

NBER WORKING PAPER SERIES

HOW MUCH DO IDIOSYNCRATIC BANK SHOCKS AFFECT INVESTMENT?
EVIDENCE FROM MATCHED BANK-FIRM LOAN DATA

Mary Amiti
David E. Weinstein

Working Paper 18890
<http://www.nber.org/papers/w18890>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2013

We would like to thank Francesco Caselli, Gabriel Chodorow-Reich, Xavier Gabaix, Mark Gertler, Takatoshi Ito, Anil Kashyap, Nobu Kiyotaki, Satoshi Koibuchi, Anna Kovner, Aart Kraay, Nuno Limao, Tamaki Miyauchi, Friederike Niepmann, Hugh Patrick, Benjamin Pugsley, and Bernard Salanie for excellent comments. We also thank Prajit Gopal, Scott Marchi, Molly Schnell and especially Preston Mui and Richard Peck for outstanding research assistance. David Weinstein thanks the Center on Japanese Economy and Business and the Institute for New Economic Thinking for generous financial support. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System, or the National Bureau of Economic Research. Any errors or omissions are the responsibility of the authors.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2013 by Mary Amiti and David E. Weinstein. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

How Much do Idiosyncratic Bank Shocks Affect Investment? Evidence from Matched Bank-Firm Loan Data

Mary Amiti and David E. Weinstein

NBER Working Paper No. 18890

March 2013, Revised February 2017

JEL No. E44,G21,G31

ABSTRACT

We show that supply-side financial shocks have a large impact on firms' investment. We develop a new methodology to separate firm-borrowing shocks from bank-supply shocks using a vast sample of matched bank-firm lending data. We decompose aggregate loan movements in Japan for the period 1990 to 2010 into bank, firm, industry, and common shocks. The high degree of financial institution concentration means that individual banks are large relative to the size of the economy, which creates a role for granular shocks as in Gabaix (2011). We show that idiosyncratic granular bank-supply shocks explain 30-40 percent of aggregate loan and investment fluctuations.

Mary Amiti

International Research

Federal Reserve Bank of New York

33 Liberty St

New York, NY 10045-0001

Mary.Amiti@ny.frb.org

David E. Weinstein

Columbia University, Department of Economics

420 W. 118th Street

MC 3308

New York, NY 10027

and NBER

dew35@columbia.edu

1 Introduction

Do idiosyncratic bank-loan supply shocks matter for aggregate investment, and if so, how much? Since the principal reason firms borrow is to finance capital expenditures, this question stands at the center of debates on the importance of financial shocks for real economic activity. Although many studies have shown that bank shocks matter for loan supply and certain types of foreign investment,¹ we know little about how important bank-loan supply shocks are in determining the overall investment rates of their borrowers and aggregate investment more generally. Moreover, studies of the relative effects of bank shocks on firm investment are hard to link to aggregate fluctuations. Our study addresses these issues by providing the first estimate of how much “idiosyncratic” financial institution shocks (i.e., shocks that are not common to all banks) matter for overall firm-level and aggregate investment rates and establishes that these shocks are an important determinant of both. In order to show this, we develop a new methodology that enables us to provide the first direct estimates of time-varying firm-borrowing and bank-supply shocks using a comprehensive, matched, lender-borrower data set covering all loans received from all sources by every listed Japanese firm over the period 1990 to 2010.

A major challenge in answering this question is to separately identify time-varying bank supply shocks from firm-borrowing shocks. We make progress in this dimension by developing a methodology that generates a unique set of bank and firm shocks (up to a numéraire) that aggregate exactly to match aggregate bank lending and total firm borrowing. Our method is based on a linear model that specifies loan growth rates as a function of bank-time fixed effects and firm-time fixed effects. We show that the bank and firm shocks we identify are appropriate for all models of lending behavior that specify loan growth from a bank to a firm as a linear combination of a bank shock, a firm shock and possibly a firm-bank interaction term. Our method produces bank and firm shocks that are identical to those obtained using weighted least squares when there are no new lending relationships. However, by exploiting a simple modification of the moment conditions, our procedure accommodates new lending relationships, producing unique bank and firm shocks that aggregate to exactly match the macro moments in the data in the presence of new lending relationships. Critically, we show that the loan growth rates must be defined in percentage terms and the data weighted by lagged lending in order to match the aggregate data. Any deviation from this procedure, such as log specifications, different weights, or unweighted estimates, will result in shocks that do not in general aggregate to the macro data.

The fact that our micro estimates are consistent with macro data allows us to develop a theoretically sound aggregation method that enables us to apply these estimates to understand aggregate lending and investment movements. We do this in two ways. First, armed with our estimates of bank supply shocks from the matched lending data, we exploit the heterogeneity in the sources of firm financing in order to identify time-varying, idiosyncratic bank-supply shocks

¹See, for example, Peek and Rosengren (1997), Peek and Rosengren (2000), Kashyap and Stein (2000), Klein, Peek, and Rosengren (2002), Paravisini (2008), and Khwaja and Mian (2008)

hitting firms. We then use these bank shocks to demonstrate that firms that borrow heavily have investment rates that are very sensitive to their lenders' supply shocks. Moreover, we show that these loan supply channels are important determinants of investment not only in financial crisis years, but in non-crisis years as well. We show that these firm-level estimates of the impact of bank shocks on firm-level investment can be consistently aggregated, which enables us to evaluate the impact of idiosyncratic bank shocks on *aggregate* investment rates of listed firms. Second, building on the work of Gabaix (2011), we develop a method for estimating "granular" bank-supply shocks, which measure the aggregate loan supply movements that arise from the idiosyncratic supply shocks of large lenders.² Regressing aggregate lending and investment rates on these granular shocks provides a second method of evaluating the macro impact of idiosyncratic bank shocks. Depending on which method one prefers, we estimate the impact of idiosyncratic bank shocks on the variance of aggregate lending and investment to be 30 to 40 percent for Japan during the 1990-2010 period.

Our work is related to a number of previous studies. Our theoretical setup nests all specifications that express loan growth rates as a linear combination of bank and firm variables. Often these models are estimated using a log specification, and we are able to show that regressing log loan growth rates on explanatory variables produces estimates that predict geometric means of loan growth rates, but these point estimates do not, in general, aggregate to produce estimates of total loan growth rates. Another important difference between our approach and the prior literature is that our estimates of bank shocks incorporate the impact of new lending relationships.

Another major advantage of our approach relative to earlier work is that we are able to estimate the shocks directly from the loan data, and hence do not need to rely on instrumental variables that are correlated with firm-borrowing and bank-supply shocks. These bank-supply shocks measure idiosyncratic movements of loan supply at the financial-institution level that cannot be explained by common loan shocks hitting all financial institutions or even by movements in loan demand from the financial institutions' borrowers. Moreover, we provide extensive evidence of the external validity of our estimates. In particular, we show that our estimates capture the impact of idiosyncratic events such as bankruptcies, capital injections, regulatory interventions, computer glitches, trading errors, and other proxy variables that previous researchers have thought important determinants of idiosyncratic bank shocks.

In order to deal with the inevitable issues arising from the use of aggregate data, several authors have worked with microdata and proxy variables for bank health to demonstrate that bank shocks can matter for bank lending and certain types of real economic activity. For example, the work of Peek and Rosengren (1997, 2000), Klein, Peek, and Rosengren (2002), Khwaja and Mian (2008),

²Gabaix (2011) coined the term "granular" because it reflects the fact that firms are not infinitesimal in size. We use the term in the same sense here to refer to the macroeconomic impact of idiosyncratic bank shocks. If all banks were infinitesimally small and had uncorrelated idiosyncratic shocks, then these shocks would not be important for understanding aggregate fluctuations. However, if banks are large or "granular", idiosyncratic shocks in one or more large institutions can move aggregate lending.

Paravisini (2008), Greenstone and Mas (2012), Amiti and Weinstein (2011), Jimenez et al. (2011), Santos (2012) and Chodorow-Reich (2013) provide bank-level or matched bank-firm level evidence that deteriorations in bank health or increases in the cost of raising capital cause banks to contract lending, raise rates, and/or have impacts on *foreign* markets or employment, but none of these papers address whether bank-supply shocks affect the overall investment rates of borrowers from these institutions at the firm or aggregate level. Thus, the question of how much these shocks matter for investment, and therefore GDP, remains unanswered. Moreover, while the existing literature makes use of instruments to identify the impacts of particular bank shocks, we are able to develop a methodology that identifies these bank shocks even in situations where it may not be possible to have measures of bank health.

Our work is also able to address a major outstanding question in the literature regarding whether bank shocks matter only following extreme events or for small firms and firms without access to other sources of capital, or whether credit crunches are a phenomenon with broader implications. For example, although Ashcraft (2005) found that the failure of healthy bank subsidiaries affected county-level output in Texas, Ashcraft (2006) argues that “these effects are likely to be very small and unworthy of concern” because “while small firms might view bank loans as special, they are not special enough for the lending channel to be an important part of how monetary policy works.” These concerns are particularly apt given the evidence that loans and other types of borrowing are substitutable. For example, Kashyap, Stein, and Wilcox (1993), Kroszner, Laeven, and Klingebiel (2007), and Adrian, Colla, and Shin (2012) show that some firms are able to substitute other forms of credit supply in the presence of loan supply shocks, and Khwaja and Mian (2008) show that bank shocks matter for small but not large firms. On the other hand, Hubbard, Kuttner, and Palia (2002) stress the difficulties that firms have substituting loans from one bank with loans from another. Consistent with both sets of studies, we find evidence that bank supply shocks do not matter for firms that borrow little to finance their capital expenditures. However, we also show that these bank-supply shocks affect investment rates of those firms that borrow heavily from banks, and these results are present even for listed firms.

Finally, our paper is also related to the work of Buch and Neugebauer (2011) and Bremus et al. (2013), who use aggregate bank loan data to construct granular bank shocks and regress them on cross-country GDP growth. However, our work differs from theirs in a number of respects. First, rather than ascribing bank shocks to loan growth rate differences across institutions, which may reflect differences due to heterogeneity in borrower characteristics across banks, our method allows us to econometrically isolate idiosyncratic bank shocks from firm-borrowing shocks and time-varying common and industry shocks. This eliminates any worry that an observed correlation between granular bank shocks and GDP might arise from large banks lending to more procyclical sectors or any factor that would cause credit demand for large institutions to covary more with GDP than credit demand for small institutions. Second, since we separate firm-borrowing and bank-supply shocks, we show that the link from the banking sector to GDP flows directly from

the affected banks to the investment decisions of their client firms. This enables us not only to be precise about the mechanism through which GDP is affected, but also to show the relative importance of the bank-lending channel in understanding investment fluctuations.

The rest of the paper is structured as follows. Section 2 develops the empirical strategy. Section 3 describes and previews the data. Section 4 provides intuition about how our methodology generates bank shock estimates and investigates their plausibility. Section 5 presents the main results regarding the impact of idiosyncratic bank shocks on firm-level investment as well as aggregate investment, and Section 6 concludes.

2 Empirical Strategy

Our econometric approach begins by specifying a fairly general empirical model that we can use to estimate the importance of each type of shock hitting the economy. In order to simplify the exposition, we will refer to financial institutions in our data as “banks” even though our data comprise banks, insurance companies, and holding companies.

2.1 Estimating Firm-Borrowing and Bank-Supply Shocks

Let L_{fbt} denote borrowing by firm f from bank b in time t . We begin by considering a class of empirical models in which we can write some measure of the growth in lending, $D(L_{fbt}/L_{fb,t-1})$, as³

$$D(L_{fbt}/L_{fb,t-1}) = \alpha_{ft} + \beta_{bt} + \varepsilon_{fbt}, \quad (1)$$

where α_{ft} denotes the “firm-borrowing channel,” β_{bt} the “bank-lending channel,” and the error term satisfies $E[\varepsilon_{fbt}] = 0$.⁴ Critically, we do not assume that the bank and firm shocks are independent. The empirical model in equation (1) can easily be understood by contemplating the standard explanations for what causes lending from a bank to a firm to vary. If lending varies because of firm-level productivity shocks, changes in factor costs, changes in investment demand, firm-level credit constraints, etc., we will measure that as arising from the firm borrowing channel, α_{ft} . Similarly, if a bank cuts back on lending because it is credit constrained, we would capture that in the bank lending channel, β_{bt} . While we remain agnostic about the specific model underlying what constitutes these shocks, we show in an online appendix that models such as Khwaja and Mian (2008) can formally be nested in this framework.⁵

³The most common approach of measuring $D(L_{fbt}/L_{fb,t-1})$ is to define it as the log change in loans ($D(L_{fbt}/L_{fb,t-1}) = \ln(L_{fbt}/L_{fb,t-1})$) but we will also consider the percentage change specification: $D(L_{fbt}/L_{fb,t-1}) = L_{fbt}/L_{fb,t-1} - 1$.

⁴We checked for autocorrelation by estimating equation (1) in percentage changes using OLS and regressing the residual on its lag. We obtain a coefficient of only -0.016 and an R^2 of 0.0003. The plot of the residual on its lag also shows no relationship and is presented in the online appendix.

⁵It is also possible to show that Chava and Purnanandam (2011) and Greenstone and Mas (2012) can be nested in this framework.

One might be tempted to identify bank and firm shocks by simply regressing loan growth, $D(L_{fbt}/L_{fb,t-1})$, on time-varying firm and bank fixed effects, but there are a number of theoretical and empirical challenges. The first is identification. The specification in equation (1) does not allow for bank-firm interaction effects that might arise in the event of moral hazard in lending or complementarities in bank and firm business lines (see Peek and Rosengren (2005), Jimenez et al. (2011), and Paravisini, Rappoport, and Schnabl (2015) for examples of these types of models). This concern raises the question of whether the firm and bank shocks identified from the estimation of equation (1) are correctly identified even in the presence of bank-firm interaction terms. The second is aggregation. We need a method to move from bank and firm shocks estimated using individual loan data to total firm, bank, and economy-wide lending behavior. Third, it is desirable to have a method that can be applied even in cases where lending relationships can be created or terminated. We now proceed to deal with each problem in turn.

Let's begin by thinking about the first problem: identification.⁶ Equation (1) does not allow for bank-firm interactions, so one might well question whether our bank and firm parameters are properly identified even if one believes that the world is better described by a more general class of interactions that allows for bank-firm interactions. We can write this class of models generally as

$$D(L_{fbt}/L_{fb,t-1}) = \alpha_{ft}^I + \beta_{bt}^I + \gamma_t Z_{fbt} + \varepsilon_{fbt}^I, \quad (2)$$

where α_{ft}^I and β_{bt}^I are firm and bank fixed effects (that are possibly different from equation (1)); Z_{fbt} is a bank-firm interaction term; and γ_t is its coefficient. In an online appendix, we also show that this equation can be derived from the structure proposed by Paravisini, Rappoport, and Schnabl (2015) as well as a structural model of loan supply and demand featuring different interest rates charged to each firm from any given bank. Thus, equation (2) nests much of the empirical finance literature.

What is the relationship between the bank and firm shocks ($\hat{\alpha}_{ft}, \hat{\beta}_{bt}$) identified by estimating equation (1) with the shocks ($\hat{\alpha}_{ft}^I, \hat{\beta}_{bt}^I$) obtained by estimating equation (2)? In order to answer this question, it is useful to build some intuition by considering two simple extreme cases. At one extreme, it is obvious that the two sets of parameter estimates would be identical if the bank-firm interaction term (Z_{fbt}) is orthogonal to the bank and firm fixed effects. In this case, we could define $\epsilon_{fbt} \equiv \gamma_t Z_{fbt} + \varepsilon_{fbt}^I$, and equation (2) would be identical to equation (1). At the other extreme, if the bank-firm interaction term were perfectly explainable by bank and firm fixed effects, we would have $Z_{fbt} = \delta_{ft}^F + \delta_{bt}^B$, where δ_{ft}^F and δ_{bt}^B are variables that vary only at the firm and bank level. The parameters in equation (1) and equation (2) would then be related by the following expressions:

⁶Throughout the paper, we assume that we have a connected set of bank-firm lending relationships, which means that no set of banks and firms can be partitioned into a nonempty subset in which none of the members of the subset have relationships with institutions outside of the subset. For example, it cannot be the case that a set of banks only lend to firms in a given industry and these firms only borrow from this subset of banks. In practice, this condition is satisfied in the data.

$\alpha_{ft} = \alpha_{ft}^I + \gamma_t \delta_{ft}^F$ and $\beta_{bt} = \beta_{bt}^I + \gamma_t \delta_{bt}^B$. In this simple case, as long as one defined the firm and bank shocks as the component in loan variation explainable by firm or bank variation, it would not matter empirically whether one defined the shocks to be $(\alpha_{ft}, \beta_{bt})$ or $(\alpha_{ft}^I + \gamma_t \delta_{ft}^F, \beta_{bt}^I + \gamma_t \delta_{bt}^B)$.

In general, neither of these conditions will be satisfied, but we can use the intuition from these examples to help us understand the relationship between the two ways of measuring shocks in equations (1) and (2). The key insight is that we can always express the interaction term as $Z_{fbt} = \delta_{ft}^F + \delta_{bt}^B + \zeta_{fbt}$. Writing the interaction term this way, however, suggests a simple way of defining firm and bank shocks such that they are invariant to the inclusion of interaction terms. If one regressed the interaction term on firm and bank fixed effects one could obtain estimates of δ_{ft}^F , δ_{bt}^B , and ζ_{fbt} , where $\hat{\zeta}_{fbt}$ will be orthogonal to both the firm and bank fixed effects by construction. Now suppose that instead of running equation (2), one estimated

$$D(L_{fbt}/L_{fb,t-1}) = \alpha_{ft}^* + \beta_{bt}^* + \gamma_t^* \hat{\zeta}_{fbt} + \epsilon_{fbt}^I. \quad (3)$$

Comparing equations (2) and (3), it is immediately apparent that the estimates must satisfy $\hat{\alpha}_{ft}^* = \hat{\alpha}_{ft}^I + \hat{\gamma}_t^* \hat{\delta}_{ft}^F$, $\hat{\beta}_{bt}^* = \hat{\beta}_{bt}^I + \hat{\gamma}_t^* \hat{\delta}_{bt}^B$, and $\hat{\gamma}_t^* = \hat{\gamma}_t$. Since both $\hat{\zeta}_{fbt}$ and ϵ_{fbt}^I are orthogonal to the firm and bank fixed effects, we can always define $\epsilon_{fbt} \equiv \gamma_t^* \hat{\zeta}_{fbt} + \epsilon_{fbt}^I$, which means that equation (3) is identical to equation (1) and we have $\hat{\alpha}_{ft}^* = \hat{\alpha}_{ft}$ and $\hat{\beta}_{bt}^* = \hat{\beta}_{bt}$. Thus, as long as one is willing to define the components of a bank-firm interaction term that vary only at the bank and firm level (i.e., $\gamma_t \delta_{bt}^B$ and $\gamma_t \delta_{ft}^F$), as components of the bank or firm shock, the inclusion or omission of a bank-firm interaction term will not affect the magnitude of bank and firm shocks.⁷

We summarize this result in the following proposition, which we prove rigorously in the appendix:

Proposition 1. *In the linear loan growth model with bank shocks, firm shocks, and an interaction term given in equation (2), WLS estimation of equation (1) and (2) produces identical estimates of the firm and bank shocks, i.e., $\hat{\alpha}_{ft}^I = \hat{\alpha}_{ft}$ and $\hat{\beta}_{bt}^I = \hat{\beta}_{bt}$ as long as the components of the interaction term that vary only at the bank or firm level are defined to be part of the bank and firm shocks.*

Proof. See Appendix A □

The intuition for why one obtains identical measures of bank and firm loan shocks in models

⁷We can see the differences in the procedures by contemplating a simple example. Suppose that the data generating process is $D_{fbt} = a_{ft} + b_{bt} + \gamma_t a_{ft} b_{bt} + \epsilon_{fbt}^I$, where a_{ft} and b_{bt} are firm and bank shocks; $Z_{fbt} \equiv a_{ft} b_{bt}$ is an interaction term; and $\gamma_t > 0$ is a parameter that reflects a tendency for healthier banks to lend more to healthier firms. We can always rewrite this model as $D_{fbt} = (a_{ft} + \gamma_t \delta_{ft}^F) + (b_{bt} + \gamma_t \delta_{bt}^B) + \gamma_t \hat{\zeta}_{fbt} + \epsilon_{fbt}^I$, where δ_{ft}^F and δ_{bt}^B are the coefficients and error term that arise from regressing Z_{fbt} on firm and bank fixed effects, and $\hat{\zeta}_{fbt} \equiv Z_{fbt} - \delta_{ft}^F - \delta_{bt}^B$. In this case, our estimates of the firm and bank shocks would equal $\hat{\alpha}_{ft} = (\hat{a}_{ft} + \hat{\gamma}_t \delta_{ft}^F)$ and $\hat{\beta}_{bt} = (\hat{b}_{bt} + \hat{\gamma}_t \delta_{bt}^B)$. Obviously, we cannot recover the estimates of a_{ft} and b_{bt} from this procedure. However, there are advantages of defining the firm and bank shocks as α_{ft} and β_{bt} as we do in this paper. In particular, our definition of the interaction term $\hat{\zeta}_{fbt}$ imposes on the data that bank-firm interactions do not affect the aggregate lending of a bank or aggregate borrowing of a firm—these impacts are defined to be parts of the bank and firm shocks.

with and without firm-bank interactions stems from the fact that when one includes an interaction term, the interaction term is only identified if some component of it is orthogonal to the firm and bank fixed effects. This is the formal reason why one cannot include a variable that only varies at the firm or bank level in a fixed effects regression with firm and bank fixed effects. It also explains why we need to define the firm and bank shocks to be the components of lending that only vary at the firm or bank level: we need to ascribe any variation in an interaction term that is perfectly explainable by a firm and a bank fixed effect to a firm and a bank shock.⁸ Failure to do so would result in the possibility of trivially affecting the estimated regression parameters (but not the fit) by including an “interaction” term that, say, only varied at the firm level. The interaction term, therefore, only matters if it contains some component that is orthogonal to the fixed effects, but the orthogonality of the interaction term means that estimation of the fixed effects will be unaffected by the inclusion of the interaction term. As long as one chooses the same *numéraire* (e.g., dropping the same variables in each regression to deal with multicollinearity and attributing the components of the interaction term that vary only at the firm level or the bank level to the firm and bank shocks), the estimated firm and bank shocks will be identical regardless of whether one estimates equation (1) or equation (2). Thus estimation of equation (1) identifies the correct bank and firm shocks given a common choice of *numéraire* regardless of whether the true model is described by equation (2).

Proposition 1 is useful for understanding the role played by bank-firm interaction terms in determining bank and firm loan shocks. While they are important to understanding why a particular bank lends more to a particular firm than another (i.e., $\gamma_t \neq 0$), bank and firm shocks (β_{bt} and α_{ft}) can be consistently estimated without using *any* of this information. This motivates working with equation (1) to obtain the estimates of the bank and firm fixed effects without modeling or estimating the impact of bank-firm interaction terms.

We now can move to tackle the second problem: aggregation. In order to have estimates of bank and firm shocks that are going to be useful for understanding macro lending behavior, we need some way of using these estimates to match the aggregate lending behavior of banks and borrowing behavior of firms. Our first step towards this goal is to prove a proposition related to the growth rates of lending for bank-client pairs with positive loans in period $t - 1$:

Proposition 2. *WLS estimation of equation (1) with $D(L_{fbt}/L_{fb,t-1}) \equiv \frac{L_{fbt} - L_{fb,t-1}}{L_{fb,t-1}}$ and $L_{fb,t-1}$ weights will produce estimates of bank and firm shocks whose loan-weighted average will exactly match the bank, firm, and economy-wide loan growth rates of loan relationships that existed in $t-1$. No other estimates of bank and firm shocks will satisfy this condition.*

⁸Note that here we are applying a definition of a firm and bank shock that stipulates that these shocks capture all variation that is either firm or bank specific as is standard in any fixed effects approach. For example, suppose that there is a troubled firm that ceases borrowing from all banks. In our formulation, we would characterize the idiosyncratic reduction in borrowing by that firm as the firm shock. However, one could also think of this as arising because no bank is willing to lend to that firm. We are therefore defining a firm shock to reflect some factor that causes *all* lending to a particular firm to shift and a bank shock as some factor that causes *all* lending by a particular bank to change.

Proof. See Appendix B □

Proposition 2 establishes that there is unique way to estimate bank and firm shocks such that they match aggregate loan growth as well as the loan growth of every firm and bank. Proposition 1 tells us that these estimates are also correct even for the class of models featuring firm-bank interactions. Taken together, we now have solved the problem of how to estimate firm and bank shocks in such a way that they aggregate to match the loan growth of pre-existing loans.⁹ Most studies use log changes instead of percentage changes as the dependent variable and often use other weighting schemes. This raises the question of how bank shocks estimated using these alternative methodologies compare with those generated in Proposition 2. We summarize the properties of these alternative methodologies in the following proposition:

Proposition 3. *OLS estimation of equation (2) with growth rates defined as percentage changes will yield predicted loan values that exactly match simple averages of individual loan growth rates for firms, banks, and the economy. WLS estimation of equation (2) using loan growth rates defined as log changes will yield predicted loan values that exactly match geometric averages of the individual loan growth rates for firms, banks, and the economy.*

Proof. See Appendix C . □

Proposition 3 provides a theoretical link between various alternative approaches to estimating equation (1). While all of these methods can be used to analyze datasets featuring matched firm bank loan growth data, they differ fundamentally in how the estimates aggregate. To keep our terminology clear, let's call L_{fbt} an "individual" loan, then defining the dependent variable as a percentage change and using OLS produces estimates that will aggregate to match the simple average of the loan growth rates of every individual loan made by a bank, received by a firm, and the economy as a whole. If one weights the data by the lagged loan level (i.e., uses the Proposition 2 method), one will obtain bank- and firm-shock estimates that aggregate to match total firm, bank, and economy growth rates of pre-existing loans. Similarly, if we define loan growth rates as log changes with no weighting, we obtain estimates that match the simple geometric averages of the loan growth rates of every individual loan made by a bank, received by a firm, and in the economy as a whole. If we weight the log changes by lagged lending we obtain *weighted* geometric averages of these loan growth rates.

We now turn to our third problem: the formation and termination of lending relationships. The termination of lending relationships is a problem when loan growth rates are expressed as log changes, but not when growth rates are expressed as percentage changes. This means we can always apply the Proposition 2 methodology if banks just terminate lending relationships but never form new ones. However, this will not be the case in general. Ideally, we would like a

⁹We thank Friederike Niepmann and Benjamin Pugsley for sharing their insights on how our procedure compares with WLS.

procedure that generates bank and firm shocks that maintain all the desirable properties of our estimates obtained in Proposition 2, but can also be applied to datasets in which there is new lending.

Fortunately, there is a simple correction to the estimates obtained by using the WLS procedure with loan growth rates measured as percentage changes. As we show in the proof of Proposition 1, we can always change the normalization by dropping the first firm and first bank from the estimation and rewriting equation (1) as:

$$D(L_{fbt}/L_{fb,t-1}) = c_t + \ddot{\alpha}_{ft} + \ddot{\beta}_{bt} + \epsilon_{fbt} \quad (4)$$

where c_t is a time fixed effect, $\ddot{\alpha}_{ft} \equiv \alpha_{ft} - \alpha_{1t}$, and $\ddot{\beta}_{bt} \equiv \beta_{bt} - \beta_{1t}$. Appendix B demonstrates that WLS estimation imposes the following sample moment equation for each firm:¹⁰

$$D_{ft}^{F+} \equiv \sum_{b \in G_{ft}} \left(\frac{L_{fbt} - L_{fb,t-1}}{L_{fb,t-1}} \right) \theta_{fb,t-1} = \hat{c}_t^W + \hat{\alpha}_{ft}^W + \sum_b \theta_{fb,t-1} \hat{\beta}_{bt}^W, \quad (5)$$

where D_{ft}^{F+} equals the growth rate of borrowing of firm f from all banks that lent it money in $t - 1$; \hat{c}_t^W , $\hat{\alpha}_{ft}^W$, and $\hat{\beta}_{bt}^W$ are the WLS estimates of c_t , $\ddot{\alpha}_{ft}$, and $\ddot{\beta}_{bt}$; G_{ft} is the set of banks lending to firm f in period $t - 1$; and

$$\theta_{fb,t-1} \equiv \frac{L_{fb,t-1}}{\sum_b L_{fb,t-1}}.$$

Similarly, the moment equation for each bank can be written as:

$$D_{bt}^{B+} \equiv \sum_{f \in G_{bt}} \left(\frac{L_{fbt} - L_{fb,t-1}}{L_{fb,t-1}} \right) \phi_{fb,t-1} = \hat{c}_t^W + \hat{\beta}_{bt}^W + \sum_f \phi_{fb,t-1} \hat{\alpha}_{ft}^W, \quad (6)$$

where G_{bt} is the set of firms that borrowed from bank b in $t - 1$,

$$\phi_{fb,t-1} \equiv \frac{L_{fb,t-1}}{\sum_f L_{fb,t-1}},$$

and D_{bt}^{B+} equals the growth rate of lending of bank b to all of its client firms that borrowed in $t - 1$.

If we want estimates of bank and firm shocks to aggregate to match *total* loan growth rates, not the loan growth rates of *pre-existing* loans, we cannot use these sample moment conditions. However, there is a very simple link between total and pre-existing loan growth rates. The total loan growth rate of a bank (D_{bt}^B) is related to the growth rates for existing loan relationships (D_{bt}^{B+}) through the following equation:

¹⁰Equations (5) and (6) follow directly from appendix equations (A12), (A17), and (A19). The moment condition for total loans given in equation (A15) is redundant because if one exactly predicts the growth rate of every bank and firm, one must match the aggregate loan growth rate.

$$\begin{aligned}
D_{bt}^B &\equiv \frac{\sum_f L_{fbt} - \sum_f L_{fb,t-1}}{\sum_f L_{fb,t-1}} = \frac{\sum_{f \in G_{bt}} L_{fbt} - \sum_f L_{fb,t-1} + \sum_{f \notin G_{bt}} L_{fbt}}{\sum_f L_{fb,t-1}} \\
&= \sum_{f \in G_{bt}} \left(\frac{L_{fbt} - L_{fb,t-1}}{L_{fb,t-1}} \right) \frac{L_{fb,t-1}}{\sum_f L_{fb,t-1}} + \frac{\sum_{f \notin G_{bt}} L_{fbt}}{\sum_f L_{fb,t-1}} \\
&= D_{bt}^{B+} + D_{bt}^{BN}
\end{aligned}$$

where D_{bt}^{BN} is new lending by the bank as a share of previous lending. Analogously, we can decompose total firm borrowing growth (D_{ft}^F) into borrowing growth from banks that lent to the firm in the previous period (D_{ft}^{F+}) and loans from new sources divided by past borrowing (D_{ft}^{FN}), i.e., $D_{ft}^F \equiv D_{ft}^{F+} + D_{ft}^{FN}$.

Thus, the reason WLS estimates do not produce bank shocks that aggregate to match total loan growth rates of banks and firms is that the WLS moment conditions require the estimated bank and firm shocks to aggregate in such a way that they match D_{bt}^{B+} and D_{ft}^{F+} and therefore omit the growth of new lending as measured by D_{bt}^{BN} and D_{ft}^{FN} . This suggests a very simple solution to the problem. If we replace the left-hand side of equations (5) and (6) with total loan growth (D_{ft}^F and D_{bt}^B) instead of loan growth to existing clients (D_{ft}^{F+} and D_{bt}^{B+}), our estimated parameters will aggregate to match the total loan growth rates of each firm and bank as well as the economy. Thus, our proposed firm and bank shocks will differ from the WLS ones in that they aggregate to match total loan growth instead of loan growth to pre-existing clients. We show this formally in the following proposition:

Proposition 4. *If we replace D_{bt}^{B+} and D_{ft}^{F+} in moment equations (5) and (6) with the bank's total lending (D_{bt}^B) and firm's borrowing growth rate (D_{ft}^F), the moment conditions become*

$$D_{bt}^B \equiv \frac{\sum_f L_{fbt} - \sum_f L_{fb,t-1}}{\sum_f L_{fb,t-1}} = \hat{c}_t + \hat{\beta}_{bt} + \sum_f \phi_{fb,t-1} \hat{\alpha}_{ft}, \quad (7)$$

$$D_{ft}^F \equiv \frac{\sum_b L_{fbt} - \sum_b L_{fb,t-1}}{\sum_b L_{fb,t-1}} = \hat{c}_t + \hat{\alpha}_{ft} + \sum_b \theta_{fb,t-1} \hat{\beta}_{bt}, \quad (8)$$

and we can uniquely identify bank and firm shocks whose loan-weighted average exactly match firm, bank, and economy-wide loan growth rates. These parameter estimates will be identical to the WLS estimates in the absence of any new lending.

Proof. See Appendix D. □

The fact that equations (7) and (8) are identical to equations (5) and (6) except in the way in which growth rates are computed can provide helpful intuition for the difference between the

WLS procedure given in Proposition 2 and the procedure in Proposition 4. Obviously, if there is no new lending, the WLS procedure described in Proposition 4 will produce the same estimates as the Proposition 4 procedure because $D_{bt}^B = D_{bt}^{B+}$ and $D_{ft}^F = D_{ft}^{F+}$ and so moment equation (5) is the same as equation (8), and moment equation (6) is identical to equation (7).

In order to obtain intuition for how the estimates obtained from the procedure presented in Proposition 4 differs from the WLS procedure of Proposition 2, let's consider WLS estimates drawn from the simplest class of cases: those with two firms and two banks and no new lending relationships. Given that our normalization requires that the bank and firm shocks for the first bank and firm are equal to zero (i.e., they are captured in the constant term), we can express the WLS estimate of the difference between the Bank 2 shock and the Bank 1 shock as¹¹

$$\hat{\beta}_{2t}^W = \frac{(D_{2t}^{B+} - D_{1t}^{B+}) - (\phi_{22} - \phi_{21})(D_{2t}^{F+} - D_{1t}^{F+})}{1 - (\phi_{22} - \phi_{21})(\theta_{22} - \theta_{12})} \quad (9)$$

where we have suppressed the $t - 1$ subscripts on the ϕ and θ parameters for notational simplicity. The terms in parenthesis are the difference in the share of lending to Firm 2 from each bank ($\phi_{22} - \phi_{21}$) and the difference in each firm's share of borrowing from Bank 2 ($\theta_{22} - \theta_{12}$). Since the loan share terms are less than one, the denominator must be positive as long as we are not in a degenerate case in which each firm only borrows from one bank, which would make it impossible to identify separate bank and firm shocks. Clearly, the partial effect of Bank 2 lending relatively more (i.e., a rise in $(D_{2t}^{B+} - D_{1t}^{B+})$) will result in a larger WLS estimate of a relative bank shock ($\hat{\beta}_{2t}^W$), but the full effect depends both on what is happening to bank lending and the interaction of each bank's loan portfolio and $(\phi_{22} - \phi_{21})$ and the relative demands of each borrower ($D_{2t}^{F+} - D_{1t}^{F+}$). For example, if Firm 2 borrows more from Bank 2 than Bank 1 ($\phi_{22} > \phi_{21}$), a rise in the relative growth in the borrowing by Firm 2 ($D_{2t}^{F+} > D_{1t}^{F+}$) will result in a smaller estimated bank shock for Bank 2 because a rise in the total borrowing of Firm 2 means that some of the increase in the differential loan supply growth of Bank 2 should be attributed to Firm 2's borrowing shock.

Now, suppose that there is some new lending in the economy. Without loss of generality assume that initially Firm 2 did not borrow from Bank 2, but then started borrowing (i.e., $D_{2t}^B > D_{2t}^{B+}$, $D_{2t}^F > D_{2t}^{F+}$, $\phi_{22} = 0$, and $\theta_{22} = 0$). We will also assume that there was no new lending between other institutional pairs so that $D_{1t}^B = D_{1t}^{B+}$ and $D_{1t}^F = D_{1t}^{F+}$. Our Proposition 4 procedure will identify the "total" bank shock by replacing the growth rates of pre-existing loans, D_{ft}^{F+} and D_{bt}^{B+} in equations (5) and (6), with $D_{bt}^B \equiv D_{bt}^{B+} + D_{bt}^{BN}$ and $D_{ft}^F \equiv D_{ft}^{F+} + D_{ft}^{FN}$. The isomorphism of equations (5) and (6) with (7) and (8) means that we can write the new estimate of the bank shock as

¹¹This equation follows directly from equation (A41).

$$\begin{aligned}
\hat{\beta}_{2t} &= \frac{(D_{2t}^B - D_{1t}^B) - (\phi_{22} - \phi_{21})(D_{2t}^F - D_{1t}^F)}{1 - (\phi_{22} - \phi_{21})(\theta_{22} - \theta_{12})} \\
&= \frac{(D_{2t}^{B+} + D_{2t}^{BN} - D_{1t}^{B+}) - (0 - \phi_{21})(D_{2t}^{F+} + D_{2t}^{FN} - D_{1t}^{F+})}{1 - (0 - \phi_{21})(0 - \theta_{12})} \\
&= \hat{\beta}_{2t}^W + \frac{D_{2t}^{BN} + \phi_{21}D_{2t}^{FN}}{1 - \phi_{21}\theta_{12}}
\end{aligned} \tag{10}$$

Comparing equations (9) and (10) reveals exactly how the Proposition 2 WLS estimate differs from the Proposition 4 estimate. The first line in equation (10) shows that the formula for the total bank shock equals the formula for the WLS bank shock with the only difference being that we use total lending growth instead of pre-existing loan growth in the formula. In the second line, we decompose total loan growth into its two components: growth of pre-existing loans and new loans as a share of pre-existing loans and also impose the condition that Firm 2 did not initially borrow from Bank 2 (i.e., $\phi_{22} = 0$, and $\theta_{22} = 0$). Finally, in the last line we rearrange terms so that we can express the total bank shock in terms of the WLS estimate of the bank shock, so we can be clear about examining a case in which there is a new loan.

We can see from equation (10) that the estimate of Bank 2's shock based on its total lending will be higher than the Proposition 2 WLS estimate for two reasons. First, if Bank 2 starts a new lending relationship, it will tend to have a larger estimated bank shock because total lending by Bank 2 rises by more than loans to pre-existing clients ($D_{2t}^{BN} > 0$), and the WLS estimate ($\hat{\beta}_{2t}^W$) does not take this into account. The magnitude of the bank shock, will depend not only on Bank 2's lending but also on the importance of this new source of lending for its new client (D_{2t}^{FN}). Bank 2's bank shock is magnified by the fact that Firm 2 started borrowing from Bank 2 but did not expand borrowing from Bank 1. This suggests that the reason for Firm 2's increase in borrowing is more about a factor affecting Bank 2 than either Bank 1 or itself. Finally D_{2t}^{FN} is multiplied by the share of Firm 2's borrowing from Bank 1. This also is intuitive because if Firm 2 is heavily dependent on Bank 1 for lending (ϕ_{21} is large), but it sources new loans from Bank 2 (rather than borrowing more from Bank 1), it indicates that the shock to the system was due to a change in Bank 2's supply of credit, not a change in Firm 2's demand.

In Section 4, we provide evidence on the validity of the bank shock estimates we obtain. In addition to estimating bank shocks using our Proposition 4 methodology, we also show how one can extend the framework to estimate a simple interaction term that enables bank shocks that vary by firm. While we know from Proposition 1 that this should not affect our estimates of the aggregate bank shock, we do consider a specification in which we allow for banks to behave differentially to different borrowers. For example, in Section 5.2.2 we show that our results are robust to allowing our estimated shocks to vary with the health of the firm and the bank.

2.2 Decomposing Bank Lending and Aggregate Lending

In the second section of Appendix D , we show that the fact that there is a unique set of firm and bank shocks that satisfy the adding-up constraints given by equations (7) and (8) enables us to obtain an exact decomposition of each bank’s aggregate lending into four terms, as follows:¹²

$$\mathbf{D}_{Bt} = (\bar{A}_t + \bar{B}_t) \mathbf{1}_B + \Phi_{t-1} \mathbf{N}_t + \Phi_{t-1} \tilde{\mathbf{A}}_t + \tilde{\mathbf{B}}_t, \quad (11)$$

where \mathbf{D}_{Bt} is a $B \times 1$ vector whose elements are each bank’s total loan growth in year t ; $(\bar{A}_t + \bar{B}_t)$ are the median firm and bank shocks in year t , which reflects any shocks that would affect all lending pairs identically in a year; $\mathbf{1}_B$ is a $B \times 1$ vector of 1’s; \mathbf{N}_t is a vector containing the median firm shock in the industry containing the firm; Φ_t is a matrix that contains as elements the weights of each loan to every borrower in time t , i.e.,

$$\Phi_t \equiv \begin{pmatrix} \phi_{11t} & \dots & \phi_{F1t} \\ \vdots & \ddots & \vdots \\ \phi_{1Bt} & \dots & \phi_{FBt} \end{pmatrix};$$

$\tilde{\mathbf{A}}_t$ is a vector with elements $\tilde{\alpha}_{ft}$, each of which equals firm shock in year t less the median firm shock in that firm’s industry in year t ; and $\tilde{\mathbf{B}}_t$ is a vector with elements $\tilde{\beta}_{bt}$, each of which equals the bank shock in year t less the median bank shock in that year.¹³

The key feature of equation (11) is that one can exactly decompose each bank’s loan growth into four elements. The first term measures “common shocks”: changes in lending that are common to all lending pairs. These shocks measure any force that would cause all lending to rise or fall by the same amount (such as due to an interest rate change). The second term is the “industry shock”: a bank-specific weighted average of the industry shocks affecting each of the bank’s borrowers. It measures changes in lending that arise because a bank might have a loan portfolio that is skewed toward borrowers in certain industries. The industry shock captures forces that might cause a bank’s lending to deviate from the typical bank’s because it specialized in lending to particular industries. We refer to the third term as the “firm-borrowing shock” or firm shock because it captures changes in a bank’s lending that arise due to the idiosyncratic changes in borrowing demand of their clients that cannot be attributed to changes in bank-loan supply.

Finally, the last term captures the idiosyncratic bank-supply shock or “bank shock” because it measures changes in a bank’s loan supply that is independent of anything related to the firms, industries, or common shocks hitting the economy. The elements of $\tilde{\mathbf{B}}_t$ equal bank b ’s supply shock in year t less the supply shock of the median bank in that year. Thus, if all banks except

¹²One could imagine alternative decompositions, but this one expresses loan growth in terms of variables that have been of particular theoretical interest.

¹³We could have defined the decompositions in equation (11) using the mean shock instead of the median. We believe the median is more appropriate because it reflects the shocks affecting the typical bank and firm, and the mean shock is much more sensitive to extreme shocks hitting small firms, which often experience enormous swings in borrowing growth. In an online appendix, we show that our results are qualitatively the same whether we define bank shocks using the mean or median.

bank b suffered a negative 10 percent shock while bank b had a negative 5 percent shock, this would be isomorphic in our framework to bank b experiencing a 10 percent positive shock and all other banks experiencing a positive 5 shock. Since the supply shocks are already purged of all factors affecting their borrowers, our measure of idiosyncratic bank shocks ($\tilde{\beta}_{bt}$) reflects what is happening at each bank *relative* to the typical bank.

Now that we have developed a methodology for decomposing bank lending into firm, bank, industry and common shocks, we can turn our attention to the task of understanding how these shocks affect aggregate lending. In order to do this, we need a little more notation. Let $w_{b,t}^B$ be the average share of bank b in total lending in year t and define $\mathbf{W}_{B,t} \equiv [w_{1t}^B, \dots, w_{Bt}^B]$. We can now use equation (11) to obtain

$$\mathbf{D}_t = \mathbf{W}_{B,t-1} \mathbf{D}_{Bt} = (\bar{A}_t + \bar{B}_t) + \mathbf{W}_{B,t-1} \Phi_{t-1} \mathbf{N}_t + \mathbf{W}_{B,t-1} \Phi_{t-1} \tilde{\mathbf{A}}_t + \mathbf{W}_{B,t-1} \tilde{\mathbf{B}}_t. \quad (12)$$

It is worth pausing a moment to contemplate the implications of equation (12). This equation decomposes aggregate loan growth, \mathbf{D}_t , into four terms based on the firm-borrowing and bank-lending channels. The first term captures the impact of common shocks on aggregate lending by measuring what happens to the lending of the typical bank-firm pair.¹⁴ The second term represents the “granular industry shock” because it captures the interaction between industry shocks and the size of the industries. The size of this term will depend on the degree of aggregation used and the variance of shocks within an industry. The third term is the “granular firm shock” because it measures the importance of firm-borrowing shocks on aggregate lending. This term will be small if demand shocks are small or if the loan share of every borrower tends to be small. Finally, we refer to the last term as the “granular bank shock” because it is a weighted average of all the idiosyncratic financial institution shocks.

Our decomposition of aggregate lending into the four channels differs in important ways from other studies. First, prior work on granular bank shocks has followed Gabaix (2011) and assumed that bank-supply and firm-borrowing shocks are uncorrelated across and between firms and banks. Equations (11) and (12) are more general in that we only need to assume that these shocks are not perfectly correlated. Second, the estimates of the bank-lending and firm-borrowing channels are consistent with the aggregate borrowing by firms and lending by banks.

Granular bank-supply shocks are likely to be particularly important for aggregate lending fluctuations if lending markets are concentrated. The reason stems from the fact that the magnitude of granular bank-supply shocks depends on two factors: the variance of idiosyncratic bank shocks ($\tilde{\mathbf{B}}_t$) and the existence of large financial institutions (i.e. some of the elements of \mathbf{W}_{Bt} are not small). As Gabaix (2011) has shown, if all institutions were sufficiently small or if their shocks were sufficiently small, then one should expect this term to be small because, on average, these shocks should cancel out due to the law of large numbers. However, as we will see in the next

¹⁴As we explain in Appendix A, our methodology does not let us separate how much of the common shock is due to firm-borrowing vs. bank-lending effects. We can only identify the sum of the two effects.

section, financial institutions are indeed quite large compared to the aggregate loan market and have loan shocks that are idiosyncratic. These facts explain why we find in our econometric section that granular bank shocks matter enormously for aggregate fluctuations.

2.3 The Link Between Firm-Level Investment-Rate Regressions and Aggregate Investment

A related class of problems surrounds how to assess the impact of idiosyncratic bank shocks on aggregate variables. We propose a “micro” and a “macro” approach. Consider a regression of the form:

$$\mathbf{y} = \mathbf{X}\Upsilon + \boldsymbol{\epsilon}, \quad (13)$$

where \mathbf{y} is the dependent variable whose elements in our case equal the firm-level investment rate ($I_{ft}/K_{f,t-1}$), where I_{ft} is firm-level investment and $K_{f,t-1}$ is the firm-level capital stock; \mathbf{X} is a matrix of explanatory variables; Υ is a vector of parameters; and $\boldsymbol{\epsilon}$ is an iid error term. In order to keep the discussion related to the literature on assessing the impact of bank shocks on firm-level investment rates, we should think of the columns of \mathbf{X} as containing firm-level variables like cash flow, Tobin’s Q, and idiosyncratic bank supplyshocks. Moreover, since bank supply shocks might be correlated with industry and other firm variables, we will need to control for these factors as well. If we estimate this equation using WLS, the estimated residuals must sum to zero, which requires that the estimated parameters, $\hat{\Upsilon}$, satisfy the following sample moment condition:

$$\mathbf{1}'\mathbf{W}\mathbf{y} - \mathbf{1}'\mathbf{W}\mathbf{X}\hat{\Upsilon} = 0, \quad (14)$$

where \mathbf{W} is a diagonal weighting matrix, and $\mathbf{1}$ is vector of ones.

It is useful to define the sum of investment over the whole time period as $I_t \equiv \sum_f I_{ft}$ and the sum of the lag of capital stock in all periods as $K_t \equiv \sum_f K_{ft}$. Thus, the average investment rate for the economy over this time period is $\frac{1}{T} \sum_t I_t/K_{t-1}$, where T denotes the number of years. Suppose we set the diagonal elements of the weighting matrix equal to $\frac{1}{T} K_{f,t-1}/K_{t-1}$. In this case $\mathbf{1}'\mathbf{W}\mathbf{y}$ will equal $\frac{1}{T} \sum_t \left\{ \sum_f [(I_{ft}/K_{f,t-1}) K_{f,t-1}/K_{t-1}] \right\} = \frac{1}{T} \sum_t I_t/K_{t-1}$ over the sample period. Thus, there exists a set of WLS parameter estimates ($\hat{\Upsilon}$) of investment rates that exactly matches the average aggregate investment rate over all years in the sample. In any time period, t , we can write the forecast of the investment rate as $\hat{y}_t(\mathbf{X}_t) \equiv \mathbf{X}_t \hat{\Upsilon}$, where \mathbf{X}_t is the matrix of explanatory variables corresponding to period t :

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_T \end{pmatrix} \quad (15)$$

Crucially, while this weighting procedure will produce estimates that match the average aggregate investment rate in the sample over the whole time period, the estimates need not match

the aggregate investment rate in each year. $\hat{y}_t(\mathbf{X}_t)$ is the theoretically correct way to aggregate estimates of bank shocks on firm-level investment rates to yield estimates of aggregate investment rates. The R^2 from regressing y_t on $\hat{y}_t(\mathbf{X}_t)$ tells us how well our micro estimates forecast aggregate investment. This also provides a simple way of assessing the impact of bank shocks on aggregate investment rates. Let \mathbf{X}_t^0 be equal \mathbf{X}_t with all of the idiosyncratic bank shock variables set to zero. $\hat{y}_t(\mathbf{X}_t^0)$, therefore, is a counterfactual prediction of what aggregate investment would have been if all of the idiosyncratic bank shocks had been zero. If we compare the R^2 from regressing y_t on $\hat{y}_t(\mathbf{X}_t)$ with the R^2 from regressing y_t on $\hat{y}_t(\mathbf{X}_t^0)$, we can assess how important bank shocks are for understanding *aggregate* investment rate fluctuations.

We can now compute two goodness of fit measures that can be informative about how much idiosyncratic bank shocks affects investment rates. The R^2 that arises from equation (13) tells us how well we fit the investment behavior of each firm. However, if we are interested in how well we fit the *aggregate* investment behavior, we want to know the R^2 from regressing y_t on $\hat{y}_t(\mathbf{X}_t)$. The R^2 from the firm-level regression need not be informative about the R^2 from the aggregate regression if, for example, there is a lot of idiosyncratic firm-level variation in investment rates that cancels as we aggregate.

One problem with this micro approach is that it can only be applied to a dataset in which we observe all of the firm-level variables contained in \mathbf{X} . Given that we only observe investment rates and lending to listed firms, we may also want a method that lets us understand how idiosyncratic bank shocks affect investment rates for the much broader sample of firms comprising the entire economy. In order to assess this, we can also conduct a macro approach that entails regressing the economy-wide investment rate on the granular lending shock terms given in the right-hand side of equation (12). Since the idiosyncratic bank shocks that make up the granular term are estimated using firm-time fixed effects, they should not be affected by changes in firm loan demand. Therefore, by varying whether we include the granular bank shock term in this macro regression, we can assess how much of the variation in economy-wide investment rates is due to granular bank shocks.

3 Data Description

3.1 Data Construction

Our data come from four sources. First, we use matched bank-firm loan data from Nikkei NEEDS FinancialQUEST for the period 1990 to 2010. Nikkei reports all short-term and long-term loans from each financial institution for every company on any Japanese stock exchange, which we sum to obtain total loans: 272,302 loans in total. Our definition of a “bank” covers all Japanese city, trust, regional, mutual banks, and insurance companies, as well as Japanese holding companies. We include loans from all financial institutions, except for the eleven that are government banks

or cooperatives. We dropped loans from government institutions such as the Development Bank of Japan and the Norinchukin Bank because we wanted to focus our results on the impact of bank-supply shocks arising from private institutions on aggregate lending.¹⁵ Our loan measure is the total borrowing from a given bank in a year, comprising all loans received from each bank for 870 to 1,633 firms per year.¹⁶ Our data cover loans from finance, insurance, and financial holding companies to manufacturing, mining, agriculture, and services (except finance and insurance). We divide the industries using the JSIC 2-digit codes, comprising 78 industries.

In general, the Japanese fiscal year runs from April in year t to March in year $t + 1$. More than 80 percent of the firms report annual loan data for the fiscal year ending in March, and the rest of the firms report loans ending in one of the other months. For most of our analysis, we include only firms that report for the year ending in March so that a year is defined over the same time period for all of our firms.¹⁷ We refer to years ending in March 31 as a “fiscal year” and denote such years by the prefix “FY” to distinguish them from calendar years. Because nine months of any fiscal year occur in the previous calendar year, one should remember that a fiscal year tends to refer to information that is lagged by one year relative to a calendar year. For example, Hokkaido Takushoku Bank, which failed in November 1997, fails in FY1998 because the 1998 fiscal year closes in March 1998.

One difficulty working with these data is tracking mergers and restructurings. Whenever a bank ceases to exist, due to either bankruptcy or merger, firms will cease reporting that bank as a source of loans. We investigated every bank in order to see if there was any report in the media of a bankruptcy or merger. If we could not find any report, we assumed that the zero loan data were accurate, but if we could find evidence of a failure or a merger, we recorded the date. Since firms sometimes reported loans coming from a restructured bank as coming from the prior institution, we recoded these loans as coming from the restructured institution if they occurred after the restructuring. In order to compute the loan growth of a new institution, we had to keep track of all the institutions that predated it. Thus, if Banks 1 and 2 merged in year t to form Bank 3, Bank 3’s loans in year $t - 1$ would be set equal to the sum of the loans of Banks 1 and 2, and the growth rate would be computed accordingly. Since we could trace the evolution of hundreds of banks in our data, we do not have any gaps associated with mergers. The entry of new banks is easy to incorporate as well because all new banks emerge in our sample as the result of mergers or the restructuring of existing institutions.¹⁸

A related issue concerns the definition of a bank. In general, we erred on the side of assuming that institutions changed whenever an institution was nationalized or privatized. For example,

¹⁵We show in an online appendix that our results are robust to excluding firms that borrow from these banks.

¹⁶The loan data do not distinguish between new loans and rolling over of existing credit lines. Our identification strategy does not depend on how banks restrict lending.

¹⁷Our results are robust to the inclusion of all months.

¹⁸The aggregate loan growth rate we estimate is the loan growth rate to listed firms that borrowed any positive total amount in period $t - 1$. All exits are incorporated.

Long-Term Credit Bank failed in 1998 and then, after an interval of nationalized control, reopened as Shinsei Bank in 2000, so the bank appears in our data as different institutions for each of these periods: LTCB, “Nationalized LTCB,” and Shinsei Bank. Similarly, bank lending activity does not change as the set of listed firms changes because we do not include a firm in our dataset in year t unless it has positive borrowing in year $t - 1$. Thus, if a firm is newly listed in year t , we only consider its contribution to aggregate loan growth from t to $t + 1$. This procedure prevents us from erroneously attributing loan growth to a change in our sample due to an increase in the number of listed firms.

To ensure sufficient observations to estimate the bank shocks, we keep only bank-year pairs that have a minimum of five loans in both t and $t - 1$. This procedure dropped 0.6 percent of the observations. The number of banks in the sample ranges from 101 to 166 depending on the year, with a smaller number of banks in the later period resulting from the wave of mergers in the 2000s.

The second source of data is the Development Bank of Japan (DBJ) database of unconsolidated reports, which provides information on a wide range of firm characteristics. We use information on investment, capital, total borrowings, bonds, cash, total assets, and the firm’s book value for our analysis (while the market value is from Nikkei). Our measure of investment is constructed as the annual difference in total tangible fixed assets plus depreciation, and the market-to-book value is the ratio of market value to shareholder equity.

Finally, we draw on two sources of aggregate information on economy-wide borrowings and investment-to-capital ratios. The flow-of-funds data from the Bank of Japan website provide data on the stock of lending to private nonfinancial corporations from private financial institutions.¹⁹ The economy-wide investment-to-capital ratio data are measured in 2000 yen for fiscal years 1990 to 2010 from the National Accounts, Economic and Social Research Institute, Cabinet Office. The summary statistics are reported in Appendix E.

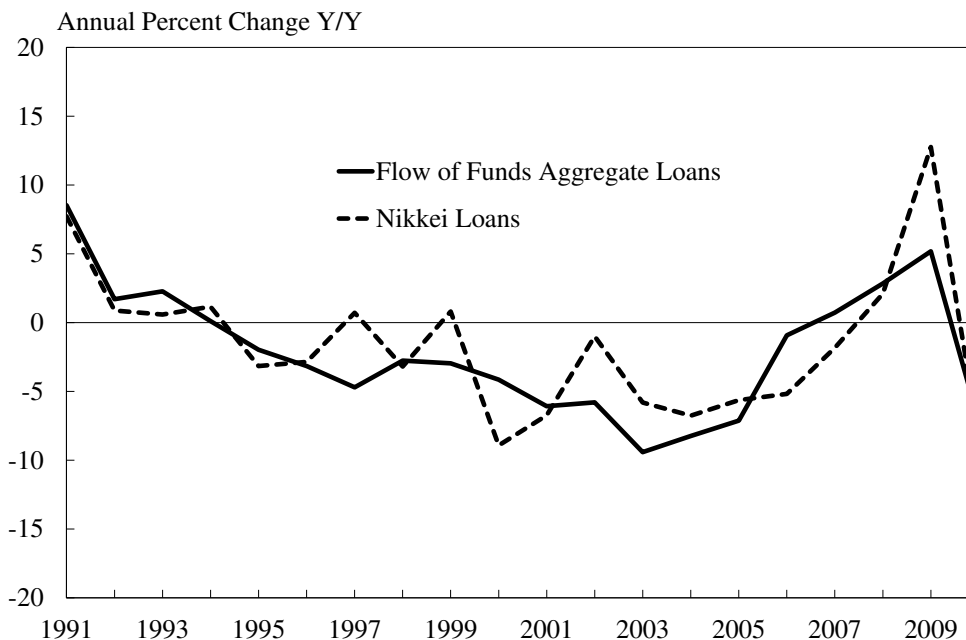
3.2 Data Preview

In this section, we highlight some key patterns in the data. First, we show that the pattern of aggregate lending from our firm-level data exhibits similar year-to-year fluctuations as those from the official economy-wide statistics, demonstrating the plausibility of using information on listed companies as a means of understanding aggregate fluctuations. Even though lending to listed nonfinancial firms in the Nikkei data accounts for only 17 percent of aggregate lending to nonfinancial enterprises in 1990 and 18 percent in 2010, the aggregate lending to listed companies from our firm-level data exhibits a similar pattern as commercial lending in the broad economy.

Figure 1 plots the annual percent change in loans to nonfinancial enterprises using the flow-of-funds lending data and the annual percent change in aggregate loans to nonfinancial enterprises

¹⁹The borrowing data are from the Bank of Japan website. The series number is FF’FOF_FFYS411L240.

Figure 1: Flow of Funds and Nikkei Debt



Note: Years are fiscal years, which roughly correspond to the calendar year plus one.

from the Nikkei firm-level database. As can be seen from this graph, aggregate corporate loans track that of listed companies extremely closely, with a correlation of 0.8. The figure makes clear that corporate borrowing fell off sharply as the bubble burst in FY1990. By FY1995, the growth rate of lending became negative and remained so for a decade.

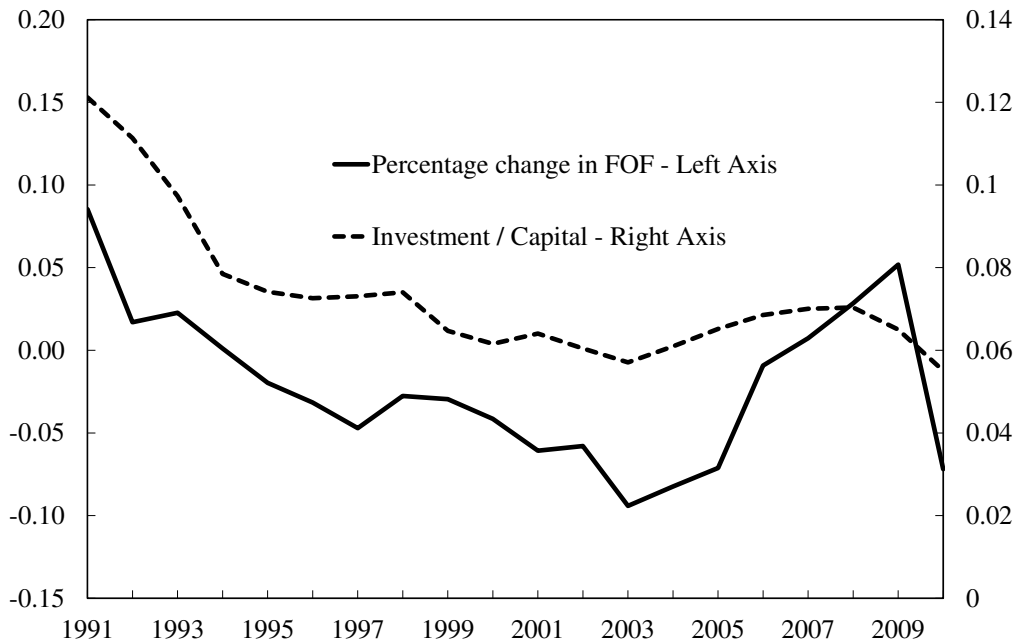
Second, we show a clear positive association between economy-wide nonfinancial corporate borrowing and economy-wide investment in the data. Figure 2 indicates that the growth rate in the stock of lending to private nonfinancial corporations from private financial institutions using the flow-of-funds data tracks the aggregate investment-to-capital ratio fairly closely. The correlation between the two series is 0.72, illustrating the tight relationship between borrowing and investment.

Third, a critical feature of our identification strategy is the exploitation of the fact that Japanese listed companies typically borrow from a large number of banks. Figure 3 presents a histogram of the number of institutions providing loans to each firm. The median firm borrowed from seven banks and 97 percent of the firms in our sample borrowed from more than one.²⁰ Moreover, since the average firm's borrowing Herfindahl index averaged 0.17, we know that the typical firm spread its borrowing out relatively evenly across many banks.²¹

²⁰This is in sharp contrast with the data underlying Khwaja and Mian (2008) in which only 10 percent of the firms borrowed from more than one bank.

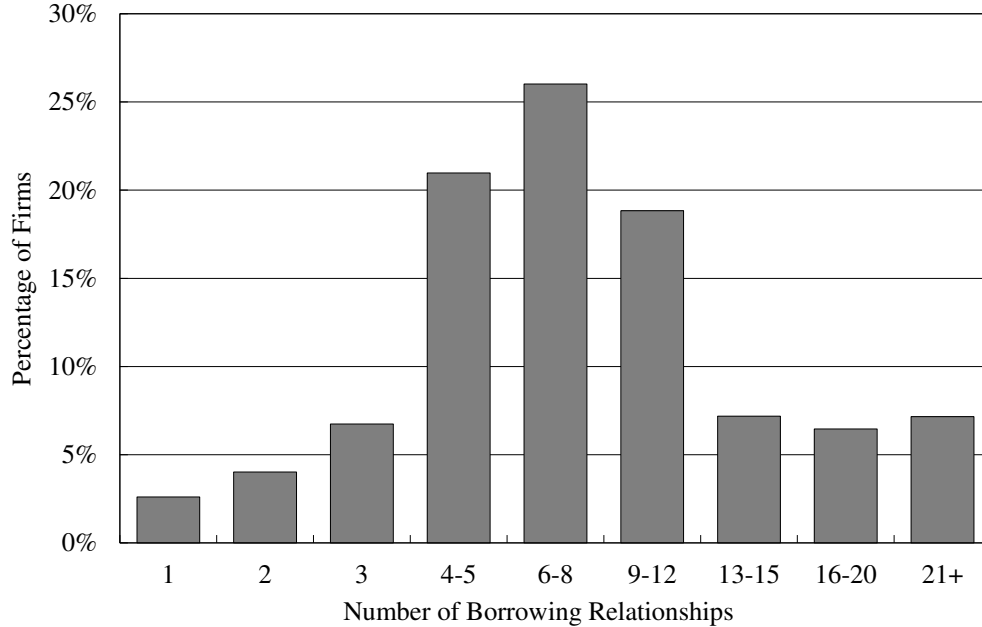
²¹The firm borrowing Herfindahl index ($= \frac{1}{F_t} \sum_f \sum_b (L_{fbt} / \sum_b L_{fbt})^2$) measures how concentrated each firm's borrowing is on average. A Herfindahl index of 1 arises if firms only borrow from one institution. A Herfindahl index of zero would arise if each firm spreads its borrowing evenly across an infinite number of banks.

Figure 2: Flow of Funds and Aggregate Investment



Note: Years are fiscal years which roughly correspond to the calendar year plus one.

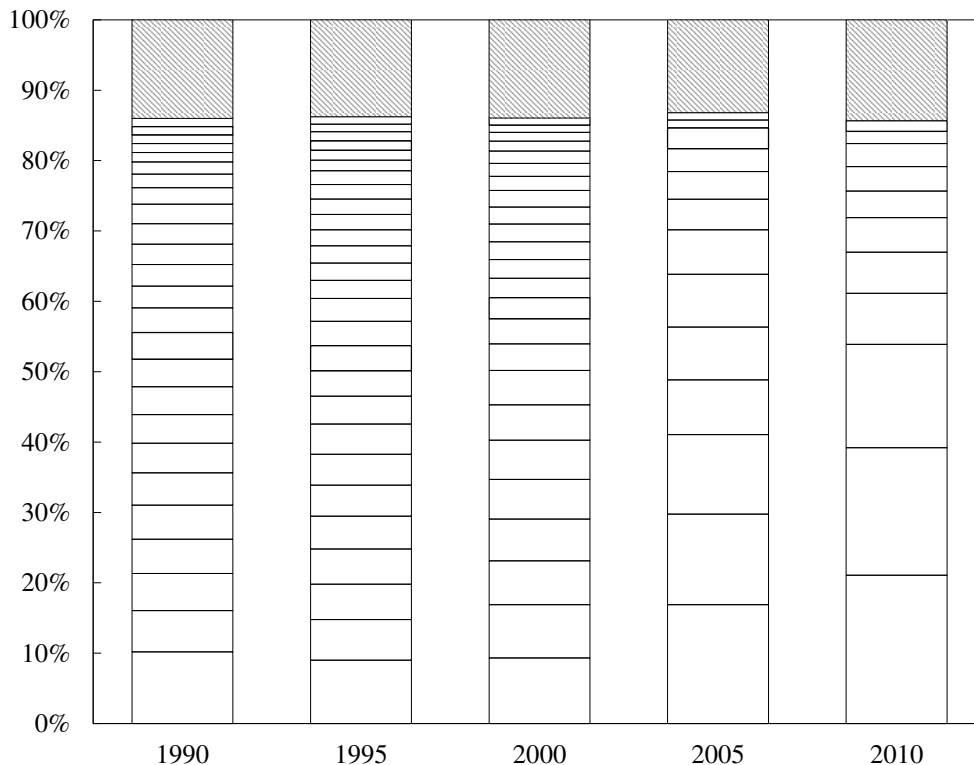
Figure 3: Distribution of Number of Loans per Firm



Finally, we need to show sufficient “granularity” in the financial sector so it can plausibly be argued that shocks to major banks in Japan are large enough that an idiosyncratic shock at one of them, such as the failure of Long-Term Credit Bank, might actually move a macroeconomic aggregate. Figure 4 shows a breakdown of total lending by bank. For each year, we depict

individual loan shares of every bank with a loan share exceeding one percent of aggregate lending to listed companies, and group those with a market share of less than one percent into the shaded region.

Figure 4: Bank Concentration



Note: Years are fiscal years, which roughly correspond to the calendar year plus one.

As the figure shows, Japanese finance has always been dominated by relatively few financial institutions. In FY1990, the three largest Japanese banks accounted for 23 percent of aggregate lending to listed companies, and this number rose to 54 percent in FY2010, with the largest financial institution, Mitsubishi UFJ Financial Group (MUFG), accounting for 21 percent of all Japanese lending.²² The concentration in lending in the world’s second largest economy at the time provides the basic motivation for our suggestion that idiosyncratic bank shocks might have macroeconomic implications. If banks are large relative to the size of the economy, idiosyncratic shocks to particular institutions could move macro aggregates if firms have difficulty substituting between different sources of finance.

Interestingly, Japan’s high levels of concentration do not make the country an outlier. For example, in the U.S., most major lenders are bank holding companies, which in 2010 jointly

²²As Amel et al. (2004) show, the growth in merger activity in Japan’s banking and insurance industry was quite similar to that in other industrial countries. The increase in concentration in Japanese finance between FY2000 and FY2010 was driven by deregulating laws related to the formation of holding companies in 1997 and, as Sakuragawa and Watanabe (2009) argue, the 2002 “Takenaka Plan,” which forced more disclosure of nonperforming loans (resulting in mergers of weak institutions).

accounted for 79 percent of all assets held by commercial banks, thrifts, and credit unions.²³ Federal Reserve data indicate that the three largest institutions in the U.S.—Bank of America, JP Morgan, and Citigroup—held 49 percent of all banking assets.²⁴ This number is remarkably close to the 54 percent number in our Japanese sample. Similarly, Buch and Neugebauer (2011) calculate bank Herfindahl indexes for many western European countries that are similar to those we obtain for Japan. Therefore, any observed large impact from idiosyncratic bank shocks in Japan cannot be attributed to a more concentrated banking sector than in other countries. Rather, the high degree of financial market concentration appears to be a feature of many developed countries.

4 Estimating the Bank Shocks

Before turning to our main results on the impact of idiosyncratic bank shocks on investment, we explore in this section the unique features of our methodology and conduct some external validation tests of our bank shock estimates to ensure that our point estimates are reasonable. We first examine how our estimates compare with standard fixed-effects estimates. Then, we explore the external validity of our estimates by checking that they are correlated with conventional measures of bank shocks.

4.1 Comparing Estimation Methods

We have already discussed how our approach is related to OLS, WLS and various other specifications in theory. We now turn to assessing how much these differences matter in practice. In order to establish a baseline to understand how much our procedure improves the fit relative to OLS, we first estimate equation (1) using OLS fixed effects to obtain estimates of the firm and bank shocks and then sum them across all client firms to obtain the bank’s growth rate.²⁵ We then can see how well the fixed-effects procedure works by comparing a bank’s actual loan growth, D_{bt}^B , with the predicted values, \hat{D}_{bt}^B . We know from Proposition 3 that these parameters will aggregate to match the unweighted average bank loan growth rate. The question we now address is how informative is the average loan growth rate for understanding a bank’s total loan growth rate. The results from this exercise suggest that using fixed effects to identify the bank-lending channel produces estimates that are almost uncorrelated with actual bank-loan growth. If we regress the bank’s predicted loan growth from the OLS regression on the actual growth, we obtain an R^2 of only 0.08. We plot these data in Figure 5 to show that this result is not driven by a particular

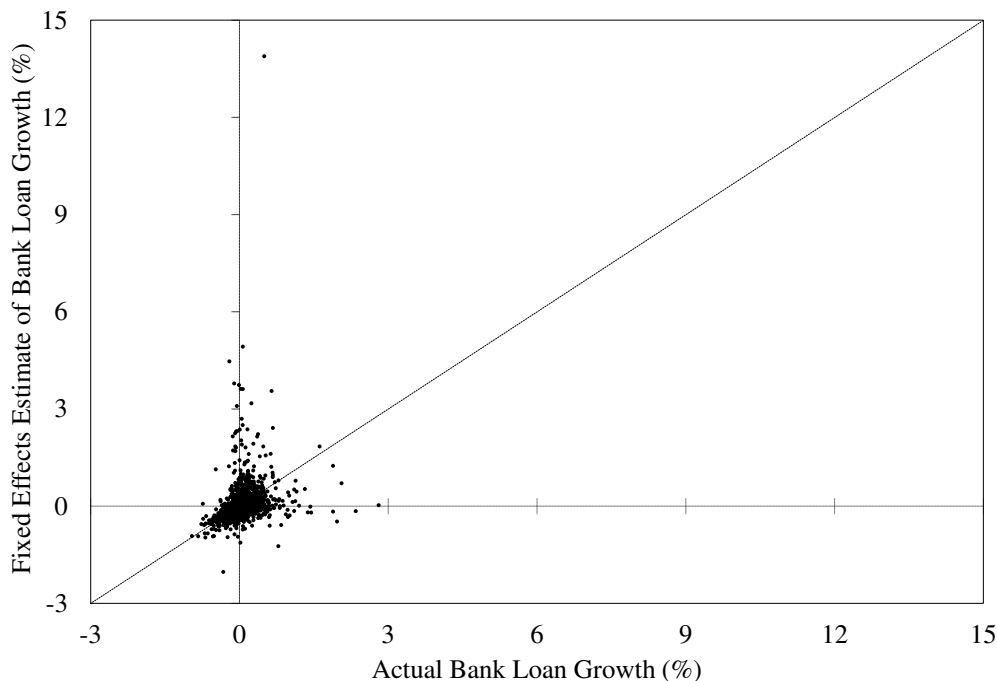
²³Federal Financial Institutions Examination Council (2011) *Annual Report 2010*, March 31.

²⁴The bank asset numbers come from the Bank Holding Company Performance Reports, which are available on the National Information Center website: (http://www.ffiec.gov/nicpubweb/content/BHCPRRPT/BHCPR_Peer.htm).

²⁵Formally, the predicted value of a bank’s loan growth is $\hat{D}_{bt}^B = \sum_{f \in G_{bt}} (\hat{c}_t + \hat{\alpha}_{ft} + \hat{\beta}_{bt})$, where the definitions of all of these parameters are given in Appendixes A and B.

institution. By contrast, if we implement our Proposition 4 methodology, the R^2 is equal to one by construction. This establishes that one needs to be quite cautious about using econometric specifications aimed at matching average growth rates to infer what is happening in aggregate at any given bank.

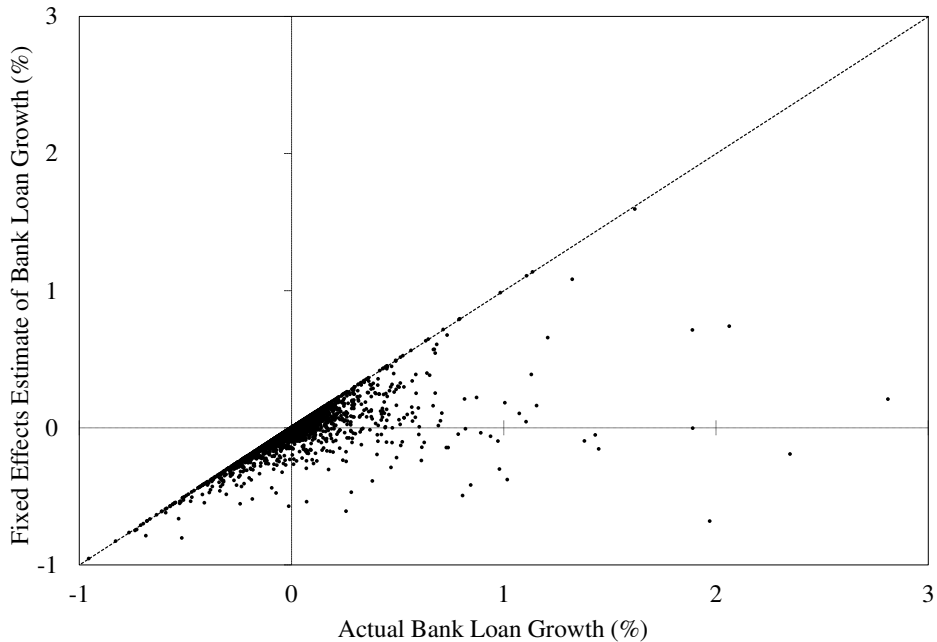
Figure 5: Predicted Bank Loan Growth Using OLS vs. Actual



Note: Predicted values are based on OLS estimates of equation (1) with loan growth expressed in percentage changes.

As we showed in Proposition 3, a log-change specification will produce estimates that aggregate to match the geometric average of the growth rate of each loan made by a bank (omitting the zeros). It is an empirical question whether this will provide parameter estimates that more closely match total bank lending. Replacing the dependent variable in equation (1) with the log change in loans in the OLS specification does not do much to improve the fit of equation (1)—the R^2 from regressing the fitted values of aggregate bank lending on actual bank lending rises to 0.25, which still leaves most of aggregate bank lending unexplained. In other words, bank and firm shocks obtained from a log change specification also do not do well to predict aggregate loan growth behavior.

Figure 6: Predicted Bank Loan Growth Using Weighted Least Squares vs. Actual



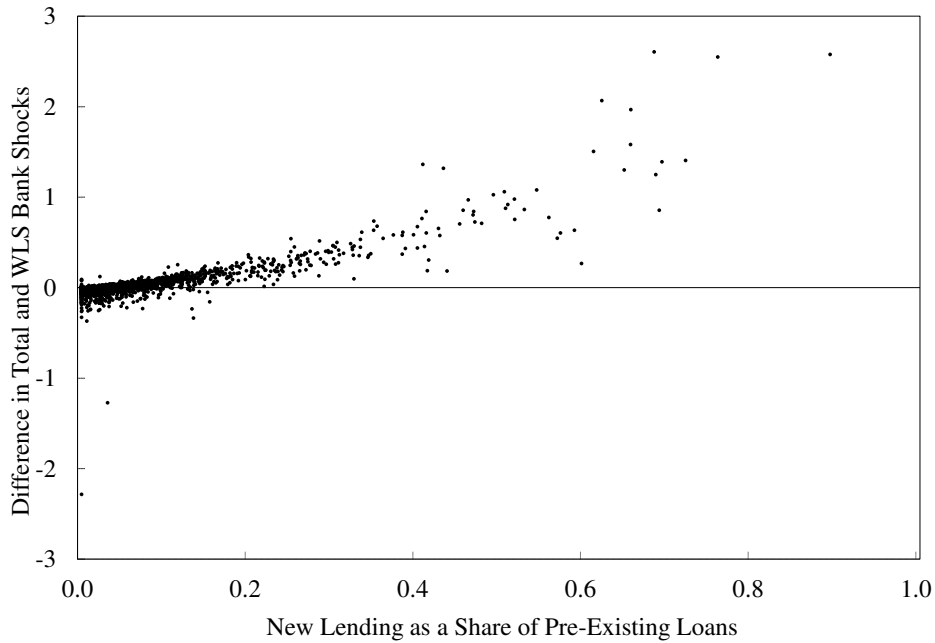
Note: Predicted values are based on weighted least squares estimates of equation (1) with loan growth defined as percentage changes, which are then aggregated according to right-hand side of equation (6).

The next step is to see how much the fit improves if we continue with the log specification but estimate equation (1) using weighted least squares. It is common to weight the data in a log specification by the size of the loan, $L_{fb,t-1}$. As Proposition 3 proved, this produces parameter estimates that aggregate to match the weighted geometric average of the growth rate of each loan made by a bank. Again, it is an empirical question how much this improves the fit. We find that using a weighted log specification only raises the R^2 to 0.36. Thus, the difficulty of identifying what is happening at the bank from individual loan information is not easily solved by using a weighted log specification.

Proposition 2 demonstrated that WLS estimates of equation (1) with the dependent variable defined as a percentage change in loan growth and using $L_{fb,t-1}$ weights produces bank and firm shock estimates that exactly match a bank's loan growth rates of pre-existing loans (D_{bt}^{B+}). Proposition 4 shows that WLS produces identical parameter estimates as our preferred methodology if there is no new lending, i.e., $D_{bt}^{B+} = D_{bt}^B$. If new lending is not important, it would be the case that WLS would produce very similar estimates as our Proposition 2 method. We assess this in Figure 6, where we plot the estimates of bank loan growth produced by WLS estimation of equation (1) (i.e., D_{bt}^{B+}) against total bank loan growth (D_{bt}^B). We see that the WLS estimates systematically underestimate bank loan growth by sometimes fairly large margins because all new lending activity is omitted. Regressing the WLS fitted values of a bank's total loans on the actual values (D_{bt}^B) produces an R^2 of only 0.51, with a slope of 0.54, which indicates that about half

of the variance in total bank loan growth comes from new loan relationships: a source of growth omitted from the WLS specification.

Figure 7: Impact of New Lending on Bank Shock Estimates



Note: Difference in total and WLS bank shocks is $(\hat{\beta}_{bt} - \hat{\beta}_{bt}^W)$ in equation (10) . New bank lending as a share of pre-existing loans is (D_{bt}^{BN}) .

A related question concerns how the estimates of bank shocks are affected by the inclusion of new lending by banks. We saw from equation (10) that the Proposition 4 estimates of bank shocks $(\hat{\beta}_{bt})$ tend to be larger than WLS estimates $(\hat{\beta}_{bt}^W)$ when, *ceteris paribus*, new lending as a share of total lending is high. We can see this effect very clearly in Figure 7, which plots the difference between our preferred estimates of bank shocks and the WLS estimates $(\hat{\beta}_{bt} - \hat{\beta}_{bt}^W)$ against new lending activity (D_{bt}^{BN}) . As the econometric theory suggests, there tend to be very small differences in the two sets of estimates when banks do not engage in much new lending activity, but the shocks can differ by a magnitude of two or more for banks that make a lot of new loans. Similarly, if we regress the bank shocks estimated using weighted least squares $(\hat{\beta}_{bt}^W)$ on bank shocks computed using our approach $(\hat{\beta}_{bt})$, we obtain an R^2 of only 0.59 and a coefficient of 0.57. Intuitively, the WLS estimates tend to be biased downwards because they do not incorporate bank shocks that result in new loans. Taken together, these results suggest that using estimates that do not aggregate and ignoring new lending activity produces noisy and biased bank shock estimates.

4.2 External Validity of Bank Shock Estimates

While our methodology generates estimates that match aggregate data by construction, we also would like some reassurance that they make sense. One way to evaluate how reasonable our estimates are is to look at their values in situations where we have a strong prior for what they should be. We pursue this in three ways. First, we examine extreme financial events affecting institutions (e.g., bank failures) and see what types of bank shocks we estimate. Second, we look at extreme bank shock estimates and examine if anything unusual happened to the institution. Third, in order to provide systematic evidence for the validity of our bank shock estimates, we examine whether they are consistent with proxy measures of bank shocks used in prior studies. We consider each in turn.

4.2.1 Extreme Events and Estimates

We have a strong intuition that financial institutions dramatically curtail lending when they are on the verge of failure. This result is clearly apparent in the bank-shock measures. Financial institutions that entered bankruptcy, as opposed to merging or receiving capital injections when financially stressed, had a mean and median idiosyncratic bank shock in the year of their failure of -9 percent. Some of the major institutions that failed had even larger negative shocks. For example, we estimate Long-Term Credit Bank (at the time, the ninth largest bank in the world) and Nippon Credit Bank had idiosyncratic bank shocks of -22 percent and -34 percent, respectively, in the years that they failed. Similarly, extreme falls in an institution’s market capitalization are also associated with negative bank shocks in the following year. For example, institutions whose market-to-book value change was in the bottom five percent of all institutions in that year, on average, had idiosyncratic bank shocks of -2.8%. Conversely, banks that received capital injections typically had *positive* bank shocks of 6.5%.

A second way of examining the plausibility of our estimates is to see what happened in the institutions with extreme bank shocks. Obviously, there are too many bank shocks to discuss each in detail, but we document what events preceded the ten largest contributors to the granular bank shock in the online appendix. A financial institution’s contribution to this channel is $w_{b,t-1}\tilde{\beta}_{bt}$ (see equation (12)), which weights each bank’s shock by its lagged share in lending. These ten shocks were preceded by some of the most memorable anecdotes of the period—results of FSA investigations into illegal activities, a computer programming error that paralyzed a bank’s ATM network, capital injections, and rogue traders—which suggests that we are correctly identifying the impacts of idiosyncratic factors that affected the loan supply of particular institutions.

4.2.2 Consistency with Prior Research

This sort of validation is based on a small sample and hard to use as a general test of validity. Fortunately, much of the existing empirical literature analyzing bank channels is explicitly or

implicitly based on the structure contained in equations (1) and (6), which lets us test whether our bank shock measures are significantly correlated with the proxy variables used in prior studies.²⁶ In other words, we can test whether our bank-shock estimates are in line with their hypotheses about what moves the bank-lending channel.

In order to see the link, we need to draw out the relationship between our framework and that in some previous studies. Since the econometrician does not observe β_{bt} and α_{ft} , prior studies have typically identified the bank channel by identifying a proxy variable, p_{bt} , which is assumed to be correlated with the bank channel, β_{bt} , but not with α_{ft} or ϵ_{fbt} in equation (1) in order to estimate $\Delta \ln L_{fbt} = \gamma p_{bt} + \alpha_{ft} + \epsilon_{fbt}$ or $D_{bt}^B = \gamma p_{bt} + \sum_f \phi_{fb,t-1} \alpha_{ft} + \sum_f \phi_{fb,t-1} \epsilon_{fbt}$. In both cases, the coefficient of interest (γ) depends crucially on the *cov*(p_{bt}, β_{bt}), *i.e.*, the extent to which the proxy variable (p_{bt}) actually covaries with the bank shock (β_{bt}).²⁷

Although allowing proxy variables to stand in for bank shocks does not allow for the identification of β_{bt} , existing empirical papers tell us how much particular proxy variables (e.g., capital adequacy ratios, market-to-book values, government credit supply, etc.) matter for lending. These studies have advantages when researchers can identify cases in which virtually all of the meaningful variation in lending behavior arises from some factor that only affects the bank. The main problem with this approach is that the use of proxy variables may not tell us about the many other types of bank shocks—arising from revelations of illegal activities, good or bad management, computer errors, etc.—that could also matter for lending. Nevertheless, if our identification strategy is correct, bank shocks in our data should be correlated with the proxy variables proposed in other studies. To put this more concretely, since we have estimates of β_{bt} and can observe the proxy variables, we can regress one on the other and test whether the assumed correlation is significant.

One of the most commonly used proxy variables is the risk-based capital ratio (*cf.* Peek and Rosengren (1997, 2000, 2005) and Amiti and Weinstein (2011)). These authors argue that, in the aftermath of the bursting of the Japanese land and stock bubbles, banks with low levels of capital adequacy were forced to cut back on lending. In other words, these studies are based on the idea that $\text{corr}(\beta_{bt}, RCR_{bt}) > 0$, where RCR_{bt} denotes the risk-based capital ratio of a bank. Similarly, Montgomery and Shimizutani (2009) have argued that the capital injections implemented between 1997 and 1999 should have caused bank lending to rise in recipient banks (because they had more capital) and therefore should be positively correlated with bank shocks. In other words, their study is based on the assumption that $\text{corr}(\beta_{bt}, \text{CapitalInjection}_{bt}) > 0$. Finally, several of these studies have also argued that deteriorations in a bank’s market-to-book value in a crisis, ΔMTB_{bt-1} , should lead to lower bank lending, and thus $\text{corr}(\beta_{bt}, \Delta MTB_{bt-1}) > 0$.

In Table 1, we examine each of these correlations using roughly the same time periods as the earlier papers. It is reassuring to see that our estimated bank shocks are correlated with proxy

²⁶For example, Chava and Purnanandam (2011) is explicitly based on equation (1).

²⁷Formally, we need the residual from projecting the proxy variable (p_{bt}) on the firm fixed effects to be correlated with β_{bt} .

variables in just the way that previous authors have suggested: risk-based capital ratios, capital injections, and changes in market-to-book values are all positively correlated with our bank-supply shock measures. Thus, we can formally reject the hypothesis that our estimates of movements in the bank lending channel are uncorrelated with actual movements.

However, it is important to also realize that, while prior work was limited to tracing out the impacts of shocks that can be tied to these proxy variables, our methodology enables us to examine the magnitude and impacts of *all* bank-supply shocks. Indeed, the low R^2 's in Table 1, suggest that the typical proxies used in the literature only capture a small part of the total variation in bank supply shocks. This is not surprising because the few events that enable the identification of γ are infrequent and only account for a small fraction of all idiosyncratic bank shocks.

Table 1: Validation of Bank-Supply Shocks

Dependent Variable: Bank Shock $_{b,t}$	(1)	(2)	(3)
Risk-Based Capital Ratio $_{b,t}$	0.015*** (0.004)		
Capital Injection $_{b,t}$		0.085*** (0.019)	
$\Delta \ln(\text{Market-to-Book Value}_{b,t-1})$			0.075** (0.032)
Year Fixed Effects	Yes	Yes	Yes
Observations	657	405	843
R^2	0.037	0.030	0.012

Notes: The risk-based capital ratio in Column 1 is the combined Tier 1 and Tier 2 risk-based capital conforming to the Basel II agreement. These data are from Peek and Rosengren (2005). In Column 2, the regressor is a dummy variable equal to 1 if bank b received a capital injection in year t as in Montgomery and Shimizutani (2009). In Column 3, the market-to-book value is computed as the average of the monthly share price multiplied by the number of shares outstanding and divided by the book value of its equity. We take the 12-month log difference of this variable. We drop the top and bottom one percentiles for bank-supply shocks and log differences of market-to-book value. These data were taken from Nikkei and from the Pacific Basin Capital Markets database. Huber-White heteroskedasticity robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

5 Main Results

Armed with our estimates of idiosyncratic firm and bank shocks, we now turn to understanding the importance of these bank shocks for lending and investment behavior. Prior work in this

area has not had a firm-specific, time-varying measure of credit constraints and therefore has focused on whether the cash-flow sensitivity of investment for some classes of firms differs from that of others. As we have argued earlier, a major advantage of our approach is that we have estimates of time-varying firm-borrowing and bank-supply changes. We will examine three paths into this problem. First, in Section 5.1 and Section 5.2 we show that idiosyncratic bank shocks have a statistically significant impact on firm-level investment. Second, in Section 5.3.1, we show that the estimates from these firm-level regressions imply that idiosyncratic bank shocks have substantial effects on aggregate investment for this sample of firms. Finally, in Section 5.3.2, we demonstrate the granular shocks we derived in Section 2.2 can explain large fluctuations in economy-wide investment rates.

5.1 Idiosyncratic Bank-Supply Shocks and Firm-Level Investment

Equation (8) provides a structure for measuring the impact of bank shocks on firm borrowing because the summation term on the right-hand side captures the impact of idiosyncratic bank supply shocks on the firm’s ability to borrow from banks.²⁸ The fact that not all firms borrow from all banks means that if loans from different banks are not good substitutes, a bank-supply shock from one bank will have a bigger impact on the investment of firms that borrow relatively more from it. To put it concretely, we use equation (8) to define the aggregate idiosyncratic bank-supply shock hitting a particular firm as

$$BankShock_{ft} = \sum_b \theta_{fbt-1} \tilde{\beta}_{bt}. \quad (16)$$

Our empirical exercise will then test whether this variable matters for firm-level investment.

Table 2 presents our findings. In the first column of Table 2, we regress the firm’s investment rate (investment divided by lagged assets) on the standard variables (cash flow and the market-to-book value) in which we also control for firm fixed effects and year fixed effects to sweep out any factors related to our sample of firms or correlations between firm-level investment and any common time-varying shock to the investment rate.²⁹ Not surprisingly, we find the standard result that there is a positive association between a firm’s investment rate and its cash flow and market-to-book value. In Column 2, we add in our industry shock variable to see if the variation in investment-to-capital ratios of a given firm is also associated with the typical change in firm-borrowing shocks within an industry. We find a positive and significant coefficient on the industry-borrowing shocks, which implies that many of the investment opportunities faced by individual firms are specific to firms in a particular industry. These might be exchange rate

²⁸The normalization in equation (8) is different from that in equation (11), which is why one expresses the idiosyncratic bank shock as $\hat{\beta}_{bt}$ and the other as $\tilde{\beta}_{bt}$, but as we show in Appendix 6, these two parameter estimates just differ by a constant.

²⁹We only include tangibles in our measure of investment and capital expenditures. Adding in intangibles, which account for, on average, three percent of total investment, leave the results unchanged.

Table 2: Firm-Level Investment Rate Regressions

Dependent Variable: Investment _{f,t} /Capital _{f,t-1}	(1)	(2)	(3)	(4)	(5)	(6)
Cash Flow _{f,t} /Capital _{f,t-1}	0.048*** (0.009)	0.047*** (0.009)	0.047*** (0.009)	0.047*** (0.009)	0.048*** (0.009)	0.048*** (0.009)
Market-to-Book Value _{f,t-1}	0.011*** (0.002)	0.011*** (0.002)	0.011*** (0.002)	0.011*** (0.002)	0.012*** (0.002)	0.012*** (0.002)
Bank Shock _{f,t}			-0.151*** (0.048)	-0.149*** (0.053)	-0.110** (0.047)	
(Bank Shock _{f,t}) x (Mean Loan-to-Asset Ratio _f)			0.732*** (0.202)	0.730*** (0.204)	0.809*** (0.199)	
(Bank Shock _{f,t}) x (Bin1 _{f,t})						0.105*** (0.037)
(Bank Shock _{f,t}) x (Bin2 _{f,t})						0.005 (0.048)
(Bank Shock _{f,t}) x (Bin3 _{f,t})						-0.088* (0.051)
(Bank Shock _{f,t}) x (Mean Bond-to-Asset Ratio _f)				-0.040 (0.463)		
Firm Shock _{f,t}					0.013** (0.006)	0.014** (0.006)
(Firm Shock _{f,t}) x (Mean Loan-to-Asset Ratio _f)					0.245*** (0.043)	0.242*** (0.043)
Industry Shock _{f,t}		0.070*** (0.017)	0.069*** (0.017)	0.069*** (0.017)	0.067*** (0.017)	0.042*** (0.005)
Fixed Effects						
Year	yes	yes	yes	yes	yes	yes
Firm	yes	yes	yes	yes	yes	yes
Observations	21,701	21,701	21,701	21,701	21,701	21,701
R ²	0.307	0.307	0.308	0.308	0.323	0.323

Notes: Standard errors (clustered by firm) are in parentheses. ***p<0.01, **p<0.05, *p<0.1. We drop the top and bottom one percentiles of each variable. The mean loan-to-asset ratio is defined for each firm as its average ratio of loans to assets over the sample period. The mean bond-to-asset ratio is similarly defined. Bin 1: mean loan-to-asset ratio > 0.14; Bin 2: mean loan-to-asset-ratio ∈ [0.08, 0.14]; Bin 3: mean loan-to-asset-ratio < 0.08.

movements that affect firms in export sectors differently from those in other sectors, or any of a host of possible variables that matter at the industry level.

The most interesting results, however, are those reported in Columns 3 to 6 of Table 2, where we add our idiosyncratic firm and bank shock variables to the specification. Since the firms in our sample are all listed nonfinancial firms that have access to both equity and bond markets, it makes sense to consider the impact of a firm-borrowing or bank-supply shock as having a different impact on a firm that borrows a lot from banks to finance its investment relative to a firm that finances investment through other means. In other words, a given bank shock is likely to have a much larger impact on the investment rate of a firm that finances, say, 80 percent of its capital through bank loans than on a firm that finances only 1 percent of its capital from loans. Therefore, in addition to including the firm and bank shocks in the specification, we also include their interaction with the firm's mean loan-to-asset ratio over the sample period.³⁰ Column 3 of Table 2 shows that the coefficient on the bank shock entering alone is negative and significant while that on the interaction term is positive and significant, indicating that loan-dependent firms' investments are affected by idiosyncratic bank shocks.³¹ In contrast with prior work on small firms, our results suggest that even listed companies, who by definition have access to non-bank financing such as equity, are affected by loan supply shocks. Interestingly, we find that negative idiosyncratic bank shocks exert a positive impact on the investment of firms that do not rely much on loans for their financing needs, which may reflect the fact that firms that don't rely much on loans for finance actually benefit relative to other firms when credit conditions tighten.³²

One concern about these results is that firms with access to forms of capital other than loans may be less susceptible to bank-supply shocks. Our results probably underestimate the effects of bank shocks because we are working with a sample of firms that all raise capital through equity markets. Within this sample, we have already noted that the effect of bank-supply shocks is most keenly felt by firms that, on average, finance a lot of their assets through loans. However, it might be the case that firms that also raise a lot of financing from bond markets are even less susceptible to bank-supply shocks. In Column 4, we interact the bank shocks with the firm's mean bond-to-asset ratio to see if firms that rely more on bond financing have less sensitivity to bank shocks. If this were the case, one would expect to see a negative coefficient on the bond interaction term. The fact that we do not find a significant coefficient on the bond interaction term suggests that the firm's loan share is the critical determinant of a bank shock on a firm's investment. While access to the bond market might serve to lower the importance of loans as a source of finance, conditional on a given loan share of financing, we obtain equally strong estimates of the impact

³⁰We cannot include the mean loan-to-asset ratio independently because we are already using firm fixed effects.

³¹The R^2 is not affected much by adding the bank shock variable, but, as we will see, this does not mean that bank shocks are not economically significant determinants of investment. The reason that the R^2 does not change much is that there is a large amount of noise in firm-level investment data. As we will see, this noise tends to cancel as we aggregate across firms.

³²This result is a feature of the model in Buera, Fattal-Jaef, and Shin (2013).

of bank-supply shocks on firms with and without high levels of bond finance.

Another possible concern is that bank shocks and firm shocks are correlated. For example, if firms that tend to have positive loan demand shocks also borrow from banks that tend to have positive supply shocks, then regressing firm outcomes on just the average bank shock will suffer from an omitted variable bias. For our sample of firms, adding firm-borrowing shocks in Column 5 hardly affects the coefficients on the bank shocks. The positive coefficient on the firm shock interacted with the mean loan-to-asset ratio implies a strong association between firm-borrowing shocks and investment for firms that are highly loan dependent, but not for firms that do not finance much of their capital expenditures through loans. The positive association almost surely arises because, in equilibrium, factors that raise a firm's marginal product of capital also raise its demand for borrowing. However, whether we control for this factor or not does not affect the strong relationship between bank shocks and firm-level investment.³³

The findings in the baseline specification in Column 5 of Table 2 imply a positive relationship between idiosyncratic bank shocks and investment for firms with loan-to-asset ratios above 0.14 (around 60 percent of firms in our sample). Meanwhile, for firms below the 0.14 threshold there is either no relationship or, if anything, a negative relationship between bank shocks and investment. Another way to see this is in Column 6, where we divide the mean loan-to-asset ratio into three bins, which we interact with bank shocks. The first (high) bin includes all firms with mean loan-to-asset ratios above 0.14, bin 2 (medium) with those between 0.08 and 0.14, and bin 3 (low) with those less than 0.08. The coefficient on the bank shocks interacted with the high dummy is positive and significant at 0.1; for those interacted with the medium dummy it is zero; and for the bank shocks interacted with the low dummy it is -0.09.

These results therefore tend to bridge the gap between studies using matched bank-firm data and extreme events that find a strong relationship between idiosyncratic bank shocks and real economic activity (e.g., Klein, Peek, and Rosengren (2002), Amiti and Weinstein (2011), Chodorow-Reich (2013)) and those that suggest firms can circumvent these shocks by accessing bond and other markets (e.g., Adrian, Colla, and Shin (2012)). Our results intuitively indicate that bank shocks matter for firms that are heavily dependent on loans, but less so for firms that are not dependent on loans. One implication of these findings is that loan-dependent firms cannot easily substitute towards other funding sources when the banks they rely on for financing cut back on their lending supply.³⁴

³³The results are robust to clustering at the firm level and to clustering by the largest bank in the firm's bank shock, as we show in Columns 1 and 2 in Table A2 in the online appendix.

³⁴One reason why firms may not substitute to other forms of financing is that that idiosyncratic bank shocks are not persistent. If we regress the bank shock on its one-period lag, with bank fixed effects, the coefficient is only -0.08 (t-stat of 3.5) with an adjusted R^2 equal to 0.03. Thus, unless firms can rapidly shift from banks that exhibit negative shocks, there is little value for them to do so since a negative shock in one year is not associated with a negative shock in future years.

5.2 Robustness

In this section, we conduct a number of robustness tests aimed at understanding whether our results are affected by lagging the firm shock, changing the set of independent variables, adjusting the sample period, and allowing for evergreening (troubled banks lending more to troubled firms). In addition, in an online appendix, we consider a number of additional robustness checks: changing the clustering of standard errors, lagging bank shocks, and dropping firms that receive government loans. None of these robustness checks qualitatively affect the results.

5.2.1 Lagged Firm Shocks and Sample Periods

Our results are robust to a number of alternative specifications. One concern might be that past firm-level shocks might be driving current bank shocks. In order to verify that this is not driving our results, in Column 1 of Table 3, we include lagged firm shocks and find that its coefficient is close to zero and its inclusion hardly affects the other coefficients. In Column 2, we include only the bank shock terms and the industry effects to see if the exclusion of all potentially endogenous variables affects the results. We find that the coefficients are almost unchanged.

Similarly, one might wonder if our results are driven only by crisis years, defined as FY1991, FY1993, FY1998, and FY2009. In Column 3, we interact the bank shocks with a crisis dummy, and find that there is no significant difference between crisis and non-crisis years.³⁵ Finally, in Columns 4 and 5, we divide the sample into two halves (FY1991–FY2000 and FY2001–FY2010) to see if the coefficients differ between the early years, when bank failures were frequent, and the later years. We find that bank shocks are significant determinants of investment in both periods and that the point estimates do not change significantly in the subsamples. Because we obtain our results for two non-overlapping subsamples, we can be confident that no single event or year is driving our results.³⁶

5.2.2 Differential Shocks to Healthy and Unhealthy Firms

It is reasonable to explore if there are ways to adapt the Proposition (4) to allow for specifications in which we explicitly allow for non-homotheticities. A number of papers (e.g. Peek and Rosengren (2005), and Jimenez et al. (2011)) have found that weak banks not only lend less but lend relatively less to healthy firms in times of trouble, a practice known as “evergreening”. Although this is not the main focus of our paper, we can also check if our results are robust to allowing banks to pass

³⁵The first years are the same as those dropped in Amiti and Weinstein (2011) because they correspond to the year of the initial bursting of the Japanese stock market bubble, the year when the *jusen* (housing finance companies) began failing, and the year when the first major banks began failing. The last crisis year corresponds to the first year of the more recent financial crisis.

³⁶Table A2 in the online appendix provides additional robustness checks of our baseline results, where we also add in lagged bank shocks in Column 3, drop firms with government loans in Columns 4 and 5, and use loan shocks constructed with weighted least squares in columns 6. Further, in Table A5, we show that our idiosyncratic bank shocks are important determinants of firm-level employment and sales.

Table 3: Firm-Level Investment Robustness

Dependent Variable:	(1)	(2)	(3)	(4)	(5)	(6)
Investment _{<i>f,t</i>} /Capital _{<i>f,t-1</i>}	lagged firm shock	only bank shock	with crisis interaction	1991-2000	2001-2010	ever-greening
Cash Flow _{<i>f,t</i>} /Capital _{<i>f,t-1</i>}	0.045*** (0.011)		0.048*** (0.009)	0.164*** (0.019)	0.042*** (0.010)	0.047*** (0.008)
Market-to-Book Value _{<i>f,t-1</i>}	0.013*** (0.002)		0.012*** (0.002)	0.009*** (0.003)	0.014*** (0.003)	0.012*** (0.002)
Bank Shock _{<i>f,t</i>}	-0.112** (0.048)	-0.148*** (0.047)	-0.105* (0.054)	-0.086 (0.081)	-0.058 (0.070)	-0.101** (0.043)
(Bank Shock _{<i>f,t</i>}) x (Mean Loan-to-Asset Ratio _{<i>f</i>})	0.760*** (0.213)	0.692*** (0.202)	0.898*** (0.218)	0.886*** (0.328)	0.624** (0.298)	0.633*** (0.178)
Crisis x (Bank Shock _{<i>f,t</i>})			0.000 (0.108)			
Crisis x (Bank Shock _{<i>f,t</i>}) x (Mean Loan-to-Asset Ratio _{<i>f</i>})			-0.418 (0.486)			
Firm Shock _{<i>f,t</i>}	0.015** (0.007)		0.013** (0.006)	-0.002 (0.010)	0.017** (0.008)	0.014** (0.006)
Firm Shock _{<i>f,t-1</i>}	0.008** (0.003)					
(Firm Shock _{<i>f,t</i>}) x (Mean Loan-to-Asset Ratio _{<i>f</i>})	0.246*** (0.046)		0.245*** (0.042)	0.388*** (0.084)	0.227*** (0.054)	0.240*** (0.042)
Industry Shock _{<i>f,t</i>}	0.075*** (0.024)	0.082*** (0.017)	0.067*** (0.017)	0.097*** (0.030)	0.050** (0.024)	0.063*** (0.017)
Fixed Effects						
Year	yes	yes	yes	yes	yes	yes
Firm	yes	yes	yes	yes	yes	yes
Observations	18,656	21,701	21,701	9,595	12,106	21,681
R ²	0.307	0.291	0.323	0.391	0.378	0.323

Notes: Standard errors (clustered by firm) are in parentheses. ***p<0.01, **p<0.05, *p<0.1. Column (3) interacts the bank shocks with crisis years (1991, 1993, 1998, 2009). Column (4) uses a restricted sample from 1991-2000, and Column (5) uses a sample restricted to 2001-2010. We drop the top and bottom one percentiles of all the variables. Column (6) uses an alternative formulation of the bank shock in which we allow the same bank to supply credit differently to healthy and unhealthy firms. Distressed firms are defined as in Hoshi, Kashyap, and Scharfstein (1990) as firms with net interest payments above operating income in years t and $t-1$, but not in year $t-2$. The results are similar if we define distressed as just having operating income below net interest payments in the previous year.

along their shocks more to healthy firms than to unhealthy ones. A simple way to nest the insights of these papers into our framework is to postulate that a bank’s supply shock differs depending on whether the client is healthy or not. We follow Hoshi, Kashyap, and Scharfstein (1990) in defining unhealthy firms as previously healthy firms whose interest payments exceed their operating income for two consecutive years.

If we split our sample according to whether the firms are healthy or not, we can obtain different estimates of $\tilde{\beta}_{bt}$ for healthy firms and unhealthy ones. In order to keep the notation simple, let $\tilde{\beta}_{bt}^H$ be the bank shock we obtain using loan data for healthy firms and $\tilde{\beta}_{bt}^U$ to be the bank shock we obtain when we only include the unhealthy firms. Define L_{bt}^H and L_{bt}^U to be bank b ’s total lending to healthy and unhealthy firms, respectively. We now can define the overall bank shock to be

$$\bar{\beta}_{bt} \equiv \frac{L_{b,t-1}^H \tilde{\beta}_{bt}^H + L_{b,t-1}^U \tilde{\beta}_{bt}^U}{L_{b,t-1}^H + L_{b,t-1}^U}. \quad (17)$$

If we believe that evergreening causes unhealthy banks to lend more to unhealthy firms, then we should expect that $\tilde{\beta}_{bt}^U > \tilde{\beta}_{bt}^H$ when $\bar{\beta}_{bt} < 0$. This is, in fact, exactly what we find in the data. The mean value of $\tilde{\beta}_{bt}^U - \tilde{\beta}_{bt}^H$ for relatively unhealthy banks is 0.06 (s.e. 0.11), indicating that relatively troubled banks lent six percent more to unhealthy firms than healthy firms after controlling for their loan demand.³⁷ Interestingly, we also find that relatively healthy banks, i.e. those institutions with bank shocks above the median lent 5.1 percent *less* to unhealthy firms (s.d. 1.2 percent), confirming the findings of Peek and Rosengren (2005) that the provision of more credit to unhealthy firms is something that unhealthy banks do, but healthy banks actually readjust their lending portfolios away from their worst clients.

Of course, our main concern is not whether we can use our methodology to replicate the results of other papers, but whether allowing banks to have different lending policies for healthy and unhealthy firms affects our main result. In order to do this, we redefined our bank shock variable to be $\sum_b \theta_{f_{bt-1}} \tilde{\beta}_{bt}^h$, where $h = U$ if the firm is unhealthy and $h = H$ if the firm is healthy. This enables us to allow a bank’s shock to vary depending on both the health of the bank and the health of the borrower. We report the result of using this regression in the last column of Table 3. The coefficient on the bank shock terms are not significantly different than those in our main specification, which indicates that our assumption that all of a bank’s borrowers receive a common bank shock is not driving our results.

5.3 How Much do Idiosyncratic Bank Shocks Matter for Investment?

Our firm-level investment regressions indicate that idiosyncratic bank shocks matter statistically for investment, but they do not tell us how much they matter *economically*. The simplest way to assess the economic magnitude is to compute a beta coefficient for the impact. For example, let’s

³⁷Because our bank shocks were estimated off a smaller sample (especially for the unhealthy firms), we had more outliers in our estimates. Therefore, we trimmed the top and bottom 5% of values of $\tilde{\beta}_{bt}^U$ and $\tilde{\beta}_{bt}^H$ before computing the mean.

consider a firm with a loan-to-asset ratio at the 75th percentile (i.e., a loan-to-asset ratio of 0.28). Since the standard deviation of the bank shock is 0.05, this implies that this firm would result in a shift in a firm’s investment rate of $[(-.11 + 0.809*0.28)*0.05 =]$ 0.006. Given that the median investment rate is 0.08, this is equivalent to a 7 percent decline in the the firm’s investment rate. If banks were severely stressed, the impact would be much larger. This type of calculation is a reasonable “proof of concept”—bank shocks can matter for investment—but it is very hard to assess what this type of calculation tells us about the importance of these shocks for aggregate fluctuations.

Similarly, we see from Table 2 that although the idiosyncratic bank shock terms are very significant statistically, the R^2 in the regression does not change much with their inclusion. One possible interpretation of this result is that idiosyncratic bank shocks do not matter much economically. Another interpretation is that the error terms in the firm-level investment equation cancel as we aggregate, rendering idiosyncratic bank shocks an important determinant of aggregate investment. We assess these two hypotheses in the next two sections.

5.3.1 Micro Estimation

Section 2.3 shows that if we weight the investment rate regression by the firm’s lagged assets we obtain estimates that aggregate to match the average aggregate investment rate over the whole time period but will not necessarily match the investment rates in each year (I_t/K_{t-1}). Thus by choosing whether to include the bank shock variables in the construction of the predicted investment rate ($\widehat{I_t/K_t}$) and seeing how this choice affects the R^2 of a regression of the investment rate on its predicted value, we can assess the importance of these bank shocks on aggregate investment fluctuations.

Table 4 presents the results from estimating the weighted firm-level investment rate regression.³⁸ The parameter estimates are qualitatively the same in the weighted regression as in the unweighted regression. Idiosyncratic bank shocks clearly matter for firms that borrow heavily, thus, weighting the data does not seem to qualitatively affect our results. Since we have weighted the regression appropriately, we know that the predicted values of the investment rate averaged over all years must match the actual average aggregate investment rate, but this need not be the case for any individual year. Therefore, if we regress I_t/K_{t-1} on its predicted value ($\widehat{I_t/K_{t-1}}$), we can assess how much of the actual variation in aggregate investment rates can be explained by different firm-level models of investment rates. As one can see by comparing the R^2 in the lower panel of the table, the inclusion of the bank shock variables in the predicted values raises the R^2 by 0.30 in the specification without firm fixed effects (the difference in R^2 between Columns i and ii)

³⁸We had to drop the year fixed effects in order to implement this procedure because if we included them the predicted values of the aggregate investment rate would match the actual values in each year. As one can see from online appendix Tables A3 and A4, dropping the year fixed effects in the unweighted specification has little impact on the remaining point estimates, so our results are not dependent on this minor change in specification. Instead of including year effects, we include the common shock to absorb economy-wide movements in lending.

Table 4: Weighted Firm-Level Investment Rate Regressions

Dependent Variable: Investment _{f,t} /Capital _{f,t-1}	(1)	(2)
Cash Flow _{f,t} /Capital _{f,t-1}	0.177*** (0.028)	0.153*** (0.022)
Market-to-Book Value _{f,t-1}	0.012*** (0.003)	0.019*** (0.003)
Bank Shock _{f,t}	0.099 (0.076)	0.084 (0.076)
(Bank Shock _{f,t}) x (Mean Loan-to-Asset Ratio _f)	1.310*** (0.290)	1.432*** (0.264)
Firm Shock _{f,t}	0.019* (0.011)	0.017* (0.010)
(Firm Shock _{f,t}) x (Mean Loan-to-Asset Ratio _f)	0.162** (0.068)	0.155** (0.067)
Industry Shock _{f,t}	0.080*** (0.025)	0.073*** (0.019)
Common Shock _t	0.507*** (0.051)	0.443*** (0.048)
(Mean Loan-to-Asset Ratio _f)	-0.105*** (0.033)	
Firm Fixed Effects	no	yes
Observations	9	21,619
R ²	0.227	0.395

Dependent Variable: Investment _t /Capital _{t-1}	(i)	(ii)	(iii)	(iv)
	with	without	with	without
	bank shocks	bank shocks	bank shocks	bank shocks
(I _t /K̄ _{t-1}) (weighted fitted values)	1.378*** (0.170)	0.891*** (0.237)	1.272*** (0.117)	0.937*** (0.194)
Observations	20	20	20	20
R ²	0.786	0.439	0.867	0.563

Note: Estimation based on the procedure described in Section 2.3.

and 0.35 in the specification with firm fixed effects (the difference in R^2 between Columns iii and iv). In other words, the firm-level regressions imply that between 30-35 percent of the variation in aggregate investment behavior of listed firms can be explained by idiosyncratic bank shocks. These results provide strong evidence that these shocks matter substantially at the aggregate level.

5.3.2 Macro Estimation

One of the limitations with the results presented so far is that they are based on a sample of listed firms. The loan data in our sample of firms account for only around 17 percent of total lending in Japan in an average year. One might reasonably wonder if the bank shocks we have identified off this sample of firms can help us understand economy-wide lending and investment. Since our methodology allows us to identify a pure idiosyncratic bank-supply shock that controls for borrower demand, the measure of granular bank shocks should be the same for the whole economy.

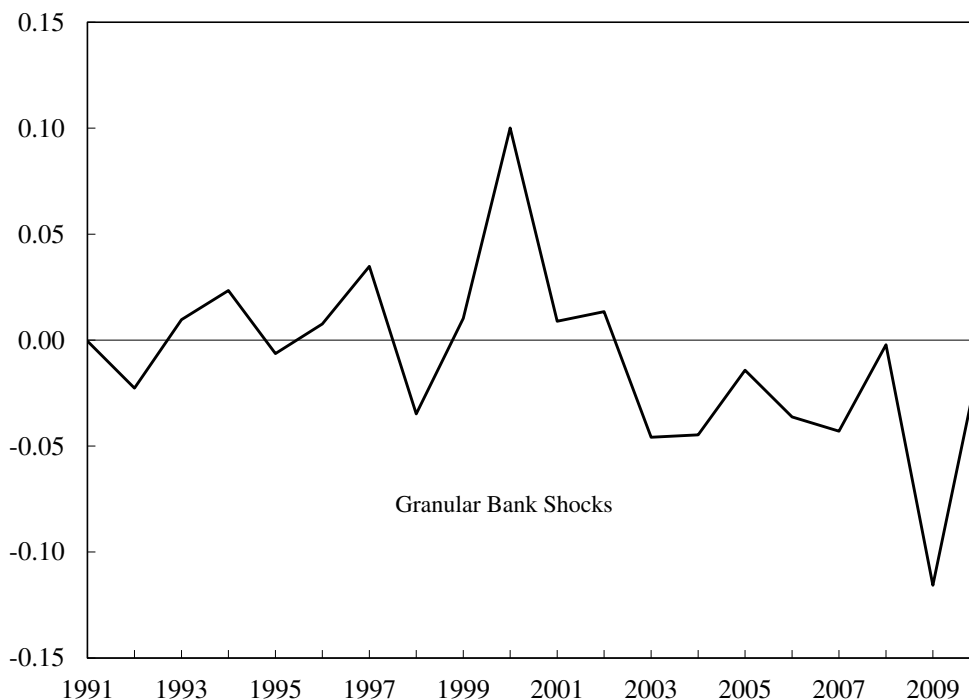
The starting point for thinking about the aggregate impact of Japanese lending is to realize that if we were to regress the left-hand side of equation (12) on each of the right-hand side terms, we would obtain coefficients of one on each term since the equation is an identity. However, this need not be the case if we replace the dependent variable with economy-wide aggregate loan growth from the flow-of-funds data. In this case, we can write the relationship between the percentage change in aggregate lending and the percentage change in lending to listed companies as $D_t^{FOF} = \mathbf{W}_{B,t-1} \mathbf{D}_{Bt} + \epsilon_t$, or

$$D_t^{FOF} = \delta + \gamma_1 (\bar{A}_t + \bar{B}_t) + \gamma_2 \mathbf{W}_{B,t-1} \Phi_{t-1} \mathbf{N}_t + \gamma_3 \mathbf{W}_{B,t-1} \Phi_{t-1} \tilde{\mathbf{A}}_t + \gamma_4 \mathbf{W}_{B,t-1} \tilde{\mathbf{B}}_t + \epsilon_t. \quad (18)$$

If we estimate equation (18), we will only obtain coefficients of one on each term if each of the terms we measured using the Nikkei data equals the terms for the whole economy and the error term is uncorrelated with any of the explanatory variables. The coefficient of interest is γ_4 , which tells us whether aggregate lending is associated with granular bank shocks. Measurement error or economy-wide variables that move more or less than the corresponding variables in the Nikkei data could cause the coefficients to differ from one. Moreover, to the extent that one thought that economy-wide lending behavior and loans to listed companies were uncorrelated, one might expect to see a low R^2 from estimating equation (18). Finally, even if we were to obtain coefficients equal to one using flow-of-funds data, it could be the case that granular bank shocks are sufficiently small that we obtain much smaller regression coefficients when executing a regression on standardized variables because a one-standard-deviation movement in granular bank supply might have a very small impact on aggregate lending. In other words, we can interpret the coefficients on each of these terms as the contribution of each variable to the flow-of-funds lending growth, while the error term corresponds to deviations in aggregate loan growth from loan growth for listed companies. Our interest is in trying to understand how important the terms we identified in equation (12) are

in determining aggregate loan fluctuations.

Figure 8: Granular Bank Shocks



Notes: Years refer to fiscal years, which roughly correspond to the calendar year plus one.

Figure 8 plots the granular bank shocks over our sample period. It is important to remember that these shocks are not the shocks to the entire banking sector (which are captured in the common shock), but instead, they represent the differential shocks to large institutions, resulting from the accumulation of idiosyncratic shocks hitting large lenders. Nevertheless, the largest movements in the series have a straightforward interpretation. In FY 1997, we see a big positive shock, which follows Prime Minister Hashimoto’s 1996 “Big Bang” financial reforms, paving the way for bank holding companies. We see a large negative granular bank shock in FY1998, which corresponds to the year in which the failure of Hokkaido-Takushoku Bank sent a shock wave through the Japanese banking system due to the willingness of the government to allow a major bank to fail. In 1998 (FY1999), large banks were given capital injections. If these policy interventions induced more lending, we would expect to see that these large institutions lent more. This is precisely what we see in the figure where we observe a big positive granular bank shock in 1999 (FY2000). The second largest granular bank shocks occur in FY2003 and FY2004 in the aftermath of the Takenaka Plan’s forced disclosure of banks’ nonperforming loans in 2002. The biggest negative granular bank shock is also easy to understand. It occurs in 2008 (FY2009): the same fiscal year as the Lehman bankruptcy. It is reasonable to suspect that the global financial meltdown in 2008 had a worse impact on large Japanese banks than on small ones who probably were much less exposed internationally. This may explain why there is a large negative granular bank shock in FY2009.

Table 5: Aggregate Effects

Dependent Variable:	Percentage Change in Flow of Funds _t			Investment _t /Capital _{t-1}		
	(1)	(2)	(3)	(4)	(5)	(6)
Common Shock _t	0.492* (0.261)	1.212*** (0.182)	1.196*** (0.180)	0.206** (0.094)	0.475*** (0.113)	1.274*** (0.303)
Industry Shock _t	0.513 (0.329)	1.391*** (0.275)	0.595*** (0.118)	-0.117 (0.183)	0.211 (0.192)	0.246 (0.223)
Firm Shock _t	0.215 (0.145)	0.318** (0.144)	0.257** (0.117)	-0.037 (0.079)	0.001 (0.060)	0.003 (0.132)
Bank Shock _t		1.170*** (0.232)	1.042*** (0.207)		0.437*** (0.124)	1.057*** (0.300)
Constant	-0.018* (0.009)	-0.003 (0.006)	0.000 (0.118)	0.076*** (0.004)	0.081*** (0.004)	0.000 (0.144)
Standardized Variables	No	No	Yes	No	No	Yes
Observations	20	20	20	20	20	20
R ²	0.420	0.782	0.782	0.303	0.675	0.675

Notes: Huber-White heteroskedasticity robust standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1. All variables, including percentage change in the FOF and aggregate investment, are standardized in (3) and (6) with a mean of 0 and a standard deviation of 1. The sample is for the 1991-2010 period. The Breusch-Godfrey LM test for autocorrelation could not reject the null of no autocorrelation.

In Table 5, we regress aggregate lending growth from the flow-of-funds data on the various components of lending growth that we identified from the Nikkei matched bank-firm data. As one can see in the first two columns of the table, aggregate lending can only be well explained if we include the (idiosyncratic) granular bank shocks.³⁹ Interestingly, while common, industry, and granular bank shocks have coefficients close to one in Column 2, the granular firm shock is attenuated, perhaps because while we have all of the large banks in our sample, we do not have unlisted firms, and hence there is more measurement error in the granular firm shock variable. Moreover, as we show in an online appendix, had we used the bank shocks using the Proposition 2 WLS methodology the explanatory power of the granular bank shocks in explaining bank lending would have been significantly reduced, presumably because these estimates do not take into account new lending activity.

Movements in aggregate lending are closely correlated with the forces we identified in the

³⁹Alternatively, we could merge together the effects of the industry and firm shocks. This does not qualitatively affect the coefficients on the bank shock variable.

Nikkei sample. When we include all regressors in the sample, the R^2 is 0.8, suggesting that most of what drives aggregate commercial lending is explainable by the factors we identified in the Nikkei sample. Moreover, the coefficients on the common shock, industry shock, and granular bank shock in Column 2 are statistically indistinguishable from one, which is the value that would obtain if these variables had been computed using data from all Japanese lending pairs. In other words, the data do not reject the hypothesis that the common, industry, and granular bank shocks identified in the Nikkei data are the same as those we would find in economy-wide data.

However, our main interest is in understanding the importance of the granular bank shock term. One metric of its importance is to consider how much of the variance in aggregate lending it explains. Comparing columns 1 and 2, we see that adding the granular bank shock term raises the R^2 of the regression by 0.36. The fact that the granular bank shock term is highly significant and moves the R^2 of the regression substantially strongly suggests that the idiosyncratic lending decisions of large financial institutions have a large impact on aggregate lending. In Column 3 of Table 5, we run a standardized regression in order to examine the relative importance of movements in the granular bank shock for understanding aggregate movements. Not surprisingly, we get a coefficient indistinguishable from one on the common shock term, indicating that a one-standard-deviation increase in lending observed in the Nikkei data is associated with a one-standard-deviation-increase in lending in Japan. What is most surprising is the importance of the granular bank shock term: a one-standard-deviation movement in this term is also associated with a one-standard-deviation movement in *aggregate* lending. This indicates that aggregate loan supply is also highly sensitive to bank shocks. In an economy composed of small financial institutions, this term would likely be zero, but the fact that it is so large implies that much of aggregate lending is explainable by the loan supply movements of large institutions.⁴⁰

We have seen that the variation in the granular bank shock is an important component of aggregate lending volatility. We now turn to estimating the impact of bank shocks on aggregate investment activity in columns 4 through 6. Our results from the previous sections have already established exogeneity of the bank shocks, a strong relationship between firm-level investment and bank shocks, an economically meaningful link between aggregate investment by listed firms and their bank shocks, and a tight link between aggregate lending and granular loan shocks. Moreover,

⁴⁰In Table A7 of the online appendix, we reproduce Table 5 using bank shock estimates based on the Proposition 2 WLS procedure instead of our proposed Proposition 4 methodology. The principle difference in the two procedures is that the Proposition 4 methodology produces estimates that aggregate to match total lending growth. We can see the problem with the WLS estimates when we try to replicate Table 5 using WLS estimates instead of the Proposition 4 estimates. If we include all the variables except the bank shocks both methods account for approximately the same amount of aggregate variance (the R^2 is 0.420 using the Proposition 4 methodology and 0.455 using the Proposition 2 WLS methodology). However, granular bank shocks have much more explanatory power when they are computed using the Proposition 4 methodology relative to the Proposition 2 WLS methodology. The latter approach only raises the R^2 by 0.2 whereas our preferred methodology raises it by 0.36. In other words, using bank shocks based on total lending by a bank (i.e., Proposition 4) rather than lending to existing clients (Proposition 2) enables us to explain over 75 percent more variation in aggregate lending. This establishes the importance of accounting for new lending when estimating the impact of granular shocks on aggregate lending.

we also know from Table 4 that idiosyncratic bank shocks account for 30-35 percent of variation in total investment among listed firms. Our next question is whether these granular bank shocks cause movements in aggregate investment in the Japanese economy as a whole.

In Table 5 we explore the relationship by regressing the Japanese economy-wide aggregate investment rate on the granular bank shocks. By construction, any common shock to a firm's investment demand will be captured in our common shock term; therefore, the granular bank shocks cannot be caused by these common shocks. The granular bank shock term in Column 5 is strongly significant and raises the R^2 of the regression from 0.303 to 0.675, which indicates that these bank shocks also account for 37 percent of the fluctuations in aggregate investment. In other words, granular shocks—idiosyncratic movements in lending to a given borrower by large banks relative to small banks—are closely associated with aggregate investment movements. Our macro estimate is surprisingly close to the 30-35 percent improvement in fit we obtained by aggregating up the bank shocks in the firm-level investment rate regressions. Despite the differences in samples and methods, we obtain very similar estimates of the impact of idiosyncratic bank shocks on aggregate investment in Tables 4 and 5. In terms of economic significance, we see in Column 6 of Table 5, where we run a standardized regression, that a one-standard-deviation movement in the granular bank shock moves the aggregate investment-to-capital ratio by one standard deviation—about as much as common shocks.⁴¹

Since we control for the firm-borrowing shocks when estimating the bank shocks, the granular bank shocks are macro variables that are not determined by fluctuations in borrower demand. The fact that they matter so much for investment suggests that the strong, causal link between loan supply and firm-level investment that we identified in Tables 2, 3, and 4 matters at the aggregate level as well. In sum, the results from the micro and macro estimation both indicate that substantial amounts of aggregate investment can be explained by idiosyncratic bank shocks.

6 Conclusion

In this paper, we develop a new methodology that enables us to estimate bank-loan and firm-borrowing shocks using a fixed-effects approach that enables researchers to directly apply estimates obtained from matched bank-firm data and firm-level investment regressions to understanding aggregate lending and investment behavior. We relate our new methodology to existing methods. We show that log specifications produce estimates that match geometric means of disaggregated lending behavior, but they do poorly in matching aggregate lending. Our method is identical to WLS fixed effects methodology when the dependent variable is defined as a percentage change and there are no new lending relationships. Importantly, our approach enables researchers to estimate bank shocks even when new loan relationships form.

⁴¹The results are qualitatively similar if we redefine the granular bank shock relative to the simple mean (see Table A4 in the online appendix).

We estimate time-varying bank shocks, firm-borrowing shocks, industry shocks, and common shocks using matched firm-bank loan data for Japan between 1990 and 2010. This decomposition enables us to establish a causal relationship between idiosyncratic bank-supply shocks and firm-level investment for those firms that are loan dependent. We find that idiosyncratic bank supply shocks are significant determinants of firm-level investment. This result is particularly surprising because our sample is comprised of listed companies that have, by definition, access to equity markets. Most importantly, we show that idiosyncratic bank-supply shocks have large effects on firm-level and economy-wide investment.

The fact that banks are large even relative to the largest economies is at the heart of the underlying mechanism for these results. The bank data in most industrial countries show that the distribution of banks is highly skewed, with a small number of large banks accounting for a large share of aggregate lending. For example, in Japan the three largest banks account for a 45 percent share of total banking assets.⁴² Because there is “granularity” in the banking data, the law of large numbers does not apply, rendering firms vulnerable to the fates of individual banks. Indeed, we show that the bank shocks account for 30 to 40 percent of the variation in Japanese aggregate lending and investment.

There are a number of implications from this study. First, our results imply that it is difficult for firms to substitute between loan sources in response to short-term changes in credit. Although the median number of bank loans per firm is seven in our sample, idiosyncratic shocks to large banks still have significant effects on these firms’ investment. Second, the fates of large financial institutions are an important determinant of investment and real economic activity. Although our study has relied on Japanese data, the key elements of the banking system that underlie these results are also prevalent in other industrialized countries, thus making it likely they would also be present in other countries. Analyzing the importance of idiosyncratic bank shocks in other countries is potentially a fertile area for future research.

References

- Adrian, Tobias, Paolo Colla, and Hyun Song Shin.** 2012. “Which Financial Frictions? Parsing the Evidence from the Financial Crisis of 2007 to 2009.” In *NBER Macroeconomics Annual 2012, Volume 27. NBER Chapters*, 159–214. National Bureau of Economic Research, Inc.
- Amel, D., C. Barnes, F. Panetta, and C. Salleo.** 2004. “Consolidation and efficiency in the financial sector: A review of the international evidence.” *Journal of Banking & Finance*, 28(10): 2493–2519.
- Amiti, Mary, and David E. Weinstein.** 2011. “Exports and Financial Shocks.” *The Quarterly Journal of Economics*, 126(4): 1841–1877.

⁴²<http://research.stlouisfed.org/fred2/series/DDOI01JPA156NWDB>

- Ashcraft, Adam B.** 2005. “Are Banks Really Special? New Evidence from the FDIC-Induced Failure of Healthy Banks.” *The American Economic Review*, 95(5): 1712–1730.
- Ashcraft, Adam B.** 2006. “New Evidence on the Lending Channel.” *Journal of Money, Credit, and Banking*, 38(3): 751–775.
- Bremus, Franziska, Claudia Buch, Katheryn Russ, and Monika Schnitzer.** 2013. “Big Banks and Macroeconomic Outcomes: Theory and Cross-Country Evidence of Granularity.” NBER Working Papers 19093.
- Buch, Claudia M., and Katja Neugebauer.** 2011. “Bank-specific shocks and the real economy.” *Journal of Banking & Finance*, 35(8): 2179–2187.
- Buera, Francisco, Roberto Fattal-Jaef, and Yongseok Shin.** 2013. “Anatomy of a Credit Crunch: from Capital to Labor Markets.” Mimeo, IMF.
- Chava, S., and A. Purnanandam.** 2011. “The effect of banking crisis on bank-dependent borrowers.” *Journal of Financial Economics*, 99(1): 116–135.
- Chodorow-Reich, Gabriel.** 2013. “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.
- Gabaix, Xavier.** 2011. “The Granular Origins of Aggregate Fluctuations.” *Econometrica*, 79(3): 733–772.
- Greenstone, Michael, and Alexandre Mas.** 2012. “Do Credit Market Shocks Affect the Real Economy? Quasi-Experimental Evidence from the Great Recession and ‘Normal’ Economic Times.” *MIT Department of Economics Working Paper No. 12-27*.
- Hoshi, Takeo, Anil Kashyap, and David Scharfstein.** 1990. “The role of banks in reducing the costs of financial distress in Japan.” *Journal of Financial Economics*, 27(1): 67–88.
- Hubbard, R.G., K.N. Kuttner, and D.N. Palia.** 2002. “Are There Bank Effects in Borrowers’ Costs of Funds? Evidence from a Matched Sample of Borrowers and Banks.” *The Journal of Business*, 75(4): 559–581.
- Jimenez, G., S. Ongena, J.L. Peydro, and J. Saurina.** 2011. “Hazardous Times for Monetary Policy: What do 23 Million Loans Say about the Impact of Monetary Policy on Credit Risk-taking.” Mimeo, European Central Bank.
- Kashyap, A.K., and J.C. Stein.** 2000. “What Do a Million Observations on Banks Say about the Transmission of Monetary Policy?” *American Economic Review*, 90(3): 407–428.
- Kashyap, Anil K., Jeremy C. Stein, and David W. Wilcox.** 1993. “Monetary Policy and Credit Conditions: Evidence from the Composition of External Finance.” *American Economic Review*, 83(1): 78–98.
- Khwaja, Asim Ijaz, and Atif Mian.** 2008. “Tracing the Impact of Bank Liquidity Shocks: Evidence from an Emerging Market.” *American Economic Review*, 98(4): 1413–1442.
- Klein, Michael, Joe Peek, and Eric Rosengren.** 2002. “Troubled Banks, Impaired Foreign Direct Investment: The Role of Relative Access to Credit.” *American Economic Review*, 92(3): 664–682.
- Kroszner, Randall S., Luc Laeven, and Daniela Klingebiel.** 2007. “Banking Crises, financial dependence, and growth.” *Journal of Financial Economics*, 84: 187–228.
- Montgomery, Heather, and Satoshi Shimizutani.** 2009. “The effectiveness of bank recapitalization policies in Japan.” *Japan and the World Economy*, 21(1): 1–25.
- Paravisini, Daniel.** 2008. “Local Bank Financial Constraints and Firm Access to External Finance.” *The Journal of Finance*, 63(5): 2161–2193.

- Paravisini, Daniel, Veronica Rappoport, and Philipp Schnabl.** 2015. "Specialization in Bank Lending: Evidence from Exporting Firms." NBER Working Papers 21800.
- Peek, Joe, and Eric S. Rosengren.** 1997. "The International Transmission of Financial Shocks: The Case of Japan." *The American Economic Review*, 87(4): 495–505.
- Peek, Joe, and Eric S. Rosengren.** 2000. "Collateral Damage: Effects of the Japanese Bank Crisis on Real Activity in the United States." *The American Economic Review*, 90(1): 30–45.
- Peek, Joe, and Eric S. Rosengren.** 2005. "Unnatural Selection: Perverse Incentives and the Misallocation of Credit in Japan." *The American Economic Review*, 95(4): 1144–1166.
- Sakuragawa, M., and Y. Watanabe.** 2009. "Did the Japanese Stock Market appropriately Price the Takenaka financial reform?" 317–340, University of Chicago Press.
- Santos, Joao C.** 2012. "Bank Corporate Loan Pricing Following the Subprime Crisis." *Review of Financial Studies*. forthcoming.

Appendix A: Proof of Proposition 1

The proof is easiest if we rewrite equation (2) in vector form so that we can be explicit about how the dummy variables enter into the equation:

$$\mathbf{D}_t^+ = \mathbf{F}_t \boldsymbol{\alpha}_t^I + \mathbf{B}_t \boldsymbol{\beta}_t^I + \mathbf{Z}_t \gamma_t + \boldsymbol{\eta}_t, \quad (\text{A1})$$

where \mathbf{D}_t^+ is a vector whose elements equal $D(L_{fbt}/L_{fb,t-1})$, i.e., the set of loan growth rates computed for all loans satisfying $L_{fb,t-1} > 0$ if loan growth is measured in percentage changes and $L_{fbt} > 0, L_{fb,t-1} > 0$ if they are measured in log changes; and $\mathbf{F}_t = (\mathbf{F}_{1t}, \dots, \mathbf{F}_{Ft})$ and $\mathbf{B}_t = (\mathbf{B}_{1t}, \dots, \mathbf{B}_{Bt})$ are matrices whose columns are given by the fixed effects, \mathbf{F}_{ft} and \mathbf{B}_{bt} , the elements of these columns equal one whenever firm $i = f$ or bank $j = b$ and equal zero otherwise; \mathbf{Z}_t is a vector of containing a firm-bank interaction variable, γ_t is the coefficient on the interaction, and $\boldsymbol{\eta}_t$ is a mean zero error.

Estimation requires us to make a few normalizations. First, since $\sum_f \mathbf{F}_{ft} = \sum_b \mathbf{B}_{bt} = \mathbf{1}$, we can drop one firm and one bank fixed effects without affecting the predicted values. Moreover, we can trivially cause the estimated coefficients to vary by changing which bank and firm we drop, so we want to be consistent about our choice of *numéraire* across specifications. We therefore rewrite this expression as

$$\mathbf{D}_t^+ = \mathbf{1} \ddot{c}_t + \ddot{\mathbf{F}}_t \ddot{\boldsymbol{\alpha}}_t + \ddot{\mathbf{B}}_t \ddot{\boldsymbol{\beta}}_t + \mathbf{Z}_t \gamma_t + \boldsymbol{\eta}_t, \quad (\text{A2})$$

where \ddot{c}_t is a time-varying parameter that captures the common shocks to lending; $\ddot{\mathbf{F}}_t$ and $\ddot{\mathbf{B}}_t$ only differ from the variables without umlauts in that the first column has been dropped and

$$\ddot{\boldsymbol{\alpha}}_t = \begin{pmatrix} \alpha_{2t}^I - \alpha_{1t}^I \\ \vdots \\ \alpha_{Ft}^I - \alpha_{1t}^I \end{pmatrix} \text{ and } \ddot{\boldsymbol{\beta}}_t = \begin{pmatrix} \beta_{2t}^I - \beta_{1t}^I \\ \vdots \\ \beta_{Ft}^I - \beta_{1t}^I \end{pmatrix}. \quad (\text{A3})$$

\ddot{c}_t^I , therefore, is the combined firm and bank shock affecting the loan from bank 1 to firm 1 ($\alpha_{1t}^I + \beta_{1t}^I$), and the elements of $\ddot{\boldsymbol{\alpha}}_t$ and $\ddot{\boldsymbol{\beta}}_t$ tell us how much the firm and bank shocks differentially affect every other loan. Some authors, such as Khwaja and Mian (2008) and Greenstone and Mas (2012), estimate a version of equation (A2) using lagged $L_{fb,t-1}$ weights, so it is useful to think about the more general proof in the case in which a researcher employs weighted least squares. Assume this is done using a diagonal weighting matrix, \mathbf{W}_{t-1} , whose diagonal elements are $w_{fb,t-1}$. Formally, we can define this matrix as

$$\mathbf{W}_{t-1} \equiv \begin{pmatrix} \mathbf{W}_{2t-1} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{W}_{Ft-1} \end{pmatrix} \text{ and } \mathbf{W}_{ft-1} \equiv \begin{pmatrix} w_{f,2,t-t} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & w_{f,B,t-t} \end{pmatrix}. \quad (\text{A4})$$

\mathbf{W}_{t-1} is the identity matrix in the case of OLS estimation, and the diagonal elements ($w_{fb,t-1}$) will equal $L_{fb,t-1}$ in the case of the commonly-used weighted least squares procedure. We can write the moment conditions as

$$E \begin{pmatrix} \mathbf{W}_{t-1} \boldsymbol{\eta}_t \\ \ddot{\mathbf{F}}_t' \mathbf{W}_{t-1} \boldsymbol{\eta}_t \\ \ddot{\mathbf{B}}_t' \mathbf{W}_{t-1} \boldsymbol{\eta}_t \end{pmatrix} = 0, \quad (\text{A5})$$

and the corresponding sample moment condition as

$$\begin{pmatrix} \mathbf{W}_{t-1} \hat{\boldsymbol{\eta}}_t \\ \ddot{\mathbf{F}}_t' \mathbf{W}_{t-1} \hat{\boldsymbol{\eta}}_t \\ \ddot{\mathbf{B}}_t' \mathbf{W}_{t-1} \hat{\boldsymbol{\eta}}_t \end{pmatrix} = 0, \quad (\text{A6})$$

where $\hat{\boldsymbol{\eta}}_t$ are the residuals obtained from estimating equation (A2).

Our second normalization rules out trivial interaction terms that are perfectly collinear with the constant term or fixed effects as this type of interaction also will not affect the predicted values. To see the problem formally, note that we can always write the interaction term as

$$\mathbf{Z}_t = \mathbf{1}\delta_t + \ddot{\mathbf{F}}_t \ddot{\boldsymbol{\delta}}_t^F + \ddot{\mathbf{B}}_t \ddot{\boldsymbol{\delta}}_t^B + \boldsymbol{\zeta}_t, \quad (\text{A7})$$

where δ_t^1 , $\ddot{\boldsymbol{\delta}}_t^F$, and $\ddot{\boldsymbol{\delta}}_t^B$ are components of \mathbf{Z}_t that only vary at the firm and bank level and $\boldsymbol{\zeta}_t \neq 0$ is an orthogonal component of the variable that does not only vary by firm or bank. Just as we did not want our estimates to vary based on which firm and bank we drop in the estimation, we also do not want to consider trivial variation arising from, for example, adding a constant to a variable, which would reduce the estimated value of c_t by the value of the constant times the estimated coefficient on the variable. Similarly, if we add a number to all the values associated with a particular bank or firm, this will just change the value of $\ddot{\beta}_{bt}$ and $\ddot{\alpha}_{ft}$ by the same amount. Since we obviously do not want to consider cases in which our estimates of bank and firm shocks move simply because we added a constant to the interaction term, we need to normalize the interaction term so that it only captures bank-firm interactions that are orthogonal to the bank and firm fixed effects. It is straightforward to do this by regressing the bank-firm interaction terms, \mathbf{Z}_t , on the firm and bank fixed effects ($\ddot{\mathbf{F}}_t$ and $\ddot{\mathbf{B}}_t$) using WLS, and defining the residual $\hat{\boldsymbol{\zeta}}_t$ as the normalized interaction term. Mechanically, this imposes the following moment conditions on $\hat{\boldsymbol{\zeta}}_t$:

$$\begin{pmatrix} \mathbf{1}' \mathbf{W}_{t-1} \hat{\boldsymbol{\zeta}}_t \\ \ddot{\mathbf{F}}_t' \mathbf{W}_{t-1} \hat{\boldsymbol{\zeta}}_t \\ \ddot{\mathbf{B}}_t' \mathbf{W}_{t-1} \hat{\boldsymbol{\zeta}}_t \end{pmatrix} = \mathbf{0} \quad (\text{A8})$$

Substituting equation (A7) into equation (A2) yields

$$\mathbf{D}_t^+ = \mathbf{1}c_t^I + \ddot{\mathbf{F}}_t \ddot{\boldsymbol{\alpha}}_t^I + \ddot{\mathbf{B}}_t \ddot{\boldsymbol{\beta}}_t^I + \hat{\boldsymbol{\zeta}}_t \gamma_t + \boldsymbol{\eta}_t, \quad (\text{A9})$$

where $c_t^I = \ddot{c}_t + \delta_t^1$, $\ddot{\boldsymbol{\alpha}}_t^I = \ddot{\boldsymbol{\alpha}}_t + \hat{\boldsymbol{\delta}}_{kt}^F$, and $\ddot{\boldsymbol{\beta}}_t^I = \ddot{\boldsymbol{\beta}}_t + \hat{\boldsymbol{\delta}}_{kt}^B$ and $\hat{\boldsymbol{\zeta}}_t = \mathbf{Z}_{kt} - \mathbf{1}\delta_{kt}^1 - \ddot{\mathbf{F}}_t \hat{\boldsymbol{\delta}}_{kt}^F - \ddot{\mathbf{B}}_t \hat{\boldsymbol{\delta}}_{kt}^B$. In other words, the identification of the bank-firm interaction terms is due to the portion of the variable that is not fully explained by the firm and bank fixed effects.

Estimation of equation (A9) requires setting the sample moments to zero:

$$\begin{aligned}
\begin{pmatrix} \mathbf{1}'\mathbf{W}_{t-1}\hat{\boldsymbol{\eta}}_t \\ \ddot{\mathbf{F}}_t'\mathbf{W}_{t-1}\hat{\boldsymbol{\eta}}_t \\ \ddot{\mathbf{B}}_t'\mathbf{W}_{t-1}\hat{\boldsymbol{\eta}}_t \end{pmatrix} &= \begin{pmatrix} \mathbf{1}'\mathbf{W}_{t-1} \left(\mathbf{D}_t^+ - \left(\mathbf{1}c_t^I + \ddot{\mathbf{F}}_t\hat{\boldsymbol{\alpha}}_t^I + \ddot{\mathbf{B}}_t\hat{\boldsymbol{\beta}}_t^I + \hat{\boldsymbol{\zeta}}_t\hat{\boldsymbol{\gamma}}_t \right) \right) \\ \ddot{\mathbf{F}}_t'\mathbf{W}_{t-1} \left(\mathbf{D}_t^+ - \left(\mathbf{1}c_t^I + \ddot{\mathbf{F}}_t\hat{\boldsymbol{\alpha}}_t^I + \ddot{\mathbf{B}}_t\hat{\boldsymbol{\beta}}_t^I + \hat{\boldsymbol{\zeta}}_t\hat{\boldsymbol{\gamma}}_t \right) \right) \\ \ddot{\mathbf{B}}_t'\mathbf{W}_{t-1} \left(\mathbf{D}_t^+ - \left(\mathbf{1}c_t^I + \ddot{\mathbf{F}}_t\hat{\boldsymbol{\alpha}}_t^I + \ddot{\mathbf{B}}_t\hat{\boldsymbol{\beta}}_t^I + \hat{\boldsymbol{\zeta}}_t\hat{\boldsymbol{\gamma}}_t \right) \right) \end{pmatrix} \\
&= \begin{pmatrix} \mathbf{1}'\mathbf{W}_{t-1} \left(\mathbf{D}_t^+ - \left(\mathbf{1}c_t^I + \ddot{\mathbf{F}}_t\hat{\boldsymbol{\alpha}}_t^I + \ddot{\mathbf{B}}_t\hat{\boldsymbol{\beta}}_t^I \right) \right) \\ \ddot{\mathbf{F}}_t'\mathbf{W}_{t-1} \left(\mathbf{D}_t^+ - \left(\mathbf{1}c_t^I + \ddot{\mathbf{F}}_t\hat{\boldsymbol{\alpha}}_t^I + \ddot{\mathbf{B}}_t\hat{\boldsymbol{\beta}}_t^I \right) \right) \\ \ddot{\mathbf{B}}_t'\mathbf{W}_{t-1} \left(\mathbf{D}_t^+ - \left(\mathbf{1}c_t^I + \ddot{\mathbf{F}}_t\hat{\boldsymbol{\alpha}}_t^I + \ddot{\mathbf{B}}_t\hat{\boldsymbol{\beta}}_t^I \right) \right) \end{pmatrix} - \gamma_t \begin{pmatrix} \mathbf{1}'\mathbf{W}_{t-1}\hat{\boldsymbol{\zeta}}_t \\ \ddot{\mathbf{F}}_t'\mathbf{W}_{t-1}\hat{\boldsymbol{\zeta}}_t \\ \ddot{\mathbf{B}}_t'\mathbf{W}_{t-1}\hat{\boldsymbol{\zeta}}_t \end{pmatrix} = \mathbf{0}
\end{aligned}$$

Substituting in equation (A8) into the last term yields

$$\begin{pmatrix} \mathbf{1}'\mathbf{W}_{t-1}\hat{\boldsymbol{\eta}}_t \\ \ddot{\mathbf{F}}_t'\mathbf{W}_{t-1}\hat{\boldsymbol{\eta}}_t \\ \ddot{\mathbf{B}}_t'\mathbf{W}_{t-1}\hat{\boldsymbol{\eta}}_t \end{pmatrix} = \begin{pmatrix} \mathbf{1}'\mathbf{W}_{t-1} \left(\mathbf{D}_t^+ - \left(\mathbf{1}c_t^I + \ddot{\mathbf{F}}_t\hat{\boldsymbol{\alpha}}_t^I + \ddot{\mathbf{B}}_t\hat{\boldsymbol{\beta}}_t^I \right) \right) \\ \ddot{\mathbf{F}}_t'\mathbf{W}_{t-1} \left(\mathbf{D}_t^+ - \left(\mathbf{1}c_t^I + \ddot{\mathbf{F}}_t\hat{\boldsymbol{\alpha}}_t^I + \ddot{\mathbf{B}}_t\hat{\boldsymbol{\beta}}_t^I \right) \right) \\ \ddot{\mathbf{B}}_t'\mathbf{W}_{t-1} \left(\mathbf{D}_t^+ - \left(\mathbf{1}c_t^I + \ddot{\mathbf{F}}_t\hat{\boldsymbol{\alpha}}_t^I + \ddot{\mathbf{B}}_t\hat{\boldsymbol{\beta}}_t^I \right) \right) \end{pmatrix} = \mathbf{0}. \quad (\text{A10})$$

Note however, that if we had not included the interaction term we would have estimated

$$\mathbf{D}_t^+ = \mathbf{1}c_t + \ddot{\mathbf{F}}_t\hat{\boldsymbol{\alpha}}_t + \ddot{\mathbf{B}}_t\hat{\boldsymbol{\beta}}_t + \boldsymbol{\epsilon}_t, \quad (\text{A11})$$

The weighted least squares estimates of the parameters (\hat{c}_t^W , $\hat{\boldsymbol{\alpha}}_t^W$, and $\hat{\boldsymbol{\beta}}_t^W$) can be obtained by setting the following sample moments to zero.

$$\begin{pmatrix} \mathbf{1}'\mathbf{W}_{t-1}\hat{\boldsymbol{\epsilon}}_t \\ \ddot{\mathbf{F}}_t'\mathbf{W}_{t-1}\hat{\boldsymbol{\epsilon}}_t \\ \ddot{\mathbf{B}}_t'\mathbf{W}_{t-1}\hat{\boldsymbol{\epsilon}}_t \end{pmatrix} = \begin{pmatrix} \mathbf{1}'\mathbf{W}_{t-1} \left(\mathbf{D}_t^+ - \left(\mathbf{1}c_t^W + \ddot{\mathbf{F}}_t\hat{\boldsymbol{\alpha}}_t^W + \ddot{\mathbf{B}}_t\hat{\boldsymbol{\beta}}_t^W \right) \right) \\ \ddot{\mathbf{F}}_t'\mathbf{W}_{t-1} \left(\mathbf{D}_t^+ - \left(\mathbf{1}c_t^W + \ddot{\mathbf{F}}_t\hat{\boldsymbol{\alpha}}_t^W + \ddot{\mathbf{B}}_t\hat{\boldsymbol{\beta}}_t^W \right) \right) \\ \ddot{\mathbf{B}}_t'\mathbf{W}_{t-1} \left(\mathbf{D}_t^+ - \left(\mathbf{1}c_t^W + \ddot{\mathbf{F}}_t\hat{\boldsymbol{\alpha}}_t^W + \ddot{\mathbf{B}}_t\hat{\boldsymbol{\beta}}_t^W \right) \right) \end{pmatrix} = \mathbf{0} \quad (\text{A12})$$

Clearly, the moment equations given by (A10) and (A12) are identical. Since we know that WLS parameter estimates are uniquely identified, we have proven Proposition 1.

Appendix B: Proof of Proposition 2

WLS estimation requires us to chose parameters that set the sample moment conditions given in equation (A10) or equivalently (A12) to be zero. We can expand the equations for each block in the matrix. For example, the element can be written as

$$\mathbf{1}'\mathbf{W}_{t-1}\hat{\boldsymbol{\epsilon}}_t = \sum_{fb \in G_t} \left(D(L_{fbt}/L_{fb,t-1}) - \hat{c}_t^W - \hat{\alpha}_{ft}^W - \hat{\beta}_{bt}^W \right) w_{fb,t-1} = 0, \quad (\text{A13})$$

where G_t is the set of fb pairs in which $D(L_{fbt}/L_{fb,t-1})$ is well defined.⁴³ This sample moment condition can be rewritten as:

⁴³In the log specification this will be whenever $L_{fbt} > 0$ and $L_{fb,t-1} > 0$. The first inequality need not hold in the percentage change specification.

$$\sum_{fb \in G_t} D(L_{fbt}/L_{fb,t-1}) \frac{w_{fb,t-1}}{\sum_{fb \in G_t} w_{fb,t-1}} = \sum_{fb \in G_t} \left(\hat{c}_t^W + \hat{\alpha}_{ft}^W + \hat{\beta}_{bt}^W \right) \frac{w_{fb,t-1}}{\sum_{fb \in G_t} w_{fb,t-1}}. \quad (\text{A14})$$

If we measure growth rates ($D(L_{fbt}/L_{fb,t-1})$) as percentage changes ($(L_{fbt}/L_{fb,t-1} - 1)$) and set $w_{fb,t-1} = L_{fb,t-1}$, which corresponds to weighting the data by $\sqrt{L_{fb,t-1}}$, the weighted average of our predicted growth rates, based on our estimated coefficients, will equal the aggregate growth rate of lending.⁴⁴ We can see this by just substituting these values for the growth rate and weights into equation (A14):

$$\begin{aligned} \sum_{fb \in G_t} \frac{L_{fbt} - L_{fb,t-1}}{L_{fb,t-1}} \frac{L_{fb,t-1}}{\sum_{fb \in G_t} L_{fb,t-1}} &= \frac{\sum_{fb \in G_t} L_{fbt} - \sum_{fb \in G_t} L_{fb,t-1}}{\sum_{fb \in G_t} L_{fb,t-1}} \\ &= \sum_{fb \in G_t} \left(\hat{c}_t^W + \hat{\alpha}_{ft}^W + \hat{\beta}_{bt}^W \right) \frac{L_{fb,t-1}}{\sum_{fb \in G_t} L_{fb,t-1}}. \end{aligned} \quad (\text{A15})$$

What about the other moment conditions? Consider, for example the row of the $\tilde{\mathbf{F}}_t'$ matrix corresponding to any firm f . In order for the moment condition for this firm to be satisfied, it must be the case that $\tilde{\mathbf{F}}_{ft}' \mathbf{W}_{t-1} \hat{\boldsymbol{\epsilon}}_t = 0$. If we write this condition without using matrix algebra and divide both sides of this moment condition by the number of loans made to firm f , $N_t^{f'}$, we have

$$\sum_{b \in G_{ft}} D(L_{fbt}/L_{fb,t-1}) \frac{w_{fb,t-1}}{\sum_{b \in G_{ft}} w_{fb,t-1}} = \sum_{b \in G_{ft}} \left(\hat{c}_t^W + \hat{\alpha}_{ft}^W + \hat{\beta}_{bt}^W \right) \frac{w_{fb,t-1}}{\sum_{b \in G_{ft}} w_{fb,t-1}}, \quad (\text{A16})$$

where G_{ft} are the set of banks lending to firm f (i.e., loan growth rates for which $D(L_{fbt}/L_{fb,t-1})$ is well defined. If we now measure growth rates as percentage changes and weight by $L_{fb,t-1}$ we have

$$\begin{aligned} \sum_{b \in G_{ft}} \frac{L_{fbt} - L_{fb,t-1}}{L_{fb,t-1}} \frac{L_{fb,t-1}}{\sum_{b \in G_{ft}} L_{fb,t-1}} &= \frac{\sum_{b \in G_{ft}} L_{fbt} - \sum_{b \in G_{ft}} L_{fb,t-1}}{\sum_{b \in G_{ft}} L_{fb,t-1}} \\ &= \sum_{b \in G_{ft}} \left(\hat{c}_t^W + \hat{\alpha}_{ft}^W + \hat{\beta}_{bt}^W \right) \frac{L_{fb,t-1}}{\sum_{b \in G_{ft}} L_{fb,t-1}}. \end{aligned} \quad (\text{A17})$$

⁴⁴One can see this by considering a regression of the form $\mathbf{D} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, where \mathbf{D} denotes the dependent variable, \mathbf{X} is the matrix of independent variables, $\boldsymbol{\beta}$ denotes the coefficient vector, and $\boldsymbol{\epsilon}$ is the error term. Weighted least squares involves premultiplying the data by a diagonal weighting matrix \mathbf{P} to yield $\mathbf{PD} = \mathbf{PX}\boldsymbol{\beta} + \mathbf{P}\boldsymbol{\epsilon}$. The WLS estimate of $\boldsymbol{\beta}$ is then $\hat{\boldsymbol{\beta}} = [(\mathbf{PX})' \mathbf{PX}]^{-1} (\mathbf{PX})' \mathbf{PD} = [\mathbf{X}' \mathbf{P}' \mathbf{PX}]^{-1} \mathbf{X}' \mathbf{P}' \mathbf{PD} = [\mathbf{X}' \mathbf{W} \mathbf{X}]^{-1} \mathbf{X}' \mathbf{W} \mathbf{D}$. Similarly, the corresponding moment condition for the error term can be written as $E[\mathbf{X}' \mathbf{P}' \mathbf{P} \boldsymbol{\epsilon}] = E[\mathbf{X}' \mathbf{W} \boldsymbol{\epsilon}] = 0$. Thus, weighting the data by a diagonal matrix \mathbf{P} , whose diagonal elements are $\sqrt{L_{fb,t-1}}$, will produce a weighting matrix \mathbf{W} whose diagonal elements are L_{fbt} .

Once again, setting the growth rate equal to the percentage growth rate of lending and weights equal to the lagged loan levels will ensure this condition is satisfied. Finally, we can verify that WLS estimates will exactly match the average firm-borrowing growth rate of each bank. To see this note that in the third block of the moment condition matrix we have for any bank b :

$$\sum_{f \in G_{bt}} D(L_{fbt}/L_{fb,t-1}) \frac{w_{fb,t-1}}{\sum_{f \in G_{bt}} w_{fb,t-1}} = \sum_{f \in G_{bt}} \left(\hat{c}_t^W + \hat{\alpha}_{ft}^W + \hat{\beta}_{bt}^W \right) \frac{w_{fb,t-1}}{\sum_{f \in G_{bt}} w_{fb,t-1}} \quad (\text{A18})$$

where G_{bt} are the set of firms borrowing from bank b (i.e., loan growth rates for which $D(L_{fbt}/L_{fb,t-1})$ is well defined. Substituting in the same definition of the loan growth rate and the weights yields

$$\begin{aligned} \sum_{f \in G_{bt}} \frac{L_{fbt} - L_{fb,t-1}}{L_{fb,t-1}} \frac{L_{fb,t-1}}{\sum_{f \in G_{bt}} L_{fb,t-1}} &= \frac{\sum_{f \in G_{bt}} L_{fbt} - \sum_{f \in G_{bt}} L_{fb,t-1}}{\sum_{f \in G_{bt}} L_{fb,t-1}} \\ &= \sum_{f \in G_{bt}} \left(\hat{c}_t^W + \hat{\alpha}_{ft}^W + \hat{\beta}_{bt}^W \right) \frac{L_{fb,t-1}}{\sum_{f \in G_{bt}} L_{fb,t-1}}. \end{aligned} \quad (\text{A19})$$

Therefore, the firm and bank shocks will exactly match the bank, firm, and economy-wide loan growth rates. This establishes the first sentence of our second proposition. Because we also know that WLS estimates are uniquely identified, we know that no other estimates of bank and firm fixed effects will have this property.

Appendix C: Proof of Proposition 3

We can prove that that OLS estimation of equation (1) produces unweighted averages of the loan growth rates by setting $w_{fbt} = 1$ and $D(L_{fbt}/L_{fb,t-1})$ equal to $L_{fbt}/L_{fb,t-1} - 1$. In this case, our moment conditions, given by equations (A14), (A16), and (A18) become

$$\frac{1}{|G_t|} \sum_{fb \in G_t} \frac{L_{fbt} - L_{fb,t-1}}{L_{fb,t-1}} = \frac{1}{|G_t|} \sum_{fb \in G_t} \left(\hat{c}_t^W + \hat{\alpha}_{ft}^W + \hat{\beta}_{bt}^W \right), \quad (\text{A20})$$

$$\frac{1}{|G_f|} \sum_{b \in G_f} \frac{L_{fbt} - L_{fb,t-1}}{L_{fb,t-1}} = \frac{1}{|G_f|} \sum_{b \in G_f} \left(\hat{c}_t^W + \hat{\alpha}_{ft}^W + \hat{\beta}_{bt}^W \right), \quad (\text{A21})$$

$$\frac{1}{|G_b|} \sum_{f \in G_b} \frac{L_{fbt} - L_{fb,t-1}}{L_{fb,t-1}} = \frac{1}{|G_b|} \sum_{f \in G_b} \left(\hat{c}_t^W + \hat{\alpha}_{ft}^W + \hat{\beta}_{bt}^W \right), \quad (\text{A22})$$

where $|G|$, $|G_f|$, and $|G_b|$ are the number of observations with positive loans in the previous period for the economy, firm f , and bank b . This establishes the first sentence of the proposition that the predicted loan growth rates using OLS estimation yields forecasts of loans that equal the unweighted average growth rates of individual loans at the firm, bank, and economy level.

We can prove the second sentence in the proposition by setting $g(L_{fbt}/L_{fb,t-1}) = \Delta \ln L_{fbt}$. In

this case, our moment conditions, given by equations (A14), (A16), and (A18) become

$$\sum_{fb \in G} \Delta \ln L_{fbt} \frac{w_{fb,t-1}}{\sum_{fb \in G} w_{fb,t-1}} = \sum_{fb \in G} \left(\hat{c}_t^W + \hat{\alpha}_{ft}^W + \hat{\beta}_{bt}^W \right) \frac{w_{fb,t-1}}{\sum_{fb \in G} w_{fb,t-1}}, \quad (\text{A23})$$

$$\sum_{b \in G_f} \Delta \ln L_{fbt} \frac{w_{fb,t-1}}{\sum_{b \in G_f} w_{fb,t-1}} = \sum_{b \in G_f} \left(\hat{c}_t^W + \hat{\alpha}_{ft}^W + \hat{\beta}_{bt}^W \right) \frac{w_{fb,t-1}}{\sum_{b \in G_f} w_{fb,t-1}}, \quad (\text{A24})$$

$$\sum_{f \in G_b} \Delta \ln L_{fbt} \frac{w_{fb,t-1}}{\sum_{f \in G_b} w_{fb,t-1}} = \sum_{f \in G_b} \left(\hat{c}_t^W + \hat{\alpha}_{ft}^W + \hat{\beta}_{bt}^W \right) \frac{w_{fb,t-1}}{\sum_{f \in G_b} w_{fb,t-1}}. \quad (\text{A25})$$

Noting that $\Delta \ln L_{fbt} = \ln(L_{fbt}/L_{fb,t-1})$, we know from Proposition 2 that these are nothing more than logs of the weighted geometric averages of one plus the loan growth rates. This establishes the second sentence of the proposition that the predicted loan growth rates using OLS estimation yields forecasts of loans that equal the geometric mean of one plus the growth rates of individual loan growth rates at the firm, bank, and economy level.

Appendix C: Proof of Proposition (4)

Identifying the Bank and Firm Shocks

It will prove useful to simplify the algebra by changing our normalization before proceeding. Importantly, this change in normalization is for mathematical convenience only, as we will renormalize the data in an economically meaningful way in Section 6. Under the normalization that was most convenient for WLS estimation, our moment equations (A19) and (A17) can be expressed as

$$\mathbf{D}_{Bt}^+ = \hat{c}_t^W \mathbf{1}_B + \mathbf{B}_t^W + \Phi_{t-1} \mathbf{A}_t^W \quad (\text{A26})$$

and

$$\mathbf{D}_{Ft}^+ = \hat{c}_t^W \mathbf{1}_F + \mathbf{A}_t^W + \Theta_{t-1} \mathbf{B}_t^W, \quad (\text{A27})$$

where \mathbf{A}_t^W and \mathbf{B}_t^W are the matrices of coefficients we obtain when we drop the first firm and bank, i.e.,

$$\mathbf{A}_t^W \equiv \begin{pmatrix} 0 \\ \hat{\alpha}_{2t}^W \\ \vdots \\ \hat{\alpha}_{Ft}^W \end{pmatrix}, \mathbf{B}_t^W \equiv \begin{pmatrix} 0 \\ \hat{\beta}_{2t}^W \\ \vdots \\ \hat{\beta}_{Ft}^W \end{pmatrix},$$

and

$$\mathbf{D}_{Ft}^+ \equiv \begin{pmatrix} D_{1t}^{F+} \\ \vdots \\ D_{Ft}^{F+} \end{pmatrix}, \mathbf{D}_{Bt} \equiv \begin{pmatrix} D_{1t}^{B+} \\ \vdots \\ D_{Bt}^{B+} \end{pmatrix}, \Theta_t \equiv \begin{pmatrix} \theta_{11t} & \dots & \theta_{1Bt} \\ \vdots & \ddots & \vdots \\ \theta_{F1t} & \dots & \theta_{FBt} \end{pmatrix}, \Phi_t \equiv \begin{pmatrix} \phi_{11t} & \dots & \phi_{F1t} \\ \vdots & \ddots & \vdots \\ \phi_{1Bt} & \dots & \phi_{FBt} \end{pmatrix}.$$

For each firm, we know that if we sum the share of its borrowing from each bank across all banks, the total must equal one. This means that $\Theta_{t-1}\mathbf{1}_B = \mathbf{1}_F$, where $\mathbf{1}_B$ and $\mathbf{1}_F$ are $B \times 1$ and $F \times 1$ vectors of 1's, respectively. This fact enables us to rewrite equations (A26) and (A27) as

$$\mathbf{D}_{Bt}^+ = (\hat{c}_t^W \mathbf{1}_B + \mathbf{B}_t^W) + \Phi_{t-1} \mathbf{A}_t^W$$

and

$$\mathbf{D}_{Ft}^+ = \mathbf{A}_t^W + \Theta_{t-1} (\hat{c}_t^W \mathbf{1}_B + \mathbf{B}_t^W).$$

Writing the moment conditions this way is tantamount to a new normalization in which we drop the constant term and define the bank shocks $\beta_{bt}^W = \hat{c}_t^W + \hat{\beta}_{bt}^W$.

We can make things more compact and simplify our notation by rewriting these equations as

$$\mathbf{D}_{Bt}^+ = \mathbf{B}_t^+ + \Phi_{t-1} \mathbf{A}_t^+ \tag{A28}$$

and

$$\mathbf{D}_{Ft}^+ = \mathbf{A}_t^+ + \Theta_{t-1} \mathbf{B}_t^+, \tag{A29}$$

where

$$\mathbf{A}_t^+ \equiv \begin{pmatrix} 0 \\ \alpha_{2t}^W \\ \vdots \\ \alpha_{Ft}^W \end{pmatrix} = \mathbf{A}_t^1, \mathbf{B}_t^+ \equiv \begin{pmatrix} \beta_{1t}^+ \\ \beta_{2t}^+ \\ \vdots \\ \beta_{Ft}^+ \end{pmatrix} = (\hat{c}_t^W \mathbf{1}_B + \mathbf{B}_t^W).$$

In other words, the moment equations given by (A19) and (A17) are equivalent to equations (A28) and (A29) up to a choice of normalization.⁴⁵ As we have already shown in Proposition 2, there is a unique solution to this system, and it is immediately apparent from equations (7) and (8) that if we solve for the parameters, \mathbf{A}_t^+ and \mathbf{B}_t^+ , we will obtain a set of bank- and firm-shock estimates that will aggregate to match every bank and firm's growth rate of *pre-existing* loans. Note, however, that if we just replaced \mathbf{D}_{Bt}^+ and \mathbf{D}_{Ft}^+ in equation (A28) and (A29) with \mathbf{D}_{Bt} and \mathbf{D}_{Ft} , we would obtain the following system:

$$\mathbf{D}_{Bt} = \mathbf{B}_t + \Phi_{t-1} \mathbf{A}_t \tag{A30}$$

and

$$\mathbf{D}_{Ft} = \mathbf{A}_t + \Theta_{t-1} \mathbf{B}_t, \tag{A31}$$

where

⁴⁵The WLS parameters obtained from the new normalization can be recovered by setting $\hat{c}_t^W = \beta_{1t}^+$ and $\hat{\beta}_{bt}^W = \beta_{bt}^+ - \beta_{1t}^+$.

$$\mathbf{A}_t \equiv \begin{pmatrix} 0 \\ \alpha_{2t} \\ \vdots \\ \alpha_{Ft} \end{pmatrix} = \mathbf{A}_t^1, \mathbf{B}_t \equiv \begin{pmatrix} \beta_{1t} \\ \beta_{2t} \\ \vdots \\ \beta_{Ft} \end{pmatrix},$$

and the parameter estimates that would solve this new system. In this case, \mathbf{A}_t and \mathbf{B}_t can be aggregated so that they exactly match every bank and firm's growth rate of *total* loans. Moreover, the isomorphism of equations (A28)-(A29) and (A30)-(A31) tells us that equations (A30)-(A31) have a unique solution. In addition, the estimated parameters that solve equations (A30)-(A31) must be identical to those that solve equations (A28)-(A29) if there are no new loans (i.e., $\mathbf{A}_t^+ = \mathbf{A}_t$ and $\mathbf{B}_t^+ = \mathbf{B}_t$ if $D_{Bt}^+ = D_{Bt}$ and $D_{Ft}^+ = D_{Ft}$). In other words, in the absence of new loans, the WLS parameter estimates and the estimates obtained by solving equations (A30) and (A31) must be identical given any common choice of normalization.

The advantage of working with equations (A30) and (A31) is that we can now obtain bank and firm-shock parameter estimates that aggregate appropriately even if there is new lending activity. In order to show how to do this we first rewrite equation (8) as

$$\begin{aligned} D_{ft}^F &\equiv \alpha_{ft} + \sum_b \theta_{fb,t-1} (\beta_{bt} - \beta_{1t}) + \beta_{1t} \\ &= \alpha_{ft} + \sum_{b \neq 1} \theta_{fb,t-1} (\beta_{bt} - \beta_{1t}) + \beta_{1t}, \end{aligned} \quad (\text{A32})$$

where we use the fact that the loan shares must sum to one, i.e., $\sum_b \theta_{fb,t-1} = 1$. Similarly, since $\alpha_{1t} = 0$, we can rewrite equation (7) for any bank, b , as

$$D_{bt}^B = \beta_{bt} + \sum_{f \neq 1} \phi_{fb,t-1} \alpha_{ft}, \quad (\text{A33})$$

and for bank 1, we have

$$D_{1t}^B = \beta_{1t} + \sum_{f \neq 1} \phi_{f1,t-1} \alpha_{ft}. \quad (\text{A34})$$

Finally, for firm 1, we have

$$D_{1t}^F = \sum_{b \neq 1} \theta_{1b,t-1} (\beta_{bt} - \beta_{1t}) + \beta_{1t}, \quad (\text{A35})$$

which can be rewritten as

$$\beta_{1t} = D_{1t}^F - \sum_{b \neq 1} \theta_{1b,t-1} (\beta_{bt} - \beta_{1t}). \quad (\text{A36})$$

If we substitute equation (A36) into equation (A32), we obtain

$$D_{ft}^F - D_{1t}^F = \alpha_{ft} + \sum_{b \neq 1} (\theta_{fb,t-1} - \theta_{1b,t-1}) (\beta_{bt} - \beta_{1t}) = \alpha_{ft} + \sum_{b \neq 1} \ddot{\theta}_{fb,t-1} \ddot{\beta}_{bt}, \quad (\text{A37})$$

where $\ddot{\theta}_{fbt} \equiv (\theta_{fbt} - \theta_{1bt})$, and $\ddot{\beta}_{bt} \equiv (\beta_{bt} - \beta_{1t})$. Subtracting equation (A34) from equation (A33) gives us

$$D_{bt}^B - D_{1t}^B = (\beta_{bt} - \beta_{1t}) + \sum_{f \neq 1} (\phi_{fb,t-1} - \phi_{f1,t-1}) \alpha_{ft} = \ddot{\beta}_{bt} + \sum_{f \neq 1} \ddot{\phi}_{fb,t-1} \alpha_{ft}, \quad (\text{A38})$$

where $\ddot{\phi}_{fbt} = (\phi_{fbt} - \phi_{f1t})$.

For every year, equations (A37) and (A38) comprise a system of $F + B - 2$ equations and $F + B - 2$ unknowns, where F denotes the number of firms and B denotes the number of banks. *In other words, the moment conditions uniquely determine the bank and firm shocks up to the choice of numéraire.* To see this, using matrix algebra, we begin with a few definitions:

$$\begin{aligned} \ddot{\mathbf{A}}_t &\equiv \begin{pmatrix} \alpha_{2t} \\ \vdots \\ \alpha_{Ft} \end{pmatrix}, \ddot{\mathbf{B}}_t \equiv \begin{pmatrix} \ddot{\beta}_{2t} \\ \vdots \\ \ddot{\beta}_{Bt} \end{pmatrix}, \ddot{\mathbf{D}}_{Ft} \equiv \begin{pmatrix} D_{2t}^F - D_{1t}^F \\ \vdots \\ D_{Ft}^F - D_{1t}^F \end{pmatrix}, \ddot{\mathbf{D}}_{Bt} \equiv \begin{pmatrix} D_{2t}^B - D_{1t}^B \\ \vdots \\ D_{Bt}^B - D_{1t}^B \end{pmatrix} \\ \ddot{\Theta}_t &\equiv \begin{pmatrix} \ddot{\theta}_{22t} & \dots & \ddot{\theta}_{2Bt} \\ \vdots & \ddots & \vdots \\ \ddot{\theta}_{F2t} & \dots & \ddot{\theta}_{FBt} \end{pmatrix}, \ddot{\Phi}_t \equiv \begin{pmatrix} \ddot{\phi}_{22t} & \dots & \ddot{\phi}_{F2t} \\ \vdots & \ddots & \vdots \\ \ddot{\phi}_{2Bt} & \dots & \ddot{\phi}_{FBt} \end{pmatrix} \end{aligned}$$

We now can rewrite equation (A37) as

$$\ddot{\mathbf{A}}_t = \ddot{\mathbf{D}}_{Ft} - \ddot{\Theta}_{t-1} \ddot{\mathbf{B}}_t \quad (\text{A39})$$

and equation (A38) as

$$\ddot{\mathbf{B}}_t = \ddot{\mathbf{D}}_{Bt} - \ddot{\Phi}_{t-1} \ddot{\mathbf{A}}_t. \quad (\text{A40})$$

If we use this fact and insert equation (A39) into equation (A40), we know that any solution to equations (8) and (A40) must satisfy

$$\ddot{\mathbf{B}}_t = \ddot{\mathbf{D}}_{Bt} - \ddot{\Phi}_{t-1} [\ddot{\mathbf{D}}_{Ft} - \ddot{\Theta}_{t-1} \ddot{\mathbf{B}}_t].$$

This equation can be rewritten as

$$\left(\mathbf{I}_{B-1} - \ddot{\Phi}_{t-1} \ddot{\Theta}_{t-1} \right) \ddot{\mathbf{B}}_t = \ddot{\mathbf{D}}_{Bt} - \ddot{\Phi}_{t-1} \ddot{\mathbf{D}}_{Ft},$$

where \mathbf{I}_{B-1} is a $(B-1) \times (B-1)$ identity matrix. We can solve for \mathbf{A}_t and \mathbf{B}_t by inverting $\left(\mathbf{I}_{B-1} - \ddot{\Phi}_{t-1} \ddot{\Theta}_{t-1} \right)$ to yield⁴⁶

$$\ddot{\mathbf{B}}_t = \left(\mathbf{I}_{B-1} - \ddot{\Phi}_{t-1} \ddot{\Theta}_{t-1} \right)^{-1} \left(\ddot{\mathbf{D}}_{Bt} - \ddot{\Phi}_{t-1} \ddot{\mathbf{D}}_{Ft} \right) \quad (\text{A41})$$

and

$$\ddot{\mathbf{A}}_t = \ddot{\mathbf{D}}_{Ft} - \ddot{\Theta}_{t-1} \left(\mathbf{I}_{B-1} - \ddot{\Phi}_{t-1} \ddot{\Theta}_{t-1} \right)^{-1} \left(\ddot{\mathbf{D}}_{Bt} - \ddot{\Phi}_{t-1} \ddot{\mathbf{D}}_{Ft} \right).$$

We then can compute β_{1t} according to the following formula:

$$\beta_{1t} = D_{1t}^B - \sum_{f \neq 1} \phi_{f1,t-1} \alpha_{ft}.$$

Decomposing the Shocks

We now turn to using our parameter estimates to show how we can decompose aggregate lending growth into bank, firm, industry, and common shocks. Our choice of *numéraire* was necessary for

⁴⁶One important assumption we make is that $\left(\mathbf{I}_{B-1} - \ddot{\Phi}_{t-1} \ddot{\Theta}_{t-1} \right)$ is invertible. This condition would be violated if, say, one firm only borrowed from one bank, and the bank had no other lending relationships.

solving the system, but it is not economically meaningful and has the unappealing property of expressing each shock relative to an arbitrary firm shock. In order to eliminate this problem, we will renormalize the system by expressing each shock relative to the median firm or bank shock. Specifically, we define the common firm shock, \bar{A}_t , as the median firm shock, and the common financial shock as the median bank shock, \bar{B}_t . We next define the firm-borrowing shock as the difference between the actual shocks and the median shock, i.e., $\dot{\mathbf{A}}_t \equiv \mathbf{A}_t - \bar{A}_t \mathbf{1}_F$, and similarly the financial institution loan shock as $\dot{\mathbf{B}}_t \equiv \mathbf{B}_t - \bar{B}_t \mathbf{1}_B$, where

$$\mathbf{A}_t \equiv \begin{pmatrix} \alpha_{1t} \\ \vdots \\ \alpha_{Ft} \end{pmatrix}, \mathbf{B}_t \equiv \begin{pmatrix} \beta_{1t} \\ \vdots \\ \beta_{Bt} \end{pmatrix}.$$

We can rewrite the system of equations as

$$\begin{aligned} D_{Ft} &= \mathbf{A}_t + \Theta_{t-1} \mathbf{B}_t \\ &= \dot{\mathbf{A}}_t + \bar{A}_t \mathbf{1}_F + \Theta_{t-1} \dot{\mathbf{B}}_t + \bar{B}_t \Theta_{t-1} \mathbf{1}_B \\ &= \dot{\mathbf{A}}_t + \Theta_{t-1} \dot{\mathbf{B}}_t + (\bar{A}_t + \bar{B}_t) \mathbf{1}_F, \end{aligned} \tag{A42}$$

where we move from the second line to the third by making use of the fact that the borrowing shares from each financial institution (Θ_{t-1}) must sum to one, i.e., $\Theta_{t-1} \mathbf{1}_B = \mathbf{1}_F$.

Just as we can decompose firm borrowing into these three shocks, we can also decompose bank lending into a similar set of three elements. If we make use of the fact that each bank's share of total lending to each firm must sum to one (i.e., $\Phi_{t-1} \mathbf{1}_F = \mathbf{1}_B$), we can rewrite equation (7) as

$$\begin{aligned} D_{Bt} &= \dot{\mathbf{B}}_t + \Phi_{t-1} \dot{\mathbf{A}}_t + (\bar{B}_t \mathbf{1}_B + \Phi_{t-1} \bar{A}_t \mathbf{1}_F) \\ &= \dot{\mathbf{B}}_t + \Phi_{t-1} \dot{\mathbf{A}}_t + (\bar{A}_t + \bar{B}_t) \mathbf{1}_B. \end{aligned} \tag{A43}$$

We now turn to isolating the industry shocks. We define the firm shock as $\tilde{\mathbf{A}}_t \equiv \dot{\mathbf{A}}_t - \text{median}_{f \in n}(\dot{\mathbf{A}}_t)$, where n denotes the industry the firm is in. Similarly, we define the bank shock as $\tilde{\mathbf{B}}_t \equiv \dot{\mathbf{B}}_t - \text{median}(\dot{\mathbf{B}}_t) = \dot{\mathbf{B}}_t$ since $\text{median}(\dot{\mathbf{B}}_t) = 0$. Finally, we can define the vector of industry-level medians as \mathbf{N}_t ; then we can rewrite firm and bank decompositions as

$$D_{Ft} = \tilde{\mathbf{A}}_t + \mathbf{N}_t + \Theta_{t-1} \tilde{\mathbf{B}}_t + (\bar{A}_t + \bar{B}_t) \mathbf{1}_F, \tag{A44}$$

and

$$D_{Bt} = \tilde{\mathbf{B}}_t + \Phi_{t-1} \tilde{\mathbf{A}}_t + \Phi_{t-1} \mathbf{N}_t + (\bar{A}_t + \bar{B}_t) \mathbf{1}_B. \tag{A45}$$

Appendix E: Summary Statistics

By Year	Mean	Median	SD	Minimum	Maximum
Percent change in Flow of Funds _{<i>t</i>}	-0.022	-0.029	0.047	-0.094	0.085
Investment _{<i>t</i>} /Capital _{<i>t-1</i>}	0.073	0.069	0.017	0.055	0.121
Common Shock _{<i>t</i>}	-0.010	-0.010	0.047	-0.142	0.066
By Bank					
Bank Shock _{<i>b,t</i>}	0.010	0.000	0.182	-0.548	0.934
Capital Injection _{<i>b,t</i>}	0.096	0.000	0.295	0.000	1.000
Risk-Based Capital Ratio _{<i>b,t</i>}	8.796	9.120	1.766	2.820	13.610
$\Delta \ln(\text{Market-to-Book Value}_{b,t-1})$	-0.099	-0.107	0.225	-0.684	0.572
By Firm					
Investment _{<i>f,t</i>} /Capital _{<i>f,t-1</i>}	0.118	0.080	0.167	-0.398	1.745
Cash-Flow _{<i>f,t</i>} /Capital _{<i>f,t-1</i>}	0.335	0.209	0.719	-1.569	15.410
Market-to-Book Value _{<i>f,t</i>}	1.696	1.295	1.377	0.234	10.367
Bank Shock _{<i>f,t</i>}	-0.011	-0.012	0.051	-0.173	0.151
Firm Shock _{<i>f,t</i>}	0.059	0.000	0.393	-0.760	3.252
Industry Shock _{<i>f,t</i>}	0.003	0.002	0.077	-0.658	2.492
Mean Loan-to-Asset Ratio _{<i>f</i>}	0.196	0.174	0.123	0.001	0.748
Mean Bond-to-Asset Ratio _{<i>f</i>}	0.048	0.029	0.057	0.000	0.359

Notes: The variables are defined as follows. By year: Percent change in Flow of Funds_{*t*} is the year-on-year percent change in the stock of lending from private financial institutions to private nonfinancial corporations. Investment_{*t*}/Capital_{*t-1*} is the total national investment in year *t* divided by the total capital stock in year *t* - 1, both measured in 2000 yen. Common Shock_{*t*} is the sum of the median firm shock and the median bank shock in each year. By bank: Bank Shock_{*b,t*} is the idiosyncratic bank shock for each bank. Capital Injection_{*b,t*} is a binary variable equal to 1 if bank *b* received a capital injection in year *t*. Risk-Based Capital Ratio_{*b,t*} is the combined Tier 1 and Tier 2 risk-based capital according to the Basel II agreement. $\Delta \ln(\text{Market-to-Book Value}_{b,t-1})$ is the lagged year-on-year log change in the market-to-book value, computed as the average of the monthly share price multiplied by the number of shares outstanding and divided by the book value of each firm's equity. By firm: Investment_{*f,t*}/Capital_{*f,t-1*} is the sum of each firm's year-on-year change in tangible fixed assets plus depreciation divided by that firm's tangible fixed assets in year *t* - 1. Cash Flow_{*f,t*}/Capital_{*f,t-1*} is firm *f*'s ordinary income plus depreciation divided by tangible fixed assets in year *t* - 1. Market-to-Book Value_{*f,t*} is the market value of each firm divided by its shareholder equity. Bank Shock_{*f,t*} is a weighted sum of the idiosyncratic bank shocks for all banks from which firm *f* was receiving a loan in year *t* - 1. Firm Shock_{*f,t*} is the idiosyncratic shock for each firm in each year. Industry Shock_{*f,t*} is the industry level shock for the industry of which firm *f* is a member. Mean Loan-to-Asset Ratio_{*f*} is firm *f*'s mean ratio of total loans to total assets over the sample period. Mean Bond-to-Asset Ratio_{*f*} is firm *f*'s mean ratio of total bonds to total assets.