

Real effects of imperfect bank-firm matching *

Luisa Farinha
Banco de Portugal
lfarinha@bportugal.pt

Sotirios Kokas
University of Essex
skokas@essex.ac.uk

Enrico Sette
Banca d'Italia
enrico.sette@bancaditalia.it

Serafeim Tsoukas
University of Glasgow
serafeim.tsoukas@glasgow.ac.uk

July 1, 2022

Abstract

Using granular bank-firm level credit data, we show that the characteristics of bank-firm matches affect firms' access to credit and real outcomes during crises. We identify a set of potential matches in pre-crisis years, and we use them to predict match formation in crisis times. We generate a measure of "imperfect matches" given by the difference between realized and predicted matches. In crisis times, imperfect matches deteriorate firm outcomes. The effects are economically important. A one standard deviation worsening in the index is associated with a drop in firms' employment, tangible assets, and survival by 0.9%, 2.7%, and 4.2%, respectively.

Keywords: Financial frictions, bank-firm match, credit, investment, employment.

JEL Classification: G21, G30, E22, E51

*We thank for their helpful comments Diana Bonfim, Barbara Casu, Andrea Caggese, Hans Degryse, Ralph De Haas, Olivier de Jonghe, Alex Kontonikas, Klaas Mulier, Jose-Luis Peydro, Nicola Pierri, Andrea Presbitero, Hormoz Ramian, Giacomo Rodano, Andrea Schertler, Victoria Vanasco, Alonso Villacorta and participants of the research seminars/conferences at the Banca d'Italia, Banco de Portugal, European Bank for Reconstruction and Development, KU Leuven, Universitat Pompeu Fabra; Universities of Birmingham, Durham, Essex, Piraeus and Sheffield, 28th Finance Forum (Nova SBE), 2nd Finance and Productivity Conference, 2019 Financial Fragmentation and Challenges for SMEs' Financing, 2021 Financial Management and Accounting Research Conference, 2021 Financial Management Association, 2021 Financial Stability, 2021 International Finance and Banking Society, 2021 Money Macro and Finance Conference, 2021 World Finance Banking Symposium, 4th Conference on Contemporary Issues in Banking, and the 2022 Rimini Centre for Economic Analysis Conference. The views expressed in this paper are those of their authors and not necessarily reflect the views of the Banco de Portugal or the Banca d'Italia.

1 Introduction

Financial intermediaries perform a key role in selecting borrowers and monitoring their activities. Evaluating firms and their investment opportunities is costly, and financial intermediaries specialize in performing these tasks by lowering the costs of acquiring and processing information. In this way, they improve the allocation of credit and foster economic growth (Levine, 2005; Beck, 2009). Yet, the way credit is allocated is not random and it crucially depends on how banks and firms match to form credit relationships. Moreover, due to a lack of granular microdata about the universe of bank-firm relationships, empirical analyses of the characteristics that drive matches and their real effects are rather scarce.

This paper provides novel evidence on how the characteristics of bank-firm matches affect firms' access to credit and real outcomes (investment, employment, survival), based on proprietary administrative data from the Portuguese Credit Register. We first explore the attributes of banks and firms that lead to match formation. We perform this test on newly formed matches in the years before the global financial crisis, a period in which banks and firms were not constrained in forming and maintaining relationships. Next, we study how frictions in the formation and persistence of bank-firm matches affect firms' uptake of credit and real outcomes. To do so we develop an index to measure the extent to which bank-firm matches in crisis times differ from those in pre-crisis times. The index compares the bank and firm characteristics that lead to matches in pre-crisis years with characteristics that result in actual matches in crisis years. We then test whether larger differences between matches formed during the crisis and those formed before the crisis, affect firms' credit availability, investment, employment growth, and chances of survival during extreme economic events. Finally, we examine whether these effects are heterogeneous across firm size or the number of lending relationships.

Existing theories that identify the determinants of bank-firm match formation, provide a useful framework for our empirical analysis. The first element is the relative size of banks and firms. Stein (2002) shows that bigger banks characterized by larger hierarchies may find it more difficult to monitor smaller borrowers. A further element is bank capital relative to firm riskiness. On one hand, banks with higher capital are more likely to match with riskier and smaller firms, as they have stronger incentives to monitor borrowers because of higher skin-in-the-game (Holmstrom

and Tirole, 1997). On the other hand, higher leverage creates stronger incentives for monitoring (Diamond and Rajan, 2001), and models of risk-shifting (Keeley, 1990; Acharya and Naqvi, 2012) predict more frequent and stable matches between low-capital banks and risky firms. Finally, competition in local credit markets is an additional, potentially relevant determinant of bank-firm matches (Allen, Carletti and Marquez, 2011).

Taking stock of these theories, we conduct our empirical analysis as follows. First, relying on comprehensive microdata on the universe of loans, bank-branches, and firms' headquarters from Portugal, we predict the determinants of bank-firm matches (at the loan level) in the three years before the global financial crisis. We find that a match between a bank and a firm is more likely to occur if the bank is large, highly capitalized, has a branch close to a firm, and the firm is less likely to default. We obtain these results even in specifications that include bank-period or firm-period fixed effects to absorb time-varying unobserved demand and supply omitted factors. These correlations may be given a causal interpretation under the assumption that time-varying firm or bank unobservables do not correlate with firm or bank controls once fixed effects are included. Even if this assumption was not verified, our findings represent novel evidence about the bank and firm characteristics associated with the formation of bank-firm relationships.

Second, we use the correlations between bank and firm characteristics that lead to match formation to compute an index which measures the difference between characteristics of matches post crisis relative to pre crisis. In doing so, we measure the difference between the observed matches in the crisis period and those the model predicts, estimated in the pre-crisis period. We document that the index increases over time, as the difference in the characteristics of banks and firms that form a match widens between 2009 and 2016. A decomposition analysis shows that the main drivers of the evolution of the index are due to changes in bank and firm characteristics and from the initiation of lending relationships, while changes in the shares of credit of different lending relationships (i.e., changes in the exposure of banks towards a firm) play a limited role. Interestingly, small and medium-sized enterprises (SMEs), which typically suffer more strongly from borrowing constraints in crisis times, almost entirely drive the wider difference in bank-firm matches formed during the crisis relative to those formed before the crisis. We document that, during the crisis, credit is lower in relationships that differ more from those that would have formed before the crisis. A deterioration in the matching index (i.e., a wider distance between matches predicted on pre-crisis

characteristics and those actually formed during the crisis) by one standard deviation decreases credit supply by between 0.1 and 0.7 percentage points, increases the likelihood of switching a lender by 1% and raises the likelihood of terminating an existing relationship by 1.5%. This is consistent with the idea that our index captures the distance between the actual matches that occur in crisis times, when banks and firms face tighter constraints in forming credit relationships, and those occurring in normal times, when these constraints are less binding and bank and firm balance sheets are stronger.

Finally, we aggregate the index at the firm level to analyze its effect on firms' real outcomes. We document that the larger the difference between the observed and predicted matches (i.e., larger differences in realized matches in crisis relative to pre-crisis times), the lower the access to credit, investment, employment, and chances of survival at the firm level. To address potential endogeneity issues concerning the index and its dynamic (e.g., changes in firm characteristics, which also correlate with firms' growth opportunities, may drive changes in the index), we exploit the exogenous and unexpected shock created by the European Banking Authority's (EBA) capital exercise in 2011. Several papers (Gropp, Mosk, Ongena and Wix, 2019; Fraisse, Lé and Thesmar, 2020; Blattner, Farinha and Rebelo, 2021) show that the capital exercise represents an exogenous shock to credit supply that induced banks to cut lending across the board, irrespective of firm characteristics, to comply with higher capital requirements.¹ Therefore, we instrument the index with the share of outstanding loans vis-a-vis EBA-affected banks. This allows us to explore exogenous variation (supply-driven) in the bank-firm matching index that is orthogonal to firm characteristics.

We find economically significant effects: a deterioration in the index by one standard deviation is associated with a drop in firms' employment, tangible assets, and survival by 0.9%, 2.7%, and 4.2%, respectively. Importantly, we document that these adverse effects occur both when the number of matches falls, and when a firm substitutes a bank but keeps the number of lending relationships constant. Moreover, we show that these effects are strongly heterogeneous depending on whether firms are small or large and on whether they have single or multiple banking relationships.

Our results withstand several robustness tests. We use several alternative specifications of

¹The EBA capital exercise was unexpectedly announced soon after the stress test conducted in July 2011 (De-gryse, Karapetyan and Karmakar, 2021). This quasi-natural experiment required four Portuguese banking groups (containing seven banks), namely CGD, Banco BPI, BCP, and ESFG, to increase the minimum Core Tier 1 ratio to 9% by the end of 2011 and to 10% by the end of 2012. Banks were selected based on size thresholds and not on the incidence of NPLs or on their capitalization.

the empirical model for the bank-firm matching process. We use different samples to investigate whether the structure of the local market or the density of big cities may drive our results.² We run a placebo test to validate our model of predicted bank-firm matches. If we compute our index for the differences between actual and predicted matches in pre-crisis times (i.e., using the in-sample predictions of the model) we find that the index does not predict credit growth, investment, employment, or probability of default. This suggests that the index has explanatory power for firm real outcomes when firms actually face borrowing constraints.

The high-quality firm-bank matched data from the Portuguese credit and firm registers represent an ideal laboratory for our analysis for three main reasons. First, the Portuguese credit registry has a very low reporting threshold, €50, which allows us to observe the entire population of bank-firm credit records. This is critical to obtain an accurate estimate of potential, observed, and predicted bank-firm matches. Second, the dataset includes a large fraction of micro and small firms. These firms are heavily bank dependent—and typically cannot substitute bank credit with market financing when their access to credit worsens. Third, the global financial crisis is arguably an exogenous shock to the Portuguese economy, having its origins in the U.S. subprime mortgage market, to which Portuguese banks were not exposed. Hence, the global financial crisis represents an exogenous shock to the ability and willingness of Portuguese banks and firms to form and maintain credit relationships.

Our findings contribute to several strands of the literature. Our work is most closely related to [Schwert \(2018\)](#), who shows, using syndicated loan data, that bank-dependent borrowers are more likely to form credit relationships with well-capitalized banks, as this allows bank-dependent firms to obtain a smoother flow of credit throughout the business cycle. We complement and extend these findings in several ways. We study the bank-firm matching in local markets to derive an index for the differences in matches between crisis and pre-crisis times, and we analyze how these differences in bank-firm matches affect firms' access to credit, investment, employment growth, and probability of default. Importantly, we observe the whole population of borrowers and lenders from a comprehensive credit registry, which is critical to obtain a full picture of bank-firm matches, extending the evidence from syndicated loan data that typically include mostly large firms. We also

²In unreported results, we use geolocalization to define the bank-firm matching. Our results remain robust to this modification.

study more broadly the drivers of match formation: capital in our set-up is one, albeit important, of several characteristics driving the creation of bank-firm matches. In this respect, we build on prior work that shows that relative bank-firm size and distance affect loan amount and price (Stein, 2002; Berger, Miller, Petersen, Rajan and Stein, 2005), but that work does not look at the probability that a bank and a firm create a relationship. Although our estimates identify predictors of a bank-firm match and warrant a causal interpretation only under relatively strict assumptions, they are informative for theories of the drivers of bank-firm relationships.

We also contribute to the large literature on relationship lending. Several papers argue that longer and stable bank-firm relationships support firms' access to credit (see for a summary Degryse, Kim and Ongena, 2009), especially in times of crises (Sette and Gobbi, 2015; Bolton, Freixas, Gambacorta and Mistrulli, 2016; Cohen, Hachem and Richardson, 2021). In particular, Darmouni (2020) quantifies the informational gap between existing and new lenders and shows that this is a key friction limiting the availability of external finance during crises. Also, in a setting of break-ups due to branch closures, Bonfim, Nogueira and Ongena (2021) show that firms do not obtain interest rate discounts when they switch banks in the aftermath of branch closures. Goncharenko, Mamonov, Ongena, Popova and Turdyeva (2022) look at how firms reallocate to other banks when their main lender fails, depending on their credit status (namely, having performing or non performing loans). Other works show that relationship lending attenuates the transmission of shocks to firms' real outcomes, allowing relationship borrowers to invest more and have higher sales and employment growth than other borrowers (Beck, Degryse, De Haas and Van Horen, 2018; Banerjee, Gambacorta and Sette, 2021).

Our work differs from the relationship-lending literature in two main ways. First, when a crisis hits, if the firm or bank characteristics deteriorate and make the match different than the one prevalent before the crisis, then firms' access to credit and their growth suffer, irrespective of whether the set of relationships remains the same, thus irrespective of whether tight relationships (long duration, large share of credit) remain active. Second, when credit relationships break, firms are allocated lower credit and grow less, especially if the new relationships differ from those prevailing in good times. Importantly, this also occurs if terminated bank-firm relationships are replaced by new ones so that the total number of relationships remains constant; credit access may still be impaired, depending on the characteristics of the new match. Our results add to the

literature on relationship lending by showing which bank and firm characteristics, on top of the commonly used proxies of relationship strength, matter for forming lending relationships and, most important affect access to credit and firm real outcomes.

Finally, the analysis of the real effects of bank-firm match quality builds on the literature on the effects of financial shocks on investment (Duchin, Ozbas and Sensoy, 2010; Cingano, Manaresi and Sette, 2016; DeJonghe, Dewachter, Mullier, Ongena and Schepens, 2020; Balduzzi, Brancati and Schiantarelli, 2018; Beck, Da-Rocha-Lopes and Silva, 2021) and employment (Chodorow-Reich, 2014; Popov and Rocholl, 2018; Bentolila, Jansen, Jiménez and Ruano, 2018; Berton, Mocetti, Presbitero and Richiardi, 2018; Dwenger, Fossen and Simmler, 2020) by showing that the worsening of bank-firm matches in crises is a specific channel through which financial shocks affect credit access and impair firm growth. Moreover, we find that the real effects of worse bank-firm matches are stronger for firms with single bank relationships, which are disproportionately smaller firms. These results are related to recent works showing that small firms are more cyclical than larger firms (Crouzet and Mehrotra, 2020). Contrary to this work based on U.S. data, we document a larger relevance of financial frictions to explain firm growth during crisis times. Last, our results on SMEs complement Chodorow-Reich, Darmouni, Luck and Plosser (2022), who document worse access to credit among SMEs relative to larger firms in U.S. data and tougher access to liquidity during the Covid-19 pandemic.

The paper is structured as follows. In section 2 we describe the empirical specifications and the estimation strategy. We present the data used in our empirical analysis in section 3. In section 4 we report the results. In section 5 we check the robustness of our findings, and we provide concluding comments in section 6.

2 Empirical specification

2.1 Bank-firm matching

Our first objective is to estimate the bank and firm characteristics associated with the formation of bank-firm matches. To this aim, we need to define potential and observed matches. As a first step, we define the relevant local credit market. According to existing theories and empirical evidence, geographical distance (particularly for small business) erodes banks' ability to acquire firm private

(soft) information because it captures proximity to the information source in various guises. More specifically, for firms in distant locations it can be more costly to borrow because of information problems (Hauswald and Marquez, 2006) and transportation costs (Acharya, Hasan and Saunders, 2006). According to these theories, banks derive cost advantages ex-ante from being physically closer to borrowers.

We calculate bank-firm matching as follows. From the Register of Financial Institutions (*Registo Especial de Instituicoes*), we observe the list of bank-branches in Portugal with information about post code and opening-closing dates for each branch. We restrict our analysis to active branches after 2006, which is the first year of our sample. Then, following Degryse and Ongena (2005) we obtain all bank-branches that operate in a four-digit post-code. These represent all potential matches for firms with headquarters in the same post-code.³ Petersen and Rajan (2002) report that the distance between U.S. bank-firms increases over time, while Granja, Leuz and Rajan (2022) find a cyclical pattern in lending distances that widen in booms and shorten in downturns. However, Bonfim, Nogueira and Ongena (2021) show that the median distance between a firm and a bank in Portugal is 1.9 kilometres.

Importantly, approximately 70% of firms employ at least one bank that has branches in the same post-code, and the remaining firms are linked with bank-branches in different post codes.⁴ Next, we merge these data using unique bank identifiers with credit registry data, and we geographically map the post codes of bank-branches and firms to calculate the observed bank-firm matches.⁵

Then, dummy $Match_{b,f,l,t}$ equals one, if the bank (b) - firm (f) match is in the credit registry within a four-digit post-code (l) at time t ; it equals zero if they do not match.⁶

With this structure, we estimate a model for the probability that a bank-firm match occurs as

³The first digit designates one of the nine post regions: Lisbon (1), Estremadura e Ribatejo (2), Beira Litoral (3), Minho e Douro Litoral (4), Trás-os-Montes e Alto Douro (5), Beira Interior (6), Alentejo (7), Algarve (8), Madeira Islands, and Azores (9). The following two digits designate postal distribution centers within a region, and the fourth digit is a designated address.

⁴In unreported results for robustness, we define the relevant market at the seven-digit postcode. In doing so, we resort to a smaller geographical area compared to the four-digit postcode. Although the number of matches is significantly smaller, the results are qualitatively unchanged.

⁵As we do not observe exactly which branch originates the loan, we aggregate all branches of a bank in a four-digit post-code, because this is our unit of analysis and final loan approval occurs at the bank's headquarters.

⁶We drop matches formed outside the four digit post-code, as these do not occur within the relevant definition of a credit market and we cannot attach local characteristics to these matches. However, we experiment with different definitions of the local credit market to check the robustness of our results.

follows:

$$\begin{aligned}
\text{Prob}(\text{match}_{b,f,l,t}) = & \lambda_1 * (\text{Firm Size}_{f,l,t} * \text{Bank Size}_{b,l,t}) + \lambda_2 * \text{Capital ratio}_{b,t} \\
& + \lambda_3 * \text{HHI}_{b,l,t} + \lambda_4 * \text{Prob}(\text{default})_{f,t-1} + \alpha_0 + \epsilon_{b,f,l,t} \quad (1)
\end{aligned}$$

where α_0 denotes different levels of time-varying and time-invariant fixed effects, and $\epsilon_{b,f,l,t}$ is a loan-level shock that captures the stochastic disturbances. The selection of the potential drivers of bank-firm matches is based on the main results of the previous literature. Several papers highlight that the matching process depends on bank-firm size (Stein, 2002), bank capital (Holmstrom and Tirole, 1997), banking competition (Allen, Carletti and Marquez, 2011), and firms' riskiness (Keeley, 1990; Acharya and Naqvi, 2012). To capture the size dependence, we use an interaction between the deciles of firm and bank size. We measure bank capital as the ratio of book equity to total assets. We include the Herfindahl-Hirschman index (HHI) of the branch concentration per bank at the four-digit post code. As for firms' riskiness, we incorporate the Banco de Portugal's estimation of the borrower's probability of default.⁷

The model in equation (1) aims at identifying the predictors of bank-firm matches to build a counterfactual for the crisis period. Our empirical approach does not focus on trying to give a causal interpretation to the effect of the firm and bank characteristics that we include in the model. Yet, the granularity of our dataset allows us to account for unobservables by including different sets of time-varying fixed effects (bank*year, firm*year, loan type, and location), because firms often have multiple bank relationships and lenders originate multiple loans within a year. Among these fixed effects, the bank*year and firm*year fixed effects are particularly important to control for time-varying demand and supply factors. Notably, we can compare how the coefficient of each driver of the bank-firm matching changes when we include different sets of fixed effects, indicating how each firm or bank observable characteristics correlates with unobservable characteristics. Provided that this fixed-effects approach is enough to fully control for omitted variables, our estimates of the effects of bank and firm drivers of match formation may be given a causal interpretation.

⁷The Banco de Portugal calculates the probability that any given firm has a significant default episode vis-à-vis the banking system using information from the central credit register and comprehensive balance sheet data (Antunes, Goncalves and Prego, 2016). This variable measures the firm's probability of default on bank debt within a one-year horizon.

In principle, we could have used a structural model to identify the determinants of bank-firm formation. Yet, we decided to resort to a reduced-form matching model for the following reasons. First, we aim at obtaining a prediction of match quality, which can be flexibly obtained through a reduced-form approach. Second, alternative approaches such as [Berry, Levinsohn and Pakes \(1995\)](#) require observing price data, and our data set does not include information on interest rates. Third, and most important, there are potentially multiple matches between banks and firms on both sides, (i.e., both banks and firms form multiple matches). To date, there is no theory of multiple matching on both sides, which can guide us in obtaining an estimable model ([Chiappori and Salanié, 2016](#)).

2.2 The imperfect-match index

We use the model in equation (1) in the pre-global financial crisis period to predict matches out of sample in the post-crisis period.⁸ This way, we can measure the extent to which matches formed in crisis years differ from those formed in pre-crisis years, and then we can check whether this difference explains access to credit and firms' real outcomes.

We construct the index through the following steps. First, we estimate equation (1) using a pre-crisis sample from 2006 until 2008 with different levels of fixed effects. Second, based on the estimated coefficients, we obtain out-of-sample predicted bank-firm matches for the post-crisis period from 2009 to 2016. Finally, we take the square of the difference between the realized matches (raw data) and the predicted likelihood that the matching occurs. So, for each bank-firm pair in the credit registry in the post-crisis period, we define the index as $(\widehat{Realized\ match} - \widehat{Predicted\ match})^2$. The matching index ranges in the (0, 1) region. Deviations from zero indicate that a match during the crisis is different from a match that would have occurred in pre-crisis times. The larger the index, greater the difference in the relative bank-firm characteristics. We use “imperfect-match index” hereafter to identify our measure of difference between matches occurring in crisis and those predicted on the basis of the parameters driving match formation in the pre-crisis period. We interpret our index as a distance or proximity to the (predicted) bank-firm matching that would have occurred before the crisis.

The underlying idea is that both firms and banks face tighter constraints in forming and main-

⁸The crisis hit Portugal after the default of Lehman Brother, which occurred in October 2008. Its effects on credit and real outcomes materialized in 2009 ([Iyer, Peydro, da Rocha-Lopes and Schoar, 2013](#)).

taining credit relationships in crisis times rather than in tranquil times. Matches closer to those prevailing in normal times are associated with a better flow of information, improved management of the credit relationship, and other factors that alleviate lending frictions during bad times. Therefore, we argue that the difference between matches in crisis times and those in normal times can affect access to credit and associated real effects.

In constructing the index, we make an implicit assumption that the matches estimated in the 2006-2008 period represent matches in a period characterized by less binding financial constraints, in which both banks and firms form matches facing limited constraints on their choices. This is a reasonable assumption because the pre-crisis period in Portugal is characterized by moderate economic growth (real GDP growth is 1.4% per year on average between 2004 and 2008) and by the absence of a housing or credit bubbles. To support this hypothesis, we construct measures of credit booms following the approach in [Greenwood, Hanson, Shleifer and Sørensen \(2022\)](#). Figure 1 shows the indicator for non-financial business credit and asset price booms.⁹ Figure 2 suggests the absence of a housing bubble in the Portuguese economy in the pre-crisis years. In both figures business and household credit are measured as a percentage of GDP. The crisis period instead features a sharp contraction in GDP (-1.6% on average between 2009 and 2013, with especially large drops in 2009 and 2012 of -3.1% and -4.1%, respectively) and a strong rise in unemployment (to 16.1% in 2013, from an average of 7.5% in the five years between 2004 and 2008). In addition, [Iyer, Peydro, da Rocha-Lopes and Schoar \(2013\)](#), using Portuguese credit registry data, document a supply-driven drop in credit, especially for smaller and riskier firms in the aftermath of the Great Recession. One might argue that the threshold to define Portugal’s pre-crisis and crisis period might be different considering the EU sovereign debt period. For this reason, we also explore the robustness of our results to measuring the crisis period starting after 2011 instead of 2009.

2.3 Estimation of firm-level effects

In this section, we analyze the effect of the imperfect-match index on firms’ real activities in the post-crisis period. To do so, we aggregate the data at the firm level, taking the weighted average

⁹The indicator signals that a country is in the “business red-zone” (i.e., it is in a period of financial over-heating), if non-financial business credit growth over the past three years is in the top quantile of the historical distribution and stock market returns over the same window are in the top tercile.

of the index across relationships using as weights the share of credit in each relationship.¹⁰

Determining a causal relation running from imperfect bank-firm matching to firms' decision-making poses an identification challenge due to endogenous matching. Banks or firms may terminate matches of worse quality, thus creating a survivorship bias. In addition, the reason a bank terminates a relationship, thus inducing a firm to seek a replacement, may correlate with firms' growth opportunities. This may be an especially relevant issue to the extent that banks with higher credit risk during good times are more likely to lend to firms with higher profit volatility (Iosifidi and Kokas, 2015). To address these concerns, we exploit an instrumental variables (IV) approach and estimate regressions of the following form:

$$Y_{f,t} = \alpha_0 + \beta_1 * \widehat{Imperfect Match}_{f,t} + \beta_2 * F_{f,t} + \mu_f + \mu_t + \epsilon_{f,t} \quad (2)$$

$$\widehat{Imperfect Match}_{f,t} = \alpha_0 + \rho * EBA borrowing share_{f,t} + \gamma * F_{f,t} + \eta_{f,t} \quad (3)$$

where the outcome variables Y are the natural logarithm for the number of employees, the natural logarithm for tangible assets, and the probability of default on bank debt within a one-year horizon. Vector F denotes a set of firm control variables that are likely to influence firms' real decisions: firm size (the logarithm of firms' real total assets); return on assets (ROA) as a profitability indicator, which is a proxy for unlisted firms' Tobin's Q (Asker, Farre-Mensa and Ljungqvist, 2015) and firms' overall indebtedness using the leverage ratio, measured as long-term debt over total assets. We expect larger firms and those in better financial shape to have higher levels of employment and tangible assets. Last, we include firm and year fixed effects.

To yield exogenous variation in the imperfect-match index, we exploit variation in firms' dependence on credit from banks affected by the European Banking Authority (EBA) capital exercise in October 2011. Specifically, we refer to the unexpected increase in bank capital requirements aimed at restoring confidence in the EU banking sector by ensuring that banks are adequately capitalized to mitigate unexpected losses. The EBA capital exercise was unexpectedly announced soon after the stress tests conducted in July 2011 (Mésonnier and Monks, 2015; Degryse, Karapetyan and Karmakar, 2021). This quasi-natural experiment required some banks to increase the minimum

¹⁰As explained, in each relationship this is the squared difference between realized and predicted matches.

levels of the Core Tier 1 ratio to 9% by the end of 2011 and to 10% by the end of 2012. Banks were selected in a quasi-random fashion— based on size, and not profitability, quality of loan portfolio, etc. The new capital requirement increased the relative cost of lending for affected banks but not for unaffected ones, allowing us to exploit a supply-driven exogenous variation in the matching index that is orthogonal to firm characteristics. [Gropp, Mosk, Ongena and Wix \(2019\)](#) show that banks reach the new required capital ratios by reducing their exposure to SME loans. In addition, [Blattner, Farinha and Rebelo \(2021\)](#) indicate that Portuguese banks involved in the capital exercise do not terminate relationships based on firm characteristics, but they cut credit across the board to meet the stringent requirements in a short time period. Importantly, they also argue that there are no liquidity trends prior to the capital exercise by estimating a dynamic firm-level difference-in-differences regression. Hence, concurrent liquidity shocks are unlikely to explain the results. The instrument shifts credit supply and in this way affects bank-firm matches, and thus the extent to which they differ from matches that would have been created in pre-crisis times.

In practice, we construct the instrument as follows:

$$EBA \text{ borrowing share}_{f,t} = \frac{\sum_{EBA} \text{Outstanding amount}_{f,t}}{\sum_{All \text{ bank}} \text{Outstanding amount}_{f,t}},$$

where the numerator is the average amount of outstanding credit of firm f from EBA-exercised banks, and the denominator is the total amount of credit from all banks.

As an extension, we test for differential effects across single- versus multiple-relationship firms. It is common in the literature to disentangle the effects of single- and multiple-relationship firms because the former group is less likely to switch lenders during extreme economic events. This is also borne out in our data, which we discuss later, as we show that approximately 70% of our firms rely on a single lender.

3 Data and summary statistics

3.1 Data description

We use proprietary administrative data from the Portuguese central bank containing detailed, high-quality, matched firm-bank information. We observe data on credit relationships and balance sheets

for both firms and banks from 2006 to 2016. The dataset spans the period before and after the sovereign debt crisis and the EBA capital exercise, and it is made up by three main sources.

The first is the Central Credit Register (CRC) of Banco de Portugal, which includes monthly loan exposures for every firm-bank pair from 2006 to 2016. This comprehensive dataset records all commercial and industrial loans to non-financial companies by all banks operating in Portugal. The threshold for reporting loan information is €50; hence, the credit register records the universe of outstanding loans to corporations and individuals. It is a legal requirement for all financial institutions granting credit in Portugal to report all loans above €50 to CRC on a monthly basis. This is an appealing characteristic of the dataset for our analysis because we can effectively construct all potential matches and observe all realized matches. This database contains information about the amount of the loan and its status, namely if it is performing, renegotiated, non-performing or defaulted.

To match all loans with the corresponding bank-specific characteristics, we use Monthly Financial Statistics data, which is a database reporting balance sheet information for financial institutions operating in Portugal. The bank-level data are monthly in frequency.

Firm balance sheets and income statements are from the Informcao Empresarial Simplicada (IES), which covers the entire universe of Portuguese non-financial firms. The firm-level data have an annual frequency. We also use firms' probability of default computed by the Banco de Portugal ([Antunes, Goncalves and Prego, 2016](#)).

As standard in the literature, we exclude companies that did not have complete records on our explanatory variables and firm-years with negative sales. To control for the potential influence of outliers, we remove observations in the 1% upper and lower tails of the distribution of the regression variables. Our panel includes 987,763 firm-bank observations from 512,446 firms. Finally, there are 453 banks active in the loan market.¹¹

3.2 Summary statistics

In this section we describe some features of firm-bank matches in our dataset. Figure 3 gives a preliminary glimpse at the evolution of potential and realized matches over the sample period. The number of potential matches by postcode declines shortly after the European sovereign debt crisis

¹¹Our sample of banks contains “caixas agricola,” which have local importance for small firms in Portugal.

in 2011 and remains low until 2017. The number of realized matches, however, drops during the sovereign debt crisis– but gradually increases after 2013.¹²

Table 1 presents descriptive statistics about the bank-branch pairs over the sample period and their relevant post codes. We observe a modest reduction in the average number of branches after 2011. This echoes the pattern in figure 3, where we report potential and realized matches on a year-by-year basis. In addition, this dynamic pattern is in line with [Bonfim, Nogueira and Ongena \(2021\)](#), who document a significant number of branch closures during this period due to cost-cutting pressures. However, the number of branches for bank post-code remains relatively stable throughout the years.

Finally, table 2 contains descriptive statistics of the variables in the empirical models. The firm-level statistics highlight that our sample includes a large fraction of SMEs.¹³ The average outstanding loan is €85,851, and the average default probability within one year is 5.5%. In table 3, we delve deeper into the data by examining the distribution of the realized firm-bank matches. Interestingly, approximately 70% of the firms in our sample match with only one bank, confirming previous findings for Portugal ([Farinha and Santos, 2002](#)).¹⁴

4 Results

4.1 Bank-firm matches

We begin our analysis by examining the determinants of bank-firm matches. Table 4 shows estimates of equation (1) using different sets of fixed effects. The findings point to a strong increase in the probability of a match as banks become larger. The estimates indicate that both small and large firms are more likely to match with a large bank. This finding is not just a mechanical effect driven by large banks having more branches. That is, we compute the set of potential matches on

¹²One might wonder whether internet banking affects the way that banks match with firms. We argue that online banking is unlikely to play a significant role within our setting for the following reasons. First, online banking is more relevant for households’ standardized products, while firms require more tailored products and specific screening and monitoring. Second, according to the Survey on Information and Communication Technologies Usage in Enterprises (IUTICE), both the adoption of internet connection for firms and the internet speed increased significantly in the later part of our sample period.

¹³A large percentage of Portuguese firms are small according to the European Commission’s criteria: only 1% are large and 85% are micro.

¹⁴There is variability across Europe in the prevalence of single as opposed to multiple banking. Portugal appears to be more similar to Belgium ([DeJonghe, Dewachter, Mullier, Ongena and Schepens, 2020](#)), and different from Italy, where multiple banking is more common ([Detragiache, Garella and Guiso, 2000](#); [Sette and Gobbi, 2015](#)).

banks active in a given post code and we assign the same weight to all active banks in a postcode, irrespective of the number of branches they have in the area. This is a new result in the literature; previous work conjectures an advantage of small banks in lending to small firms ([Berger, Miller, Petersen, Rajan and Stein, 2005](#)). Interestingly, the result that larger banks are more likely to form matches holds when we progressively saturate the model with combinations of firm, bank, and period fixed effects, although it drops in magnitude.

Moving to bank characteristics, the capital ratio is negatively related with the probability of matching. However, when we include bank fixed effects, the coefficient turns positive and significant. This is in line with previous studies ([Schwert, 2018](#)) suggesting that bank capital is a measure of risk-bearing capacity. Banks that are better capitalized are more likely to form credit relationships. In addition, we show that the concentration of the local credit market is negatively associated with the probability that matches form between banks and firms. This finding, which holds when we control for location fixed effects, is consistent with expectations that firms are less likely to match with banks in more concentrated markets. Access to credit in highly concentrated markets is presumably more costly—because of higher interest rates or collateral requirements.

In terms of riskiness, firms more likely to default are less likely to form matches. Once again, this is consistent with the notion that high-risk firms have limited access to external financing and thus are more credit constrained ([Jiménez, Ongena, Peydró and Saurina, 2014](#)). Overall, the controls in the model to predict bank-firm matches that occur in normal times have the expected signs.

We dig further into the baseline results to look for heterogeneous effects across firm size, which is an important dimension of the lending technology banks choose ([Degryse, Kim and Ongena, 2009](#)). Specifically, in [table 5](#) we report results from specifications including an interaction between the dummy for small firms and bank capital and firms' probability of default. We observe that the point estimate on the interaction between small firms and bank capital is positive and significant (at the 1% level).

Bank capital is positively correlated with the probability of a match particularly for small firms, and with riskier firms, irrespective of firm size. This is important, as it suggests that stronger banks are more likely to match with riskier firms, pointing to an allocation of risk toward banks that have a higher risk-bearing capacity. These findings are consistent with [Schwert \(2018\)](#), who works on a

very different sample of syndicated loans, which typically go to large firms, from U.S.-based banks and borrowers.

4.2 Imperfect-match index

In this section we construct the index of predicted matches, conditional on the bank and firm characteristics that correlate with the probability of bank-firm formation. As discussed in subsection 2.2, the index of imperfect matching measures how much the realized matches in crisis years (2009-2016) differ from those predicted on the basis of the estimates from the pre-crisis period (2006-2008).¹⁵ The model to predict matches (the specification shown in column IV in table 4) in pre-crisis years has a high goodness of fit, as we can explain 50% of the variation in the probability that a match occurs.

Figure 4 shows the evolution of the index during the crisis period, when Portugal was hit by both the global financial crisis and the subsequent Eurozone sovereign debt crisis. We observe that the index worsens over time, especially between 2010 and 2012, reaching a peak in 2015, before improving somewhat in 2016. In figure 5, we break down the index for small and large firms.¹⁶ Small (large) firms are those whose assets are in the lower (upper) quartile of the total assets distribution. Interestingly, we show that the index for large firms remains relatively stable over time, but it deteriorates for their smaller counterparts from 2010 onward.

To better understand what drives the variation in the imperfect-match index, we use the approach in Baily, Hulten and Campbell (1992) to decompose the change in the imperfect-match index at the firm-level into four main components. The first is the contribution of the bank and firm fundamentals, (i.e., the change in the predicted match) at the bank-firm level (block 1); the second is the change in the share of credit, keeping the index constant (block 2); the third is the contribution of new lending relationships (block 3); and the fourth is the contribution of the termination of existing ones (block 4). Formally:

¹⁵The imperfect match 2009 refers to the indicator computed using parameters over 2006-2008. In a robustness check we also estimate match formation using variables between 2006 and 2011, and we predict matches out of sample from 2012 onward (see table A1).

¹⁶Higher index values indicate larger deviation from the predicted match, and thus a worse match.

$$\begin{aligned}
\Delta MI_{ft} = & \sum_b (MI_{f,b,t} - MI_{f,b,t-1}) * sharecredit_{f,b,t} \\
+ & \sum_b MI_{f,b,t-1} * (sharecredit_{f,b,t} - sharecredit_{f,b,t-1}) \\
& + \sum_b (MI_{f,b,t} * sharecredit_{f,b,t}) \\
& - \sum_b (MI_{f,b,t-1} * sharecredit_{f,b,t-1})
\end{aligned}$$

where MI_{ft} is the matching index aggregated at the firm-level and $MI_{f,b,t}$ is the index at the bank-firm level. The first two terms are computed on relationships that are in place in consecutive periods (“incumbents”). The term $\sum_b (MI_{f,b,t} - MI_{f,b,t-1}) * sharecredit_{f,b,t}$ is the “within” component, as it measures the effect of changing the characteristics of banks and firms, keeping the share of credit constant (as of time t). The term $\sum_b MI_{f,b,t-1} * (sharecredit_{f,b,t} - sharecredit_{f,b,t-1})$ is the “between” component, as it measures the effect of changing the share of credit, keeping the quality of the index constant (as of time $t-1$). The final blocks correspond to the extensive margin. The term $\sum_b (MI_{f,b,t} * sharecredit_{f,b,t})$ captures the contribution of new relationships, which is computed for relationships in place at t (“new entrants”)– but that were not in place in $t - 1$. Finally, the term $\sum_b (MI_{f,b,t-1} * sharecredit_{f,b,t-1})$ stands for the contribution of exit, which is only for relationships in place at $t-1$ but not in t .

Table 6 reports the contribution of each of the four components in explaining the variation in the imperfect-match index. Overall, the index changes from 0.160 in 2009 to 0.190 in 2016 (see the top panel of table 6), indicating a worsening of match quality during the crisis years (see also figure 4). Changes in the index come mainly from changes in bank and firm characteristics (block 1) and from the opening of new bank-firm relationships (block 3). The contribution of changes in the shares of credit (block 2) and of the termination of existing relationships is marginal (block 4). This suggests that new relationships opened during the crisis period are on average worse (i.e., less similar to those prevailing in normal times), than existing ones, leading to a deterioration in the index. The small contribution of the termination of existing relations suggests that the credit

relationships that end during the crisis are similar to those that remain in place or are characterized by smaller shares of credit (as of 2009), leading to a marginal contribution of the termination of relationships to the overall change in the match quality index.

4.3 Matching quality and access to credit: Bank-firm evidence

Next, we explore the real effects of imperfect bank-firm matching. As a starting point, we validate the economic rationale of the matching index by analyzing how it affects the access to credit, and relationship survival at the bank-firm level. In table 7 we report estimates of a regression of the (log) outstanding loan amount on the matching index. We find evidence that the matching index negatively affects the supply of bank credit across all models. In other words, higher index values, which imply larger deviations from the pre-crisis match, lead to a significant drop in the credit that firms obtain during a downturn at the credit-relationship level.¹⁷ This finding is robust to including different combinations of fixed effects. Importantly, this finding, at the bank-firm level does not depend on the composition of bank lending to firms or the number of branches at the four-digit level; rather, it is solely affected by the characteristics of the match. Based on the reported specifications, we control for time-varying unobservable supply factors (bank*year fixed effects), demand factors (firm*year fixed effects), and more restrictive bank-firm unobserved characteristics (bank*firm fixed effects). For example, in column III of table 7, we compare the same bank lending to different firms (within variation), but we control for location and common shocks. The effect is economically significant: a one-standard-deviation deterioration in match quality (this is captured by an increase in the index) is associated with a drop in credit of between 0.1 and 0.7 percentage points, depending on the specification. This is a sizable effect because it corresponds to between 1/20 and 1/4 of a standard deviation in the outstanding loan amount. In terms of loan volume, this corresponds to a drop of between €9,000 and €45,000, a non-trivial amount as the average loan outstanding amount is €85,851.

In table 8, we examine whether and how the matching quality influences the probability of switching lenders or terminating a lending relationship. For the definition of firms switching lenders, we closely follow Ioannidou and Ongena (2010) and Bonfim, Nogueira and Ongena (2021). Specif-

¹⁷As the imperfect-match index is a generated fitted regressor, we alleviate possible measurement bias in the construction of the standard errors by replicating table 7 with bootstrap standard errors with 300 replications (see table A3 in the online appendix).

ically, we define a new credit relationship as a *Switch* when we observe in the firm’s credit registry a new loan from a bank that the firm does not have a lending relationship with during the previous twelve months. However, this does not differentiate between firms that replace banks keeping the number of lenders constant and those that add a new banking relationship. To disentangle the former group from the latter, we create *Termination of lending*. This is a dummy that equals one if the bank terminates an existing relationship, and zero otherwise. Columns I to III present results for switching lenders, followed by termination of lending in columns IV to VI. The results show that higher quality matches are less likely to be associated with either a switch or an outright termination. This is a sanity check for the imperfect-match index, which indeed identifies as imperfect matches those more likely to end.

4.4 Real effects of imperfect matching

We now explore the real effects of imperfect firm-bank matching during the crisis. To carry out this test, we aggregate the credit registry data at the firm-year level. As mentioned, changes in the match index may be endogenous to firm or bank developments. To address this concern, we use an IV approach based on exogenous variation from the EBA capital exercise initiated in 2011 and affecting a subset of banks in Portugal. Specifically, following the strategy detailed in sub-section 2.3, we use the share of credit granted by capital-exercised banks (Gropp, Mosk, Ongena and Wix, 2019; Blattner, Farinha and Rebelo, 2021) as a plausible exogenous instrument for the matching index.¹⁸ We identify banks that are affected by the EBA exercise and link this information to the credit registry.¹⁹ In this manner, we exploit only variations in the imperfect match index that are due to the EBA borrowing share. The intuition is that the EBA capital exercise leads affected banks to reach the new capital requirements by reducing SMEs’ lending. This reduction likely hits harder the firms that are attached to EBA-affected banks because the new capital requirement increased the relative cost of lending for the affected banks but not for unaffected ones. Gropp, Mosk, Ongena and Wix (2019) and Blattner, Farinha and Rebelo (2021) show that increasing

¹⁸The minimum Core Tier 1 ratio increased to 10% in 2012, and banks had until the end of year to comply. At the same time, banks subject to the EBA stress tests were also subject to stricter capital requirements. These additional capital requirements were offset by curtailing lending.

¹⁹In Portugal, the EBA’s rules affected four banking groups (containing seven banks), namely CGD, Banco BPI, BCP, and ESFG.

capital requirements reduces bank lending.²⁰

Table 9 shows the results of the IV model. The first-stage estimates in panel A show that the instrument has the expected effect on the perfect-match index: following the EBA capital exercise the imperfect-match index deteriorates. The instrument is highly statistically significant and it does not seem to suffer from weak-instruments biases. In panel B, we show the second-stage estimates, where the main outcome variables are the natural logarithm of the number of employees (columns I to III) and natural logarithm of tangible assets (columns IV to VI). Given that we conduct our analysis at the firm level, we incorporate firm and year fixed effects to control for time-invariant characteristics and macroeconomic effects, respectively. We run the tests on the whole sample and on sub-samples accounting for both single and multiple bank relationships.

Starting with column I, we observe that imperfect matches exert a negative and highly significant effect on firm employment. The point estimate suggests that a one-standard deviation increase (worsening) in matching quality is associated with a drop in firm employment of 0.9 percentage points, which is about 90 % of a standard deviation. When we split our sample to distinguish between firms with single and multiple relationships (columns II and III), we find that former entirely drives the effects: match quality matters for determining employment decisions at single-bank firms, but firms with multiple relationships remain unaffected by changes in match quality. In the following columns of table 9, we rerun the same regressions but use firms' fixed tangible assets as the outcome variable and find that the main results persist. Specifically, a one-standard-deviation worsening in the imperfect-match index is associated with a reduction in firms' tangible assets of 2.7 percentage points, which is 1.2 times a standard deviation and again an economically significant effect.

We confirm the robustness of our findings to using an OLS estimator. The results are in table A4 and are both quantitatively and qualitatively similar to the baseline IV estimates. We continue to find that large deviations from the perfect match are associated with declines in both employment and tangible assets. Interestingly, the OLS coefficients are somewhat larger in absolute value than the ones obtained using IV, suggesting that match quality deteriorates, especially for firms with worse investment and growth opportunities.

²⁰We find consistent results in our sample. Table A2 shows that firms borrowing from EBA banks experience a lower credit growth relative to other firms.

Next, we analyze the extent to which a change in the total number of lenders or a substitution of banks drive the variation in match quality, keeping the number of lenders constant. In the former scenario, it would be difficult to disentangle the effect of a change in the imperfect-match index from that of a contraction in credit stemming from the termination of a credit relationship. In principle, firms may be negatively affected by the termination of credit relationships if they cannot promptly substitute them by increasing credit from other banks or by starting new relationships, which may prove difficult during economic downturns. We explore this question by estimating the same IV model on the sample of firms that switch lenders, keeping the total number of lenders constant. This is a powerful test because it allows us to examine the effect of the quality of the match controlling for the total number of credit relationships, and thus, in principle, credit availability. We show the results of this exercise in table 10. As is evident in panel A, our instrument is valid, powerful, and in line with the first-stage results in table 9. In the second stage (panel B), the estimated coefficient on the instrumented imperfect-match index is negative and statistically significant at 1% across all specifications. The point estimates are somewhat smaller than those in the baseline models, but statistical significance is stronger in the case of investment for multiple lenders (column VI).

Focusing on the differential role of relationship lending, we show that firms with single relationships remain strongly affected. However, we also find, in contrast to table 10, a negative and significant effect on the level of employment and investment rate for firms with multiple relationships that switch to new lenders during the crisis. This finding sheds light on the real effects of switching banks during adverse economic events. For instance, [Ioannidou and Ongena \(2010\)](#), among others, document the loan-pricing advantage of switching lenders. We show that, in crisis times, switching lenders may come at the cost of forming a worse match, leading to lower credit access with associated negative real effects on investment and employment. Our findings contribute to the relationship lending literature. They suggest that the relative bank-firm characteristics matter for the credit relationships in crisis times and not the anticipated loss of information that reduces credit for firms looking for a new relationship.

Finally, we test the effects of the imperfect bank-firm matches in crisis times on firms' survival prospects. Given that firm closures are a major concern during extreme events (including the current Covid-19 crisis), we replicate the models from tables 9 and 10 in table 11, using the probability of default as the dependent variable. We find that the chances of failure are positively

associated with the imperfect-match index. Moreover, this finding is stronger for firms with single bank relationships. Overall, imperfect bank-firm matches not only harm employment and growth, but also reduce firms' survival rate.

5 Robustness

We conduct various robustness tests for the results in the previous section. We summarize the results of these tests below and show them in the online appendix.

5.1 Alternative specifications

We examine whether our baseline results remain unchanged when we employ an additional set of fixed effects, a different estimation method, and further firm- and bank-specific variables. We re-estimate the empirical models from table 4 and report the results in table A5. Column I includes industry-location-size and year fixed effects, column II shows results from a probit model, column III clusters standard errors at the bank-firm level, column IV includes further firm characteristics, and column V incorporates additional bank characteristics. In summary, the results of the baseline model are qualitatively and quantitatively the same.

5.2 Removing Lisbon and Porto

In order to confirm that our results are not driven by how we define the geographical area of potential and realized matches, we remove Lisbon and Porto from the sample. The rationale is that these are the two largest cities in Portugal, for which the definition of postcodes at the four-digit level may yield excessively small areas. We reproduce the baseline model that validates the bank-firm matching after removing the two cities from our sample and report the estimates in table A6. Next, we run the real effect test excluding Lisbon and Porto; the results shown in table A7 are unchanged.

5.3 Placebo test: Real effects in pre-crisis period

In this section we validate the choice of using the pre-crisis period as a counterfactual scenario to predict bank-firm matches in the crisis period. If imperfect matches had a real effects already in

the pre-crisis period, then firms likely faced constraints in forming matches with banks.

To test whether the imperfect matching index has real effects in the pre-crisis period, we aggregate the credit registry data at the firm-year level and conduct a test by regressing firms' access to credit and growth perspectives on the imperfect matching index. We construct the imperfect matching index for the pre-crisis period and then measure how much the realized matches in the pre-crisis years differ from those predicted for the same period. Notably, the index for the pre-crisis period has a mean equal to 0.03 and less sizeable variation of 0.06. This suggests that the imperfect match index has good in-sample prediction properties.

Table A8 shows results. Column I reports the estimates for the placebo test on the log of outstanding amount that a firm receives. The coefficient estimate on the matching index exhibits no statistical significance, and a magnitude that is close to zero. In columns II to IV, we replace the dependent variable with firm's employment, investment, and survival. Interestingly, all coefficients are statistically insignificant and close to zero, except for the regression of the investment rate, in which the imperfect match index is marginally significant at 10%. Overall, this test shows that the matching index in the pre-crisis period did not explain differences in firms' access to credit and growth perspectives: in pre-crisis years with fewer credit and liquidity frictions, events that create a bank-firm mismatch would have small consequences-since it could be unwound. This test validates the choice of the pre-crisis period as a counterfactual scenario for the crisis-period.

6 Conclusion

This paper studies the drivers of bank-firm matches and how changes in match characteristics affect firms' access to credit and firm growth, as measured by employment, investment rate, probability of default. The paper defines a set of potential matches based on banks active in the local credit market where a firm is based. Next, it estimates a model to study the drivers of bank-firm matches using data from the years preceding the global financial crisis. This period is characterized by moderate credit growth and the absence of credit or housing bubbles (Greenwood, Hanson, Shleifer and Sørensen, 2022); it can thus represent a benchmark for bank-firm match formation in normal times. This allows us to study how bank-firm matches in crisis times differ from model predictions and test whether these differences translate into lower credit growth, firm investment, employment,

and firm survival.

We find that observed matches between banks and firms are more frequent if banks are large (irrespective of firm size), bank capital is higher, the local credit market is less concentrated, and firms are less likely to default. Next, we document that in crisis times, matches that differ from predicted ones (i.e., matches that involve firms and banks with characteristics different from those that correlate with match formation in good times) are associated with lower firm credit growth, lower employment, lower investment, and higher probability of default. Importantly, these findings survive after controlling for the potential endogeneity of match-quality changes, and they hold keeping the number of bank relationships constant. This indicates that the effect of imperfect matches on firm outcomes is not just due to a drop in the number of lending relationships.

Our results extend the findings of the literature on relationship lending, opening up the black box of bank-firm matches. We find that the relative characteristics of banks and firms that lead to match formation matter for the efficiency of lending relationships in bad times. When matches are broken and substituted with new ones, the ensuing loss of soft information is amplified when new matches differ more in terms of relative characteristics, from the matches prevailing in good times. Even when matches are still in place, their quality may change because of changes in characteristics of banks and firms. This affects access to credit and firm growth.

We also provide insights to understand why disruptions in credit markets have such large and pervasive effects on economic growth. Even if firms are able to substitute banks in downturns, there is a loss in terms of access to credit and firm growth, when the relative characteristics of matched banks and firms differ from those that lead to match formation in good times. Finally, our findings suggest that policy interventions targeting bank and firm characteristics that lead to stable matches in lending relationships are critical to ensure a steady flow of credit to the real economy in crises times.

References

- Acharya, V., Hasan, I. and Saunders, A.: 2006, Should banks be diversified? Evidence from individual bank loan portfolios, *The Journal of Business* **79**(3), 1355–1412.
- Acharya, V. and Naqvi, H.: 2012, The seeds of a crisis: A theory of bank liquidity and risk taking over the business cycle, *Journal of Financial Economics* **106**(2), 349–366.
- Allen, F., Carletti, E. and Marquez, R.: 2011, Credit market competition and capital regulation, *The Review of Financial Studies* **24**(4), 983–1018.
- Antunes, A., Goncalves, H. and Prego, P.: 2016, Firm default probabilities revisited, *Economic Bulletin and Financial Stability Report*, Banco de Portugal .
- Asker, J., Farre-Mensa, J. and Ljungqvist, A.: 2015, Corporate investment and stock market listing: A puzzle?, *The Review of Financial Studies* **28**(2), 342–390.
- Baily, M. N., Hulten, C. and Campbell, D.: 1992, Productivity dynamics in manufacturing plants, *Brookings Papers on Economic Activity. Microeconomics* **1992**, 187–267.
- Balduzzi, P., Brancati, E. and Schiantarelli, F.: 2018, Financial markets, banks’ cost of funding, and firms’ decisions: Lessons from two crises, *Journal of Financial Intermediation* **36**, 1–15.
- Banerjee, R. N., Gambacorta, L. and Sette, E.: 2021, The real effects of relationship lending, *Journal of Financial Intermediation* **48**, 100923.
- Beck, T.: 2009, The econometrics of finance and growth, *Palgrave Handbook of Econometrics* pp. 1180–1209.
- Beck, T., Da-Rocha-Lopes, S. and Silva, A. F.: 2021, Sharing the pain? Credit supply and real effects of bank bail-ins, *The Review of Financial Studies* **34**(4), 1747–1788.
- 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.
- Bentolila, S., Jansen, M., Jiménez, G. and Ruano, S.: 2018, When credit dries up: Job losses in the great recession, *Journal of the European Economic Association* **16**(3), 650–695.
- Berger, A. N., Miller, N. H., Petersen, M. A., Rajan, R. G. and Stein, J. C.: 2005, Does function follow organizational form? Evidence from the lending practices of large and small banks, *Journal of Financial Economics* **76**(2), 237–269.
- Berry, S., Levinsohn, J. and Pakes, A.: 1995, Automobile prices in market equilibrium, *Econometrica* **63**(4), 841–890.
- Berton, F., Mocetti, S., Presbitero, A. F. and Richiardi, M.: 2018, Banks, firms, and jobs, *The Review of Financial Studies* **31**(6), 2113–2156.
- Blattner, L., Farinha, L. and Rebelo, F.: 2021, When losses turn into loans: The cost of undercapitalized banks, *accepted in American Economic Review* .
- Bolton, P., Freixas, X., Gambacorta, L. and Mistrulli, P. E.: 2016, Relationship and transaction lending in a crisis, *The Review of Financial Studies* **29**(10), 2643–2676.
- Bonfim, D., Nogueira, G. and Ongena, S.: 2021, Sorry, we’re closed: Bank branch closures and corporate credit conditions, *Review of Finance* **25**(4), 1211–1259.
- Chiappori, P.-A. and Salanié, B.: 2016, The econometrics of matching models, *Journal of Economic Literature* **54**(3), 832–61.
- Chodorow-Reich, G.: 2014, The employment effects of credit market disruptions: Firm-level evidence from the 2008–9 financial crisis, *The Quarterly Journal of Economics* **129**(1), 1–59.
- Chodorow-Reich, G., Darmouni, O., Luck, S. and Plosser, M.: 2022, Bank liquidity provision across the firm size distribution, *Journal of Financial Economics* **144**(3), 908–932.
- Cingano, F., Manaresi, F. and Sette, E.: 2016, Does credit crunch investment down? New evidence on the real effects of the bank-lending channel, *The Review of Financial Studies* **29**(10), 2737–2773.
- Cohen, J., Hachem, K. and Richardson, G.: 2021, Relationship lending and the great depression, *The Review of Economics and Statistics* **103**(3), 505–520.

- Crouzet, N. and Mehrotra, N. R.: 2020, Small and large firms over the business cycle, *American Economic Review* **110**(11), 3549–3601.
- Darmouni, O.: 2020, Informational frictions and the credit crunch, *The Journal of Finance* **75**(4), 2055–2094.
- Degryse, H., Karapetyan, A. and Karmakar, S.: 2021, To ask or not to ask? Bank capital requirements and loan collateralization, *Journal of Financial Economics* **142**(1), 239–260.
- Degryse, H., Kim, M. and Ongena, S.: 2009, *Microeconometrics of Banking Methods, Applications, and Results*, Oxford University Press.
- Degryse, H. and Ongena, S.: 2005, Distance, lending relationships, and competition, *The Journal of Finance* **60**(1), 231–266.
- DeJonghe, O., Dewachter, H., Mullier, K., Ongena, S. and Schepens, G.: 2020, Some borrowers are more equal than others: Bank funding shocks and credit reallocation, *Review of Finance* **24**(1), 1–43.
- Detragiache, E., Garella, P. and Guiso, L.: 2000, Multiple versus single banking relationships: Theory and evidence, *The Journal of Finance* **55**(3), 1133–1161.
- Diamond, D. W. and Rajan, R. G.: 2001, Liquidity risk, liquidity creation, and financial fragility: A theory of banking, *Journal of Political Economy* **109**(2), 287–327.
- Duchin, R., Ozbas, O. and Sensoy, B.: 2010, Costly external finance, corporate investment, and the subprime mortgage credit crisis, *Journal of Financial Economics* **97**(3), 418–435.
- Dwenger, N., Fossen, F. and Simmler, M.: 2020, Firms’ financial and real responses to credit supply shocks: Evidence from firm-bank relationships in Germany, *Journal of Financial Intermediation* **41**, 100773.
- Farinha, L. and Santos, J.: 2002, Switching from single to multiple bank lending relationships: Determinants and implications, *Journal of Financial Intermediation* **11**, 124–151.
- Fraisse, H., Lé, M. and Thesmar, D.: 2020, The real effects of bank capital requirements, *Management Science* **66**(1), 5–23.
- Goncharenko, R., Mamonov, M., Ongena, S., Popova, S. and Turdyeva, N.: 2022, Quo vadis? Evidence on new firm-bank matching and firm performance following, *Evidence on New Firm-Bank Matching and Firm Performance Following “Sin” Bank Closures (February 6, 2022)*.
- Granja, J., Leuz, C. and Rajan, R. G.: 2022, Going the extra mile: Distant lending and credit cycles, *The Journal of Finance* **77**, 1259–1324.
- Greenwood, R., Hanson, S. G., Shleifer, A. and Sørensen, J. A.: 2022, Predictable financial crises, *The Journal of Finance* **77**(2), 863–921.
- Gropp, R., Mosk, T., Ongena, S. and Wix, C.: 2019, Banks response to higher capital requirements: Evidence from a quasi-natural experiment, *The Review of Financial Studies* **32**(1), 266–299.
- Hauswald, R. and Marquez, R.: 2006, Competition and strategic information acquisition in credit markets, *The Review of Financial Studies* **19**(3), 967–1000.
- Holmstrom, B. and Tirole, J.: 1997, Financial intermediation, loanable funds, and the real sector, *The Quarterly Journal of Economics* **112**(3), 663–691.
- Ioannidou, V. and Ongena, S.: 2010, Time for a change: loan conditions and bank behavior when firms switch banks, *The Journal of Finance* **65**(5), 1847–1877.
- Iosifidi, M. and Kokas, S.: 2015, Who lends to riskier and lower-profitability firms? Evidence from the syndicated loan market, *Journal of Banking & Finance* **61**, S14–S21.
- Iyer, R., Peydro, J.-L., da Rocha-Lopes, S. and Schoar, A.: 2013, Interbank liquidity crunch and the firm credit crunch: Evidence from the 2007–2009 crisis, *The Review of Financial Studies* **27**(1), 347–372.
- Jiménez, G., Ongena, S., Peydró, J.-L. and Saurina, J.: 2014, Hazardous times for monetary policy: What do twenty-three million bank loans say about the effects of monetary policy on credit risk-taking?, *Econometrica* **82**(2), 463–505.
- Keeley, M. C.: 1990, Deposit insurance, risk, and market power in banking, *The American Economic Review* pp. 1183–1200.
- Levine, R.: 2005, Finance and growth: Theory and evidence, *Handbook of Economic Growth* **1**, 865–934.

- Mésonnier, J. and Monks, A.: 2015, Did the EBA capital exercise cause a credit crunch in the Euro area?, *International Journal of Central Banking* **11**, 75–117.
- Petersen, M. A. and Rajan, R. G.: 2002, Does distance still matter? The information revolution in small business lending, *The Journal of Finance* **57**(6), 2533–2570.
- Popov, A. and Rocholl, J.: 2018, Do credit shocks affect labor demand? Evidence from employment and wages during the financial crisis, *Journal of Financial Intermediation* **36**, 16–27.
- Schwert, M.: 2018, Bank capital and lending relationships, *The Journal of Finance* **73**(2), 787–830.
- Sette, E. and Gobbi, G.: 2015, Relationship lending during a financial crisis, *Journal of European Economic Association* **13**(3), 453–481.
- Stein, J. C.: 2002, Information production and capital allocation: Decentralized versus hierarchical firms, *The Journal of Finance* **57**(5), 1891–1921.

Figure 1: Business credit and asset prices growth

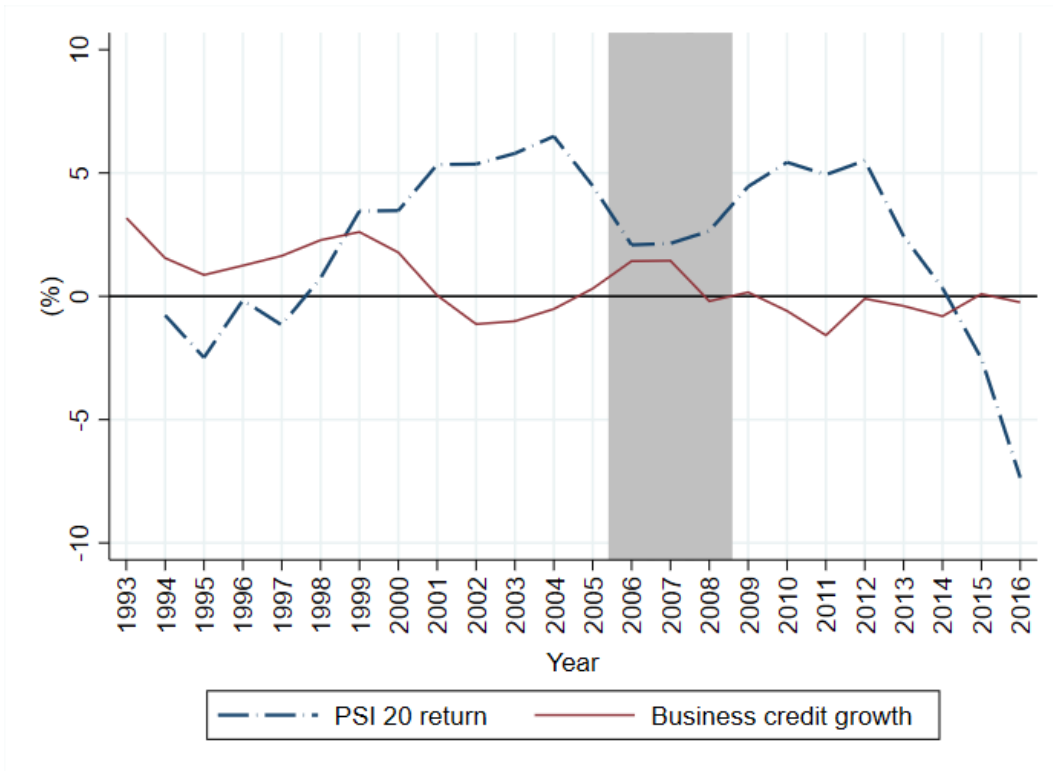


Figure 2: Household credit and household prices growth

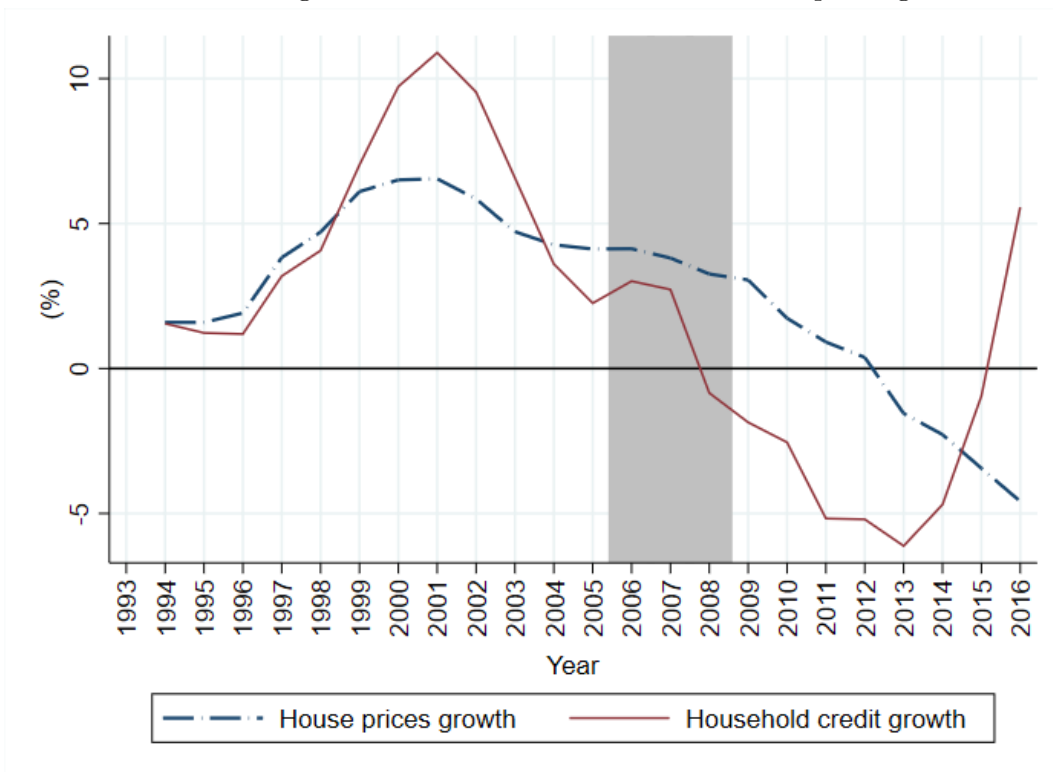


Figure 3: Realized and potential bank-firm matches

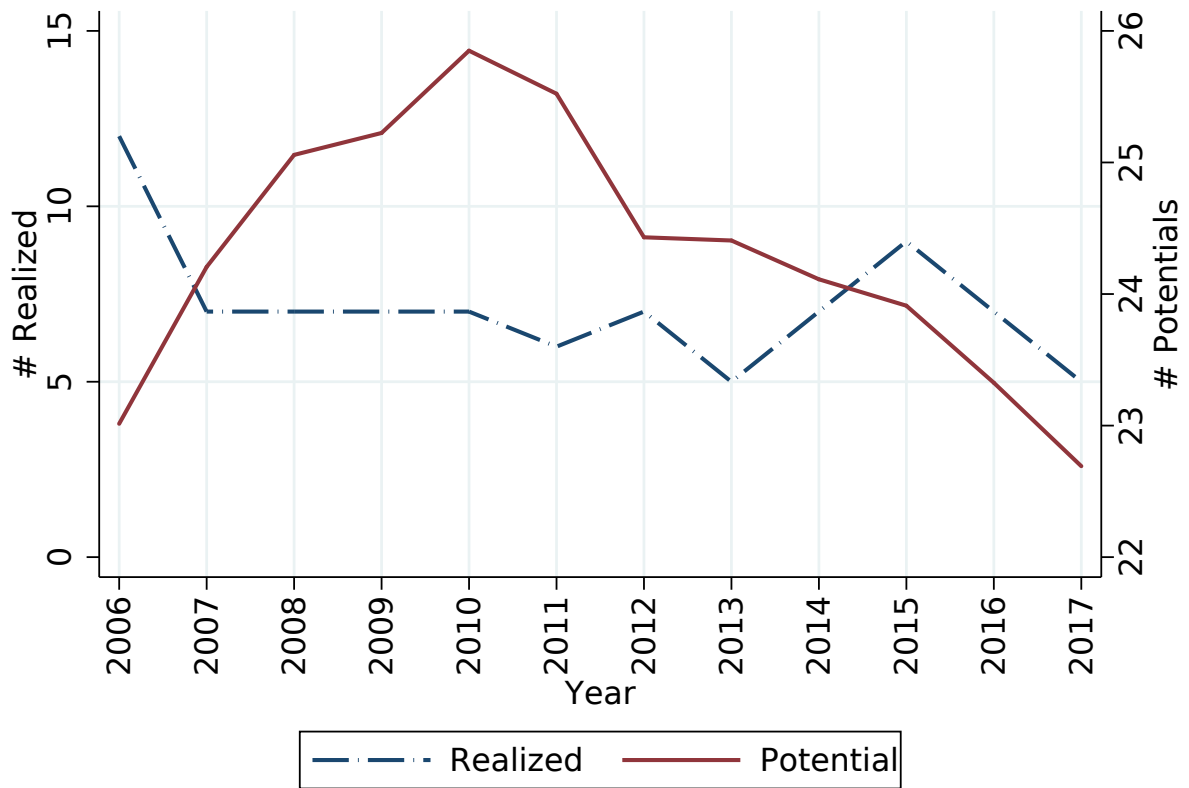


Figure 4: Imperfect-match index

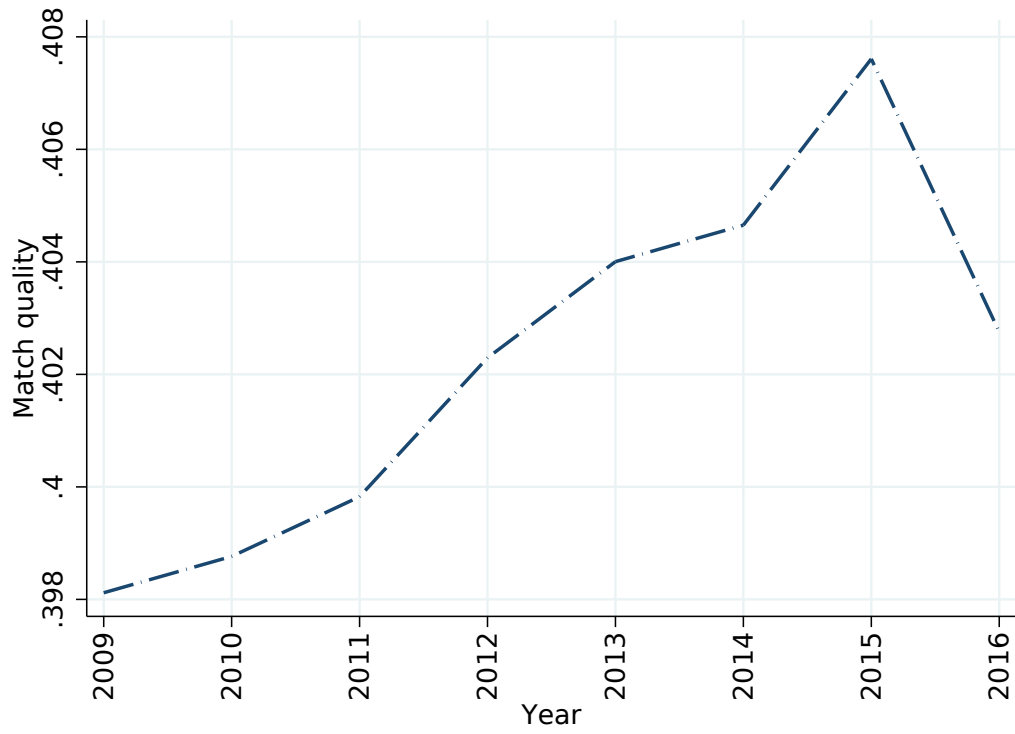


Figure 5: Imperfect-match index for small ($\leq 25^{th}$) and large ($\geq 75^{th}$) firms



Tables

Table 1: Bank-branch information

Year	# of banks	# of branches	# of branches per bank	# of branches per bank-zipbase
2006	410	3,184	653	4
2007	415	3,357	673	4
2008	429	3,457	690	4
2009	423	3,492	692	4
2010	412	3,492	689	4
2011	402	3,328	600	3
2012	393	3,279	590	3
2013	387	3,235	570	3
2014	379	3,117	535	3
2015	362	3,041	512	3
2016	338	2,977	490	3
2017	328	2,921	469	3

The table reports the evolution of bank-branch information.

Table 2: Summary statistics

	Obs.	Mean	Std.	Min	Max.
<i>Panel A: Bank-firm matching level</i>					
New relationship (0,1)	5,647,211	0.127	0.333	0.000	1.000
Number of possible matches	5,647,211	20.935	14.664	1.000	74.000
Number of realized matches	5,647,211	1.241	0.652	1.000	12.000
Large firm-Large bank	5,647,211	0.242	0.428	0.000	1.000
Small firm-Small bank	5,647,211	0.247	0.431	0.000	1.000
Small firm-Large bank	5,647,211	0.258	0.438	0.000	1.000
Large firm-Small bank	5,647,211	0.252	0.434	0.000	1.000
Imperfect-match index	2,937,273	0.151	0.148	0.010	0.811
Outstanding amount (EUR)	418,535	85,851	178,389	1.000	800M
Termination of lending (0,1)	802,049	0.279	0.449	0.000	1.000
Switching lender (0,1)	802,049	0.229	0.420	0.000	1.000
Capital ratio	5,647,211	0.250	0.501	-0.728	5.017
Ln(bank assets)	5,647,211	9.407	1.994	2.779	13.916
Ln(deposits)	5,647,211	4.278	2.616	0.074	9.276
Bank cash	5,647,211	4.413	2.906	-4.186	8.252
HHI (branches)	5,647,211	0.567	0.261	0.125	2.234
<i>Panel B: Firm level</i>					
Imperfect-match index	312,444	0.156	0.060	0.010	0.811
Prob(default)	618,067	0.055	0.065	0.000	0.905
Ln(number of employees)	544,480	1.429	1.161	0.000	9.624
Ln(fixed tangible assets)	551,318	10.371	2.238	-4.605	22.257
EBA borrowing share	619,241	0.113	0.685	0.000	1.000
Ln(firm assets)	619,241	12.155	1.874	-4.605	23.262
Ln(ROA)	619,241	-3.528	1.802	-43.185	14.128
Ln(leverage)	619,241	-0.319	0.945	-17.658	16.559
Ln(turnover)	572,829	12.072	1.801	-4.605	22.988
Ln(total expenses)	612,432	12.074	1.785	-4.605	23.000

The table provides basic descriptive statistics. See online appendix [A1](#) for precise definitions of the variables.

Table 3: Total number of realized matches within firm-year

#	Freq.	Percent	Cum.
1	690,488	69.83	69.83
2	178,496	18.05	87.89
3	65,424	6.62	94.5
4	29,012	2.93	97.44
5	14,250	1.44	98.88
6	6,840	0.69	99.57
7	2,814	0.28	99.85
8	944	0.10	99.95
9	297	0.03	99.98
10	130	0.01	99.99
11	44	0.00	100
12	24	0.00	100
Total	988,763	100	
Unique number of banks: 453			
Unique number of firms: 512,446			

The table reports the distribution of the total number of realized matches in the final sample.

Table 4: Bank-firm matching

	I	II	III	IV	V	VI	VII
Large_large	0.119*** [267.193]	0.102*** [209.097]	0.031*** [31.051]	0.031*** [31.051]	0.033*** [32.985]	0.017*** [16.247]	0.013*** [8.096]
Small_large	0.089*** [210.030]	0.083*** [80.384]	0.012*** [8.797]	0.012*** [8.797]	0.026*** [26.676]		0.012*** [6.973]
Small_small	-0.025*** [-68.130]	-0.015*** [-15.093]	-0.015*** [-14.801]	-0.015*** [-14.801]		-0.018*** [-17.806]	-0.038*** [-33.249]
Capital ratio	-0.024*** [-58.586]	-0.012*** [-26.859]	0.003*** [2.762]	0.003*** [2.762]	0.004*** [3.178]		0.012*** [6.326]
HHI	-0.016*** [-37.826]	-0.014*** [-8.225]	-0.012*** [-7.470]	-0.012*** [-7.470]		-0.008*** [-4.817]	-0.014*** [-8.602]
Prob(default)	0.044*** [18.452]	-0.060*** [-10.751]	-0.060*** [-10.864]	-0.060*** [-10.864]		-0.060*** [-10.974]	-0.060*** [-11.122]
Observations	5,013,829	5,011,739	5,011,739	5,011,739	5,010,697	5,011,739	3,049,146
R-squared	0.038	0.082	0.099	0.099	0.118	0.111	0.467
Year FE	Y	Y	Y	Y			Y
Firm FE		Y	Y	Y		Y	
Bank FE			Y	Y	Y		
Location FE				Y	Y	Y	Y
Firm*Year FE					Y		
Bank*Year FE						Y	
Firm*Bank FE							Y
Cluster SE	Robust	Robust	Robust	Robust	Robust	Robust	Robust

The table reports coefficients and t -statistics (in brackets). We estimate the regression: $Prob(match)_{b,f,l,t} = \lambda_1 * (Firm\ Size_{f,l,t} * Bank\ Size_{b,l,t}) + \lambda_2 * X_{b,f,l,t} + a_0 + \epsilon_{b,f,l,t}$. We estimate all specifications using a linear probability model, where the dependent variable is a dummy that equals one if the bank (b) - firm (f) matching is identified in the credit registry within a four-digit post-code (l) at time t , and zero if they do not match. To capture size dependence, we use interactions between firm and bank size. For instance, “Large_large” is the interaction between a large firm (above the median of the distribution of total assets) and a large bank (above the median of the distribution of total capital). To avoid the dummy trap, we exclude the “Large_small” interaction. We measure bank capital as the ratio of book equity to total assets. Herfindahl-Hirschman index (HHI) is the branch concentration per bank at the four-digit post code. As for firms’ riskiness, we incorporate the Banco de Portugal’s estimation of the borrower’s probability of default. We include fixed effects as noted in the lower part of the table to control for different levels of unobserved heterogeneity. The *, ** and *** marks denote the statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 5: Bank-firm matching: Heterogeneous effect

	I	II	III
Large_large	0.032*** [31.892]	0.032*** [31.837]	0.023*** [22.864]
Small_large	0.011*** [8.186]	0.003** [1.970]	-0.009*** [-5.720]
Small_small	-0.016*** [-15.644]	-0.024*** [-21.245]	-0.026*** [-22.064]
Capital ratio	-0.001 [-1.107]	0.003** [2.433]	-0.029*** [-23.619]
HHI	-0.013*** [-7.602]	-0.012*** [-7.454]	-0.012*** [-7.150]
Prob(default)	-0.061*** [-11.066]	-0.134*** [-18.457]	0.084*** [10.038]
Small_firm * Capital_ratio	0.008*** [9.440]		
Small_firm * Prob(default)		0.172*** [16.924]	
Large_firm * Prob(default)			-0.185*** [-16.120]
Large_firm * High_capital			-0.005*** [-5.757]
High_capital * Prob(default)			-0.082*** [-12.514]
Large_firm * High_capital * Prob(default)			0.025** [2.414]
Observations	5,011,739	5,011,739	5,011,739
R-squared	0.099	0.099	0.099
Year FE	Y	Y	Y
Firm FE	Y	Y	Y
Bank FE	Y	Y	Y
Location FE	Y	Y	Y
Cluster SE	Robust	Robust	Robust

The table reports coefficients and t -statistics (in brackets) for firms with single and multiple lending relationships. We estimate the regression: $Prob(match)_{b,f,l,t} = \lambda_1 * (Firm\ Size_{f,l,t} * Bank\ Size_{b,l,t}) + \lambda_2 * X_{b,f,l,t} + a_0 + \epsilon_{b,f,l,t}$. We estimate all specifications using a linear probability model, where the dependent variable is a dummy that equals one if the bank (b) - firm (f) match is identified in the credit registry within a four-digit post code (l) at time t , and zero if they do not match. To capture size dependence, we use interactions between firm and bank size. For instance, “Large_large” is the interaction between a large firm (above the median of the distribution of total assets) and a large bank (above the median of the distribution of total bank assets). To avoid the dummy trap we exclude the “Large_small” interaction. We include fixed effects as noted in the lower part of the table to control for different levels of unobserved heterogeneity. The *, ** and *** marks denote the statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 6: Decomposition of the changes in the imperfect-match index

Decomposition of the change in the imperfect-match index between 2009 and 2016		
<i>Mean of imperfect-match index (Year=2009): 0.160</i>		
<i>Mean of imperfect-match index (Year=2016): 0.190</i>		
Components	Absolute difference	Proportion (%)
Firm and bank characteristics (Block 1)	0.0272	87.37
Changes in the share of credit (Block 2)	-0.0000	-0.09
New relationships opened (Block 3)	0.0042	13.52
Relationships closed (Block 4)	-0.0002	-0.80
Overall	0.0312	100

The table reports the decomposition of the change in the imperfect-match index between 2009 and 2016. Each line reports the average of each block across firms.

Table 7: Imperfect-match index and credit supply

	I	II	III	IV	V
Imperfect match	-4.563***	-0.340***	-0.706***	-1.429***	-1.427***
	[-130.159]	[-4.642]	[-4.755]	[-6.258]	[-6.251]
# of bank-branches					0.058 [0.537]
Observations	258,627	130,398	31,043	38,698	38,698
R-squared	0.108	0.651	0.704	0.708	0.708
Year FE	Y	Y		Y	Y
Bank FE	Y	Y			
Firm FE		Y			
Location FE		Y	Y	Y	Y
Firm*Year FE			Y		
Bank*Year FE			Y		
Firm*Bank FE				Y	Y
Cluster SE	Robust	Robust	Robust	Robust	Robust

The table reports coefficients and t -statistics (in brackets). We estimate the regression: $Y_{b,f,t} = \alpha_0 + \beta_1 * Imperfect - match_{b,f,t} + \epsilon_{b,f,t}$. In all specifications the dependent variable is the outstanding amount of credit of the realized firm (f) - bank (b) matches that operate in the same four-digit postcode (l) at time (t). The main explanatory variable is the imperfect-match index. For the index, larger deviations from zero indicate a wider difference of observed matches from those prevailing in good times. We include fixed effects as noted in the lower part of the table to control for different levels of unobserved heterogeneity. The *, ** and *** marks denote the statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 8: Switching lenders and terminating relationships: Loan-level evidence

	I	II	III	IV	V	VI
Dependent variable	Switching lender			Termination of lending		
Imperfect match	0.019*** [5.810]	0.054*** [14.242]	0.053*** [14.039]	0.057*** [17.187]	0.101*** [25.928]	0.100*** [25.608]
Capital ratio	-0.048*** [-4.410]	-0.040*** [-3.149]		-0.051*** [-4.612]	-0.044*** [-3.332]	
HHI branch	-0.039*** [-3.817]	-0.045*** [-3.630]	-0.020 [-1.466]	-0.043*** [-4.062]	-0.051*** [-3.983]	-0.022 [-1.580]
Prob(default)	0.149*** [6.513]			0.159*** [6.848]		
Ln(firm assets)	0.013*** [6.089]			0.016*** [7.579]		
Ln(bank assets)	0.022*** [4.195]	0.036*** [5.657]		0.025*** [4.596]	0.038*** [5.752]	
Observations	297,301	252,610	252,567	297,301	252,610	252,567
R-squared	0.443	0.452	0.455	0.435	0.444	0.448
Year FE	Y			Y		
Firm FE	Y			Y		
Bank FE	Y	Y		Y	Y	
Location FE	Y	Y	Y	Y	Y	Y
Firm*Year FE		Y	Y		Y	Y
Bank*Year FE			Y			Y
Cluster SE	Robust	Robust	Robust	Robust	Robust	Robust

The table reports coefficients and t -statistics (in brackets). We estimate the regression: $Y_{b,f,t} = \alpha_0 + \beta_1 * Imperfect - match_{b,f,t} + \epsilon_{b,f,t}$. We estimate all specifications using a linear probability model, where the dependent variables are switching lenders (columns I to III) and termination of lending (columns IV to VI). The main explanatory variable is the imperfect-match index that is the difference between the observed and predicted matches. For the index, larger deviations from zero indicate a wider difference of realized matches from those prevailing in good times. We include fixed effects as noted in the lower part of the table to control for different levels of unobserved heterogeneity. The *, ** and *** marks denote the statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 9: The real effects of the imperfect-match index: Firm-level evidence

		Panel A: First stage					
		I	II	III	IV	V	VI
Dependent variable	Imperfect match						
Group	Full sample	Single	Multiple	Full sample	Single	Multiple	
EBA borrowing share	0.003*** [4.352]	0.004*** [4.390]	-0.009 [-1.455]	0.001*** [4.096]	0.001** [2.465]	-0.003 [-0.969]	
		Panel B: Second stage					
		I	II	III	IV	V	VI
Dependent variable	Ln(# of employees)			Ln(fixed tangible assets)			
Group	Full sample	Single	Multiple	Full sample	Single	Multiple	
$\widehat{Imperfect\ Match}$	-5.300*** [-10.105]	-5.339*** [-9.996]	0.292 [0.033]	-16.318*** [-14.892]	-16.767*** [-14.890]	-0.407 [-0.014]	
Firm control variables	Y	Y	Y	Y	Y	Y	
Observations	134,267	115,359	21,297	131,204	112,528	20,967	
R-squared	0.936	0.935	0.254	0.908	0.908	0.258	
Year FE	Y	Y	Y	Y	Y	Y	
Firm FE	Y	Y	Y	Y	Y	Y	
LM-test for under identification	974	952	5.771	903	882	1.29	
P-value for under identification	0.000	0.000	0.016	0.000	0.000	0.256	
F-stat for weak identification	116.85	983	3.018	114.3	909	0.668	
Weak identification 10% CR	16.38	16.38	16.38	16.38	16.38	16.38	
Cluster SE	Robust	Robust	Robust	Robust	Robust	Robust	

The table reports coefficients and t -statistics (in parenthesis) using a 2SLS regression. We estimate the following IV set-up: $Y_{f,t} = \alpha_0 + \beta_1 * \widehat{Imperfect\ Match}_{f,t} + \beta_2 * F_{f,t} + \mu_f + \mu_t + \epsilon_{f,t}$, with the first step: $\widehat{Imperfect\ Match}_{b,f,t} = \alpha_0 + \rho * EBA\ borrowing\ share_{f,t} + \gamma * F_{f,t} + \eta_{b,f,t}$. In both steps, we include firm-level controls for firm size, ROA, and leverage. For our instrument, we follow [Gropp, Mosk, Ongena and Wix \(2019\)](#) and construct it as $EBA\ borrowing\ share_{f,t} = \frac{\sum_{EBA} Outstanding\ amount_{f,t}}{\sum_{All\ bank} Outstanding\ amount_{f,t}}$, where the numerator is the average amount of outstanding credit of firm f from EBA exercised banks, and the denominator is the total amount of credit from all banks. We report the first-stage regressions in panel A. The LM statistic is distributed as chi-square under the null that the equation is unidentified. The F -stat is distributed as chi-square under the null of exogeneity. We include fixed effects as noted in the lower part of the table to control for different levels of unobserved heterogeneity. The *, **, *** marks denote the statistical significance at the 10%, 5%, and 1% level, respectively.

Table 10: The real effects of the imperfect-match index: Firm-switcher-level evidence

		Panel A: First stage					
		I	II	III	IV	V	VI
Dependent variable	Imperfect match						
Group	Full sample	Single	Multiple	Full sample	Single	Multiple	
EBA borrowing share	0.003*** [6.155]	0.003*** [4.863]	0.001*** [5.876]	0.002*** [3.234]	0.003*** [4.607]	0.001*** [5.550]	
		Panel B: Second stage					
		I	II	III	IV	V	VI
Dependent variable	Ln(# of employees)			Ln(fixed tangible assets)			
Group	Full sample	Single	Multiple	Full sample	Single	Multiple	
$\widehat{Imperfect\ Match}$	-3.441*** [-5.346]	-3.380*** [-4.502]	-5.313*** [-3.648]	-6.932*** [-5.543]	-7.465*** [-5.039]	-6.242** [-2.567]	
Firm control variables	Y	Y	Y	Y	Y	Y	
Observations	57,909	50,149	7,734	58,071	50,325	7,723	
R-squared	0.292	0.313	0.202	0.247	0.269	0.249	
Year FE	Y	Y	Y	Y	Y	Y	
Firm FE	Y	Y	Y	Y	Y	Y	
LM-test for under identification	162.8	141.6	38.66	162.5	155.3	24.18	
P-value for under identification	188.6	148.4	34.17	166.3	129.7	30.61	
F-stat for weak identification	0.000	0.000	0.000	0.000	0.000	0.000	
Weak identification 10% CR	191.6	151	34.53	168.5	131.6	30.81	
Cluster SE	16.38	16.38	16.38	16.38	16.38	16.38	

The table reports coefficients and t -statistics (in parenthesis) using a 2SLS regression only for firms that switch banks. We estimate the following IV set-up: $Y_{f,t} = \alpha_0 + \beta_1 * \widehat{Imperfect\ Match}_{f,t} + \beta_2 * F_{f,t} + \mu_f + \mu_t + \epsilon_{f,t}$, with the first step: $Imperfect\ Match_{b,f,t} = \alpha_0 + \rho * EBA\ borrowing\ share_{f,t} + \gamma * F_{f,t} + \eta_{b,f,t}$. In both steps, we include firm-level controls for firm size, ROA, and leverage. For our instrument, we follow [Gropp, Mosk, Ongena and Wix \(2019\)](#) and construct it as $EBA\ borrowing\ share_{f,t} = \frac{\sum_{EBA} Outstanding\ amount_{f,t}}{\sum_{All\ bank} Outstanding\ amount_{f,t}}$, where the numerator is the average amount of outstanding credit of firm f from EBA exercised banks, and the denominator is the total amount of credit from all banks. We report the first-stage regressions in panel A. The LM statistic is distributed as chi-square under the null that the equation is unidentified. The F -stat is distributed as chi-square under the null of exogeneity. We include fixed effects as noted in the lower part of the table to control for different levels of unobserved heterogeneity. The *, ** and *** marks denote the statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 11: Imperfect-match index and the probability of default: Firm-level evidence

		Panel A: First stage					
		I	II	III	IV	V	VI
		All firms			Only for firms that switched lenders		
Dependent variable		Imperfect match					
Group		Full sample	Single	Multiple	Full sample	Single	Multiple
EBA borrowing share		0.001*** [8.443]	0.001*** [32.648]	0.000 [1.431]	0.001*** [29.620]	0.001*** [29.964]	0.001 [1.431]
		Panel B: Second stage					
		I	II	III	IV	V	VI
		All firms			Only for firms that switched lenders		
Dependent variable		Prob(default)					
Group		Full sample	Single	Multiple	Full sample	Single	Multiple
$\widehat{Imperfect\ Match}$		0.274*** [7.377]	0.269*** [6.972]	-0.854 [-0.688]	0.203*** [5.118]	0.194*** [4.823]	3.113 [0.611]
Firm control variables		Y	Y	Y	Y	Y	Y
Observations		148,238	128,056	22,543	55,172	47,092	9,736
R-squared		0.697	0.685	0.0647	0.708	0.719	-11.898
Year FE		Y	Y	Y	Y	Y	Y
Firm FE		Y	Y	Y	Y	Y	Y
LM-test for under identification		198.8	110.3	40.65	142.6	192.4	0.395
P-value for under identification		0.000	0.000	0.000	0.000	0.000	0.530
F-stat for weak identification		114.8	62.42	40.99	81.19	108.8	0.382
Weak identification 10% CR		16.38	16.38	16.38	16.38	16.38	16.38
Cluster SE		Robust	Robust	Robust	Robust	Robust	Robust

The table reports coefficients and t -statistics (in parenthesis) using a 2SLS regression. We estimate the following IV set-up: $Y_{f,t} = \alpha_0 + \beta_1 * \widehat{Imperfect\ Match}_{f,t} + \beta_2 * F_{f,t} + \mu_f + \mu_t + \epsilon_{f,t}$, with the first step: $\widehat{Imperfect\ Match}_{b,f,t} = \alpha_0 + \rho * EBA\ borrowing\ share_{f,t} + \gamma * F_{f,t} + \eta_{b,f,t}$. In both steps, we include firm-level controls for firm size, ROA, and leverage. For our instrument, we follow [Gropp, Mosk, Ongena and Wix \(2019\)](#) and construct it as $EBA\ borrowing\ share_{f,t} = \frac{\sum_{EBA} Outstanding\ amount_{f,t}}{\sum_{All\ bank} Outstanding\ amount_{f,t}}$, where the numerator is the average amount of outstanding credit of firm f from EBA exercised banks, and the denominator is the total amount of credit from all banks. We report the first stage regressions in panel A. The LM statistic is distributed as chi-square under the null that the equation is unidentified. The F -stat is distributed as chi-square under the null of exogeneity. We include fixed effects as noted in the lower part of the table to control for different levels of unobserved heterogeneity. The *, ** and *** marks denote the statistical significance at the 10%, 5%, and 1% levels, respectively.

Online appendix

A1 Definitions of the variables used

- *New relationship*: is a dummy that equals one if the firm has a loan from a bank with which it had no relationship previously, and zero otherwise.
- *Match*: is a dummy that equals one if the bank (b) - firm (f) matching is identified in the credit registry within a four-digit post-code (l) at time t , and zero if they do not match.
- *Imperfect match*: is an index made up by the average predicted match - the average realized match, weighted by the share of credit in each relationship.
- *Termination of Lending*: is a dummy that equals one if the bank has terminated an existing relationship, and zero otherwise.
- *Switching*: is a dummy that equals one if we observe in the firm's credit registry a new loan from a bank with which it did not have a lending relationship during the previous twelve months, and zero otherwise.
- *Large firm*: is a dummy that equals one if the firm's real total assets are above the median size of all firms, and zero otherwise.
- *Small firm*: is a dummy that equals one if the firm's real total assets are below the median size of all firms, and zero otherwise.
- *Large bank*: is a dummy that equals one if the bank's total assets are above the median assets of all banks, and zero otherwise.
- *Small bank*: is a dummy that equals one if the bank's total assets are below the median assets of all banks, and zero otherwise.
- *ROA*: denotes the firm's return on assets.
- *Leverage*: is the firm's long term debt.
- *Prob (default)*: is measured as the probability that any given firm will have a significant default episode within a one-year horizon using information from the central credit register and comprehensive balance sheet data.
- *Capital ratio*: is the bank's ratio of book equity to total assets.
- *EBA bank*: is a dummy that equals one for all banks in Portugal included in the 2011 EBA capital exercise, and zero otherwise.
- *EBA borrowing share* $_{f,t} = \frac{\sum_{EBA} \text{Outstanding amount}_{f,t}}{\sum_{All\ bank} \text{Outstanding amount}_{f,t}}$; where the numerator is the average amount of outstanding credit of firm f from EBA exercised banks, and the denominator is the total amount of credit from all banks.
- *HHI*: denotes the Herfindahl-Hirschman index.

Table A1: Imperfect match index and loan outstanding amount: EU sovereign debt crisis

	I	II	III	IV	V
Imperfect match (2011)	-4.861***	-0.125***	-0.761***	-1.394***	-1.387***
	[-103.527]	[-3.984]	[-3.984]	[-4.140]	[-4.115]
# of bank-branches					0.304
					[1.029]
Observations	179,923	77,304	20,524	20,549	20,549
R-squared	0.106	0.681	0.717	0.730	0.730
F-stat	10.718	1.539	15.871	17.141	9.041
Year FE	Y	Y		Y	Y
Bak FE	Y	Y			
Firm FE		Y			
Locatios FE		Y	Y	Y	Y
Firm*Year FE			Y		
Bak*Year FE			Y		
Firm*Bak FE				Y	Y
Cluster SE	Robust	Robust	Robust	Robust	Robust

The table reports coefficients and t -statistics (in brackets). We estimate the regression: $Y_{b,f,t} = \alpha_0 + \beta_1 * Imperfect Match_{b,f,t} + \beta_2 * F_{f,t} + \mu_f + \mu_t + \epsilon_{b,f,t}$. We estimate all specifications using OLS, where the dependent variable is the outstanding amount of credit of the realized firm (f) - bank (b) matches that operate in the same four-digit postcode (l) at time (t). The main explanatory variable is the imperfect match index that is calculated as the difference between the observed and predicted matches. For the index, larger deviations from zero indicate a wider difference of realized matches from those prevailing in good times. The *, ** and *** marks denote the statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A2: EBA exercise and outstanding credit

	I	II
EBA exercise	-0.571*** [-2.951]	-0.614** [-2.653]
Capital ratio	0.493 [0.955]	1.965*** [8.186]
HHI	-0.024 [-1.334]	-0.008 [-0.442]
Ln(deposits)	-0.000 [-0.013]	-0.001* [-1.975]
Bank size	-0.336** [-2.593]	0.020 [0.118]
Observations	407,556	407,553
R-squared	0.020	0.046
F-stat	24.11	27.03
Year FE	Y	Y
Bank FE		Y
Cluster SE	Bank	Bank

The table reports coefficients and t -statistics (in brackets). We estimate the regression: $Y_{b,f,t} = \alpha_0 + \beta_1 * EBAexercise_{b,f,t} + \beta_2 * F_{f,t} + \mu_f + \mu_t + \epsilon_{b,f,t}$. We estimate all specifications using OLS, where the dependent variable is the outstanding amount of credit of the realized firm (f) - bank (b) matches that operate in the same four-digit postcode (l) at time (t). We include fixed effects as noted in the lower part of the table to control for different levels of unobserved heterogeneity. The *, ** and *** marks denote the statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A3: Imperfect match index: Bootstrap SE

	I	II	III	IV	V
Imperfect match	-4.563***	-0.340***	-0.760**	-1.429***	-1.427***
	[-96.576]	[-2.911]	[-2.310]	[-4.008]	[-3.862]
# of bank-branches					0.058
					[0.386]
Observations	258,627	130,398	31,043	38,698	38,698
R-squared	0.104	0.651	0.704	0.708	0.708
Wald (P-value)	0.000	0.004	0.002	0.000	0.001
Year FE	Y	Y		Y	Y
Bank FE	Y	Y			
Firm FE		Y			
Location FE		Y	Y	Y	Y
Firm*Year FE			Y		
Bank*Year FE			Y		
Firm*Bank FE				Y	Y
Cluster SE	Bootstrap	Bootstrap	Bootstrap	Bootstrap	Bootstrap

The table reports coefficients and t -statistics (in brackets). The dependent variable is the outstanding amount for the realized firm (f) - bank (b) matches that operate in the same four-digit postcode (l) at time (t). The main explanatory variable is the imperfect match index that is calculated as the difference between the observed and predicted matches. For the index, larger deviations from zero indicate a wider difference of realized matches from those prevailing in good times. We include fixed effects as noted in the lower part of the table to control for different levels of unobserved heterogeneity. The *,** and *** marks denote the statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A4: Imperfect-match index and real effects: Firm-level OLS estimates

	I	II	III	IV	V	VI
Dependent variable	Ln(# of employees)			Ln(fixed tangible assets)		
Group	Full sample	Single	Multiple	Full sample	Single	Multiple
Imperfect match	-5.581*** [-136.321]	-6.074*** [-136.138]	-4.394*** [-36.860]	-11.736*** [-143.236]	-12.816*** [-141.528]	-8.026*** [-37.357]
Firm control variables	Y	Y	Y	Y	Y	Y
Observations	279,000	257,691	21,309	267,530	246,548	20,982
R-squared	0.119	0.136	0.085	0.146	0.164	0.109
Year FE	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	Y
Cluster SE	Robust	Robust	Robust	Robust	Robust	Robust

The table reports coefficients and t -statistics (in parenthesis) using a 2SLS regression. We estimate the following set-up: $Y_{f,t} = \alpha_0 + \beta_1 * Imperfect\ Match_{f,t} + \beta_2 * F_{f,t} + \mu_f + \mu_t + \epsilon_{f,t}$. We include firm-level controls for firm size, ROA, and leverage. The main explanatory variable is the imperfect match index that is calculated as the difference between the observed and predicted matches. For the index, larger deviations from zero indicate a wider difference of realized matches from those prevailing in good times. We include fixed effects as noted in the lower part of the table to control for different levels of unobserved heterogeneity. The *, **, and *** marks denote the statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A5: Bank-firm matching: Alternative tests

	I	II	III	IV	V
Large_large	0.032*** [37.742]	0.114*** [28.096]	0.031*** [30.983]	0.031*** [29.326]	0.023*** [22.830]
Small_large	0.023*** [28.400]	0.003 [0.858]	0.012*** [9.002]	0.015*** [10.536]	0.004*** [3.175]
Small_small		-0.196*** [-74.567]	-0.015*** [-15.271]	-0.012*** [-10.861]	-0.015*** [-14.803]
Capital ratio		0.149*** [13.841]	0.003*** [2.787]	0.004*** [3.003]	0.006*** [5.016]
HHI		-0.032*** [-15.603]	-0.012*** [-8.018]	-0.013*** [-7.715]	-0.013*** [-7.727]
Prob(default)		0.197*** [17.281]	-0.060*** [-11.248]	-0.057*** [-9.020]	-0.060*** [-10.868]
Ln(turnover)				-0.001** [-1.980]	
Ln(total expenses)				0.008*** [9.250]	
Ln(deposits)					0.006* [1.901]
Bank cash					0.000*** [23.257]
Observations	5,645,040	4,977,513	5,011,739	4,616,007	5,011,739
R-squared	0.097		0.099	0.100	0.099
X-sq (Probit)		203174			
Year FE	Y	Y	Y	Y	Y
Firm FE	Y		Y	Y	Y
Bank FE	Y	Y	Y	Y	Y
Location FE	Y	Y	Y	Y	Y
Industry*Location*Size*Year FE	Y				
Cluster SE	Robust	Robust	Bank*Firm	Robust	Robust

The table reports coefficients and t -statistics (in brackets). We estimate the regression: $Prob(match)_{b,f,l,t} = \lambda_1 * (Firm\ Size_{f,l,t} * Bank\ Size_{b,l,t}) + \lambda_2 * X_{b,f,l,t} + a_0 + \epsilon_{b,f,l,t}$. We estimate all specifications using a linear probability model, where the dependent variable is a dummy that equals one if the bank (b) - firm (f) match is identified in the credit registry within a four-digit post code (l) at time t , and zero if they do not match. To capture size dependence, we use interactions between firm and bank size. For instance, “Large_large” is the interaction between a large firm (above the median of the distribution of total assets) and a large bank (above the median of the distribution of total capital). To avoid the dummy trap we exclude the “Large_small” interaction. We include fixed effects as noted in the lower part of the table to control for different levels of unobserved heterogeneity. The *, ** and *** marks denote the statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A6: Imperfect-match index and credit supply: Excluding Lisbon and Porto

	I	II	III	IV	V
Imperfect match	-4.409***	-0.279***	-0.744***	-1.328***	-1.3270***
	[-118.74]	[-3.58]	[-4.69]	[-5.49]	[-5.48]
# of bank-branches					0.115
					[0.543]
Observations	218,083	110,594	26,267	32,708	32,708
R-squared	0.106	0.653	0.705	0.711	0.711
F-stat	20.411	12.84	22.000	30.11	15.20
Year FE	Y	Y		Y	Y
Bank FE	Y	Y			
Firm FE		Y			
Location FE		Y	Y	Y	Y
Firm*Year FE			Y		
Bank*Year FE			Y		
Firm*Bank FE				Y	Y
Cluster SE	Robust	Robust	Robust	Robust	Robust

The table reports coefficients and t -statistics (in brackets). We estimate the regression: $Y_{b,f,t} = \alpha_0 + \beta_1 * Imperfect\ match_{b,f,t} + \mu_f + \mu_t + \epsilon_{b,f,t}$. We estimate all specifications using OLS, where the dependent variable is the outstanding amount of credit of the realized firm (f) - bank (b) matches that operate in the same four-digit postcode (l) at time (t). The main explanatory variable is the imperfect match index that is calculated as the difference between the realized and predicted matches. For the index, larger deviations from zero indicate a wider difference of realized matches from those prevailing in good times. We include fixed effects as noted in the lower part of the table to control for different levels of unobserved heterogeneity. The *,** and *** marks denote the statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A7: Imperfect-match index and real effects: Firm-level evidence excluding Lisbon and Porto

		Panel A: First stage					
		I	II	III	IV	V	VI
Dependent variable		Imperfect match					
Group		Full sample	Single	Multiple	Full sample	Single	Multiple
EBA borrowing share		0.002*** [13.759]	0.002*** [13.920]	0.000 [0.462]	0.002*** [12.790]	0.002*** [13.006]	0.000 [0.273]
		Panel B: Second stage					
		I	II	III	IV	V	VI
Dependent variable		Ln(# of employees)			Ln(fixed tangible assets)		
Group		Full sample	Single	Multiple	Full sample	Single	Multiple
$\widehat{Imperfect\ Match}$		-8.470*** [-6.547]	-8.742*** [-6.768]	-59.992 [-0.483]	-32.894*** [-10.489]	-34.067*** [-10.885]	-107.162 [-0.256]
Firm control variables		Y	Y	Y	Y	Y	Y
Observations		115,346	99,243	18,056	112,683	96,799	17,777
R-squared		0.933	0.255	-10	0.911	0.91	-28.93
Year FE		Y	Y	Y	Y	Y	Y
Firm FE		Y	Y	Y	Y	Y	Y
LM-test for under identification		194.1	198.5	0.217	167.8	173.4	0.075
P-value for under identification		0.000	0.000	0.641	0.000	0.000	0.783
F-stat for weak identification		189.3	193.8	0.213	163.6	169.2	0.074
Weak identification 10% CR		16.38	16.38	16.38	16.38	16.38	16.38
Cluster SE		Robust	Robust	Robust	Robust	Robust	Robust

The table reports coefficients and t -statistics (in parenthesis) using a 2SLS regression. We estimate the following IV set-up: $Y_{f,t} = \alpha_0 + \beta_1 * \widehat{Imperfect\ Match}_{f,t} + \beta_2 * F_{f,t} + \mu_f + \mu_t + \epsilon_{f,t}$, with the first step: $\widehat{Imperfect\ Match}_{b,f,t} = \alpha_0 + \rho * EBA\ borrowing\ share_{f,t} + \gamma * F_{f,t} + \eta_{b,f,t}$. In both steps, we include basic firm-level controls for firm size, ROA, and leverage. For our instrument, we follow [Gropp, Mosk, Ongena and Wix \(2019\)](#) and construct it as $EBA\ borrowing\ share_{f,t} = \frac{\sum_{EBA} Outstanding\ amount_{f,t}}{\sum_{All\ bank} Outstanding\ amount_{f,t}}$, where the numerator is the average amount of outstanding credit of firm f from EBA exercised banks, and the denominator is the total amount of credit from all banks. We report the first stage regressions in panel A. We drop Lisbon and Port from the sample. The LM statistic is distributed as chi-square under the null that the equation is unidentified. The F -stat is distributed as chi-square under the null of exogeneity. We include fixed effects as noted in the lower part of the table to control for different levels of unobserved heterogeneity. The *, ** and *** marks denote the statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A8: Placebo test in pre-crisis period: Imperfect-match index and firm outcomes

	I	II	III	IV
Dependent variable	Ln(amount)	Ln(# of employees)	Ln(fixed tangible assets)	Prob(default)
Imperfect match	0.434 [0.893]	0.032 [0.878]	0.174* [1.879]	0.003 [0.850]
Firm control variables	Y	Y	Y	Y
Observations	36,663	89,547	94,670	100,802
R-squared	0.657	0.962	0.935	0.759
Year FE	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y
Cluster SE	Robust	Robust	Robust	Robust

The table reports coefficients and t -statistics (in brackets) when we aggregate the credit registry data at the firm year. We estimate the regression only for the non-crisis period (2006-2008): $Y_{f,t} = \alpha_0 + \beta_1 * Imperfect - match_{f,t} + \epsilon_{f,t}$. The dependent variables are reported in the second line. The main explanatory variable is the imperfect-match index. We include firm and year fixed effects as noted in the lower part of the table. The *,** and *** marks denote the statistical significance at the 10%, 5%, and 1% levels, respectively.