered treatment timing. The dependent variable in columns (1)-(4) measures whether on a net basis, a bank increases its number of branches in a locality in a year (the number of newly opened branches exceeds the number of closed branches). In columns (5)-(8), the number of existing bank branches are normalized by the local population in 1,000 persons and the dependent variable measures whether a bank opens a new branch in a locality in a year. Table A1 contains all definitions and Table 1 the summary statistics for each variable. *Country * Treatment Event* clustered robust *p*-values are reported in parentheses. ***, **, * correspond to the 1%, 5%, and 10% level of significance, respectively.

Dependent variable →	Net branch opening				New branch opening			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Information sharing * No. branches other banks	0.165**	0.395***						
	(0.014)	(0.002)						
No. branches other banks	0.039***	0.663***						
	(0.002)	(0.000)						
Information sharing * No. branches own bank			-0.070***	-0.125***				
			(0.000)	(0.002)				
No. branches own bank			0.007*	-0.389***				
			(0.090)	(0.000)				
Information sharing * No. branches other banks per 1,000 population					1.477**	3.241**		
					(0.023)	(0.042)		
No. branches other banks per 1,000 population					-0.471**	10.814***		
					(0.015)	(0.000)		
Information sharing * No. branches own bank per 1,000 population							-1.190**	-2.089**
							(0.017)	(0.036)
No. branches own bank per 1,000 population							0.402**	-9.698***
							(0.015)	(0.000)
Locality * Year * Treatment Event Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bank * Year * Treatment Event Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Locality * Bank * Treatment Event Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes
R-squared	0.628	0.712	0.628	0.727	0.501	0.624	0.501	0.624
Observations	833,916	833,916	833,916	833,916	200,274	200,274	200,274	200,274

Table A5

Information Sharing and the Clustering of Bank Branches: Controlling for Local Branch Density, Population Growth, and Economic Growth

This table reports linear probability regressions to estimate the impact of the introduction of information sharing on bank branch clustering. The dependent variable measures whether a bank opens a new branch in a locality in a year. Observations from all countries that introduce information sharing in the same calendar year are grouped together and combined with the observations from not (yet) treated (control) countries within a six-year window around the introduction year. Following Cengiz et al. (2019) these event-specific data sets are then stacked to estimate a single coefficient. Table A1 contains the definitions and Table 1 the summary statistics for all variables. *Country* * *Treatment Event* clustered robust *p* -values are reported in parentheses. ***, **, * correspond to the 1%, 5%, and 10% level of statistical significance, respectively.

Dependent variable →	New branch opening							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Information sharing * No. branches other banks	0.021***	0.019***	0.017***	0.020***				
	(0.000)	(0.000)	(0.000)	(0.000)				
No. branches other banks	0.009***	0.010***	0.002	0.011***				
	(0.000)	(0.000)	(0.253)	(0.000)				
Information sharing * No. branches own bank					-0.068***	-0.071***	-0.073**	-0.070***
					(0.002)	(0.001)	(0.042)	(0.001)
No. branches own bank					0.019**	0.024***	-0.012*	0.021**
					(0.014)	(0.007)	(0.094)	(0.018)
Information sharing	-0.021	-0.0120	-0.008	-0.021	0.111**	0.119**	0.085*	0.109**
	(0.327)	(0.379)	(0.692)	(0.368)	(0.014)	(0.011)	(0.087)	(0.016)
No. bank branches per 1,000 population	-0.009			-0.022	0.074***			0.077***
	(0.849)			(0.660)	(0.001)			(0.002)
Population growth rate		-0.001		-0.001		-0.001		-0.001
		(0.548)		(0.548)		(0.609)		(0.629)
Night-time light growth rate			-0.004**	0.011**			-0.004***	0.006
			(0.013)	(0.031)			(0.004)	(0.282)
R-squared	0.014	0.015	0.005	0.016	0.016	0.016	0.009	0.017
Observations	200,274	165,948	830,636	165,528	200,274	165,948	830,636	165,528

Table A6

Information Sharing and Simultaneous Openings and Closures of Bank Branches

This table reports linear probability regressions to estimate the impact of the introduction of information sharing on bank branch clustering. The dependent variable measures whether a bank opens a new branch in a locality ranked in the top quartile based on the total number of bank branches. Observations from all countries that introduce information sharing in the same calendar year are grouped together and combined with the observations from not (yet) treated (control) countries within a six-year window around the introduction year. Following Cengiz et al. (2019) these event-specific data sets are then stacked to estimate a single coefficient. Table A1 contains the definitions and Table 1 the summary statistics for all variables. *Country * Treatment Event* clustered robust *p*-values are reported in parentheses. ***, **, * correspond to the 1%, 5%, and 10% level of statistical significance, respectively.

Dependent variable →	New branch opening				New branch opening in top quartile (no. branches)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Information sharing * Closed branch	0.078***	0.053			0.064***	0.047***		
	(0.005)	(0.109)			(0.000)	(0.004)		
Information sharing * Closed branch in bottom quartile (no. branches)			0.075***	0.047*			0.058***	0.039***
			(0.001)	(0.059)			(0.000)	(0.003)
Closed branch	-0.037***	-0.037***			-0.028***	-0.025***		
	(0.000)	(0.004)			(0.000)	(0.001)		
Closed branch in bottom quantile (no. branches)			-0.053***	-0.035***			-0.034***	-0.023***
			(0.000)	(0.000)			(0.000)	(0.001)
Locality * Year * Treatment Event Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Locality * Bank * Treatment Event Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes
R-squared	0.428	0.537	0.428	0.537	0.522	0.642	0.522	0.642
Observations	833,916	833,916	833,916	833,916	833,916	833,916	833,916	833,916

Table A7

Information Sharing and the Clustering of Bank Branches in Different Sized Towns and Cities

This table reports linear probability regressions to estimate the impact of the introduction of information sharing on bank branch clustering in localities with different population sizes. The dependent variable measures whether a bank opens a new branch in a locality in a year. Observations from all countries that introduce information sharing in the same calendar year are grouped together and combined with the observations from not (yet) treated (control) countries within a six-year window around the introduction year. Following Cengiz et al. (2019) these event-specific data sets are then stacked to estimate a single coefficient. Table A1 contains the definitions and Table 1 the summary statistics for all variables. *Country * Treatment Event* clustered robust *p*-values are reported in parentheses. ***, **, * correspond to the 1%, 5%, and 10% level of statistical significance, respectively.

In cities with a population of:	*Less than 50,000*		*50,000 to 250,000*		*More than 250,000*	
Dependent variable →			New branch opening			
	(1)	(2)	(3)	(4)	(5)	(6)
Information sharing * No. branches other banks	0.377***		0.575***		0.625***	
	(0.010)		(0.001)		(0.002)	
No. branches other banks	0.820***		0.915***		0.500***	
	(0.000)		(0.000)		(0.000)	
Information sharing * No. branches own bank		-0.129***		-0.119***		-0.150***
		(0.000)		(0.002)		(0.000)
No. branches own bank		-0.575***		-0.513***		-0.211***
		(0.000)		(0.000)		(0.000)
Locality * Year * Treatment Event Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Bank * Year * Treatment Event Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Locality * Bank * Treatment Event Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.702	0.729	0.663	0.700	0.684	0.697
Observations	234,790	234,790	98,766	98,766	95,777	95,777

Table A8

Information Sharing and the Geographical Clustering of Bank Branches: Controlling for Other Country-Level Reforms

This table reports linear probability regressions to estimate the impact of the introduction of information sharing on bank branch clustering. The dependent variable measures whether a bank opens a new branch in a locality in a year. Observations from all countries that introduce information sharing in the same calendar year are grouped together and combined with the observations from not (yet) treated (control) countries within a six-year window around the introduction year. Following Cengiz et al. (2019) these event-specific data sets are then stacked to estimate a single coefficient. Table A1 contains the definitions and Table 1 the summary statistics for all variables. *Country * Treatment Event* clustered robust *p* -values are reported in parentheses. ***, **, * correspond to the 1%, 5%, and 10% level of statistical significance, respectively.

Dependent variable →	New branch opening									
X →	No. branches other banks					No. branches own banks				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Information sharing * X	0.396***	0.386***	0.383***	0.110***	0.100***	-0.125***	-0.122***	-0.121***	-0.045***	-0.034***
	(0.002)	(0.003)	(0.002)	(0.000)	(0.006)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
EU membership * X	-0.158***				0.040	0.032				-0.021
	(0.324)				(0.793)	(0.401)				(0.547)
Competition policy * X		0.101***			-0.020		-0.025			-0.018
		(0.009)			(0.818)		(0.100)			(0.345)
Small-scale privatisation * X			0.146***		0.137**			-0.047***		-0.043**
			(0.001)		(0.015)			(0.007)		(0.027)
Pro-competition bank regulation * X				-0.044***	-0.043*				-0.029*	-0.029**
				(0.005)	(0.058)				(0.054)	(0.035)
X	0.662***	0.444***	0.133	0.777***	0.319	-0.389***	-0.333***	-0.214***	-0.287***	-0.091
	(0.000)	(0.000)	(0.443)	(0.000)	(0.150)	(0.000)	(0.000)	(0.001)	(0.000)	(0.290)
Locality * Year * Treatment Event Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bank * Year * Treatment Event Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Locality * Bank * Treatment Event Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.712	0.712	0.713	0.721	0.721	0.727	0.727	0.727	0.731	0.732
Observations	833,916	833,916	833,916	650,601	650,601	833,916	833,916	833,916	650,601	650,601