(Unobserved) Heterogeneity in the bank lending channel: Accounting for bank-firm interactions and specialization*

Bryan Gutierrez  Alonso Villacorta  Lucciano Villacorta
University of Minnesota  UC Santa Cruz  Central Bank of Chile

March 2023

Abstract

The bank lending channel is heterogeneous across firms. Using matched bank-firm lending data, we develop a framework that identifies firm-demand and bank-supply credit shocks. Bank shocks are allowed to have heterogeneous effects across (unobserved) firm groups, in a setup in which group membership is left unrestricted. We decompose credit growth dynamics into time-varying firm and time-varying bank-group effects, where grouping minimizes a least-squares criterion. Our results show significant heterogeneity in the bank lending channel: i) the effect of bank shocks varies considerably across the identified firm groups, ii) we quantify the importance of bank-firm relationships: aggregate credit growth drops by up to 20% when bank-firm relationships are randomly allocated, iii) the impact of bank shocks on firm investment is not significant and imprecisely estimated when homogeneity in the bank lending channel is assumed, while it is precisely estimated to an elasticity of 4 when heterogeneous effects are allowed.

JEL Classification: C23, E22, E51, G21.

Keywords: Bank lending channel, credit supply identification, unobserved heterogeneity, grouped-fixed effects.

*The views expressed in this paper are our own and do not necessarily coincide with those of the Banco Central de Chile. We would like to thank Stephane Bonhomme, Steve Davis, Galina Hale, Victoria Ivashina, Richard Thakor, Manuel Luy, Monika Piazzesi and seminar audiences at the University of Minnesota, the Central Bank of Chile, and UC Santa Cruz for helpful comments and discussions. Contact e-mails: avillaco@ucsc.edu.
1 Introduction

How important is the bank lending channel? That is, how much do bank credit supply shocks impact firms’ borrowing capacity, investment, and overall real activity?\(^1\) Is the bank lending channel heterogenous across firms? That is, do banks propagate their credit supply shocks differentially to different types of firms? How valuable are bank-firm relationships for the propagation of credit shocks? These questions are key to understand the propagation of financial shocks on real activity, so they stand at the center of the macrofinance literature.

Answering these questions empirically is challenging though. Credit fluctuations may be driven by demand or supply factors, and identifying bank shocks is difficult since events that influence the supply of credit are likely to influence the demand for credit as well. Recent developments in the literature exploit bank-firm credit data and, by observing multiple connections between firms and banks, are able to control for firm demand factors and extract exogenous variations in bank shocks from variations in lending. The method of Khwaja and Mian [2008] (KM hereafter) relies on using firm fixed effects to control for firm demand factors when assessing the impact of a bank shock, measured as a bank differential exposure to a liquidity shock during a natural experiment. This useful methodology has been widely extended to assess the impact of other types of bank shocks, in different event studies, on firm outcomes (see for example, Chodorow-Reich [2014], Paravisini et al. [2015], Jiménez et al. [2017], Bentolila et al. [2018], Jiménez et al. [2020]). Amiti and Weinstein [2018] (AW hereafter) extend this idea and exploit the time series variation of the data through two-way fixed effect regressions to account for unobservable time-varying bank and firm factors, which are then used to assess their impact of firm investment. While these methodologies provide a tractable way to deal with unobserved heterogeneity regarding firm and bank credit factors, they rely on a substantive assumption regarding the transmission of credit shocks, that is, that shocks to each bank are propagated equally to all its connected firms, and vice-versa, so that the within firm comparison fully absorbs firm-specific changes in credit demand. The absence of interactions between bank and firm factors in the KM or AW framework then restrict the possibility of an heterogenous transmission of shocks and the specific matching patterns between banks and firms.

However, the transmission of credit shocks may be heterogeneous due to interactions between bank and firm characteristics arising from lending specialization or relationship lending. Theories of relationship lending (Petersen and Rajan [1994], Peek and Rosengren [2005], Detragiache et al. [2000],

\(^1\) Throughout the paper we refer to bank credit supply shocks (or simply banks shocks) as shifts (contractions or expansions) in the supply of bank credit that capture changes in the health of banks’ balance sheets (for example, driven by bank-specific liquidity shocks, monetary policy shocks that affect banks ability to attract deposits, or changes in capital regulation) that are unrelated to firm demand for credit or creditworthiness.
Degryse and Ongena (2005) point towards a special treatment of bank to certain groups of firms: Banks are constantly faced with a portfolio allocation task in which they have to monitor and select their expansion or contraction of lending across borrowers. Such portfolio decision will depend on the borrowers’ characteristics and on the specific way each bank assesses them that, in particular, could differ across banks due to differences in bank information acquisition, lending specialization, or business models. In particular, recent empirical studies provide evidence of such differential response depending on some observable characteristics (Paravisini et al. [2021], Ivashina et al. [2022], Blickle et al. [2023]). For example, using detailed information on bank-firm credit data from Peru, Paravisini et al. [2021] show that bank specialization in export markets is a crucial determinant for the propagation of credit shocks. Also, using data from Peru, Ivashina et al. [2022] show that bank shocks affect differentially firms depending on their loan type demand (e.g. asset based loans or cash-flow based loans). And, Blickle et al. [2023] document loan specialization across industries among large U.S. banks. But, what if the econometrician cannot access such specific data about bank or firm specialization? Or, what if the heterogeneity is based on an unobserved margin? How can we learn about the heterogeneity in the transmission of shocks and its implications when there exists interactions between unobserved bank and firm factors?

In this paper, we develop a framework that identifies bank-supply and firm-demand credit shocks from bank-firm credit data allowing for interactions between bank and firm unobserved factors that could lead to heterogeneous transmission of credit shocks. In particular, the effect of credit shocks is allowed to vary across types or groups of firms, while maintaining the assumption of homogenous transmission within types of firms. The approach relies on the idea that from the perspective of the bank and its relationship with firms, there is a discrete number of types in which firms are classified regarding the relevant characteristics that would lead to a special treatment or an heterogenous transmission of a credit shock. Besides the heterogenous transmission across groups, the model allows for firm-specific demand shocks that affect equally all relationships. The heterogeneous effects of bank shocks, the firm-specific demand shocks, and the firm group membership are left unrestricted as in a “fixed-effects” (or “grouped fixed-effects”) fashion and are estimated from the data. Our methodology decomposes credit growth dynamics into time-varying firm and a time-varying bank-group effects, where grouping minimizes a least-squares criterion.

If groups were observed by the econometrician, then we could identify bank shocks by applying the KM or AW framework for each group, that is, by measuring the systematic differences in the lending of banks to the same set of firms belonging to the same group. However, since groups are unobserved, we combine these ideas with machine learning techniques that help in clustering observations (Hahn and Moon [2010], Bonhomme and Manresa [2015], Ando and Bai [2016], Bonhomme et al. [2019]). Typi-
cally, clustering techniques classify observations into groups depending on some dissimilarity measure based on observable characteristics. Our goal instead is to cluster firms based on an heterogeneous unobserved response to an unobserved credit shock. We follow Bonhomme and Manresa [2015] and adjust their proposed “grouped fixed-effects” estimator to our bank lending framework by incorporating the insights from KM and AW. Intuitively, firms whose differences in their borrowing patterns from the same specific banks are most similar are grouped together in estimation.

We apply our estimation framework to credit registry data from Peru from 2005 to 2017. We observe every bank-firm lending relationship for corporate firms, which under our framework allows us to identify bank-supply and firm-demand credit shocks allowing for unobserved heterogeneous effects. Additionally, we obtain the financial statements information from the Peruvian Stock Exchange to measure firm investment. Combining both sources of information, we are able to study the real effects of the supply lending shocks on firm investment across time.

Our novel framework allows us to uncover new results on the importance of the bank lending channel that cannot be uncovered without a more flexible treatment of unobserved heterogeneity. First, we provide novel evidence on the heterogeneous effect of bank shocks depending on the firm group. We show that the lending patterns of firms connected with each bank shows considerably heterogeneity across our identified groups. Such heterogeneity in lending patterns between groups of firms and banks leads to sizable biases of the estimated standard model under homogenous effects. This is further reflected in the mean square error of the estimated model, which improves significantly when the model allows for two types of firms (heterogenous effects) relative to one type (homogenous effects). Interestingly, for some banks, their transmission of banks shocks may have different signs across types of firms: For certain years, there are banks with an estimated positive bank shock (expansion of their credit supply) with a group of firms while a negative bank shock (contraction in their credit) with other group.

Second, the heterogeneity in the impact of bank shocks lead to bank-firm interaction effects that begs the question whether banks and firms create relationships and sorting patterns that help expand (or contract) the overall aggregate lending. Our framework allows us to estimate the entire loan growth distribution corresponding to a counterfactual reallocation of relationships between banks and firms. In particular, we estimate as a counterfactual the aggregate credit growth when bank-firm relationships are randomly allocated. With this exercise, our aim is to assess the contribution of sorting for the propagation of credit shocks, that is, a bank-firm match lending channel. We find that for most of the years aggregate credit growth is enhanced by the observed bank-firm network relative to a random allocation of relationships: achieving a more than 15% higher credit growth rate in 2016 and 2017,
an a growth rate that is 3% to 10% higher for most years, with the exception of 2008-09 in which the observed network produces a 5% lower growth rate.\footnote{Under a model with homogenous effects, this exercise would lead to no change in aggregate credit growth by construction.}

Third, we explore the real effects of bank supply shocks on firm investment. We use our identified bank shocks, which are heterogenous by groups, to analyze whether firms’ investment is sensitive to their lenders’ supply shocks. For this exercise, banks shocks are aggregated at the firm level as in AW. We find that when bank shocks are estimated assuming homogenous effects, we obtain imprecise and insignificant effects of bank shocks on investment. Instead, when we consider heterogeneity in the effect of bank shocks across groups of firms, we find a significant impact on firm investment. The estimation of the elasticity of bank shocks on investment is about 4 (we get a similar result when the number of groups is chosen to be 2 to 5).

Fourth, we study the transmission of an observed bank credit shock in an event study as in Paravisini et al. 2015. The bank shock is measured by the banks exposure to foreign funding shortage experienced during the capital flow reversal in Peru in 2008. We find significant differences in the transmission of such credit shock across our identified firms’ groups.

Finally, we show explore the observable characteristics among our estimated groups. The groups present significant differences across a variety of observable characteristics like the average collateral sizes posted on their loans, the share of loans denominated in dollars relative to soles, their average exports’ volume and growth rate, and in their probability to be connected to banks with high foreign liabilities. Instead, we do not find significant differences in their risk scores. Interestingly, we observe that export value and collateral size are variables that significantly sort the groups. This type of sorting is consistent with the specialization on export destinations markets emphasized by Paravisini et al. [2021] and on the lending contracts demanded by firms as by Ivashina et al. [2022].

**Outline.** Section 2 provides an overview of one of the main framework used to identify banks shocks exploiting credit registry data, and discusses the identification assumption of homogenous effects. Section 3 extends the framework to allow for heterogeneous effects of banks shocks across firm groups, and discusses identification. Section 4 describes the estimation methodology. Section 5 describes the data, and section 6 presents the results. Section 7 concludes.

### 2 A review of the standard framework: a model of homogeneous effects

The empirical model consists of $N$ firms and $B$ banks that interact in different moments in time $t = 1,...,T$. If at some point in time a firm $f$ and a bank $b$ interact, then a network (relation) $D_{f bt} = 1$ is
formed, otherwise $D_{fbt} = 0$.

Let $y_{f,b,t}^*$ be the potential growth rate of a loan between bank $b$ and firm $f$ at time $t$ if the link $(f, b, t)$ exists. Let's consider the following linear specification with two-sided heterogeneity for $y_{f,b,t}^*$

$$y_{f,b,t}^* = \alpha_{f,t} + \beta_{b,t} + \epsilon_{f,b,t},$$

where $\alpha_{f,t}$ is time-varying firm-specific unobserved heterogeneity, $\beta_{b,t}$ is time-varying bank-specific unobserved heterogeneity, and $\epsilon_{f,b,t}$ is an idiosyncratic unobserved shock that varies across firms, banks and time. That is, the decomposition assumes that $\beta_{b,t}$ capture all common factors across all loans from bank $b$ to all its connected firms, and $\alpha_{f,t}$ is the common factor across all loans firm $f$ gets from all its connected banks.

We observe realizations of $y_{f,b,t}^*$ only when $D_{fbt} = 1$. Let's define the observed growth rate of a loan between bank $b$ and firm $f$ as $y_{f,b,t} = D_{f,b,t}y_{f,b,t}^*$.

A common and crucial assumption in the literature (e.g., as used in Amiti and Weinstein [2018]) is the following conditional mean independence assumption:

**Assumption 1.** $E[\epsilon_{f,b,t}|D, \alpha, \beta] = 0$ where $D, \alpha$ and $\beta$ denotes the entire vector of $D_{f,b,t}$, $\alpha_{f,t}$ and $\beta_{b,t}$ for all $f, b, t,$ respectively.

This assumption states that the formation of networks (relationships between banks and firms) is strictly exogenous once we conditioned on the firm-specific unobserved heterogeneity and the bank-specific unobserved heterogeneity. That is, idiosyncratic shocks to the growth rate of a loan do not affect the formation of relationships between firms and banks.

Given the exogeneity of the network implied by assumption 1, we can estimate model (1) conditioning on the observed network $D$ without the need of jointly modeling the distribution of $\{y_{f,b,t}^*, D_{f,b,t}\}$. Under assumption 1, multiple connections between firms and banks for every period $t$ allow to identify the common factors $\alpha_{f,t}, \beta_{b,t}$ as time-varying firm and bank fixed effects, respectively. However, interpreting them as credit-demand or credit-supply factors rely on an specific assumption of homogeneity in the transmission of shocks.

Assumption 1 might fail if the creation of a new link depends on the specific relation of a bank with a group of firms. For instance, Paravisini et al. [2021] provide evidence that firms that export to a particular market are more likely to start a relationship with a bank that specializes in that market.

**Identification under homogeneous effects.** The main assumption is that: i) any credit supply shock that affects bank $b$ is transmitted equally to all its connected firms, and ii) any credit demand shock
that affects firm $f$ is transmitted equally to all its connected banks. Under these assumptions, any differential borrowing of the same firm $f$ with two different banks could not be systematically driven by a credit demand shock. By taking differences of equation (1) for the same firm in two banks $b_0, b_1$ we get

$$y_{f,b,t} - y_{f,b_0,t} = (\beta_{b,t} - \beta_{b_0,t}) + (\epsilon_{f,b,t} - \epsilon_{f,b_0,t}),$$

and by taking averages over the common group of firms that borrow from both banks $b$ and $b_0$ denoted $I(b, b_0)$, and given assumption 1, we get

$$E_{f \in I(b,b_0)} [y_{f,b,t} - y_{f,b_0,t}] = \beta_{b,t} - \beta_{b_0,t}. \quad (2)$$

Since the homogeneity assumption implies that this difference cannot be driven by firm demand, we have that $\beta_{b,t} - \beta_{b_0,t}$ captures the specific credit supply shock to bank $b$ relative to $b_0$. Similarly, the homogeneity assumption implies that the fixed effect $\alpha_{f,t}$ can be interpreted as a firm demand specific shock.

### 3 Evidence of heterogeneity

If there exists homogeneity in the transmission of bank shocks, then all systematic differences in the borrowing of different firms across two banks will be explained by the bank shock $(\beta_{b,t} - \beta_{b_0,t})$. And, in particular, the average borrowing of different groups of firms should change similarly across different banks, that is, the left hand side in equation (2) should be the same for any grouping of firms $I(b, b_0)$ that borrow from both banks.

Figure 1 presents the sample analog of the left hand side of (2) in our data set for different grouping of firms. We consider only firms that borrow from the four main banks in our data set, and we show the results when grouping firms by different percentiles of their overall loan growth.

Panel 1a shows the average loan growth rate (relative to bank 1) across all firms that are connected to the four banks for the year 2017. According to equation (2), under homogenous effects, such differential borrowing would capture each specific bank shock $\beta_{b,t} - \beta_{b=1,t}$, so we would conclude that bank 2 experienced a negative bank shock of about -12% (relative to bank 1), while banks 3 and 4 experienced very small banks shocks. However, the next panels show that such differential borrowing across banks
changes considerably across the different firm groupings. Panel 1b shows the differential average growth rates for firms above/below the median. We can see that on average firms in both groups borrow less from bank 2 than from bank 1, but the magnitude is considerably different (around -10% vs -16%). More strikingly are the results for bank 3, we can see that firms in the below-median “blue” group experienced an even higher drop from their borrowing from bank 3 (relative to bank 1), while firms in the above-median “red” group experienced an increase in their borrowing. Panels 1c and 1d further classify firms into finer groups, dividing firms into three and four quantiles, respectively. All panels indicate that there exists heterogeneity in the responses by groups, which suggest that the additive (homogenous) effects of bank and firm shocks implied by the linear specification in section 2 is not satisfied in the data.

4 A model with heterogenous effects

Let $t = 1, ..., T$ be the index for time, $f = 1, ..., N_F$ the index for firms, and $b = 1, ..., N_B$ the index for banks. Let $G$ be the number of groups (which is unknown and fixed), and let $\mathcal{G} = \{g(1), ..., g(N_F)\}$ be
any grouping of firms into the $G$ groups. Then, for each $f$, we have $g(f) \in \{1, ..., G\}$.

We denote with $y_{f,b,t}$ the growth rate of loans between bank $b$ and firm $f$ at time $t$, which we express as:

$$y_{f,b,t} = \alpha_{f,t} + \beta_{b,g(f),t} + \epsilon_{f,b,t},$$

where $\alpha_{f,t}$ is a firm time-varying effect (fixed across banks), $\beta_{b,g(f),t}$ is a bank-group time-varying effect (fixed across firms within groups), and $\epsilon_{f,b,t}$ is the unit specific error term.

**Assumption 2.** $E[\epsilon_{f,b,t}|D, \alpha, \beta_g] = 0$ where $D$, $\alpha$ and $\beta_g$ denotes the entire vector of $D_{f,b,t}$, $\alpha_{f,t}$ and $\beta_{b,g(f),t}$ for all $f,b,t$, respectively.

This assumption states that the formation of networks is strictly exogenous once we conditioned on the firm-specific unobserved heterogeneity and the bank-group unobserved heterogeneity. That is, idiosyncratic shocks to the growth rate of a loan of firms that belongs to the same type/group do not affect the formation of relationships between firms and banks. As opposed to 1, assumption 2 allows for endogenous networks of firms that belongs to different groups. For instance if the relevant grouping is due to market-specialization, the assumption allows for firms that export to different markets to form links with different banks (specialized in different markets) as in Paravisini et al. [2021].

**Assumptions on bank shocks heterogeneity.** We make the following assumptions regarding how credit supply shocks to banks are transmitted to their connected firms.

**Assumption 3.** Heterogeneity across groups: We assume that bank credit supply shocks can be heterogenous across groups of firms.

This heterogeneity across groups is left unrestricted.

**Assumption 4.** Homogeneity within groups: We assume that bank credit supply shocks are homogenous within groups of firms.

The assumption of homogeneity within groups allow us to exploit within group variation to identify (interpret) firm-demand shocks with $\alpha_{f,t}$ and bank credit supply shocks with $\beta_{b,g(f),t}$.

**Identification/estimation under known groups.** If groups were known, a comparison between the loan growth rate of firms in the same group $g$ borrowing from two banks $b$ and $b_0$, we have:

$$E_{f \in g, f \in I(b,b_0)} [y_{f,b,t} - y_{f,b_0,t}] = \beta_{b,g,t} - \beta_{b_0,g,t}, \ \forall g \in \{1, ..., G\}.$$  

Given our assumption of homogeneity within groups, we have that this differential borrowing of firms from the two banks cannot be explained by firm demand factors since all of these firms belong to the
same group for which their demand effects are transmitted homogeneously. That is, we can interpret \( \beta_{b,g,t} - \beta_{b_0,g,t} \) as being purely driven by a credit supply shock.

After the normalization \( \beta_{b_0,g,t} = 1 \) and identification of each bank supply shock \( \beta_{b,g,t} \), we can calculate for each firm its loan growth rate relative to the bank supply shock and average over all its connected banks, denoted \( I(f) \), which from 3 leads to

\[
E_{b \in I(f)} [y_{f,b,t} - \beta_{b,g,t}] = \alpha_{f,t},
\]

which would then identify the firm \( f \) credit demand shock.

If groups were known, then a standard estimation of time-varying bank and firm fixed effects separately for each group would provide consistent estimates of these objects.

**Identification/estimation under unknown groups.** Following Bonhomme and Manresa [2015], for a pre-defined number of groups \( G \) we define our grouped estimator as the solution of:

\[
(\hat{\alpha}, \hat{\beta}, \hat{\gamma}) = \arg\min_{(\alpha, \beta, \gamma) \in \mathbb{R}^{NF \times NB \times \Gamma_G}} \sum_{t=1}^{T} \sum_{f=1}^{N_F} \sum_{b=1}^{N_B} (y_{f,b,t} - \alpha_{f,t} - \beta_{b,g(f),t})^2,
\]

(4)

where the minimum is taken over all possible groupings \( \gamma = \{g(1), ..., g(N_F)\} \) of the \( N_F \) firms into \( G \) groups, firm-specific time effects, and bank-group-specific time effects.

For given values of \( \alpha_{f,t} \) and \( \beta_{b,g(f),t} \), the optimal group assignment for each firm is:

\[
\hat{g}(f|\hat{\alpha}_{f,t}, \hat{\beta}_{b,g(f),t}) = \arg\min_{g} \sum_{t=1}^{T} (y_{i,t} - \alpha_{f,t} - \beta_{b,g(f),t})^2,
\]

(5)

where we take the minimum \( g \) in case of a non-unique solution. The estimator of \( \alpha_{f,t} \) and \( \beta_{b,g(f),t} \) in 4 can be written as:

\[
(\hat{\alpha}, \hat{\beta}) = \arg\min_{(\alpha, \beta) \in \mathbb{R}^{NF \times NB \times \Gamma_G}} \sum_{t=1}^{T} \sum_{f=1}^{N_F} \sum_{b=1}^{N_B} (y_{f,b,t} - \alpha_{f,t} - \beta_{b,g(f),t})^2,
\]

(6)

where \( \hat{g}(f|\hat{\alpha}_{f,t}, \hat{\beta}_{b,g(f),t}) \) is given by 5, and the estimate of \( g(f) \) is simply \( \hat{g}(f|\hat{\alpha}, \hat{\beta}) \).

**Proposition 1.** Assume that there exists a finite number of groups \( G \) and a grouping function \( g(f) \) for which assumptions 3 and 4 hold. Then, the estimator \( (\hat{\alpha}, \hat{\beta}, \hat{\gamma}) \) in (4) provide consistent estimates of \( g(f) \) and \( (\alpha_{f,t}, \beta_{b,g(f),t}) \) as defined in (3).
5 Computation

We build on Bonhomme and Manresa [2015] and the econometrics literature that develop techniques to group observations (Bai [2009], Ando and Bai [2016], Bonhomme et al. [2019]). Typically, these techniques cluster the data into groups depending on some dissimilarity measure based on some observable characteristics. Our estimator in 4 instead optimally groups firms according to their dissimilarities on how firms respond to an unobserved shock, since the bank shocks we assume are heterogenous are unobserved. We extend the method in Bonhomme and Manresa [2015] to implement it to our framework.

We use the following algorithm that performs a simple iterative strategy to minimize 6.

Algorithm

1. Set the number of groups: $G$.

2. Set $s = 0$. Guess initially some group assignment $g^{(s=0)}(f) \in \{1, \ldots, G\}$.

   - For our initial guess, we estimate homogenous time-varying firm and bank effects, and then group firms using “kmeans” over the estimated $\hat{\alpha}_{f,t}$’s.

3. For given $g^{(s)}(f)$, estimate firm-time and bank-group-time fixed effects $\hat{\beta}_{b,g}^{(s)}(f,t)$ and $\hat{\alpha}_{f,t}^{(s)}$

   $\left(\hat{\alpha}^{(s)}, \hat{\beta}^{(s)}\right) = \arg \min_{\alpha_{f,t}, \beta_{b,g}^{(s)}(f,t)} \sum_{t=1}^{T} \sum_{f=1}^{N_F} \sum_{b=1}^{N_B} \left(y_{f,b} - \alpha_{f,t} - \beta_{b,g}^{(s)}(f,t)\right)^2$ (7)

4. For given $\hat{\alpha}^{(s)}, \hat{\beta}^{(s)}$, select optimal group assignment: For all $f = 1, \ldots, N_F$

   $g^{(s+1)}(f) = \arg \min_{g \in \{1, \ldots, G\}} \sum_{t=1}^{T} \sum_{b=1}^{N_B} \left(y_{f,b} - \hat{\alpha}_{f,t}^{(s)} - \hat{\beta}_{b,g}^{(s)}\right)^2$ (8)

5. Set $s = s + 1$ and go to Step 3 until numerical convergence.

Optimal number of groups Although our procedure can be applied for different values of $G$, an important discussion in practice is to determine the optimal number of groups in the data, $G^*$. Building on Almagro and Manresa [2021], we propose the following N-fold cross-validation procedure for our algorithm:

1. For a given year $t$, set the number of groups and repetitions: $G$ and $M$, respectively.
2. \( \forall G \) and \( \forall m \in \{1, \ldots, M\} \), split the sample in \( N \) parts (folds): \( P_t (G, m) \equiv \{P_{t,1} (G, m), P_{t,2} (G, m), \ldots P_{t,N} (G, m)\} \).

3. Take \( P_{t,k} (G, m) \) as the testing sample, and the remaining parts of the sample, \( P_{t,-k} (G, m) \), as the training sample.

4. Using the training sample, \( P_{t,k} (G, m) \), estimate the group structure and the parameters of interest, which total number is denoted by \( J \). After that, compute the out-of-sample mean squared error, \( MSE \), for the training sample, \( P_{t,k} (G, m) \), as follows:

\[
MSE_t (G, m, P_{t,k}) = \frac{1}{J} \sum_{j \in P_{t,k}} (y_{f,t,j} - \hat{\alpha}_f (G, m) - \hat{\beta}_{b,g(f),t,j} (G, m))^2.
\] (9)

5. We calculate an out-of-sample \( R^2 \), to measure the performance of the heterogeneous model \( G > 1 \) with respect to the homogenous case \( G = 1 \) for each of the repetitions:

\[
R^2_t (G > 1, m, P_{t,k}) = 1 - \frac{MSE (G > 1, m, P_{t,k})}{MSE (G = 1, m, P_{t,k})}.
\]

6. We take the average for each of the folds and the repetitions in the sample

\[
R^2_t (G > 1) = \frac{1}{N M} \sum_{m=1}^{M} \sum_{k=1}^{N} R^2_t (G > 1, m, P_{t,k}).
\]

7. We take the average over time of the out-of-sample mean squared error for each group

\[
R^2 (G > 1) = \frac{1}{T} \sum_{t=1}^{T} R^2_t (G > 1).
\]

8. We calculate an out-of-sample \( R^2 \), to measure the performance of the heterogeneous model \( G > 1 \) with respect to the homogenous case \( G = 1 \)

\[
R^2 (G > 1) = 1 - \frac{MSE (G > 1)}{MSE (G = 1)}.
\]

9. Choose the value of \( G^* \) as the one that maximizes the out-of-sample \( R^2 \)

\[
G^* = \max \{R^2 (G > 1) \mid R^2 (G > 1) > R^2 (G = 1)\}.
\]
For our data, we find that the optimal number of groups is $G^* = 2$. In particular, the model with $G = 2$ relative to the homogenous case of $G = 1$ explains 6% more of the variability in the data. In Figure 2, we display the distribution of the out-of-sample R-squared for each of the folds and the repetitions. The measure compares the explanatory power of the $G = 2$ (heterogenous case) with respect to the $G = 1$ (homogenous case). A positive value implies that the heterogeneous case model has a better out-of-sample explanatory power than the homogeneous case. In general, we can observe how the heterogeneous model with $G = 2$ is most likely to perform better than the traditional homogeneous case with $G = 1$: on average, changing an homogeneous model for an heterogeneous one helps to avoid losing 5% of explanatory power in the data.

Figure 2: Estimated Values of $R^2_t (G^* = 2, m, P_{t,k})$

5.1 Properties of our estimator

We study the properties of our estimator through a Monte Carlo analysis. We generated 1000 replications generating data from the model specified in 3. For the data generating process we set: the number of groups to $G = 5$, the number of banks $N_B = 10$ (the number of banks in our data set is 17), the variance of the error term $\sigma_e = 0.01$, and we analyze the properties of the estimator for different values of $N_F$. In a first exercise, we choose a very large number of firms to analyze the consistency of the estimator, and, on a second exercise, we choose $N_F$ to proxy the number of firms in our dataset.

Figure 3 illustrates the results of the simulation analysis. Since there are $G \times N_B = 50$ credit supply shocks that are simulated, the figure presents a k-density over these 50 generated parameters and the k-density of the estimated ones. We can see that, for both the very large $N_F$ and the one calibrated to the data, that the estimation tracks closely the the true parameters.
6 Data

Credit Registry Data: The source of information is named Registro de Crédito de Deudores (RCD) and belongs to the administrative registries of the Peruvian financial regulator, Superintendencia de Banca, Seguros y AFPs (SBS). The data allows us to observe the loan balance of every corporate firm at a given bank from 2005 to 2018. We consider that a firm is corporate if according to SBS is classified as “Corporativo” or “Gran Empresa”. Considering the definitions of such classifications, our sample of firms correspond to those that obtained at least 5 million dollars in annual sales. Corporate firms represent approximately 55% of the total amount of commercial loans. Additionally to the information of loan balance, we can observe the total amount of loan guarantees, the credit rating, a measure of the size, the type of loan, and the currency of the loan.

Financial Statements: We collect information of the financial statements from the Peruvian Stock Exchange, Superintendencia del Mercado de Valores. We observe the balance sheet for all the firms that list on the Peruvian Stock Exchange and also a group of firms that does not list, but report voluntarily to such institution. Our main variable of interest from the balance sheet is the value of fixed capital. The information corresponds from 2007-2017.

7 Results

7.1 Evidence of heterogeneity

We start our analysis by examining how does credit growth varies across banks. In Figure 4 we display the average credit growth of four representative banks (the main 4 banks in our dataset) across different firm groups. The y-axis in the graph shows the average credit growth rate of each firm
type with each bank, while the x-axis indicates a given bank. A differential average credit growth of two banks indicate the presence of a bank shock.\textsuperscript{3} Panel (a) shows the growth rates when we consider only one group (that is we assume all firms are of the same type). Under the homogeneity assumption, an horizontal line would imply the absence of a bank-specific supply shock among these four banks.\textsuperscript{4} More specifically, the steepness of the credit growth between two banks measure the relative importance of the bank supply shock specific to each bank. The graph shows that there are differences between banks: the average credit growth for bank 1 and bank 2 is approximately 6%, while for bank 3 is 2% and for bank 4 is 8%.

Figure 4: Average credit growth by bank $E[y_{f,b,t} | b, t]$ (for $t = 2017$)

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure4.png}
\caption{Average credit growth by bank $E[y_{f,b,t} | b, t]$ (for $t = 2017$)}
\end{figure}

In contrast, panel (b) shows the average credit growth rate for two different groups, where groups are identified by our algorithm. The figure shows that there exists significant heterogeneity in the patterns of credit growth across groups. The analysis is similar than for Panel (a), but in this case, we should consider the credit growth within groups. Assuming homogeneity hides significant differences that can only be uncover by allowing for heterogeneity. First, compared to panel (a), there are significant differences across groups for almost all the banks included in the graph. For example, for bank 2 we can observe that the average credit growth for group 1 ($g = 1$) is around 24%, while for group 2 ($g = 2$) is close to -15%. Although it might seem that the differences in credit growth across groups for bank 1 are minor, they are sizable. The average credit growth for group 1 ($g = 1$) is approximately 3%, while for group 2 ($g = 2$) is 11%. For the remaining banks, the differences are also significant.

\textsuperscript{3}This figure is intended for illustration. For the identification of a bank shock we should compare the differential average credit across the group of firms that borrow from both banks.

\textsuperscript{4}The level of the line would capture an aggregate factor across the four banks that could be capturing either aggregate demand or supply factors. It cannot be attributed specifically to supply effects under the assumptions in section 2.
**Bank shocks estimates.** Now we proceed to dissect the results that we obtain by employing our empirical specification. Compared to our analysis before, the estimation of bank shocks exploits variation in the differences in credit growth across banks only for the set of connected firms. Figure 5 displays the bank supply shocks estimates for the homogeneous and heterogeneous cases, which follows from the algorithm described in 5 for $G = 1$ and $G = 2$, respectively. The black bars in the figure represent the estimates under the homogeneous case ($G = 1$), while the blue and red bars denote the estimates for each of the groups of the heterogeneous case when the number of groups is equal to $G = 2$. Similar to our previous analysis, the estimates vary significantly across groups. For example, for bank 2, we can see that, under the homogeneous case, we find a negative estimated bank shock, while when we allow for heterogeneity, we find that for group 1 the bank shock is positive, while for group 2 is negative. Thus, assuming homogeneity generates an overestimation of the bank supply shocks for some firms and an underestimation for other firms.

![Figure 5: Bank shocks estimates by bank $\hat{\beta}_{f,b,t}$ (for $t = 2017$)](chart.png)

### 7.2 Quantification of the value of bank-firm relationships

We conduct a random reallocation exercise for every firm-bank observation in our data. Our procedure is the following: First, we obtain from our empirical strategy the decomposition of the average credit growth as the sum of the firm demand shock $\hat{\alpha}_{f,t}$ and the heterogeneous bank supply shocks $\hat{\beta}_{n,g(f),t}$ which are heterogenous by groups. Second, for each firm-bank observation, we randomize without replacement the bank that is lending to each firm (keeping the same number of connections). Third, we calculate the counterfactual credit growth for each of the observations as the sum of the estimated firm demand shock and the counterfactual estimated heterogeneous bank supply shock associated with the simulated random network. Fourth, we calculate the average credit growth across all the observations. Finally, we repeat this exercise 10,000 times to calculate the distribution of the average credit growth.
Figure 6 shows the distribution of the average credit growth for all the simulations for year 2017. The y-axis displays the density of the distribution, while the x-axis represents how much the aggregate credit growth changes if we randomly reallocate the bank-relationships. The median of the distribution suggests that the average credit growth across firms would fall by 20% if we randomly reallocate the bank-relationships. This results highlight the importance of the value of bank-firm relationships: the observed sorting of bank-firm relationships enhance aggregate credit growth. Note that, under a model with homogenous effects \((G = 1)\), this exercise would lead to no change in aggregate credit growth by construction.

Figure 6: Percentage change in average credit growth in the random network relative to the observed network

![Histogram](image)

Note. Distribution of the difference between the average growth rate of the simulated random network and the actual average credit growth rate in the data.

### 7.3 The impact of credit supply shocks on firm investment

In this part, we reexamine the real effects of credit supply shocks on firm investment. We measure investment as the growth rate in total tangible fixed assets plus depreciation. First, we calculate the credit supply shock at the firm level by weighting the bank-specific supply shocks that the firm has in a given bank by the share that each loan represents for the firm.

\[
\text{Supply}_{f,t} = \sum_b \theta_{f,b,t-1} \hat{\beta}_{b,g(f),t} \quad \text{with} \quad \theta_{f,b,t-1} = \frac{L_{f,b,t-1}}{\sum_b L_{f,b,t-1}},
\]

where \(L_{f,b,t-1}\) denote borrowing by firm \(f\) from bank \(b\) in time \(t - 1\).

Second, after obtaining the firm level measure, we estimate a linear regression where the dependent variable is investment. We control the regression by a set of fixed effects, and with our estimates of the demand firm shock.
Investment_{f,t} = c + \phi_f^F + \phi_{G(f),t}^G + \beta_1 \Delta f_{f,t} + \beta_2 \text{Supply}_{f,t} + \epsilon_{f,t}

Figure 7 shows the estimated elasticity of the supply shock on investment (\hat{\beta}_2), under the estimation for different number of groups (as represented in the X-axis). The results reveal that when banks shocks are estimated under the homogeneous effects, although the effect is positive, the estimated effect is very imprecise and is not statistically significant. As we saw in the previous sections, assuming homogeneous effects generates a bias in the estimates of the credit supply shocks. As such estimates are aggregated at the firm level, the bias will remain at the firm level. Instead, when assuming heterogeneous effects (G > 1), the estimated impact of banks shocks on firm investment is statistically significant. Our estimates for the heterogeneous case suggest that a 1% change in the credit supply increases investment by around 5%.

7.4 Case Study

In this subsection, we use our group fixed effects estimator to study heterogeneous responses of credit growth to an observed bank (supply shock). To do so, we combine our group fixed effects estimator with the methodology developed in Khwaja and Mian [2008] and we applied it to Peruvian data following Paravisini et al. [2015].

Consider the following linear specification for the log of credit as in Paravisini et al. [2015]:

\[ \log L_{f,b,t} = \eta_{f,b} + \alpha_{f,t} + \theta X_{b} Post_{t} + \epsilon_{f,b,t} \]

where \( \log L_{f,b,t} \) is the log of the average outstanding debt of firm \( i \) with bank \( b \) during the intervals \( t = Pre, Post \), where \( Pre \) and \( Post \) periods correspond to the 12 months before and after July 2008,
\( \eta_{f,b} \) is a time-invariant firm-bank fixed effect that controls for time-invariant unobserved bank-firm characteristics, \( \alpha_{f,t} \) is a time-varying firm fixed effect, \( X_{b,t} \) is an observable variable that varies at the bank level and takes the value of one if a bank \( b \) has high foreign liabilities and zero otherwise. The linear specification leads to the following model for credit growth:

\[
y_{f,b,t} \equiv \log L_{f,b,Post} - \log L_{f,b,Pre} = \tilde{\alpha}_f + \theta X_{b} + \nu_{f,b}
\]

As in Paravisini et al. 2015, \( \theta \) is identified by comparing the differential borrowing of the firms that borrows from high foreign liabilities banks and from low foreign liabilities banks. The multiple connection structure of the data allows us to eliminate demand shocks to the firms under the assumption of homogeneity and help us to tease out the marginal effect of a bank shock.

Using our identified groups, we estimate a similar model allowing for heterogeneous effects (at the group level) of the same observed bank shock:

\[
y_{f,b,t} = \tilde{\alpha}_f + \theta_{g(f)} X_{b} + \tilde{\nu}_{f,b}
\]

In this specification, \( \theta_{g(f)} \) is identified by comparing the differential borrowing of the firms that belong to group \( g \) and borrows from high foreign liabilities banks and from low foreign liabilities banks. Figure 8 displays our main results. In Panel A, we display the coefficients that we estimate using the heterogeneous model. The case \( G = 1 \) replicates the homogeneous case as obtained in Paravisini et al. [2015]. The heterogeneous estimates reveal that there are groups of firms which have been significantly negative impacted for being connected with high foreign liabilities banks, but there are also firms that increased their borrowing during this period. Complementarily, we calculate the average treatment effect by weighting the estimating effect by the number of observations in each group as in Arellano and Bonhomme [2012] or Gibbons et al. [2018]. Panel B, shows the results and displays that the average treatment effect decreases in magnitude when considering heterogeneity but not considerably.
To further investigate the effects behind the difference in groups we explore the average observable characteristics when the number of groups is 3 (G=3). In Table 1 we observe that on average the most affected group had less collateral, exported more, had higher ratio of loan dollarization, and had more loans with foreign liability banks. From Panel B of Figure 8 we observed that the least affected group had a higher credit growth with high foreign liability banks than with low foreign liability banks. This
result might be consistent with the fact that high foreign liability banks reallocate credit towards safer firms to mitigate reductions in their credit growth.

Table 1: Observable characteristics and group structure

<table>
<thead>
<tr>
<th>Group (G = 3)</th>
<th>Collateral Size (million of soles)</th>
<th>Risk Score (from 0 to 4)</th>
<th>Export Value (million of dollars)</th>
<th>Loan Dollarization (share)</th>
<th>Loans w/ Liabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Most Affected Group</td>
<td>7.34</td>
<td>0.16</td>
<td>31.71</td>
<td>74%</td>
<td></td>
</tr>
<tr>
<td>Middle Affected Group</td>
<td>11.65</td>
<td>0.18</td>
<td>29.66</td>
<td>71%</td>
<td></td>
</tr>
<tr>
<td>Least Affected Group</td>
<td>8.24</td>
<td>0.18</td>
<td>13.78</td>
<td>67%</td>
<td></td>
</tr>
</tbody>
</table>

7.5 Observable characteristics and group structure

We now explore the observable characteristics among the estimated groups. In order to display a variability among the variables, we display the information for the case in which the total number of groups is 3 (G = 3). Panel A shows the evolution of the credit growth over time for each of the estimated groups. The figure depicts that group number 1 (g = 1) is consistently growing faster in the time series than the other groups. It also displays that group number 2 (g = 2) is growing at a similar rate than group number 3 (g = 3) over time. In addition, Panel B shows other relevant observable characteristics besides credit growth. We observe that collateral size and export value are variables that significantly sort the groups. Note that this is consistent with the collateral loan channel studied by Ivashina et al. [2022] and also with the specialization of Paravisini et al. [2021]. However, there are two relevant difference with our approach: first, we employ a non-parametric model account for unobservable factors that are not relevant, but might not be observed by the econometrician. For instance, if international trade data is not observable and we consider that credit shocks are heterogeneous across exporters, with our framework we can recover groups that differ in terms of their trade status and we can account for such heterogeneity at the group-bank level. Second, the source of heterogeneity might come from different factors and the group estimator will capture. For example, let’s consider that bank shocks are heterogeneous across collateral size and trade status. If we employ the group-bank estimator, we will be able to recover such combination of heterogeneous effects.
8 Conclusions

This paper develops an empirical framework to identify and estimate heterogeneous effects of time-varying bank shocks from bank-firm credit data. Our methodology provides a flexible framework to study the importance of the bank lending channel when there is heterogeneity in the transmission of bank shock to different types of firms. The heterogeneous transmission of bank shocks is particularly relevant in the presence of bank-firm relationships or bank specialization. To allow for heterogeneity in a flexible yet parsimonious way, we rely on the idea that from the perspective of the banks and their relationship with firms, there is a discrete number of “types” of firms, and hence, the propagation of bank shocks to the firm-level growth rate of loans depends on the firm’s type. These interactions between banks and groups of firms naturally arise when banks specialize in a particular market where a group of firms operate. If banks develop skills, expertise, or technology in evaluating projects in a specific sector, geographical market, or economic activity, it may create complementarities between specialized banks and firms that operates in that particular market.

To implement this idea, we combine state-of-the-art panel data techniques that allow for time-varying group fixed effects (Bonhomme and Manresa [2015]) with the two-sided fixed effects frame-
work used to disentangle demand and supply shocks from bank-firm credit data (as in Amiti and Weinstein [2018]). We show in simulations that our proposed estimator behaves correctly and consistently estimates the bank’s shocks under heterogeneity when the number of firms operating with multiple banks is sufficiently large (as large as in our empirical example). When applying our flexible framework to credit registry data from Peru, we find significant heterogeneity in the bank lending channel as we find that bank shocks are transmitted differently across identified groups of firms. Given this heterogeneous transmission of bank shocks, we then quantify the importance of the bank-firm relationship (or specialization) in terms of aggregate credit. Interestingly, we find that aggregate credit growth drops by 20% when bank-firm relationships are randomly allocated. Finally, we investigate the impact of bank shocks on firm investment. When we neglect the possibility of heterogeneous effects in the estimation of the bank’s shocks from bank-firm loan level data, the resulting bank shock does not affect firm investment. However, once we apply our methodology and recover bank shocks that have heterogeneous effects on loans, we find that the estimated bank shocks have an important effect on firm investment.
References


