Formality as a Financing Decision

David Morales* Daniela Puggioni†

Job Market Paper

November 2026

Latest version available here

Abstract

Can bank loans make informal firms formalize? We match the universe of bank loans, the social security records, and the economic censuses in Mexico to track formality, loans, and firm transitions from 2005 to 2025. Three facts emerge. Three quarters of firms with a bank loan are informal; informal firms pay higher rates, provide more collateral on collateralized loans, and obtain smaller loans; and firms move between formality and informality in both directions. We develop a model in which firms that need a loan and are at the margin between being formal and being informal change their status when the relative terms or availability of loans change. Leveraging bank credit supply shocks, we find that a ten percent credit supply expansion widens the financial advantage of formal firms and, as the model predicts, raises the number of firms that formalize by 850 (0.5 percent) and lowers the number of firms that become informal by 600 (0.4 percent) per five-year window in our sample markets. We demonstrate that bank loans affect not only formal firms but also informal ones, and are a lever for formalization.

Keywords: informality, firm dynamics, financial frictions, credit supply, formalization.
JEL codes: O17, G21, L25

*Toulouse School of Economics, david.morales@tse-fr.eu. I am deeply grateful to Matteo Bobba for his guidance. I thank Milo Bianchi, Oscar Fentanes, Eugenia Gonzalez-Aguado, Jean-Marie Lozachmeur, and Costas Meghir for valuable feedback. I also thank participants at the TSE BID Workshop, TSE Macro Workshop, TSE Public Workshop, TSE Finance Workshop, TSE Yinghua He Empirical Micro Workshop, Yale Labor Breakfast, Yale Development Breakfast, NOVAfrica PhD Workshop, and ENTER Jamboree for helpful comments and discussions. I am grateful to the teams at Banco de México's EconLab and the *Laboratorio de Microdatos del INEGI* for their support. I acknowledge financial support from the French National Research Agency (ANR), grant RECHC – ANR-20-CE26-0011-01. All errors are our own.

†Banco de México, dpuggionih@banxico.org.mx

1 Introduction

The effects of the financial sector on formal firms are well known in the literature.[1] However, little is known about its effect on the informal sector. This gap reflects the scarcity of data linking informality and finance. This paper overcomes these data limitations, uncovers stylized facts linking informality and finance, and addresses the following research question: Can bank loans affect the formality decision of firms? We show that credit supply expansions reach informal firms but the improvement in interest rates and in the volume of credit is larger for formal firms. This asymmetric effect increases the financial advantage of formal firms and changes formality transitions in both directions. The reason is that, for firms that need a loan, the gains from better loan terms in the formal sector can offset the tax costs of formality, making formality a financial decision.

This paper matches the universe of bank loans to social security administrative records in Mexico at the firm level to uncover the formality of borrowers, the loan position of formal firms, and the difference in loan terms between formal and informal enterprises. The matched dataset contains nearly 3.3 million firms, 1.3 million borrowing and 2.2 million formal, across 119 banks between 2005 and 2025. However, informal firms without a bank loan cannot be observed in the matched dataset by construction. To overcome this problem, we merge these two datasets with the economic censuses, which is only possible at the market-level, to observe all types of firms over time: formal and informal, borrowers and non-borrowers. To our knowledge, no other dataset contains formality, bank loans, and firm transitions in the same local labor markets[2] over twenty years. As a result, we can study a relatively unexplored relationship: the effect of bank loans on informal firms. The relevance of this relationship comes from the prevalence of the informal sector. In Mexico, 88 percent of firms were informal in the 2024 Economic Census.

With these data, we document three stylized facts. First, contrary to a common assumption, informal firms borrow from banks: around three quarters of the firms holding a bank loan in Mexico are informal.[3] Therefore, informal firms are directly exposed to bank loan conditions. However, because the informal sector is so large, only 5 percent of informal

[1] This literature neither considers informality nor distinguishes firms by it. For instance, credit supply shocks move the borrowing, investment, and employment of firms (Khwaja and Mian, 2008; Chodorow-Reich, 2014; Amiti and Weinstein, 2018; Huber, 2018; Greenstone et al., 2020), credit constraints hold back firm growth in developing economies (Banerjee and Duflo, 2014), and financial frictions misallocate capital across firms (Buera et al., 2011; Midrigan and Xu, 2014; Moll, 2014).

[2] A Local Labor Market is constructed by the Bank of Mexico using commuting flows. The methodology and information are published in the Bank of Mexico's market analysis.

[3] A firm is informal when it pays no social security contributions for any worker. In the loan data, this means that the firm has no record as an employer in the social security registry. In the census, that the establishment reports zero contributions and positive income. This definition is used throughout the paper. The data section discusses the definition, including the treatment of the intensive margin and of firms with no income or no employees.

firms hold a bank loan, against 13 percent of formal firms.

Second, informal firms face worse loan terms than their formal counterparts. At the median, an informal firm pays an interest rate 0.7 percentage points higher, needs 3 percentage points more collateral on collateralized loans, and obtains loans that are less than a third of those of a formal firm (33.2 thousand dollars smaller). These differences hold among borrowers of similar size, in workers and revenue, and through most of the distribution.[4]

Third, formality status is not permanent and it can change, potentially in response to financial conditions. In Mexico, between the 2019 and 2024 Economic Censuses, 21.8 percent of formal firms became informal and 3.5 percent of informal firms formalized. In contrast to most previous work that treats formality as an absorbing state from which firms cannot transition into informality (Meghir et al., 2015; Ulyssea, 2018), we document two-way formality transitions.[5]

To understand the plausible mechanisms that link formality decisions to the financial sector, we build a parsimonious partial equilibrium model in which firms choose between formality, informality, and exit. Formality carries taxes and higher operating costs but better financial terms (lower marginal loan cost and more pledgeable collateral), while informality avoids taxes and has lower operating costs but faces worse financial terms and a higher exit probability. Therefore, the formality decision trades off the tax cost against the financial benefits of being formal. The model predicts that a relative improvement in the terms or availability of loans for formal firms with respect to the informal ones should increase formalization and decrease formal-to-informal transitions, for firms that need additional loans. Once the output price responds to entry, the model also predicts that exit rises and that only formal entry rises. Formal entrants and deeper loans lower the output price and displace the least productive incumbents of both sectors.

The model predictions are tested in the empirical section, where the challenge to claim causality is endogeneity in the form of reverse causality (formality affecting financial conditions rather than the reverse) and unobserved common characteristics (more productive firms both formalize and borrow). To overcome this challenge, we instrument changes in loans with bank credit supply shocks.

This method disentangles the changes in credit within firms that borrow from several banks into supply (bank fixed effects) and demand (firm fixed effects) (Khwaja and Mian, 2008; Amiti and Weinstein, 2018), leaving the market of interest outside of the estimation sample. Afterwards, it computes the predetermined bank shares in the market of interest. The final instrumental variable is the aggregation of the shift (bank fixed effects), weighted by the share (bank market shares), as in the quasi-experimental Shift-Share framework of

[4]The exception is collateral at the 90th percentile.

[5]The formal-to-informal and informal-to-formal transitions are not an artifact of the definition of informality. They hold under alternative definitions: accounting for the intensive margin, using the social security employer number, or using the tax identifier (Appendix A).

Borusyak et al. (2022).[6] The instrument is strong, with a first-stage F-statistic above 265, and a one-standard-deviation shock roughly doubles five-year market credit.

A credit supply shock reaches both formal and informal firms, but asymmetrically, and increases the financial advantage of being formal. Four years after a one-standard-deviation credit-supply shock, formal firms have 25 percent lower collateral ratios, 19 percent more credit, and interest rates about half a percentage point lower. For informal firms, the four-year credit response is smaller (16 percent), while neither the change in interest rates (-0.1 percentage points) nor in collateral ratios (-13 percent) is statistically significant. The number of firms holding a loan, in contrast, rises similarly in both sectors over five years, by about 3,100 among formal firms (3 percent) and by about 7,900 among informal firms (2.4 percent). In sum, a credit shock increases credit more for formal than for informal firms, interest rates fall only for formal firms, and the reductions in collateral and the increases in borrowers do not differ significantly between sectors.

The model predictions are corroborated. A credit supply shock that improves the financial advantage of being formal increases formalization and reduces formal-to-informal transitions. A ten percent credit supply expansion raises the number of firms that formalize by 850 (0.5 percent) and lowers the number that become informal by 600 (0.4 percent) per five-year window in our sample markets. The results suggest that policies that affect the relative loan advantage of formal firms, from credit guarantees to bank competition, can increase the share of formal firms. The formality-composition results are robust to balance, pre-trend, and placebo tests, the removal of the largest lender and of the largest markets, and a shock-level design with re-identified shocks. Beyond the switching margins, for a ten percent increase in credit supply, the stock of formal firms rises by about 3,500 firms (0.5 percent) and formal entry by about 1,700 firms (0.7 percent) across the markets in our sample. At the same time, exit increases by about 8,300 firms (0.5 percent), mostly among informal firms, which the model attributes to competition in the goods market.

On the real side, the model predicts an increase in formal employment, coming from formal entry and from the relaxation of constraints for formal borrowers. It also predicts a fall in informal value added, coming from the formalization of the most productive informal firms and the exit of the least productive ones. As predicted, but less precisely estimated, a ten percent credit expansion raises formal employment across the sample markets by about 34,000 workers (0.2 percent) and the formal wage bill by about 0.9 million dollars per day (0.2 percent), and lowers value added in the informal sector by about 0.8 billion dollars (0.5 percent), with no change in the formal one. For scale, at the average market, a ten percent credit expansion adds about 62 registered workers and raises the formal wage bill

[6]Inference uses the exposure-robust standard errors of Adão et al. (2019). The design rests on roughly seventeen effective bank shocks. No single bank-window carries more than 17 percent of the exposure weight, and the results survive rebuilding the instrument without the largest lender, at a cost in first-stage strength.

by about 1,600 dollars a day. The model has no prediction for wages, since it is in partial equilibrium and assumes a fixed wage. The estimates are consistent with this assumption: the mean daily wage across employers is essentially unchanged, falling by about one cent (0.1 percent). If anything, the small fall is consistent with the entry of smaller formal employers that pay below the market average.

The paper's contribution lies in showing that bank credit reaches informal firms and moves the formality margin. In particular, that a credit supply expansion changes the composition of firms toward formality, through more formalization and fewer moves to informality. The responses of employment, wages, and value added are smaller and less robust, and we report them as secondary evidence. Moreover, the paper contributes to three literatures.

To the credit supply literature it adds the margin that previous work could not incorporate: informality. This literature identifies credit supply shocks from multi-bank firms (Khwaja and Mian, 2008), from market-sector cells that also include single-bank firms (Degryse et al., 2019), or from granular bank shocks (Amiti and Weinstein, 2018), and finds real effects on employment and output in advanced economies (Chodorow-Reich, 2014; Huber, 2018; Greenstone et al., 2020) and credit constraints on small firms in developing ones (Banerjee and Duflo, 2014). In Mexico, Morais et al. (2019) find strong real effects of credit supply shocks driven by foreign monetary policy, and Gutierrez et al. (2023) estimate that a one-standard-deviation credit supply shock raises formal employment growth by 1.4 percentage points. Using household survey data, Bruhn and Love (2014) argue that opening a bank branch raises informal business ownership, employment, and income.[7] Because administrative records do not capture informal firms and household surveys do not follow the firm, its loans, or its formality status over time, this literature cannot assess whether credit affects the loans of informal firms or their formality decision, whereas our matched data can.

To the informality literature it adds a financial lever that moves the formality margin. In general, informal firms are small, unproductive, and rarely formalize (La Porta and Shleifer, 2014; Ulyssea, 2020). In Mexico, Levy (2018) and Busso et al. (2012) position informality as the main source of misallocation and low productivity, while Alvarez and Ruane (2024) find that its growth had little effect on aggregate productivity. On the modeling side, Ulyssea (2018) separates the intensive and extensive margins of informality, Meghir et al. (2015) and Bobba et al. (2022) integrate the formality choice into job search, and Haanwinckel and Soares (2021) rationalize informality through compensating differentials. The margins studied empirically are mainly the fiscal and administrative costs of formality. In Brazil, lowering taxes formalizes existing informal firms with a low elasticity (Rocha et al.,

[7] In Brazil, formal employment increased after a reform that expanded credit (Fonseca and Van Doornik, 2022). Melcangi (2024) shows that even unconstrained firms increase their cash reserves and reduce employment in a credit crisis.

2018), and stricter enforcement raises formal employment (Almeida and Carneiro, 2012). In Mexico, simplifying registration has modest effects, mainly on wage earners opening a business (Bruhn, 2011; Kaplan et al., 2011), and policies that changed the relative cost of formality, such as non-contributory health insurance or the 2019 border zone tax cut, reduced the formalization of small firms or left formal employment unchanged (Bosch and Campos-Vazquez, 2014; Campos-Vazquez et al., 2020).[8] We explore the financial margin at the market level in a causal way and find that formality status is sensitive to loan terms. Guided by the data, our model allows firms to switch formality status in both directions, whereas most previous models do not allow a formal firm to become informal.[9]

To the literature on firm dynamics with financial frictions it adds quasi-experimental evidence on how financing terms shape the formality choice. Building on Hopenhayn (1992) and on the models that integrate financial frictions into firm dynamics (Evans and Jovanovic, 1989; Cooley and Quadrini, 2001; Gomes, 2001; Albuquerque and Hopenhayn, 2004; Clementi and Hopenhayn, 2006; Buera et al., 2011; Arellano et al., 2012; Midrigan and Xu, 2014; Moll, 2014), the branch of the literature that links formality to finance introduces the channel in which firms formalize to gain access to credit (Straub, 2005), connects the size of the informal economy to the enforcement of credit contracts (Antunes and Cavalcanti, 2007; Quintin, 2008), and calibrates the financial disadvantage of informality in economies with heterogeneous firms to trace it to output, productivity, and tax revenue (D'Erasmo and Boedo, 2012; D'Erasmo, 2016; Franjo et al., 2022; Erosa et al., 2023). The empirical evidence is thinner. Firm evasion is lower where credit information and branches reach further (Beck et al., 2014), industries more dependent on finance have fewer informal jobs where credit is more present (Gandelman and Rasteletti, 2016), and the gains from formalizing in Brazil do not run through access to credit (Fajnzylber et al., 2011). None of this evidence observes the loan terms of informal firms or follows their formality status over time, and the calibrated models typically exclude informal firms from credit (Franjo et al., 2022). Instead of calibrating the formal-informal gap in interest rates, collateral, and loan size, we measure it in the loan registry and change it with quasi-experimental variation, confirming the direction these models predict.

The paper proceeds as follows. Section 2 describes the data and Section 3 the descriptive stylized facts. Section 4 builds the model, Section 5 develops the identification strategy, and Section 6 presents the results. Section 7 concludes.

[8] Field experiments that reimburse the registration cost find similarly small effects (De Mel et al., 2013).

[9] Adom (2025) calibrates a dynamic model for Nigeria in which financial constraints delay the formalization of informal firms, whereas Dix-Carneiro et al. (2026) study the effects of trade on informality in Brazil.

2 Data and Motivating Evidence

This work uses three microdata sources and the Bank of Mexico definition of local labor markets in Mexico.[10] The first data source is the credit registry and includes the universe of bank loans to formal and informal firms, where amounts, costs, and collateral requirements are observed. The second source is the social security records that measure formal employment and wages. The third dataset is the economic census, which is a quinquennial enumeration of formal and informal establishments with fixed facilities, on which we can follow establishments over time.

The credit registry is generated by the National Banking and Securities Commission (CNBV) from the regulatory report R04C that every bank has to file by law.[11] These data include the universe of commercial bank loans to businesses, from 2005 to date, at a monthly frequency. The characteristics of each loan are included, such as the amount, the interest rate, the collateral requirement, and the anonymized identity of the firm. After cleaning, it includes 119 banks and 1.1 million borrowing firms, between 2005 and 2025.

The Mexican Social Security Institute (IMSS) records cover the universe of registered employers and their workers, formal by definition. Wages are daily base salaries, which include cash payments, bonuses, commissions, and payments in kind.[12]

The economic census, conducted by the National Institute of Statistics and Geography (INEGI), surveys establishments with fixed physical facilities every five years, and the analysis uses the 2009, 2014, 2019, and 2024 waves.[13] An establishment in the economic census is informal if it reports zero social security contributions and positive income, formal otherwise. The economic census allows us to identify entry, exit and formality transitions of establishments, as well as their value added.

The 777 local labor markets in the country are constructed by the Bank of Mexico using commuting flows. Credit is observed in 705 of them, and the estimation sample keeps the 556 markets in which the instrument covers at least half of baseline credit, which translates into 1,394 market-window observations across the census windows.

To combine the datasets, we first match the credit registry and the social security

[10] The data was accessed through the Econlab at Banco de Mexico. The EconLab collected and processed the data as part of its effort to promote evidence-based research and foster ties between Banco de Mexico's research staff and the academic community. Inquiries regarding the terms under which the data can be accessed should be directed to econlab@banxico.org.mx.

[11] There is an aggregate, public version of these data available on the commission's website. We used the individual non-public version, accessed through the Bank of Mexico secured environment, EconLab, where all the information is anonymized.

[12] Wages are bottom-coded at the minimum wage and top-coded at twenty five times the Mexico City reference minimum. The unit of observation is a job rather than a person, so a worker holding two registered jobs contributes twice to our employment counts.

[13] The census is representative of localities with more than 2,500 inhabitants and non-agricultural sectors. Activities without fixed physical facilities, such as street vending, are excluded.

records at the firm level. This match makes it possible to identify the formality of the borrower, which is informal if it does not appear in the social security records in the same market and sector, formal otherwise.[14] The matched data contains 3.3 million firms. Notice that the credit-social security merged data includes formal firms with and without loans, and informal firms with loans; however, it does not capture informal firms that do not have a bank loan. To overcome this problem we merge the previous matched datasets to the economic census at the market-level. By doing this we are able to identify all the formal and informal establishments, with or without loans, at the market-level, and to follow them over time.[15]

The formality definition is the same in the census, in the credit registry, and throughout the paper: a firm is formal when it pays social security contributions for at least one worker.[16] This treatment of firms without registered workers as informal is the one used for Mexico by Levy (2018) and Busso et al. (2012). Two clarifications. First, a one-person firm, i.e., a self-employed owner, is formal only if the owner pays social security contributions on his or her own behalf. More broadly, an informal-to-formal transition is a firm that starts paying social security contributions for its first employee or employees, the owner included, and a formal-to-informal transition is a firm that stops paying them for its last ones. Second, we focus on the extensive margin of informality: only 0.7 and 5.3 percent of firms operated on the intensive margin in 2014 and 2019, respectively (Table A.3),[17] so we deliberately exclude this margin from the analysis.

The final dataset excludes agriculture, mining, construction, transport, financial services, and government. Borrowers whose main address is outside of the country are excluded. The monthly credit registry and social security records are made annual by keeping only the January snapshot. We keep only observations with valid locality and activity codes. All the merges are made using anonymized versions of the tax number (RFC), the employer social security registration number (IMSS), and borrower ID. All the data construction and

[14] A formal firm whose registered workers are recorded in another market or sector is classified as informal by this rule. Thus, by classifying a formal firm as informal, the rule, if anything, understates the differences in loan terms by status. The classification enters only the descriptive statistics on loan terms and the annual results by borrower status. The five-year estimates classify establishments with the census.

[15] Figure A.1 compares the informality share of firms with bank loans from two sources: the credit and social security matched registries and the economic census. The registry classification is used only for loan pricing, credit volumes, and collateral. The census classification is used for the establishment counts, the transitions, and the value-added results. The average and state levels differ mainly due to the differences in the definition of establishment and firm.

[16] Establishments with zero contributions and zero income are classified as formal. The positive income requirement keeps establishments that report no income in the reference year out of the informal count, since their zero contributions cannot be read as evasion. Any of the firms with zero income that does operate informally is counted as formal, so the rule, if anything, understates informality.

[17] The low prevalence of the intensive margin in Mexico could partly be explained by the widespread use of outsourcing before the 2021 reform (Estefan et al., 2024; Colonna and Aldeco Leo, 2025).

definitions are collected in Appendix A.

The combination of these three datasets allows us to observe credit, formality, and firm dynamics in the same markets over twenty years, which, to our knowledge, no other dataset allows. The effects of bank loans on informal firms are relatively unexplored, mainly due to the unavailability of the data. In contrast, the uniqueness of these data allows us to uncover the fact that bank loans reach the informal sector, and the conditions under which this happens, which is a central argument of this paper.

3 Descriptive statistics

Table 1: Descriptive statistics on firms, credit, and formality

	Formal	Informal	All
Panel A. Firms and bank credit: Economic Census 2024			
Firms	658,065	4,810,115	5,468,180
Share of firms (percent)	12.0	88.0	100.0
Credit incidence (percent of firms)	13.3	5.2	6.2
Share of borrowers (percent)	26.0	74.0	100.0
Panel B. Credit terms by borrower status: credit registry, 2005-2025			
Unique borrowers	210,895	1,114,533	1,269,287
Authorized credit per borrower, median (thousand USD)	47.9	14.7	17.2
Interest rate, median (percent)	15.5	16.2	16.0
Relationships without collateral (percent)	80.2	88.1	86.2
Collateral to authorized credit, median	0.45	0.48	0.47

Notes: Panel A: 2024 Economic Census. An establishment is informal if it reports zero social security contributions and positive income. Panel B: Credit registry matched with the social security records, January snapshots, 2005 to 2025. A borrower is formal if it appears in the social security records. Firms that borrow under both statuses in different years (56,141) are counted in both columns of the unique borrower row. Authorized credit is the sum of authorized lines across the borrower's banks, in nominal pesos at 19 pesos per dollar. Interest rates are annualized gross rates at the bank relationship level, winsorized at the 1st and 99th percentiles within year. Collateral is observed from 2010; the ratio divides collateral value by the authorized line among relationships with positive collateral. Table A.6 reports the full distributions.

This section provides general descriptive statistics of the matched dataset. Three stylized facts stand out for their relevance and implications. The most important and poorly documented descriptive fact from the data is that informal firms are not outside the banking system, a fact corroborated by two different sources. Table 1 reveals that in 2024, three

quarters of firms with bank loans were informal (census). This means that changes in credit conditions affect informal firms directly.

The second stylized fact is that the loan contracting conditions are harsher for informal firms. At median values, an informal firm pays a 0.7 percentage point higher interest rate, needs 3 percentage points more collateral on collateralized loans, and obtains loans that are less than a third of those of a formal firm (14.7 against 47.9 thousand dollars, 33.2 thousand smaller). The share of relationships without collateral is, however, 7.9 percentage points higher for informal firms.

Table 2: Loan terms of informal relative to formal borrowers of similar size and revenue

	(1)	(2)	(3)	N
Interest rate (percentage points)	3.26***	1.94***	1.40***	2,502,283
	(0.02)	(0.02)	(0.02)	
Log authorized credit	−1.040***	−0.609***	−0.509***	2,091,011
	(0.006)	(0.005)	(0.005)	
Log collateral per peso of credit	0.244***	0.032*	0.045**	378,561
	(0.017)	(0.017)	(0.018)	
No collateral (percentage points)	6.31***	4.31***	3.92***	2,474,489
	(0.09)	(0.09)	(0.09)	
Year fixed effects	Yes	Yes	Yes	
Log revenue and log employees		Yes	Yes	
Market by sector by year fixed effects			Yes	

Notes: Each number reports the coefficient of separate regressions on an indicator variable that equals one if the borrower is informal. The depend variable is the one indicated in the first column. The regression in column (1) includes year fixed effects; column (2) adds log revenue and log employees; column (3) adds market by sector by year fixed effects. Data: the credit registry matched to the social security records, January snapshots, 2005 to 2025. The sample consists of the bank relationships for which the bank reported the borrower's revenue and number of employees, 49 percent of formal and 38 percent of informal relationship years. The interest rate is the annualized gross rate of the relationship, in percentage points. Authorized credit is the sum of the borrower's lines, in logs, one observation per borrower and year. Collateral per peso of authorized credit is in logs and uses collateralized relationships only. No collateral is an indicator for a relationship without posted collateral, in percentage points. The two collateral outcomes start in 2010, when the registry begins to report collateral. The interest rate, revenue, and employees are winsorized at the 1st and 99th percentiles within year. Standard errors in parentheses, clustered by firm. N is the number of observations in columns (1) and (2); column (3) drops singleton cells and uses 2,500,879, 2,089,110, 376,705, and 2,473,166 observations. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

These gaps persist among borrowers of similar size. Table 2 compares formal and informal borrowers of similar revenue and number of employees, as reported by the bank,

in the same market, sector, and year. The results show that after controlling for the local labor market, the sector, the year, and size (measured by the number of employees and revenue), informal borrowers pay interest rates 1.4 percentage points higher, obtain about 40 percent (51 log points) less authorized credit, post 4.6 percent more collateral per peso of credit on collateralized loans, and are 3.9 percentage points more likely to borrow without collateral. Overall, these differences translate into a financial advantage for formal firms.

Table 3: Establishment transitions across censuses

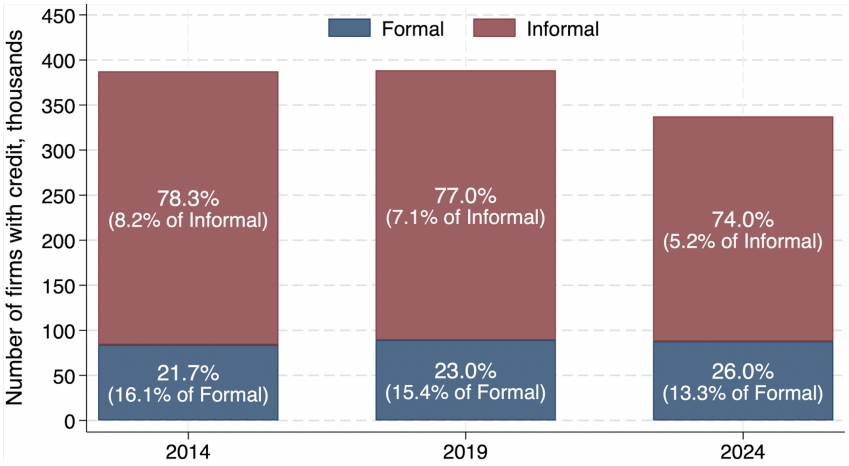
		Destination		
		Formal	Informal	Exit
2014 Origin	Formal	122,784	23,526	56,994
	(% of Formal)	60.4	11.6	28.0
	Informal	263,565	2,232,216	1,531,660
	(% of Informal)	6.5	55.4	38.0
2019 Origin	Formal	324,052	138,041	170,946
	(% of Formal)	51.2	21.8	27.0
	Informal	146,071	2,547,652	1,473,395
	(% of Informal)	3.5	61.1	35.4

Notes: An establishment is informal if it reports zero social security contributions and positive income. Origin rows count the establishments present in the origin census; therefore, the 2019 origin totals include establishments that entered between 2014 and 2019 and are not in the 2014 destination columns. Exit means the establishment is not observed in the following census. Source: Economic Censuses 2014, 2019 and 2024. The table covers the two most recent five-year windows.

The third stylized fact, displayed in Table 3 and Figure 1, is that the formality status of a firm is dynamic and can change in both directions. Between the 2019 and 2024 Economic Censuses, 3.5 percent of informal firms formalized and 21.8 percent of formal firms became informal and the pattern holds in every census window. In contrast to previous literature, where the formal status is an absorbing state,[18] we allow formality transitions to happen in both directions. This suggests that formality is not a one-time decision, is continuously re-optimized and is sensitive to changes in the economic environment, including potentially the bank credit conditions.

[18]See, for example, Meghir et al. (2015); Imbert and Ulyssea (2026); Ulyssea (2018).

Figure 1: Borrowing shares by formality



Notes: An establishment is classified as informal if it reports zero social security contributions and positive income. The main numbers represent the percentage of firms with loans in a given year with that formality status. The numbers in parentheses are the percentage of firms with that formality status in a given year with a loan. The source is the Economic Census (2014, 2019, 2024).

These three stylized facts guide the modeling decisions, the empirical strategy, and the main arguments of the paper. The central idea is simple: certain firms (those almost indifferent between being formal and being informal and that need a bank loan) decide to be formal or not based on the different loan conditions they would face as formal or informal. Therefore, for these firms formality is a financing decision.

Beyond these three facts, five others matter. First, informality is highly prevalent, 88 percent of the 5.5 million establishments are informal. Second, loan incidence is higher among formal firms, 13.3 percent in the formal sector against 5.2 percent in the informal one. Third, the market is the unit of analysis, and the median market contains 2,319 firms, 50 borrowers, and 2.5 million dollars of authorized loans. Fourth, informal firms hold less capital per worker at every age and invest less per worker (Figure A.2). Finally, 85 percent of borrowing firms in 2024 hold loans from a single bank (Table A.9). This last fact is consequential for the empirical strategy and motivates a bank-level robustness check that also includes single-bank firms.

Table 4: Estimation sample: local markets across five-year windows

	Mean	SD	p10	Median	p90	N
Panel A. Pooled levels or flows at the start of the window, per market						
All firms	8,874.6	29,241.7	415	2,319	20,316	1,394
Formal	1,094.3	3,872.7	13	138.5	2,695	1,394
Informal	7,780.3	25,628.2	393	2,159.5	17,436	1,394
Firms with bank credit (census)	572.8	1,793.1	3	119	1,331	1,394
Switchers, informal-to-formal	253.4	879.8	3	38.5	589	996
Switchers, formal-to-informal	229.0	814.4	3	37	537	996
Borrowers (registry)	556.8	3,714.0	4	50.5	919	1,394
Total authorized credit†	431.5	5,578.9	0.1	2.5	195.6	1,394
Collateral ratio	2.21	55.35	0.00	0.04	0.31	1,394
Interest rate (percent)	14.86	4.56	8.88	14.91	20.30	1,394
IMSS workers	26,655	142,866	37	1,316	47,088	1,394
Panel B. Instrument						
Shift-Share instrument (standardized)	0.248	1.197	−0.893	−0.261	2.063	1,394
Coverage	0.921	0.110	0.762	0.967	1.000	1,394

Notes: All statistics pool the 1,394 market window observations (556 markets, windows 2009-2014, 2014-2019, 2019-2024), unweighted. N smaller where undefined. Census is the source for firms by status and firms with bank credit. The registry is the source for borrowers, credit, collateral ratio and interest rate (credit-weighted market averages). † Million dollars at 19 pesos per dollar. Coverage: share of baseline credit held by banks with an estimated shock. The Shift-Share instrument is standardized over all market windows with a constructible instrument (2,356 market windows, Table A.2), not over the estimation sample alone, which is why its mean and standard deviation here differ from zero and one. Coverage is the share of baseline credit held by banks with an estimated shock; the estimation sample keeps market windows with coverage of at least 0.5. Five-year changes in every outcome and the full versions by source are reported in Tables A.7 and A.8.

4 A parsimonious model of formality and finance

The objective of this section is to build a parsimonious model that rationalizes the link between bank credit conditions and firms' formality decisions. The model is a firm model in the tradition of Hopenhayn (1992), with two operating regimes as in Ulyssea (2018) and costly external finance as in Gomes (2001). The empirical facts of Section 3 guide the modeling decisions. For instance, we allow firms to switch their formality status back and forth, as observed in the data, and informal firms to access bank loans under sector-specific loan terms. The model identifies which firms respond when financial conditions change and signs their response. Moreover, it serves as the stepping stone for the empirical strategy.

4.1 Environment

Time is discrete and lasts two periods, $t \in \{1, 2\}$. The economy is populated by two types of risk-neutral entrepreneurs: incumbents and potential entrants. An incumbent entrepreneur owns a firm that belongs to a (formal or informal) sector and holds $k_1 > 0$ installed capital at period one, while a potential entrant entrepreneur is endowed with financial wealth $k_0 \geq 0$ and may or may not start a firm in sector $s \in \{\text{informal}, \text{formal}\}$ in period two. Every entrepreneur discounts period two at factor $\beta \in (0, 1)$ and consumes its dividends and the depreciated capital, which depreciates at rate $\delta \in (0, 1)$. We assume that output is the *numéraire* good, the wage w is exogenous to the firm,[19] and both the output price and the pre-tax wage are the same for formal and informal firms.[20] These assumptions allow us to focus on the formality transitions caused by financial conditions. We relax the assumption of an exogenous output price in Subsection 4.4 to capture the competition in the goods market induced by changes in financial conditions.

An active firm in sector s with capital k and labor ℓ produces

$$y \;=\; \theta\, k^{\alpha} \ell^{\nu}, \qquad \alpha, \nu > 0, \qquad \alpha + \nu < 1, \tag{1}$$

where log productivity combines transitory shocks and a permanent component ω, drawn at birth from $\ln \omega \sim N(0, \sigma_\omega^2)$ and known to the firm, such that follow an AR(1) process[21]

$$\ln \theta_1 \;=\; \ln \omega + \varepsilon_1, \qquad \ln \theta_2 \;=\; \rho \ln \theta_1 + (1 - \rho) \ln \omega + \varepsilon_2, \tag{2}$$

where the innovations $\varepsilon_t \overset{iid}{\sim} N(0, \sigma_\varepsilon^2)$, $\rho \in (0, 1)$, and Φ and ϕ denote the cumulative distribution and density functions of the standardized innovation $\varepsilon_t / \sigma_\varepsilon$, which is standard normal (Φ and ϕ carry no subscript, unlike the financing cost Φ_s and the spread ϕ_s introduced below). As shown below, the expectation of the profitability index $x \equiv \theta^{1/(1-\nu)}$ is the only object of this process that enters the firm's decisions.

A formal firm pays a proportional tax $\tau_y \in [0, 1)$ on revenue and a payroll tax $\tau_\ell \geq 0$ on its wage bill. An informal firm evades both taxes. Therefore, the only advantages of formality are financial, together with a survival advantage introduced below, and the only

[19] Holding the wage fixed is a partial equilibrium simplification. The simplification is supported by the fact that the formal wage barely responds to credit shocks in our annual estimates. Moreover, the aggregate wage movements are absorbed by the year fixed effects of the empirical design.

[20] Equal pre-tax wages and output prices across sectors are a normalization to compare the average firm. A sector-specific wage or price would enter only through Ω_s, and hence through the wedge $T = \Omega_f / \Omega_i$, which the analysis treats as a parameter. Every result below holds for any T that satisfies Assumption 1. We cannot measure wages or prices for informal firms in this data.

[21] The process is common across sectors. We abstract from sector differences in productivity dynamics. In our model a sector-specific innovation variance would only rescale Ω_s in Equation (7) through $\hat{x}$ and would not alter the described mechanisms.

advantages of informality are fiscal savings and lower fixed costs.[22] Static profits are

$$\pi_s(\theta, k) = \max_{\ell \geq 0} (1 - \tau_{y,s})\, \theta k^{\alpha} \ell^{\nu} - (1 + \tau_{\ell,s})\, w\, \ell, \tag{3}$$

with $\tau_{y,f} = \tau_y$, $\tau_{\ell,f} = \tau_\ell$ and $\tau_{y,i} = \tau_{\ell,i} = 0$. The model deliberately excludes the intensive margin of informality, i.e., firms paying payroll taxes on only part of their workforce. This decision is data-driven. As Table A.3 shows, only 0.7 and 5.3 percent of establishments used the intensive margin in 2014 and 2019, respectively.

Wages and taxes are paid within each period out of contemporaneous revenue and wealth. Investment, in contrast, must be paid one period before it becomes productive. Therefore, a firm raises external funds, denoted by $a \geq 0$, to cover any difference between the cost of desired capital and its internal funds.[23] External finance entails a sector-specific cost, so a firm raising an amount a incurs a financing cost of the form

$$\Phi_s(a) \;=\; \phi_s\, a, \qquad \phi_i \;>\; \phi_f \;>\; 0, \tag{4}$$

which is repaid, together with the principal, out of the following period's resources. The parameter ϕ_s is the present value of the spread between the marginal cost of external funds and the risk-free rate in each sector, so the loan enters the firm's problem only through this cost. Notice that debt lasts one period. Firms borrow at the transition stage to install capital, subject to the budget $k \leq n + a$, and no debt is carried further, at the cost of abstracting from leverage dynamics. There is no default in the model, so ϕ_s is a pure intermediation wedge. In the data the spread also contains a default premium, which the model does not separate. The spread in the informal sector is higher because informal firms cannot pledge audited accounts, registered assets, or formal payroll as information collateral, in line with the data. Besides the sector difference in price, there is a difference in the amount of external funds available to a firm, which is limited by a collateral constraint on the capital being installed,

$$a \;\leq\; b_s\, k, \qquad 0 < b_i < b_f < 1. \tag{5}$$

This constraint means that lenders only finance a fraction b_s of the installed capital. For formal firms this fraction is higher because being formal raises pledgeability.

[22] For tractability, we exclude the size-dependent audit cost of informal firms used by Ulyssea (2018). Including it would make formality always optimal for large firms, regardless of financial conditions. For smaller firms, for which this cost is negligible, the exclusion is inconsequential: the financial advantage of formality and the conclusions of the model still hold.

[23] External finance includes a wide range of funding sources. This paper focuses on bank loans and abstracts from the optimal choice of financing instruments. More broadly, the financing-cost function can be interpreted as a reduced-form representation of other sources of costly external finance, provided that their relevant features are captured by the same specification.

Hence, the pair (ϕ_s, b_s) summarizes in reduced form the financial conditions faced by sector s.[24] Both have a counterpart in the data. Loan rates move with ϕ_s, which prices the spread over the risk-free rate. The collateral ratio, i.e. collateral per peso of credit, is the inverse of b_s for a loan secured by the capital being installed. A lender that requires collateral worth $1/b_s$ per peso lent caps the loan at $b_s k$. Hence a lower collateral ratio is a higher b_s. We use this mapping for the direction of the changes, not for their level, since most loans provide no collateral and, when they do, the median collateral covers less than half of the authorized credit. The annual results of Section 6 measure the effects of credit supply shocks on these two data counterparts of the model's financial parameters by sector. Given their importance, all the comparative statics below are taken with respect to these parameters.

Regarding costs, a firm faces entry, $c_s^e \geq 0$, operating, $c_s \geq 0$ with $c_f > c_i \geq 0$, and switching costs, $c_{if}, c_{fi} \geq 0$. Additionally, in period two and before producing, a sector-specific exogenous death probability materializes, $\lambda_s \in [0, 1)$ with $\lambda_i \geq \lambda_f$.[25]

Let $o \in \{e, i, f\}$ denote the origin of a firm: the current sector, informal (i) or formal (f), for an incumbent, and out (e) for a potential entrant. Notice that e also denotes the destination of a firm that exits. Let n_o be the internal funds of a firm before paying the fee of its destination. The internal funds are last period's profit net of the fixed cost plus the resale value of the depreciated capital for an incumbent, and the initial wealth for a potential entrant:

$$n_o = \begin{cases} \pi_o(\theta_1, k_1) - c_o + (1-\delta)k_1, & \text{incumbent of sector } o \in \{i, f\}, \\ k_0, & \text{potential entrant, } o = e. \end{cases} \tag{6}$$

The destination d leaves $n_{o,d} = n_o - \kappa_{od}$ inside the firm, where κ_{od} is the fee of moving from o to d: c_d^e for an entrant, c_{if} or c_{fi} for a switching incumbent, and zero otherwise. An exiting incumbent simply consumes n_o and a non-entrant consumes k_0. In the firm's problem below, n denotes the funds inside the firm, $n = n_{o,d}$.

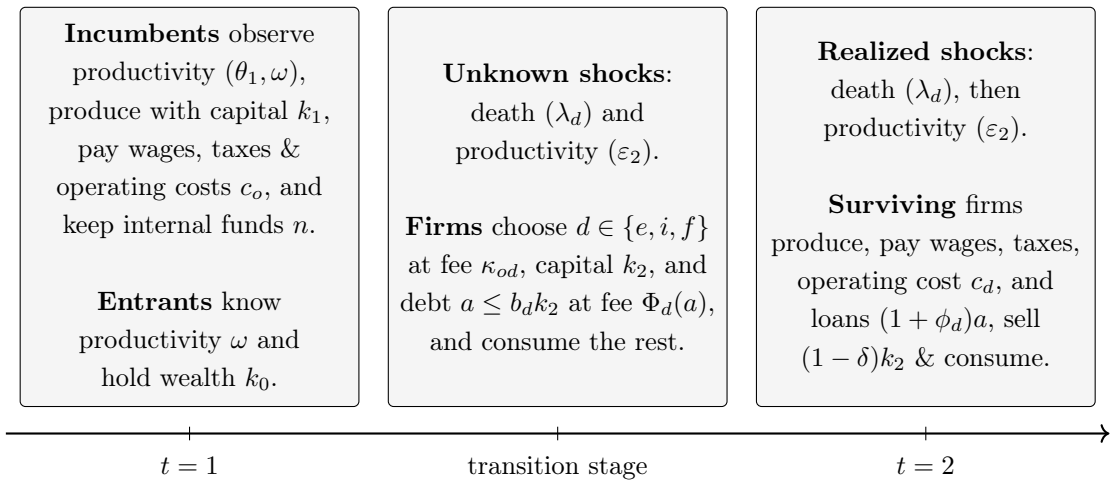
Figure 2 summarizes the timing. In period one, incumbents observe their productivity (θ_1, ω), produce with their installed capital k_1, and pay wages, taxes, and the operating cost c_o. Incumbents therefore enter the transition stage with the internal funds n_o of Equation (6). Potential entrants reach the transition stage with their productivity and

[24] The financial primitives are taken as given rather than derived from deeper frictions: limited enforcement generates endogenous borrowing limits (Albuquerque and Hopenhayn, 2004; Clementi and Hopenhayn, 2006), costly intermediation and default generate endogenous spreads (Cooley and Quadrini, 2001; Arellano et al., 2012), and D'Erasmo and Boedo (2012) embed endogenous credit in an economy with informality. Each such theory delivers the two objects assumed here, a quantity limit and a price wedge, and is consistent with the sector ranking observed in the data. Taking them as primitives, as in Midrigan and Xu (2014) and Moll (2014), keeps the mapping from the quasi-experimental evidence transparent.

[25] The inequality $\lambda_i \geq \lambda_f$ is consistent with the exit rates observed in the data.

wealth (ω, k_0). At the transition stage, where neither knows the productivity innovation $\varepsilon_2 \sim N(0, \sigma_\varepsilon^2)$ nor the death shock (probability λ_d), each firm or potential firm chooses its next period sector, $d \in \{\text{out}, \text{informal}, \text{formal}\}$. Notice that out means exit for an incumbent and staying out for a potential entrant. The sector decision may entail an entry or switching fee, which the firm pays at this stage. Firms that continue in a sector then choose their capital k_2 and external funds $a \leq b_d k_2$ at cost $\Phi_d(a)$, and the funds not invested are consumed. In period two, the death shock is realized. The owners of destroyed firms consume $(1-\delta)k_2$ net of the loan repayment. Surviving firms observe ε_2, produce, pay wages, taxes, the operating cost, c_d, and the loan with its fee, $(1+\phi_d)a$. At the end of period two, the depreciated capital $(1-\delta)k_2$ is sold and consumed. The world ends.

Figure 2: Timing of the two-period model.



4.2 Capital and investment decisions

After solving the static labor problem of the firm, we can re-express its profits as

$$\pi_s(\theta, k) = \Omega_s \, x \, k^\gamma, \tag{7}$$

where $\gamma \equiv \alpha/(1-\nu) \in (0,1)$, $x \equiv \theta^{1/(1-\nu)}$ is the profitability index, and $\Omega_s \equiv (1-\nu)\,(1-\tau_{y,s})^{\frac{1}{1-\nu}} \left[\frac{\nu}{(1+\tau_{\ell,s})w}\right]^{\frac{\nu}{1-\nu}}$ collects the technology, wage, and taxes paid in sector s. This expression is useful because profits are linear in x, so the single sufficient statistic in every decision below is the expectation of the profitability index, $\hat{x} \equiv E[x_2 \mid \mathcal{I}_1]$, where $\mathcal{I}_1$ is the information set at the transition stage. All the statements are formally presented in Appendix B and all proofs are collected in Appendix B.13.

Consider now the problem of capital and external funds at the transition stage of a firm that decided to operate next period in sector s, with internal funds $n \geq 0$ and expected profitability index $\hat{x}$:

$$G_s(\hat{x}, n) \equiv \max_{k \geq 0} \; -k - \Phi_s\big(a\big) + \beta\Big[(1-\lambda_s)\,\Omega_s \hat{x}\, k^\gamma + (1-\delta)k\Big], \quad a = \max\{0, k-n\} \leq b_s k. \tag{8}$$

Therefore, the value of operating in sector s is $U_s(\hat{x}, n) = n + G_s(\hat{x}, n) - \beta(1 - \lambda_s)c_s$. In other words, the entrepreneur consumes the internal funds n minus the resources put into the firm and expects profit net of the fixed cost if it survives plus the depreciated capital resale.

Define the cost of capital as $R^0 \equiv 1 - \beta(1 - \delta)$ when it is acquired with internal funds and as $R_s \equiv R^0 + \phi_s$ when capital is paid for with external funds borrowed under sector s conditions. Then, the corresponding optimal capital targets under these two prices are

$$\bar{k}_s(\hat{x}) = \left[\frac{\gamma\beta(1 - \lambda_s)\Omega_s\hat{x}}{R^0}\right]^{\frac{1}{1-\gamma}} \quad > \quad \tilde{k}_s(\hat{x}) = \left[\frac{\gamma\beta(1 - \lambda_s)\Omega_s\hat{x}}{R_s}\right]^{\frac{1}{1-\gamma}}. \tag{9}$$

Table 5 summarizes the solution to the capital-financing problem and distinguishes four types of firms. Cash-flow sensitivity refers to the marginal response of optimal capital to internal funds, $\partial k_s^*/\partial n$, holding $\hat{x}$ and s fixed.

Table 5: Four regions of optimal capital and external funds

Region	Internal funds n against k_s^*	Capital k_s^*	Borrowing a_s^*	Cash-flow sensitivity
R1: Unconstrained	$n \geq \bar{k}_s$	$\bar{k}_s(\hat{x})$	0	0
R2: Constrained non-borrower	$\tilde{k}_s \leq n < \bar{k}_s$	n	0	1
R3: Borrower	$(1 - b_s)\tilde{k}_s \leq n < \tilde{k}_s$	$\tilde{k}_s(\hat{x})$	$\tilde{k}_s - n$	0
R4: Collateral-constrained	$n < (1 - b_s)\tilde{k}_s$	$\frac{n}{1 - b_s}$	$b_s k_s^*$	$\frac{1}{1 - b_s}$

The first type (R1) is an unconstrained firm with sufficient internal funds to finance its internally financed capital target ($n \geq \bar{k}_s$). This firm does not borrow, and its cash-flow sensitivity is zero. The second type (R2) is a constrained non-borrower. Its internal funds are insufficient to attain the internally financed capital target, but borrowing is not optimal. The firm therefore invests all its internal funds, so its cash-flow sensitivity is one because each additional unit of internal funds increases capital by one unit. The third type (R3) is a borrower that attains the capital target associated with the higher cost of external finance. It borrows the difference between its chosen capital expenditure and its internal funds. Its cash-flow sensitivity is zero because additional internal funds reduce borrowing without changing its capital choice. The fourth type (R4) is a collateral-constrained firm. This firm borrows up to its collateral limit and invests less than its externally financed capital target. Its cash-flow sensitivity is $1/(1 - b_s) > 1$. Each additional unit of internal funds supports $1/(1 - b_s)$ additional units of capital, including $b_s/(1 - b_s)$ units financed through additional borrowing.

The derivatives of the value of operating in sector s, U_s, with respect to the financial parameters (ϕ_s, b_s) measure the value of a change in financial conditions for the firms of each sector (Lemma B.5):

$$\begin{aligned} \frac{\partial U_s}{\partial \phi_s} &= -\, a_s^*(\hat{x}, n) \;\leq\; 0, \\ \frac{\partial U_s}{\partial b_s} &= \mu_s\, k_s^* \;=\; \mathbf{1}_{\mathrm{R4}} \Big[\gamma\beta(1-\lambda_s)\Omega_s \hat{x}\, {k_s^*}^{\gamma-1} - R_s\Big] \frac{n}{(1-b_s)^2} \;\geq\; 0, \end{aligned} \tag{10}$$

where $\mu_s \geq 0$ is the multiplier on the collateral constraint (5). The first inequality is strict exactly when the firm borrows and the second exactly when the collateral constraint binds. Finally, the cross-derivatives are zero: $\partial U_s/\partial \phi_{s'} = \partial U_s/\partial b_{s'} = 0$ for $s' \neq s$.

4.3 Sector decision

Recall that the origin o and the destination d take values in {e = out, i = informal, f = formal}, and that the destination leaves internal funds $n_{o,d} = n_o - \kappa_{od}$ inside the firm, as stated in Equation (6). Then, the firm's formality decision compares

$$\Delta_o(\hat{x}, n_o) \;\equiv\; U_f(\hat{x}, n_o - \kappa_{of}) - U_i(\hat{x}, n_o - \kappa_{oi}). \tag{11}$$

Notice that the empirically relevant comparison is that of borrower firms (R3) at origin and destination, whose desired capital exceeds their internal funds. The statement below is for those firms. Appendix B gives the conditions under which the threshold structure extends to the other branches. For borrowing firms at origin and destination, the following expression holds

$$\Delta_o(\hat{x}, n_o) = A\, \hat{x}^{\frac{1}{1-\gamma}} \;-\; B\, n_o \;-\; C_o, \tag{12}$$

with

$$A \equiv (1-\gamma)\, \gamma^{\frac{\gamma}{1-\gamma}} \left[M_f^{\frac{1}{1-\gamma}} R_f^{-\frac{\gamma}{1-\gamma}} - M_i^{\frac{1}{1-\gamma}} R_i^{-\frac{\gamma}{1-\gamma}} \right], \qquad B \equiv \phi_i - \phi_f \;>\; 0, \tag{13}$$

$$C_o \equiv (1+\phi_f)\, \kappa_{of} - (1+\phi_i)\, \kappa_{oi} + \beta\big[(1-\lambda_f)c_f - (1-\lambda_i)c_i\big]. \tag{14}$$

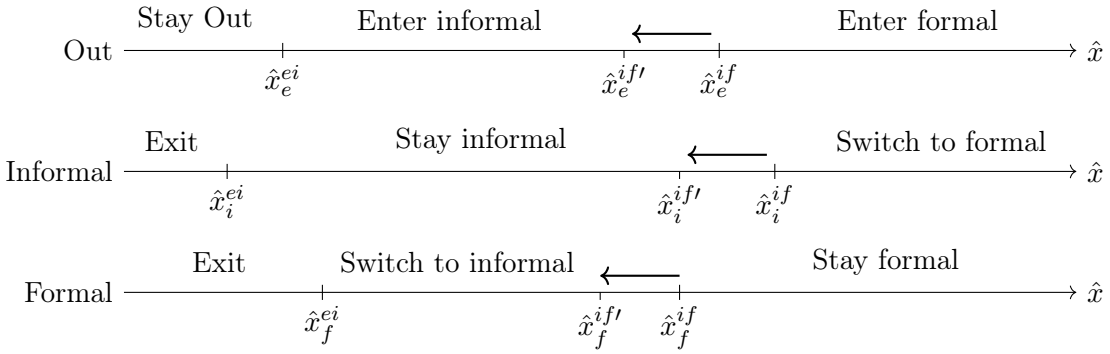
where $M_s \equiv \beta(1-\lambda_s)\Omega_s$. The coefficient A is the difference in the value of growth between formal and informal loan terms. The coefficient B is the spread gap, so each peso of internal funds is worth B more inside an informal firm, which borrows at the higher spread. The constant C_o collects the fee and operating cost differences of formality for a firm of origin o.

Assumptions 1-3 are the conditions under which the following results on sector decisions hold. In particular, Assumption 1 states that, for a firm that has to borrow, the financial advantage of formality outweighs its fiscal cost net of the survival advantage. Assumption 2 rules out the case in which formality is preferred at every productivity level and requires

that entering sector d and switching to d' at once cost more than entering d' directly. Finally, Assumption 3 states that a firm productive enough to be at the formality threshold is productive enough to prefer operating informally to exiting (or to staying out, for an entrant).

More generally, for each origin o and pair of destinations d, d' there exists a unique productivity threshold $\hat{x}_o^{dd'}$ at which the firm is indifferent between d and d'. At a common level of internal funds, the thresholds are ordered on the expected productivity line such that $\hat{x}_i^{ei} \leq \hat{x}_e^{ei} < \hat{x}_f^{ei} < \hat{x}_f^{if} < \hat{x}_e^{if} < \hat{x}_i^{if}$ (Proposition B.7 in Appendix B states the conditions). Figure 3 summarizes these thresholds and previews the comparative statics below.

Figure 3: Sector thresholds by origin state and the effect of better formal credit terms



Notes: Each line orders the destinations of an origin state by expected profitability. Solid ticks are the thresholds at baseline. As an example, dashed ticks and arrows show where the formality thresholds for firms that borrow move after an improvement in formal loan terms that exceeds the improvement in informal terms ($\Delta\phi_f < \Delta\phi_i \leq 0$ or $\Delta b_f > \Delta b_i \geq 0$) once weighted by the firm's borrowing in each sector ($a_f^*|\Delta\phi_f| > a_i^*|\Delta\phi_i|$, Proposition B.10). The analogous condition holds for collateral. The movements of the thresholds between exit and informality are not drawn, for clarity, and no threshold moves for firms that do not borrow at a fixed output price.

So far we have established two separate implications of the model. The first is an anatomy of financially constrained firms based on the firm's expected productivity and its internal funds. The second is an endogenous sector sorting based on the expected productivity of the firm. However, the relevant predictions of the model are in the interaction of these two implications. To see this, notice that for borrowers and collateral-constrained firms the sector productivity thresholds depend on the financial conditions (parameters ϕ_s and b_s), that is, $\hat{x}_o^{dd'}(\phi_i, b_i, \phi_f, b_f)$ for all $o, d, d' \in \{\text{e, i, f}\}$, whereas for unconstrained firms they do not.

Table 6, which states Proposition B.10, reports the sign and size of each threshold's response to the financial conditions for borrowing firms. Let $D_o \equiv Bn_o + C_o > 0$ and let a_s^* be evaluated at the relevant margin. Then, for borrowing firms, the thresholds respond to financial conditions as follows:

Table 6: Response of the sector thresholds to financial conditions

Threshold	$\phi_f \uparrow$	$\phi_i \uparrow$	$b_f \uparrow$	$b_i \uparrow$
Formality margins: $\hat{x}_e^{if}, \hat{x}_i^{if}, \hat{x}_f^{if}$	$\dfrac{(1-\gamma)\,a_f^*}{D_o} > 0$	$-\dfrac{(1-\gamma)\,a_i^*}{D_o} < 0$	$-\dfrac{\mu_f\,k_f^*}{N_o} \leq 0$	$\dfrac{\mu_i\,k_i^*}{N_o} \geq 0$
Informal entry/exit: $\hat{x}_e^{ei}, \hat{x}_i^{ei}$	0	$\dfrac{a_i^*}{G_i'} > 0$	0	$-\dfrac{\mu_i\,k_i^*}{G_i'} \leq 0$
Formal exit: $\hat{x}_f^{ei}$	0	$\dfrac{a_i^*}{G_i'} > 0$	0	$-\dfrac{\mu_i\,k_i^*}{G_i'} \leq 0$

Notes: The first row reports $\partial \ln(\text{threshold})/\partial(\text{parameter})$, with $N_o \equiv \hat{x}\,\partial\Delta_o/\partial\hat{x}$ evaluated at the threshold ($= D_o/(1-\gamma)$ on R3×R3). The second and third rows report level derivatives, with $G_i' \equiv \partial G_i/\partial\hat{x} = M_i\,k_i^{*\gamma}$ the slope of the informal value (with $k_i^* = \tilde{k}_i(\hat{x})$ on R3, $k_i^* = n/(1-b_i)$ on R4) and a_i^* the optimal debt. Both are evaluated at internal funds $n_{o,i}$. The collateral terms $\mu_s k_s^*$ are the derivatives $\partial U_s/\partial b_s$ of Equation (10), positive only when the collateral constraint binds (R4).

The table tells us how the thresholds move with changes in the financial conditions. However, the empirical results measure changes in transitions, not in thresholds. To translate thresholds into transitions we must consider the distribution of productivity shocks. To illustrate this, consider an informal incumbent with permanent productivity ω and internal funds n. Since $\hat{x}$ increases with $\ln\theta_1 = \ln\omega + \varepsilon_1$, the firm formalizes when θ_1 exceeds the productivity cutoff $\bar{\theta}_{if}(\omega, n)$ that corresponds to $\hat{x}_i^{if}(n)$, so the probability of an informal-to-formal switch is

$$\mathbb{P}(i \to f \mid \omega, n) = 1 - \Phi(z_{if})\,, \qquad z_{if} \equiv \frac{\ln\bar{\theta}_{if}(\omega, n) - \ln\omega}{\sigma_\varepsilon}, \tag{15}$$

which is the fraction of informal incumbents with permanent component ω and funds n whose realized period-one productivity θ_1 exceeds the cutoff. Analogous expressions hold for every other transition. For the formalization rate of borrowers,

$$\frac{\partial\,\mathbb{P}(i \to f \mid \omega, n)}{\partial\phi_f} \;=\; -\underbrace{\frac{\phi(z_{if})}{\sigma_\varepsilon}}_{\text{Density}}\;\underbrace{\frac{1-\nu}{\rho}}_{\theta_1\text{ Units}}\;\underbrace{\frac{(1-\gamma)\,a_f^*}{D_i}}_{\partial\ln\hat{x}_i^{if}/\partial\phi_f} \;<\; 0. \tag{16}$$

The symmetric expression in ϕ_i, the collateral analogs, and the other transitions are in Proposition B.11. Two implications of the link between thresholds and transitions matter for the empirical exercise. First, the sign of every transition response is determined by the sign of the threshold shift (for instance, $\partial\ln\hat{x}_i^{if}/\partial\phi_f$ in Equation (16)). That is, a lower threshold increases the transitions into that destination. Second, the response of a transition rate to a change in a loan term is proportional to the density of firms near the threshold, $\phi(z)/\sigma_\varepsilon$, and to the borrowing of those marginal firms near the threshold, a_s^*, as Equation (16) displays. Therefore, self-financed firms do not contribute to the response, and

the elasticities are large where many constrained firms concentrate close to their thresholds. The aggregate response is an incidence-weighted integral concentrated among small, young, low-net-worth firms.[26]

In the model a firm transition is caused by one of two reasons: a productivity shock or a threshold shift due to financial conditions. In the data both are observationally equivalent. Hence, it is important to highlight that the empirical strategy uses the data to discipline the model, since it disentangles productivity changes from changes in financial conditions. The details come in the next section, but for now it is enough to know that the empirical strategy separates changes in credit driven by demand (productivity, ε_1 in the model) from those driven by supply shocks (the loan terms ϕ_s and b_s in the model), and then explains firm transitions caused exclusively by changes in financial conditions.

The model yields testable predictions of the effects of changes in financial conditions on firm transitions. In particular, we test the signs against the results. The model predicts that a relative improvement of the loan terms for formal firms with respect to those faced by informal firms lowers every formality threshold (the arrows of Figure 3).[27] This means that relatively cheaper or more abundant credit in the formal sector, once the improvement in each sector is weighted by the firm's borrowing there, increases formal entry, raises informal-to-formal switching, and reduces formal-to-informal switching. Anticipating the results, the estimates below confirm this prediction. A bank credit supply shock improves the loan terms of formal borrowers by more and, as the model predicts, formalization rises and formal-to-informal transitions fall.

The model so far delivers predictions for the formality thresholds and entry. Realistic predictions about exit, however, require competition in the goods market. The next subsection introduces this competition by letting the output price respond to the quantity produced.

4.4 Competition in the goods market and exit

Until now, the model has assumed an exogenous output price, which rules out competition in the goods market. However, this assumption is particularly relevant for the exit margins, where the credit-driven entry of new firms increases competition and can drive incumbent firms out. This subsection relaxes this assumption by allowing the local output price to

[26] Appendix B.11 works out what the Shift-Share coefficient identifies through the lens of the model.

[27] When only formal terms improve, every formality threshold falls unconditionally. When both sectors improve, the thresholds fall whenever the formal improvement is larger once weighted by the firm's borrowing in each sector, $a_f^*|\Delta\phi_f| > a_i^*|\Delta\phi_i|$ by the first row of Table 6, and analogously for collateral. At the entry and informal-to-formal margins, formal borrowing is larger ($a_f^* > a_i^*$), so a larger improvement in formal loan terms ($|\Delta\phi_f| > |\Delta\phi_i|$) suffices there. In contrast, at the formal-to-informal margin, formal borrowing is not necessarily larger ($a_i^* - a_f^* = \tilde{k}_i - \tilde{k}_f + c_{fi}$, Corollary B.13), so the threshold falls if and only if the formal improvement exceeds the informal one by the factor a_i^*/a_f^*.

respond to the number and size of firms as in Hopenhayn (1992), which introduces goods market competition in reduced form.[28] This extension is a generalization of the baseline case and adds a price channel to the credit effects discussed in the baseline model.

Let the common output price be $P > 0$ with inverse demand $P = \mathcal{P}(Q)$, where Q is the aggregate output after the transition stage and $\mathcal{P}$ is a continuous and strictly decreasing function. The equilibrium price exists and is unique (Proposition B.16). The values of Q and P are known at the transition stage because the idiosyncratic shocks average out with a continuum of firms. Thus, each firm takes P as given and the static profits become $\pi_s(\theta, k; P) = \Omega_s (P\theta)^{\frac{1}{1-\nu}} k^{\gamma}$, where the price enters in the same position as productivity (Lemma B.15). Notice that every object of the baseline model (the capital and borrowing policy, the value of each sector, and all the thresholds above) is unchanged once the profitability index is measured in price-adjusted units, $P^{\frac{1}{1-\nu}}\hat{x}$. Equivalently, in units of the profitability index every threshold j is divided by $P^{\frac{1}{1-\nu}}$, that is, $\hat{x}_j(P) = \hat{x}_j(1)\, P^{-\frac{1}{1-\nu}}$. Therefore, a lower price raises all the thresholds in the same proportion. Proposition B.16 characterizes the equilibrium response.

In this new setting, a credit supply expansion still increases the number of firms, labor, and output, which is the credit effect described in the baseline model. The added price effect of this extension is the decrease in price that follows the rise in output and that in turn raises every threshold in productivity units. For the informal entry, formal exit, and informal exit thresholds the price effect dominates (price thresholds), while for the formal entry, formal-to-informal, and informal-to-formal thresholds the credit effect dominates (credit thresholds) as long as local demand is sufficiently elastic (Proposition B.16). The credit effect is weaker for the price thresholds for two reasons. First, the price thresholds, in contrast to the credit thresholds, depend only on the informal loan terms (Proposition B.10), which the credit shock moves less, as the results show. Second, the credit effect depends on the borrowing intensity of the marginal firm (Lemma B.5), which is smaller at the price thresholds, where marginal firms have lower productivity and less capital and, therefore, borrow less.

To see this formally, let a credit supply shock of intensity ζ lower the spread of sector s by the proportion χ_s^{ϕ} and raise its collateral capacity by the proportion χ_s^{b}. The shock raises the borrower's value of operating in sector s by $g_s \equiv \mathrm{d}U_s/\mathrm{d}\zeta = \chi_s^{\phi}\phi_s a_s^* + \chi_s^{b} b_s \mu_s k_s^* \geq 0$ (Lemma B.5). Notice that, as in the baseline case, the credit effect at the credit thresholds is negative when the formal improvement dominates, $g_f > g_i$ (Corollary B.13). Then, the log shift of threshold j for borrowers is the sum of the credit and price effects, which Table 7 summarizes.

[28] Monopolistic competition in the local market as in Melitz (2003) delivers the same reduced form, i.e. a demand index that falls with the mass of competitors and enters profits like productivity.

Table 7: Credit and price effects on the sector thresholds

Threshold j		Credit effect s_j	Total effect
Credit thresholds:	$\hat{x}_e^{if}, \hat{x}_i^{if}, \hat{x}_f^{if}$	$-\dfrac{g_f - g_i}{N_o} < 0$	$s_j + \psi/(1-\nu) < 0$
Price thresholds:	$\hat{x}_e^{ei}, \hat{x}_i^{ei}, \hat{x}_f^{ei}$	$-\dfrac{g_i}{\hat{x}_j\, G_i'} \leq 0$	$s_j + \psi/(1-\nu) > 0$

Notes: s_j is the log shift of threshold j in index units at a fixed output price. $\psi/(1-\nu)$ is the price effect, where $\psi \equiv -\mathrm{d}\ln P^*/\mathrm{d}\zeta > 0$ is the proportional fall of the equilibrium price (Proposition B.16). N_o and G_i' are the slopes of Table 6. g_i is evaluated at internal funds $n_{o,i}$. The signs of the total effect hold when local demand is sufficiently elastic (Proposition B.16). In productivity units the total shift is $[(1-\nu)s_j + \psi]/\rho$ for incumbents and $(1-\nu)s_j + \psi$ for entrants. Credit thresholds: formal entry, informal-to-formal, and formal-to-informal transitions. Price thresholds: informal entry, informal exit, and formal exit.

The credit effect of Table 7 combines the derivatives of Table 6 with respect to ϕ_s and b_s into a single credit shock ζ. For the total effect, a threshold therefore rises if and only if the price effect dominates the credit effect, i.e., $\psi > (1-\nu)\,|s_j|$. This is the case only for price thresholds. In terms of firm transitions, the three predictions that matter for the empirical results are: (1) the increase in formal and informal exit, (2) no increase in informal entry, and (3) the rise in formal entry and formalizations and the decrease in formal-to-informal transitions due to the dominance of the credit effect.

The model also predicts the change in value added through the inflows and outflows of firms after a credit expansion. The informal sector loses more firms through formalization and exit and gains fewer (no additional entrants due to the price effect and fewer formal-to-informal transitions). These flows cause total informal value added to fall. In the formal sector the effect on value added is ambiguous. There are more inflows (formalization and entry) and fewer formal-to-informal outflows, but also more exits. The effect on total formal value added depends on whether the exit outflow outweighs the inflows. Proposition B.14 also characterizes the productivity of the marginal entrants and their subsequent exit.

5 Empirical strategy: Shift-Share identification with exogenous shocks

We are interested in estimating the impact of credit supply shocks on firm-level transitions (formality status, entry, and exit) and other outcomes. The problem is that OLS estimates would be contaminated by reverse causality and omitted-variable bias: markets with more

informal-to-formal transitions may demand more credit, and an unobserved technology shock may simultaneously raise firm entry and credit volumes.

Consider the ideal experiment where additional funding is randomly allocated across banks, raising the credit supply of the selected banks. Because the selection is random, it is uncorrelated with firms' demand for credit and, hence, with firm transitions and other outcomes. In this setting, identification requires, as an exclusion restriction, that the random allocation affect outcomes only through credit in equilibrium.

We approximate these conditions using a quasi-experimental design, following Amiti and Weinstein (2018) and Greenstone et al. (2020). First, we decompose yearly bank-firm credit growth into demand and supply components. We exclude from the sample the market m that we are interested in (the Leave-One-Out step) to eliminate the mechanical bias that would arise from estimating $\hat{g}_{bt}$ on data that include market m's own observations. Specifically, we estimate

$$\Delta \ln L_{bft} = a_{ft}^{-m} + g_{bt}^{-m} + \epsilon_{bft}^{-m}, \tag{17}$$

where Δ denotes the one-year log change, a_{ft}^{-m} is a firm-year fixed effect that absorbs credit demand, and g_{bt}^{-m} is a bank-year fixed effect that recovers the bank-specific supply shock. The regression is weighted by the total credit authorized from the bank to the firm over the five preceding years and the estimated bank-year effects are normalized to have zero credit-weighted mean within each year. Finally, we add the yearly shocks over the five-year window, $\hat{g}_{bt,5}^{-m} = \sum_{k=0}^{4} \hat{g}_{b,t-k}^{-m}$, to match the five-year differences of the outcome variables.[29]

Notice that the estimation relies on the yearly credit variation across banks within a firm, so that only firms borrowing from two or more banks in a given year contribute to the estimation.[30] In what follows, we maintain the standard assumption that the supply shock identified on multi-bank firms also applies to single-bank firms. Relaxing this assumption yields similar results, as shown in the robustness exercise in Section 6, where we replace firm-year fixed effects with market-sector-year fixed effects as in Degryse et al. (2019). This change allows single-bank firms to also contribute to identification at the cost of potentially absorbing demand shifters into the estimated supply shocks. To validate that the bank effects capture supply rather than demand, Table C.1 shows that bank shocks predict within-firm credit changes in the excluded market (Khwaja and Mian, 2008). A one log point higher bank shock raises the firm's credit from that bank by 0.13 log points relative to its credit from its other banks. At the firm level, the bank shocks affect both single-bank and multi-bank firms, with different intensities (1.1 and 9.2 log points per standard deviation of the share-weighted bank shock), supporting the extension of the

[29] The zero credit-weighted mean of the yearly shocks implies that summing over a year without a valid estimated shock is equivalent to assigning the average shock to that year.

[30] From the pool of borrower firm-years, 13 percent hold loans from more than one bank and 87 percent borrow from a single bank. In terms of loan volume, 66 percent comes from multi-bank firms, while 34 percent comes from single-bank firms.

supply interpretation to single-bank firms.

Second, we construct bank shares at baseline in each market,

$$s_{bm(t-5)} = \frac{L_{bm(t-5)}}{\sum_{b'} L_{b'm(t-5)}}, \tag{18}$$

and define the Leave-One-Out Shift-Share instrument as

$$\hat{Z}_{mt}^{LOO} = \sum_{b \in \mathcal{B}_{mt}} s_{bm(t-5)}\, \hat{g}_{bt,5}^{-m}, \tag{19}$$

where $\mathcal{B}_{mt}$ denotes the set of banks with at least one estimated shock in the window. Following the exogenous-shifter framework of Borusyak et al. (2022), identification requires (i) many uncorrelated shocks with dispersed exposure and (ii) that the shifters $\hat{g}_{bt,5}^{-m}$ be as-good-as-randomly assigned across banks conditional on the included controls.

The as-good-as-random requirement would fail if banks receiving larger supply shocks systematically lend in markets with atypical transition dynamics. This would occur, for example, if an oil boom simultaneously raised firm entry in oil-intensive markets and lending by oil-specialized banks in those same markets. In this case, the estimated coefficient would partly capture sectoral demand rather than credit supply. We address these concerns by including controls that may correlate with both the shifter and the outcome (most directly, the market's sectoral composition of employment and of the lending banks' portfolios) and by conducting a battery of robustness checks (Section 6).

The many-uncorrelated-shocks condition requires exposure to be dispersed across many effectively independent bank shocks. In our context, the Mexican banking system is highly concentrated. Table 8 shows that the effective sample size, $\text{ESS} = 1/\sum_{b,t} w_{bt}^2$, in the pooled period is 16.6 and in the 2014-2019 window a single bank accounts for half of the exposure weight. We argue, nevertheless, that this concentration is not a threat to our identification since the point estimates do not rest on any single bank: no bank-window carries more than 17 percent of the pooled exposure weight. Moreover, the point estimates are robust to dropping the largest bank altogether, at the cost of precision (Appendix C).

The left panel of Figure 4 displays the instrument on a map, where neighboring markets can be subject to very different magnitudes of the shock, which is where the identifying variation lies. Moreover, there is no visible geographical relation between the instrument and the baseline informality share (right panel), a relation that the balance tests below examine more formally.[31]

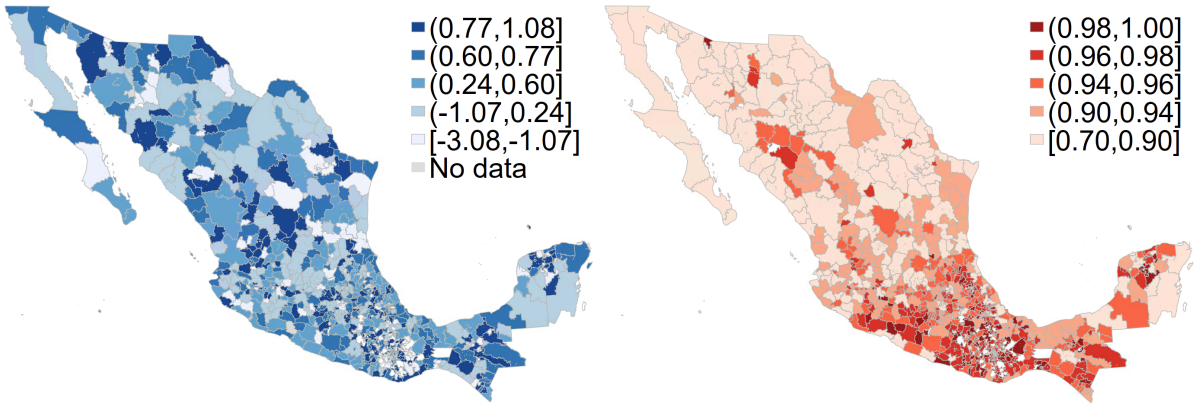
[31] Figure A.3 plots the map of realized credit growth.

Table 8: Descriptive statistics of the standardized bank shocks

	Exposure-weighted			Unweighted					
Window	Mean	Std. dev.	IQR	Mean	Std. dev.	IQR	No. of banks	ESS	W
2009-2014	0.20	0.59	1.12	−0.23	0.56	1.07	26	8.38	0.18
2014-2019	0.14	1.53	3.05	1.07	0.88	1.00	42	3.40	0.51
2019-2024	−0.35	0.39	0.53	−0.64	0.49	0.38	61	7.94	0.29
Pooled	0.00	1.00	1.17	0.00	1.00	1.19	67	16.56	0.17

Notes: The table describes the accumulated five-year bank supply shocks $\hat{\gamma}_{bt,5}$, estimated on the full sample rather than the leave-one-out counterparts used by the instrument, whose only difference is the exclusion of one market at a time. All statistics are computed over the bank-window cells with positive exposure in the estimation sample of the main specification (1,394 market-windows, three five-year windows ending in 2014, 2019, and 2024). Within each block (exposure-weighted and unweighted), shocks are standardized to zero mean and unit variance over those cells using the corresponding weights, so the pooled means and standard deviations equal 0 and 1 by construction. IQR is the difference between the 75th and 25th percentiles. The weight w_{bt} is bank b's share of the baseline credit of the estimation sample markets in window t, normalized to sum to one within each window for the window rows and across all bank-window cells for the pooled row. ESS $= 1/\sum w_{bt}^2$ is the effective sample size of the shocks (Borusyak et al., 2022). W denotes the largest normalized weight. Nb. banks counts banks with an estimated shock and positive exposure in each window. The pooled row reports distinct banks (67 banks, across 129 bank-window cells).

Figure 4: The identifying variation across local markets



Notes: The left panel maps the Shift-Share instrument across local labor markets. The right panel maps the baseline informality share. Gray areas have no credit data.

Once the instrument is built, the main estimating equation that relates local credit growth to outcome variables at the market-level is:

$$\Delta_5 y_{mt} = b\,\Delta_5 \ln(C_{mt}) + X'_{m(t-5)}\Gamma + \eta_t + \varepsilon_{mt}, \tag{20}$$

where $\Delta_5 y_{mt}$ is the five-year change in the outcome (such as the log number of firms by formality status, entry, wages, or value added); $\Delta_5 \ln(C_{mt})$ is the five-year log change in credit in market m, instrumented by $\hat{Z}_{mt}^{LOO}$, which is standardized to unit variance; and η_t captures the year fixed effects. The vector $X_{m(t-5)}$ contains the controls (used in every market-level regression), which are: the sectoral composition of employment and of bank credit in the market (the manufacturing and commerce shares, with services and other activities as the omitted group); the credit-weighted average interest rate; the log numbers of banks and of loans; and the sum of exposure shares, which is the share of credit granted by banks with a valid estimated supply shock, entered as a main effect and interacted with the year fixed effects. This last control accounts for the fact that exposure shares sum to less than one when some banks have no valid estimated supply shock, as prescribed by Borusyak et al. (2022) in the presence of incomplete shares. All controls vary at the market-level and are fixed at their values at the start of each five-year window. Regressions are weighted by the market's baseline credit $C_{m(t-5)}$.

Notice that the specification is estimated in five-year differences, with updating exposure shares that take their value at the start of each window, rather than in levels with market fixed effects. We choose this specification because, when shares update across windows, unit fixed effects do not isolate shock variation over time, whereas the first-difference specification does, and the updating shares keep the first-stage strong (Borusyak et al., 2022).The corresponding first-stage is

$$\Delta_5 \ln(C_{mt}) = \pi \hat{Z}_{mt}^{LOO} + X'_{m(t-5)}\Pi + \eta_t + u_{mt}. \tag{21}$$

Some outcomes are available yearly and therefore we analyze their one-year changes, instead of five-year changes, for instance, credit, interest rates, and collateral requirements. For these cases, the empirical strategy above applies after replacing every five-year index, change, and accumulation with its one-year counterpart.

The coefficient of interest is b in Equation (20), which is the elasticity of the outcome with respect to credit at the market level, identified by exogenous bank-level supply shocks. The interpretation is that a one percent increase in credit supply implies an outcome change of $\hat{b}$ percent if the outcome is in logs, or of $\hat{b}/100$ percentage points if the outcome is measured in percent, as the interest rate. Adão et al. (2019) argue that conventional standard errors over-reject in this setting because market residuals are correlated. That is because markets with similar bank shares are exposed to the same shocks regardless of geographic distance. Following these authors' suggestion, we report exposure-robust standard errors for the IV coefficients. First-stage coefficients are reported with standard errors clustered by market. The first-stage F computed with exposure-robust standard errors remains above 100 in the main estimation samples.

To interpret b through the model, consider the reduced-form regression of the outcome

on the instrument,

$$\Delta_5 y_{mt} = a\,\hat{Z}^{LOO}_{mt} + X'_{m(t-5)}\Gamma_a + \eta_t + e_{mt}, \tag{22}$$

so that $\hat{b} = \hat{a}/\hat{\pi}$ holds with one instrument and the same controls, weights, and sample in Equations (20), (21), and (22). In model terms, the instrument is a shift in the loan terms (ϕ_s, b_s) of the model (Appendix B.11). In the model, the numerator a is a density-weighted sum of threshold shifts, whereas the denominator π is the loan deepening of existing borrowers plus the loans created at entry, eliminated at exit, or repriced at transitions. The numerator is large where the productivity of many firms concentrates near the thresholds. In the data, loan deepening is the dominant component of the denominator, as the annual results show. Importantly, the scale of the instrument enters the numerator and the denominator alike, so the interpretation of b as the effect per unit of credit does not depend on it.

6 The effects of credit supply shocks

The analysis starts with the reduced-form estimates on credit and loan terms at an annual frequency. Figure 5 reports the effect of the standardized instrument on log credit h years before and after the shock. The contemporaneous coefficient is 0.140 (s.e. 0.031). Therefore, a one-unit increase in the standardized instrument, one-standard-deviation of the annual supply shock, raises market credit by 14 log points, about 15 percent, in the year of the shock.[32] This means that for a market with the median baseline credit of the sample, 2.5 million US dollars, the supply of credit increases by roughly 375 thousand dollars. For scale, the median daily formal wage bill across markets is 12.8 thousand dollars, so the increase amounts to about one month of formal wages, or 8 percent of the annual formal wage bill, and to 1.6 percent of the annual revenues that borrowing firms report to their banks (median ratio across market years, 2016-2025).

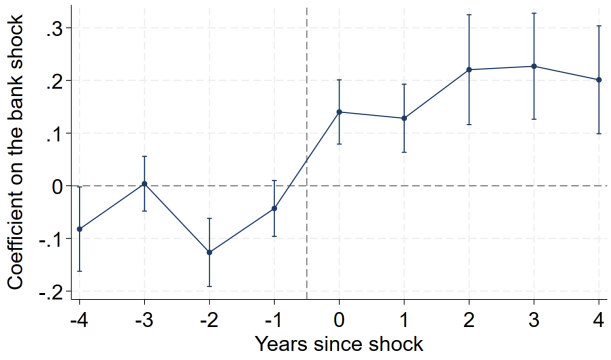
The effect is persistent over time, reaching a coefficient of 0.220 (0.053) two years after the shock and remaining at 0.201 (0.052) four years after. The rise after the first year comes from loan sizes, not from new relationships. Credit per loan keeps rising through the third year, while the numbers of loans and borrowers fall at the time of the shock and are back to zero from the second year on. The model is consistent with the persistence of the effect. The mechanism is that a larger loan increases capital, profits, and next period's internal funds, relaxing the collateral constraint and allowing a larger loan.

The pre-shock coefficients are not all zero, but they have the opposite sign of the responses. This means that credit growth was lower, and rates and collateral higher, in the

[32] one-standard-deviation here is one unit of the standardized annual instrument, computed over all market years with a constructible instrument; the five-year section uses the credit-weighted standard deviation of the instrument within its estimation sample.

years before a positive shock. The shocks thus reached markets whose credit conditions had been tightening, not markets already on the path we attribute to the shock. This same pattern is observed in the five-year placebos for value added. Part of the annual response may therefore be a recovery, so we use these estimates for timing and for the comparison between formal and informal borrowers, which nets out market-wide movements. We take the magnitudes from the five-year design, where the lagged change of the outcome can be controlled for.

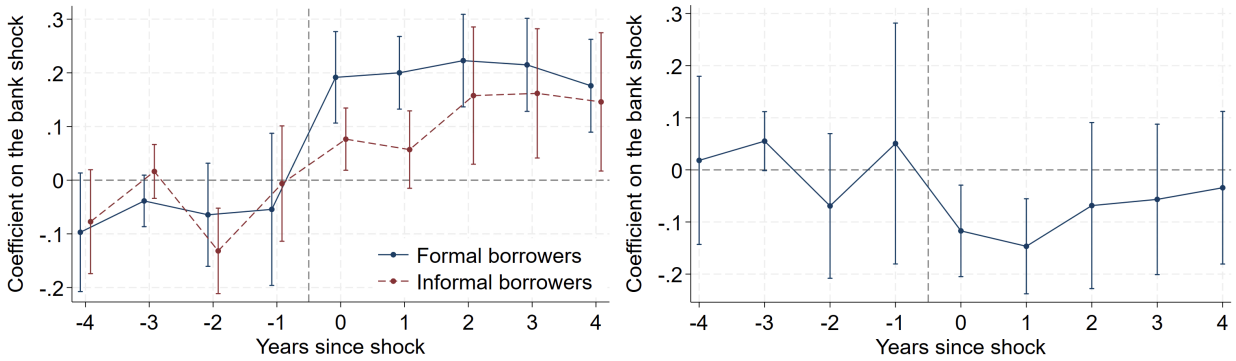
Figure 5: Credit response to a bank credit supply shock



Notes: Reduced-form coefficients on the standardized bank shock from separate annual market-level regressions of log credit at each horizon h. The dependent variable is the change in log credit from $t-1$ to $t+h$ for $h \geq 0$ and from $t+h-1$ to $t+h$ for $h < 0$. Regressions include the annual analogs of the main controls, are weighted by lagged market credit, and use market years with instrument coverage of at least 0.5. 95 percent confidence intervals, s.e. clustered by market. Coefficients and standard errors in Table C.2.

Split by formality, the credit response is larger in the formal sector (Figure 6). The contemporaneous effect on the formal sector is 0.192 log points (*s.e.* 0.043), about 21 percent, which is two and a half times the effect on the informal sector, 0.076 (0.030). This difference, informal minus formal, is negative and significant at the one percent level in the first two years: -0.117 (*s.e.* 0.045) in the year of the shock and -0.147 (0.047) one year after. This differentiated effect supports the idea that a bank supply shock changes the relative financial benefits of formality, plausibly prompting firms to re-optimize their formality decision.

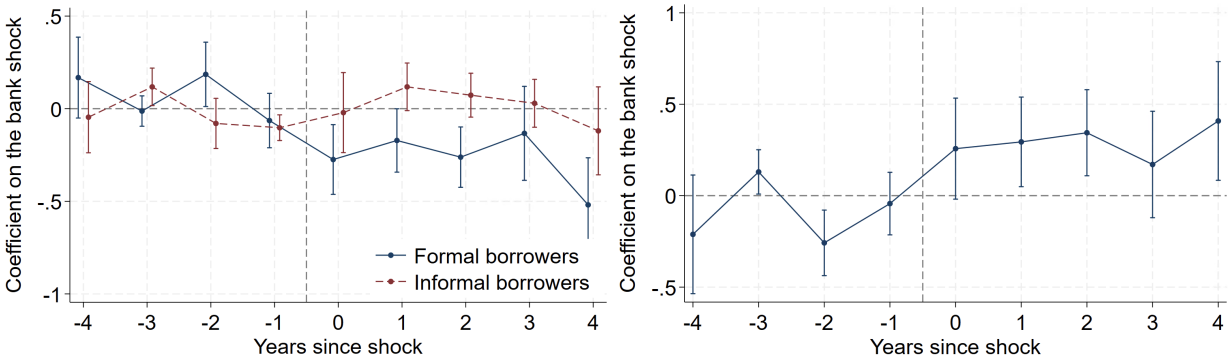
Figure 6: Credit response by formality status of the borrower



Notes: Same specification as Figure 5. Left: log credit to formal and to informal borrowers. Right: difference between the two, informal minus formal, from one regression per horizon on the difference in log credit. 95 percent confidence intervals. Coefficients in Tables C.2 and C.4.

It has been established that a credit supply shock has an asymmetric effect on the volume of credit to formal and informal firms. Now, we investigate whether the supply shock has a differentiated effect on the loan terms, such as interest rates and collateral requirements. Figure 7 displays the response of the lending rate, which falls only for formal firms. For a formal borrower, the rate falls contemporaneously by 0.27 percentage points (s.e. 0.10) and by 0.52 (0.13) four years later. The latter is a 3.4 percent fall of the median formal rate of 15.5 percent. In contrast, for an informal borrower the rate is unresponsive to the shock, -0.02 (0.11) the same year and -0.12 (0.12) four years later. In consequence, the difference, informal minus formal, plotted on the right panel, widens by 0.26 (0.14) percentage points the same year, 0.29 (0.13) one year after and 0.41 (0.17) four years later, more than half of the baseline gap of 0.7 percentage points, significant at the 5 percent level at one, two and four years. Before the shock, the coefficients of this difference have no consistent sign, with two significant values of opposite sign (0.13, s.e. 0.06, at $h = -3$ and -0.26, s.e. 0.09, at $h = -2$); after it, they are positive at every horizon.

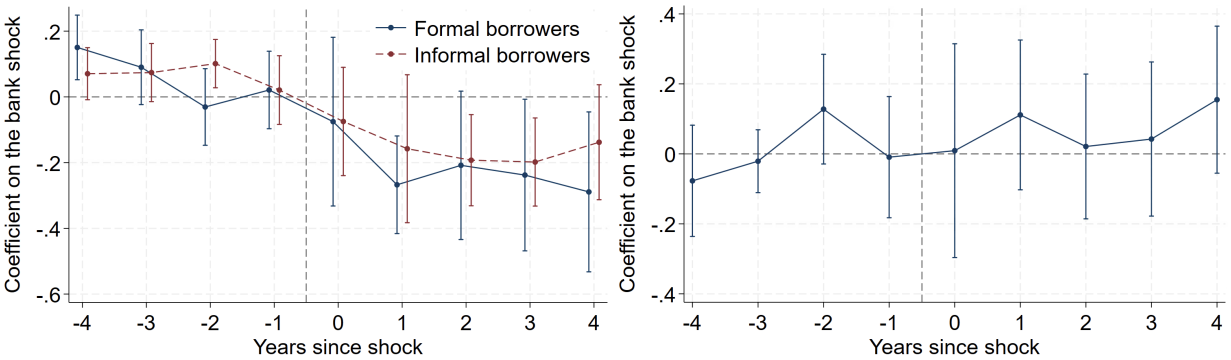
Figure 7: Interest rate response by borrower formality



Notes: Same specification as Figure 5, for the credit-weighted lending rate in percentage points. Left: loans to formal and to informal borrowers in the market. Right: difference between the two, informal minus formal, from one regression per horizon. 95 percent confidence intervals. Coefficients in Tables C.2 and C.4.

Figure 8 repeats the exercise for collateral requirements and finds that the decrease in collateral after the shock is indistinguishable between formal and informal borrowers. For a formal borrower, the collateral per peso of credit falls by roughly a quarter: by 0.27 log points (s.e. 0.08) one year after the shock and by 0.29 (0.12) four years later. For informal borrowers the fall is marginally smaller, 0.19 (0.07) at two years and 0.14 (0.09) at four. However, this difference is not statistically significant, as displayed in the right panel. Hence, the shock relaxes collateral for both groups, with no detectable difference by formality.

Figure 8: Collateral response by borrower formality



Notes: Same specification as Figure 5, for log collateral provided per peso of credit, observed from 2010. Left: loans to formal and to informal borrowers in the market. Right: difference between the two, informal minus formal, from one regression per horizon. 95 percent confidence intervals. Coefficients in Tables C.2 and C.4.

Table C.3 reports additional results. In particular, the credit supply response is concentrated in the intensive margin (credit per loan), with no effects on the number of loans, borrowers, or loans per borrower in either formality group. The responses by formality

therefore come from existing borrowers rather than from borrowers changing formality status. Additionally, the share of loans without collateral falls by less than one percentage point, so the collateral relief is not a shift toward uncollateralized lending. Formal employment rises by about half a percent two to three years after the shock, significant only at the ten percent level, and wages do not respond.

Taken together, these results indicate that, after a positive credit supply shock, formal borrowers obtain more and cheaper credit, whereas informal borrowers absorb less of the shock, with a delay, and at unchanged rates. In contrast, collateral requirements fall for both groups, with no significant difference between them. The evidence of a differentiated effect supports the idea that a bank supply shock improves the relative financial benefit of being formal, plausibly causing a re-optimization of the formality decision, and increasing the number of informal firms formalizing, while reducing the number of formal-to-informal transitions. To test this conjecture, supported by the model, we change the time span of the analysis to five-year changes, the time window available to follow firms over time, and present the main results of the IV estimates on formality and firm outcomes.

The first stage displayed in Table 9 shows that the Shift-Share instrument strongly predicts the amount of standing loans. The coefficient on the standardized instrument is 1.11 (s.e. 0.07). The instrument is standardized over all constructible market windows. Within the estimation sample its credit-weighted standard deviation is 0.63, so a one-standard-deviation shock raises five-year market credit by about 70 log points, roughly a doubling.

The model predicts that an improvement in the financial conditions of formal firms relative to informal ones, such as the one just documented, should increase the number of informal firms formalizing and decrease the formal-to-informal transitions. The transition estimates put the predictions to the test by estimating the response of the four origin-destination possibilities of surviving firms. The results confirm the model predictions: for a ten percent increase in credit supply, 850 more firms formalize (0.5 percent) and 600 fewer leave formality (0.4 percent) per five-year window across the markets in our sample. Additionally, $1,000$ more firms remain formal (0.4percent), while informal stayers do not respond (-0.06 percent, about $1,600$ fewer, not significant). Through the lens of the model, the increase in formalization and the decrease in formal-to-informal transitions can happen simultaneously only if the improvement in loan terms is larger in the formal sector (Corollary B.13). Thus, the estimated transitions and the estimated differences in loan terms are mutually consistent through the mechanism of the model.

Table 9: Shift-Share IV estimates of the effect of bank shocks on firm transitions

	Dependent variable: Five-year log change in the number of firm transitions			
	Informal→Formal	Formal→Informal	Formal→Formal	Informal→Informal
$\widehat{\Delta_5 \ln(C)}$	0.052^{***}	-0.039^{***}	0.040^{***}	-0.006
	(0.017)	(0.014)	(0.009)	(0.004)
Mean	314	290	489	4,601
Obs.	957	958	970	996
Markets	522	527	535	548

Notes: Columns count surviving firms by formality status at the beginning and end of each five-year window (origin→destination), in logs. Two-stage least squares estimates at the market (LLM) × Economic Census five-year window level. Dependent variables are five-year log changes. Markets with a zero count in a window are excluded. The endogenous regressor is the five-year log change in market credit, instrumented by the leave-one-market-out shift-share credit-supply shock, standardized to unit variance. All columns include year fixed effects and the baseline (start-of-window) controls: market sector shares and credit-weighted sector shares of the lending banks' portfolios; the credit-weighted average interest rate; log numbers of banks and of loans; one minus the sum of the bank shares that enter the instrument; and the coverage share (the share of baseline credit granted by banks with a valid estimated supply shock), entered as a main effect and interacted with the year fixed effects. The sample is restricted to market-years with coverage of at least 0.5 (Table C.12 varies this threshold). Mean is the unweighted mean of the dependent variable in levels at the end of the window over the estimation sample; baseline levels are in Table 4. Exposure-robust (AKM) standard errors following Adão et al. (2019) in parentheses. Stars use the corresponding p-values. The first-stage F statistic ranges from 287 to 290 across columns. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Regarding the stock of firms and the exit and entry margins, the estimates in Table 10 imply that credit supply changes the formality composition of firms while barely changing their total number. In particular, a one percent increase in credit supply over the five-year window raises the stock of formal firms by 0.05 (s.e. 0.01) percent, while the informal stock does not move (-0.002, not significant). The total count rises by 0.004 percent (not significant), which is the share-weighted sum of the two changes. At the mean market, where nearly nine out of ten firms are informal, a ten percent credit expansion adds about six formal firms (0.5 percent of the formal stock), with no change in the informal stock. Aggregated over the 556 markets in the estimation sample, the effect amounts to roughly 3,500 additional formal firms.

Table 10: Shift-Share IV estimates of the effect of bank shocks on the number of firms

	Dependent variable: Five-year log change in the number of firms								
	Existing			**Entering**			**Exiting**		
	Total	Formal	Informal	Total	Formal	Informal	Total	Formal	Informal
$\widehat{\Delta_5 \ln(C)}$	0.004	0.050***	−0.002	0.015	0.067***	0.012	0.048***	0.038***	0.055***
	(0.005)	(0.013)	(0.006)	(0.009)	(0.018)	(0.010)	(0.012)	(0.014)	(0.012)
Mean	10,069	1,250	8,823	4,366	458	3,917	3,170	349	2,832
Obs.	1,394	1,389	1,394	996	979	996	996	965	996
Markets	556	552	556	548	538	548	548	531	548

Notes: See Table 9 for specification and inference details. The first-stage F statistic ranges from 266 to 290 across columns. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The mechanism operates mainly through the strong formal entry response (0.067). Additionally, the informal entry does not move (0.012, n.s.) and exit rises in both sectors, with a stronger effect in the informal sector (0.048 total, 0.038 formal, 0.055 informal). At the average market, a ten percent credit expansion is associated with about 15 (roughly 8,300 across the 548 markets in the sample) additional exits, 0.5 percent of exits. Both patterns are what the model predicts once the output price responds to entry (Section 4.4). Exit and informal entry are price thresholds, i.e., thresholds at which the price effect of the credit expansion dominates its credit effect. The intuition is that the formal entrants and the deeper loans of borrowers in both sectors raise output and lower its price, which pushes the least productive incumbents of both sectors out and keeps marginal informal entrants out. Therefore, from the model's perspective, the exit margin is a competition component of the credit expansion, where formal entrants and borrowers with deeper loans displace the least productive incumbents in a Schumpeterian way (Schumpeter, 1942).

A supply-driven ten percent credit expansion adds about 3.1 firms (0.7 percent) to the 458 formal entrants in the average market. Equivalently, a one-standard-deviation expansion in the credit supply shock, which roughly doubles market credit, raises formal entry by about 4.6 percent. For comparison, this effect is of the same order as the five percent increase in registered businesses and start-ups that followed SARE, Mexico's fast-track registration program, which cut the time to register a business from about a month to two days (Bruhn, 2011; Kaplan et al., 2011).

The number of firms with bank credit rises similarly in both sectors (0.042 formal and 0.034 informal per one percent of credit, statistically indistinguishable, Table C.5). Together with the annual evidence on loan terms above, this implies that a credit supply shock reaches a similar share of firms in both sectors. However, formal borrowers get more and cheaper credit, whereas informal borrowers absorb less of the shock, with a delay,

and at unchanged rates. In contrast, the collateral reduction does not differ significantly between sectors.

Table 11: Shift-Share IV estimates of the effect of bank shocks on formal employment and wages, and on value added

	Dependent variable: five-year log change in						
	Formal sector only (IMSS)			All firms (Economic Census)			
	Workers	Wage bill	Mean wage	Avg. VA	Total VA	VA formal	VA informal
$\widehat{\Delta_5 \ln(C)}$	0.020***	0.024**	−0.012***	0.029*	0.032*	0.028	−0.049**
	(0.005)	(0.009)	(0.004)	(0.015)	(0.017)	(0.031)	(0.021)
Mean†	30,967	12.7m	204	1.06m	22.5bn	16.8bn	5.7bn
Obs.	1,386	1,386	1,386	1,394	1,394	1,394	1,394
Markets	552	552	552	556	556	556	556

Notes: See Table 10 for specification and inference details. The first-stage F statistic is 266 in every column. The three labor-market columns are from IMSS records and therefore cover formal (registered) employment: "workers" is the market total of IMSS-registered workers, "wage bill" the market total of daily contribution-base wages, and "mean wage" the market average of employers' mean daily wages, each employer weighted equally. Thus, no accounting identity links the first three columns. The four value-added columns are from the Economic Census and cover all firms: "avg. VA" is average per-firm value added, "total VA" aggregate value added, and the last two columns split total value added by the census formality classifier (informal: establishment with zero social security contributions and positive income). †Means in levels, in MXN (m = million, bn = billion): the mean daily wage is 204 MXN (≈ 10.7 USD); the mean daily wage bill 12.7 million MXN (≈ 0.67 million USD); average per-firm value added 1.06 million MXN (≈ 56,000 USD) and total market value added 22.5 billion MXN (≈ 1.2 billion USD), from census values recorded in thousands of current pesos. All conversions at 19 MXN/USD. Census worker counts, available only in the 2019 and 2024 waves, are reported in Table C.6. The instrument is weak on that single-window subsample. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Finally, we study the effects of the credit supply shock on the labor market and on value added. Table 11 shows that credit expansions raise formal employment and the formal wage bill, while the mean wage across employers falls slightly and value added falls in the informal sector. At the average market, a ten percent credit expansion adds 62 (0.20 percent) workers registered in the social security records, formal by definition. The formal wage bill rises accordingly, by about 1,600 US dollars a day (0.24 percent, the daily pay of about 150 workers at the mean wage), and the mean daily wage across employers, which weights each employer equally, falls marginally, by about one cent on a mean of 10.7 US dollars (0.12 percent), consistent with the entry of smaller formal employers that pay below the market average. Average value added per firm rises by about 160 US dollars a year (0.29 percent) and total value added by 3.8 million US dollars a year (0.32 percent), both

significant at the ten percent level. By formality, informal-sector value added falls by 1.5 million US dollars a year (0.49 percent, significant at the five percent level) while formal-sector value added does not move (2.5 million US dollars, 0.28 percent, not significant).

These responses are consistent by the model. Formal labor demand rises after the credit shock because there are more formal firms and because they install more capital (Lemma B.1). Formal borrowers that deepen their loans hire more, which raises formal employment and the formal wage bill. The formalization of the most productive informal firms and the exit of the least productive ones through the price effect lower informal value added (Subsection 4.4). The model holds the wage fixed, an assumption consistent with the essentially unchanged mean wage.

Two caveats apply to these results, which we therefore read as a complement to the firm-margin results rather than as stand-alone findings. First, the value-added responses are sensitive to market trends: with the outcome's own lagged five-year change as a control, the aggregate and formal coefficients reverse sign while the informal decline strengthens (Table C.8). The result we rely on is therefore the reallocation of production away from the informal sector. Second, the design cannot speak to informal employment: census worker counts exist only for 2019 and 2024, a single window on which the instrument is weak (first-stage F of between 1.9 and 2.4). Table C.6 reports those estimates for completeness.

Two central lessons are drawn from the results presented in this section. First, when banks supply additional credit, they serve formal and informal firms but on better terms for the formal sector. Second, the formality decision of some formal and informal firms can be affected by relative changes in loan terms. Combined, these two lessons imply that the differentiated effect of the bank shocks translates, as predicted by the model, into changes in formality decisions. Firms close to the formality margin and in need of credit respond to the improvement in loan conditions for formal firms, inducing informal firms to formalize and discouraging formal firms from becoming informal.

The next paragraphs examine the robustness of the five-year results on three fronts: the identifying assumptions, the construction of the instrument, and the estimation sample. The main conclusion on formality survives these tests. A full account is given in Appendix C.

On the identifying assumption, we provide evidence that the instrument is as good as randomly assigned conditional on the controls. Following Borusyak et al. (2022) and Goldsmith-Pinkham et al. (2020), in Table C.7 we regress the controls five and ten years before the baseline on the instrument. None of the coefficients is distinguishable from zero at the ten percent level. Moreover, Table C.8 adds the dependent variable's own lagged five-year change as a control, where only aggregate and formal value-added responses reverse sign. Finally, Table C.9 replaces each outcome with its five-year change before the shock window. The stock of firms shows a small positive pre-period correlation, common to both

sectors, whereas the estimated effects load entirely on formal firms.

Wages and value added show negative placebo coefficients: markets that received larger shocks had been losing wages and value added in the previous window, so credit supply expanded where real conditions had deteriorated. Conditional on the controls, the shocks are therefore not as good as randomly assigned with respect to the prior trajectory of these outcomes, and their contemporaneous responses may partly reflect recoveries rather than the credit expansion, which is why no conclusion rests on them. The formality composition result does not share this problem: the placebo coefficients for the stock of firms are small and common to both sectors, and the formal firm response is unchanged when its own lagged change is controlled for.

On the instrument construction, three variants of the instrument leave the results qualitatively unchanged: the conventional version without the Leave-One-Out correction (Table C.10), a renormalization of the instrument by coverage (Table C.11), and coverage thresholds of zero and 0.8 (Table C.12). Additionally, Table C.13 drops the largest lender from the shocks and exposure shares. The point estimates are similar, at the cost of a lower first stage (F of 35), indicating that they do not rely on a single bank.

On the estimation sample, excluding the three largest metropolitan areas (Mexico City, Monterrey and Guadalajara) leaves the formality composition results intact (Table C.14). Lastly, Table C.16 conducts the analysis at the shock (bank) level, instead of at the share (market) level, where the credit supply shocks are re-identified with market-sector-year fixed effects, instead of firm-year fixed effects, as in Degryse et al. (2019). This specification allows single-bank firms to also contribute to the estimation. Although nothing mechanical links this design to the market-level one, it reproduces the composition of the response.

As a whole, the main results are sound and are not the product of any single methodological choice. They survive the balance, pre-trend, and placebo tests; every reconstruction of the instrument; the removal of the largest lender and of the largest markets; and a shock-level design with re-identified shocks. Weights and sample composition matter only for precision and the labor and value-added responses are the margins that depend on the specification, which is why the paper does not lean on them.

7 Conclusion

This paper matches the universe of bank loans, the administrative social security records, and the economic censuses to build a unique dataset that contains formality, bank loans, and firm transitions in the same local labor markets (as defined by community zones) over twenty years. As a result, we can study how bank loans affect informal firms, a relationship barely examined in the literature. With these data, we document three stylized facts. First, informal firms are inside the banking system, in contrast to what is commonly assumed.

Three quarters of firms with credit were informal in 2024. Second, informal firms face worse loan terms than their formal counterparts. Comparing the median borrower of each group, informal firms pay a 0.7 percentage point higher interest rate, need 3 percentage points more collateral on collateralized loans, and obtain 33.2 thousand dollars smaller loans than a formal firm, less than a third of the formal median. Third, firms move regularly between formality and informality, in both directions. Of the establishments observed in 2019, 3.5 percent of informal ones had formalized and 22 percent of formal ones had become informal by 2024. These facts motivate the hypothesis that we formalize in a model and test with a Shift-Share design: firms close to indifference between being formal or informal and in need of a loan change their formality status after a relative change in the terms or availability of loans.

We build a model in which firms choose whether to be formal or informal. Their formality decision trades off the tax cost against the financial benefits of being formal. The model predicts that an improvement in the terms or availability of loans for formal firms, relative to informal ones, increases formalization and decreases formal-to-informal transitions, for firms that need additional loans. Moreover, the model predicts that, once the output price responds to entry, the exit of the least productive incumbents rises in both sectors and informal entry does not increase. The empirical analysis, which uses bank credit supply shocks in a Shift-Share design that exploits the variation in bank presence across local labor markets, is consistent with these predictions. Intuitively, the same nationwide lending change of a bank affects more the markets where that bank has a larger share of credit. Improving the formal advantage in the terms and availability of loans increases formalization and reduces formal-to-informal transitions. Four years after a one-standard-deviation credit-supply shock, formal firms have 25 percent lower collateral ratios, 19 percent more credit, and interest rates about half a percentage point lower. For informal firms, the four-year credit response is smaller (16 percent), while neither the change in interest rates (-0.1 percentage points) nor in collateral ratios (-13 percent) is statistically significant. Moreover, the collateral reduction does not differ significantly between the two sectors. Consistent with the model, a credit supply expansion that improves the formal financial advantage shifts the composition of firms toward formality: a ten percent increase in credit yields about 850 (half a percent) more formalizations and 600 (0.4 percent) fewer returns to informality per five-year window across the markets in our sample. At the average market, the same expansion adds 62 formal workers (0.2 percent) and 0.6 million US dollars a year of formal wage bill (0.2 percent), leaves the mean wage essentially unchanged (a fall of one cent a day, 0.1 percent), and lowers informal value added by 1.5 million US dollars a year (0.5 percent).

The results suggest that policies that affect the relative loan advantage of formal firms, from credit guarantees to bank competition, can shift the share of formal firms. Three

implications follow. First, a financial policy is also a formalization policy. What moves the formality margin in our setting is the difference in loan terms between the two sectors. For instance, a public guarantee program for registered firms would increase the financial advantage of formality and raise formalization, while equalizing loan terms across sectors would have the opposite effect. Second, the lever is sizable. The supply shocks we observe are large, one standard deviation roughly doubles market credit, and the effect on formal entry is of the same order as that of Mexico's registration reform. Third, credit-driven formalization reallocates production. Informal value added falls and exits rise among informal incumbents, so part of the formal increase is reallocation rather than creation. These implications hold for local shocks. The main limitation of this work is the absence of general equilibrium effects in the loan and labor markets. Competition in the goods market is covered only in reduced form, through a local demand curve. For instance, a large credit expansion would move wages and the loan terms offered by every lender. In this context, our market level elasticities, which include any local spillover to firms without bank credit, do not translate one for one into what a national policy would deliver. Addressing these limitations with a model that includes wage and loan market responses is an avenue for future research to better understand the effects of the financial sector on informality.

References

Idossou Marius Adom. A structural analysis of firms' transition from informality to formality in Nigeria. *Journal of Development Economics*, page 103684, 2025.

Rodrigo Adão, Michal Kolesár, and Eduardo Morales. Shift-share designs: Theory and inference. *The Quarterly Journal of Economics*, 134(4):1949–2010, 2019. ISSN 0033-5533. doi: 10.1093/qje/qjz025. URL https://doi.org/10.1093/qje/qjz025.

Rui Albuquerque and Hugo A Hopenhayn. Optimal lending contracts and firm dynamics. *The Review of Economic Studies*, 71(2):285–315, 2004.

Rita Almeida and Pedro Carneiro. Enforcement of labor regulation and informality. *American Economic Journal: Applied Economics*, 4(3):64–89, 2012.

Jorge A Alvarez and Cian Ruane. Informality and aggregate productivity: The case of Mexico. *European Economic Review*, 167:104791, 2024.

Mary Amiti and David E Weinstein. How much do idiosyncratic bank shocks affect investment? Evidence from matched bank-firm loan data. *Journal of Political Economy*, 126 (2):525–587, 2018.

Antonio R Antunes and Tiago V de V Cavalcanti. Start up costs, limited enforcement, and the hidden economy. *European Economic Review*, 51(1):203–224, 2007.

Cristina Arellano, Yan Bai, and Jing Zhang. Firm dynamics and financial development. *Journal of Monetary Economics*, 59(6):533–549, 2012.

Abhijit V Banerjee and Esther Duflo. Do firms want to borrow more? Testing credit constraints using a directed lending program. *Review of Economic Studies*, 81(2):572–607, 2014.

Thorsten Beck, Chen Lin, and Yue Ma. Why do firms evade taxes? The role of information sharing and financial sector outreach. *The Journal of Finance*, 69(2):763–817, 2014.

Matteo Bobba, Luca Flabbi, and Santiago Levy. Labor market search, informality, and schooling investments. *International Economic Review*, 63(1):211–259, 2022.

Stephen Bond and Costas Meghir. Dynamic investment models and the firm's financial policy. *The Review of Economic Studies*, pages 197–222, 1994.

Kirill Borusyak, Peter Hull, and Xavier Jaravel. Quasi-experimental shift-share research designs. *The Review of Economic Studies*, 89(1):181–213, 2022. ISSN 0034-6527. doi: 10.1093/restud/rdab030. URL https://doi.org/10.1093/restud/rdab030.

Mariano Bosch and Raymundo M Campos-Vazquez. The trade-offs of welfare policies in labor markets with informal jobs: The case of the "Seguro Popular" program in Mexico. *American Economic Journal: Economic Policy*, 6(4):71–99, 2014.

Miriam Bruhn. License to sell: the effect of business registration reform on entrepreneurial activity in Mexico. *The Review of Eeconomics and Statistics*, 93(1):382–386, 2011.

Miriam Bruhn and Inessa Love. The real impact of improved access to finance: Evidence from Mexico. *The Journal of Finance*, 69(3):1347–1376, 2014.

Francisco J Buera, Joseph P Kaboski, and Yongseok Shin. Finance and development: A tale of two sectors. *American Economic Review*, 101(5):1964–2002, 2011.

Matías Busso, María Victoria Fazio, and Santiago Levy Algazi. (In) formal and (Un) productive: the productivity costs of excessive informality in Mexico. Technical report, IDB working paper series, 2012.

Raymundo M Campos-Vazquez, Victor Delgado, and Alexis Rodas. The effects of a place-based tax cut and minimum wage increase on labor market outcomes. *IZA Journal of Labor Policy*, 10(1):1–24, 2020.

Gabriel Chodorow-Reich. 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, 2014.

Gian Luca Clementi and Hugo A Hopenhayn. A theory of financing constraints and firm dynamics. *The Quarterly Journal of Economics*, 121(1):229–265, 2006.

Agustina Colonna and Lorenzo Aldeco Leo. Outsourcing, labor regulations and profit-sharing: Evidence from mexico. Technical report, Working Papers, 2025.

Thomas F Cooley and Vincenzo Quadrini. Financial markets and firm dynamics. *American Economic Review*, 91(5):1286–1310, 2001.

Suresh De Mel, David McKenzie, and Christopher Woodruff. The demand for, and consequences of, formalization among informal firms in Sri Lanka. *American Economic Journal: Applied Economics*, 5(2):122–150, 2013.

Hans Degryse, Olivier De Jonghe, Sanja Jakovljević, Klaas Mulier, and Glenn Schepens. Identifying credit supply shocks with bank-firm data: Methods and applications. *Journal of Financial Intermediation*, 40:100813, 2019.

Pablo N D'Erasmo. Access to credit and the size of the formal sector. *Economía*, 16(2): 143–199, 2016.

Pablo N D'Erasmo and Hernan J Moscoso Boedo. Financial structure, informality and development. *Journal of Monetary Economics*, 59(3):286–302, 2012.

Rafael Dix-Carneiro, Pinelopi K Goldberg, Costas Meghir, and Gabriel Ulyssea. Trade and domestic distortions: The case of informality. *Econometrica*, 94(2):573–618, 2026.

Andrés Erosa, Luisa Fuster, and Tomás R Martinez. Public financing with financial frictions and underground economy. *Journal of Monetary Economics*, 135:20–36, 2023.

Alejandro Estefan, Roberto Gerhard, Joseph P Kaboski, Illenin O Kondo, and Wei Qian. Outsourcing policy and worker outcomes: Causal evidence from a mexican ban. Technical report, National Bureau of Economic Research, 2024.

David S Evans and Boyan Jovanovic. An estimated model of entrepreneurial choice under liquidity constraints. *Journal of Political Economy*, 97(4):808–827, 1989.

Pablo Fajnzylber, William F Maloney, and Gabriel V Montes-Rojas. Does formality improve micro-firm performance? Evidence from the Brazilian simples program. *Journal of Development Economics*, 94(2):262–276, 2011.

Julia Fonseca and Bernardus Van Doornik. Financial development and labor market outcomes: Evidence from Brazil. *Journal of Financial Economics*, 143(1):550–568, 2022.

Luis Franjo, Nathalie Pouokam, and Francesco Turino. Financial frictions and firm informality: A general equilibrium perspective. *The Economic Journal*, 132(645):1790–1823, 2022.

Néstor Gandelman and Alejandro Rasteletti. The impact of bank credit on employment formality: Evidence from Uruguay. *Emerging Markets Finance and Trade*, 52(7):1661–1678, 2016.

Paul Goldsmith-Pinkham, Isaac Sorkin, and Henry Swift. Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8):2586–2624, 2020.

Joao F Gomes. Financing investment. *American Economic Review*, 91(5):1263–1285, 2001.

Michael Greenstone, Alexandre Mas, and Hoai-Luu Nguyen. Do credit market shocks affect the real economy? Quasi-experimental evidence from the Great Recession and "normal" economic times. *American Economic Journal: Economic Policy*, 12(1):200–225, 2020.

Emilio Gutierrez, David Jaume, and Martín Tobal. Do credit supply shocks affect employment in middle-income countries? *American Economic Journal: Economic Policy*, 15 (4):1–36, 2023.

Daniel Haanwinckel and Rodrigo R Soares. Workforce composition, productivity, and labour regulations in a compensating differentials theory of informality. *The Review of Economic Studies*, 88(6):2970–3010, 2021.

Hugo A Hopenhayn. Entry, exit, and firm dynamics in long run equilibrium. *Econometrica*, pages 1127–1150, 1992.

Kilian Huber. Disentangling the effects of a banking crisis: Evidence from German firms and counties. *American Economic Review*, 108(3):868–898, 2018.

Clément Imbert and Gabriel Ulyssea. Rural migrants and urban informality: Evidence from Brazil. *Econometrica*, 94(3):911–939, 2026.

David S Kaplan, Eduardo Piedra, and Enrique Seira. Entry regulation and business start-ups: Evidence from Mexico. *Journal of Public Economics*, 95(11-12):1501–1515, 2011.

Asim Ijaz Khwaja and Atif Mian. Tracing the impact of bank liquidity shocks: Evidence from an emerging market. *American Economic Review*, 98(4):1413–1442, 2008.

Rafael La Porta and Andrei Shleifer. Informality and development. *Journal of Economic Perspectives*, 28(3):109–126, 2014.

S Levy. Under-rewarded efforts: The elusive quest for prosperity in Mexico. washington, dc: Inter-american development bank, 2018.

Costas Meghir, Renata Narita, and Jean-Marc Robin. Wages and informality in developing countries. *American Economic Review*, 105(4):1509–1546, 2015.

Davide Melcangi. Firms' precautionary savings and employment during a credit crisis. *American Economic Journal: Macroeconomics*, 16(1):356–386, 2024.

Marc J Melitz. The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6):1695–1725, 2003.

Virgiliu Midrigan and Daniel Yi Xu. Finance and misallocation: Evidence from plant-level data. *American Economic Review*, 104(2):422–458, 2014.

Benjamin Moll. Productivity losses from financial frictions: Can self-financing undo capital misallocation? *American Economic Review*, 104(10):3186–3221, 2014.

Bernardo Morais, José-Luis Peydró, Jessica Roldán-Peña, and Claudia Ruiz-Ortega. The international bank lending channel of monetary policy rates and QE: Credit supply, reach-for-yield, and real effects. *Journal of Finance*, 74(1):55–90, 2019.

Erwan Quintin. Contract enforcement and the size of the informal economy. *Economic Theory*, 37(3):395–416, 2008.

Rudi Rocha, Gabriel Ulyssea, and Laísa Rachter. Do lower taxes reduce informality? Evidence from Brazil. *Journal of Development Economics*, 134:28–49, 2018.

Joseph A Schumpeter. *Capitalism, Socialism and Democracy*. Harper & Brothers, New York, 1942.

Stéphane Straub. Informal sector: The credit market channel. *Journal of Development Economics*, 78(2):299–321, 2005.

Gabriel Ulyssea. Firms, informality, and development: Theory and evidence from Brazil. *American Economic Review*, 108(8):2015–2047, 2018.

Gabriel Ulyssea. Informality: Causes and consequences for development. *Annual Review of Economics*, 12(1):525–546, 2020.

A Data

The credit registry and the IMSS employer records are processed inside Bank of Mexico's EconLab, and the economic census inside INEGI's microdata laboratory. The data at the individual level cannot be exported, so we are able to merge the two sources only at the market-level. Table A.1 lists the definitions.

Table A.1: Variable definitions by source

Object	Source and definition
Formality, registry base	A borrower is formal in a given year when its RFC appears in the IMSS employer records of that year in the same market and sector; otherwise it is informal. Firms in IMSS without a registry record are formal firms without bank credit.
Formality, census	EC, an establishment is informal if it reports zero social security contributions (`J300A` $= 0$) and positive income, formal otherwise.
Employment	Economic Census (EC), total persons engaged including owners and unpaid workers (`H001A`). IMSS, workers registered in January; in the matched base, the larger of the IMSS count and the registry's own count.
Wages	IMSS, January wages, bottom- and top- coded under IMSS rules (minimum wage and 25 times the reference wage).
Value added, wage bill and capital	EC, value added (`A131A`), remunerations (`J000A`) and end-of-year fixed assets at book value (`Q000A`).
Credit incidence, census	EC, the firm reports bank credit (`O511A`).
Credit, collateral and rate	R04C, January snapshot of every loan: authorized line, posted collateral and annual rate in percent, aggregated to borrower, bank, market, sector and year. The collateral field starts in 2010 and is empty in 2025.
Loan rates by status	R04C, unweighted over borrower years, by the IMSS registration of the borrower.
Firm dynamics	EC waves five years apart; exit is absence from the next wave, switching is measured among survivors, and entrants are establishments that were not observed in the previous wave.
Markets and sectors	Localities are mapped to municipalities and then to the public local labor market classification (777 markets). Sectors follow SCIAN; agriculture, mining, construction, transport, financial services and government are excluded.
Units	Pesos as reported in each source; dollar figures in the notes use 19 pesos per dollar.

The credit and social security registries enter as the January file of each year from 2005

to 2025. Loans with missing locality or activity code take the codes from other loans of the same borrower in the same file. Records that still lack a locality, an activity, an authorized line or a total exposure are dropped, as are records with a country code other than Mexico. Activities are mapped to SCIAN, and agriculture, mining, construction, transport, financial services and government are excluded. Localities are mapped to municipalities and municipalities to markets. The loan records are collapsed to one cell per borrower, bank, market, sector and year, with amounts summed, rates averaged and loans counted. In the social security records we drop workers with a missing or zero wage. A borrower with no IMSS employer record in a year is informal that year; IMSS firms without a registry record are formal firms without bank loans.

Table A.9 shows three changes in reporting volume, in 2010, 2017, and 2023. We do not adjust the series across them. These changes have limited consequences for the empirical design, for two reasons. First, the bank shocks are estimated from continuing bank-firm relationships, with firm-year fixed effects, and normalized to a zero credit-weighted mean within each year. This means that the bank shock estimates are immune to changes in the set of reported borrowers and to changes common to all banks. Second, the market-level regressions include year fixed effects, which absorb reporting changes common to all markets.

Table A.2 follows the registry from the raw January files to the estimation sample. Panel A counts institutions, borrowers and borrower years over 2005-2025 after cleaning; Panel B counts the market windows that enter the five-year regressions.

Table A.2: From the raw credit registry to the estimation sample

Panel A. Credit registry					
Banks †					123
Banks after cleaning‡					119
Borrowing firms, registry†					1,141,464
Borrower firms, matched base§					1,269,287
Borrower years, matched base, 2005-2025					5,222,713
Local markets					777
Local markets with registry credit					705
Panel B. Local markets by census year					
	2009	2014	2019	2024	All
Markets with an instrument	501	606	643	606	2,356
With coverage ≥ 0.5	295	408	498	533	1,734
Estimation sample¶	–	398	477	519	1,394

† January files 2005-2025, after dropping records with no locality, activity code, authorized credit line, or total exposure. ‡ We excluded agriculture, mining, construction, transport, financial services, and government sectors, as well as cases where activity or locality was not available or firms outside of the country. § Firms are anonymized and matched by the EconLab using the employer social security registration number (IMSS), tax identifier (RFC), and credit registry identifier. A firm present in several markets or sectors counts once per cell. ¶ Markets with coverage ≥ 0.5 in which the five year changes in outcomes, credit controls are all observed. 2009 can not compute a five years change and therefore is empty. Coverage is the share of a market's baseline credit held by banks with a shock. The 2,356 market years belong to 683 markets, the estimation sample to 556.

Table A.3: Establishment transitions across censuses with the intensive margin

		Destination			
		Formal	Informal	Mix	Exit
2014 Origin	Formal	94,053	20,182	10,990	48,971
		(54.0%)	(11.6%)	(6.3%)	(28.1%)
	Informal	111,773	2,232,216	151,792	1,531,660
		(2.8%)	(55.4%)	(3.8%)	(38.0%)
	Mix	10,522	3,344	7,219	8,023
		(36.1%)	(11.5%)	(24.8%)	(27.6%)
2019 Origin	Formal	197,196	54,785	20,666	108,369
		(51.8%)	(14.4%)	(5.4%)	(28.4%)
	Informal	84,059	2,547,652	62,012	1,473,395
		(2.0%)	(61.1%)	(1.5%)	(35.4%)
	Mix	48,735	83,256	57,455	62,577
		(19.3%)	(33.0%)	(22.8%)	(24.8%)

Note: Percentages are calculated within each origin category. Formal: 100% of workers have social security (SS); Informal: 0%; Mix: strictly between 0% and 100%.

Table A.4: Establishment transitions across censuses with social security number definition

		Destination		
		Formal	Informal	Exit
2019 Origin	Formal	84,146	73,597	93,991
		(33.4%)	(29.2%)	(37.3%)
	Informal	142,684	2,261,091	1,331,140
		(3.8%)	(60.5%)	(35.6%)

Note: Percentages are calculated within each origin category. An establishment is classified as informal if it reports no social security (IMSS) number. The table excludes 813,508 observations (17%) with missing responses to the question about the social security (IMSS) registration number.

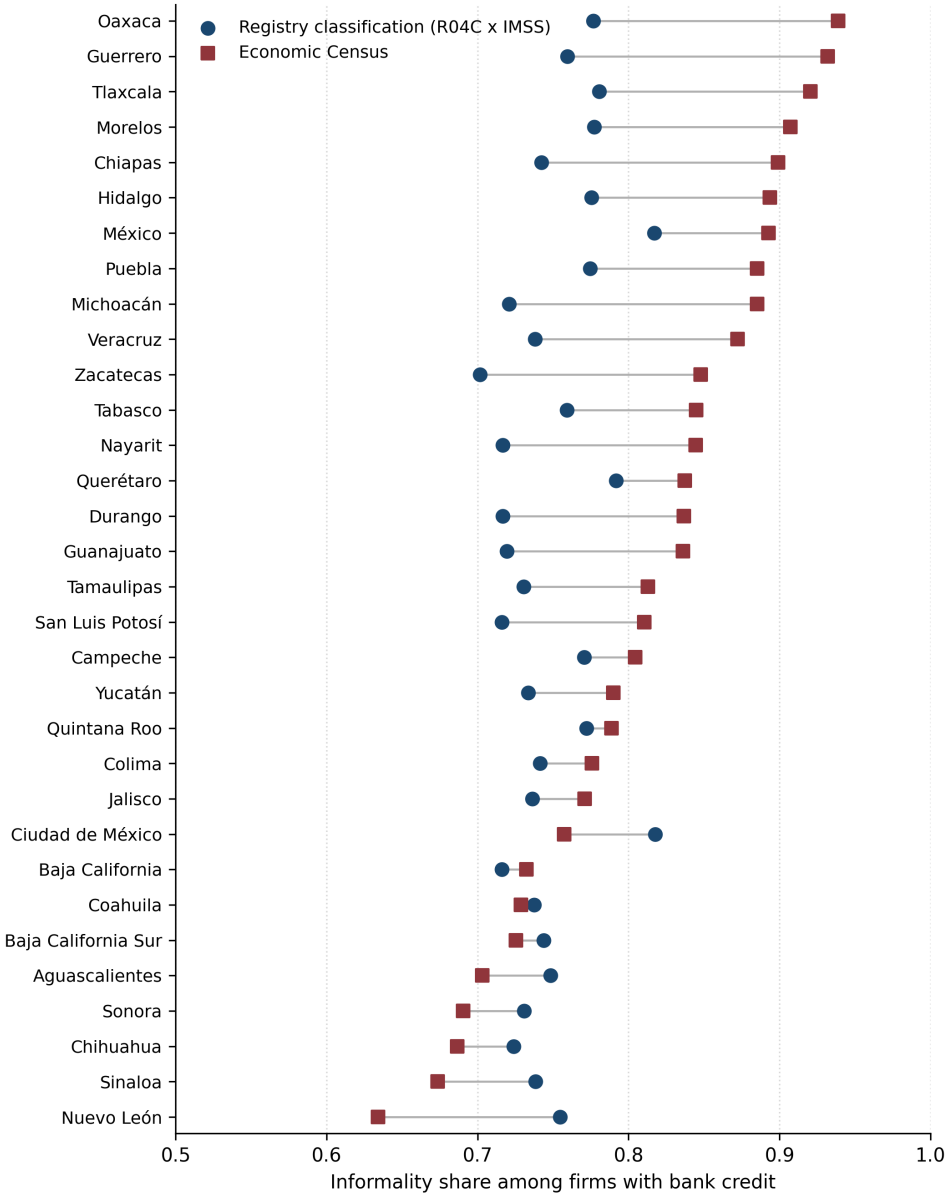
Table A.5: Establishment transitions across censuses with tax administrative number definition

		Destination		
		Formal	Informal	Exit
2019 Origin	Formal	749,911	284,018	475,078
		(49.7%)	(18.8%)	(31.5%)
	Informal	424,885	1,439,976	1,162,103
		(14.0%)	(47.6%)	(38.4%)

Note: Percentages are calculated within each origin category. An establishment is classified as informal if it reports no tax (RFC) number. The table excludes 264,186 observations (5%) with missing responses to the question about the tax (RFC) number.

The census series in Figure A.1 uses the census IMSS registration question, the closest counterpart to the administrative match; the establishment counts and transition tables elsewhere in the paper instead classify an establishment as informal when it reports zero social security payments earning income, and the two measures should not be read as the same object.

Figure A.1: Registry classification of borrower informality against the census



Notes: Share of firms with bank credit classified as informal, by state, sorted by the census value. Registry: share of the 474,237 firms borrowing in the 2018 R04C whose RFC has no IMSS employer registration. This validation universe predates the sector and market restrictions of the analysis sample, whose January 2018 count is 404,842 borrowers (Table A.9) with the same informal share (76.7 against 76.4 percent). Census: establishments not registered with IMSS (question h424a) among those reporting bank credit, 2019 census that includes 2018 information. The classification, used only for loan pricing, credit volumes, and collateral, gets the level of borrower informality approximately right (76 percent overall against an average of 81 percent across states in the census) but compresses the cross-state gradient: it overstates informality where the census puts it lowest (Nuevo León, Sinaloa, Chihuahua, Sonora) and understates it where the census puts it highest (Oaxaca, Guerrero, Tlaxcala, Morelos). The average and state levels differ mainly due to the differences in the definition of establishment and firm.

Table A.6: Descriptive statistics on firms, credit, and formality

	Formal	Informal	All
Panel A. Firms and bank credit: Economic Census 2024			
Firms	658,065	4,810,115	5,468,180
Share of firms (percent)	12.0	88.0	100.0
Firms with bank credit	87,788	249,588	337,376
Credit incidence (percent of firms)	13.3	5.2	6.2
Share of borrowers (percent)	26.0	74.0	100.0
Panel B. Credit terms by borrower status: credit registry, pooled 2005-2025			
Borrower years	1,089,934	4,132,779	5,222,713
Borrowers, 2024	83,334	171,635	254,969
Unique borrowers, 2005-2025	210,895	1,114,533	1,269,287
Banks per borrower, mean	1.35	1.14	1.19
Borrowers with two or more banks (percent)	23.9	10.2	13.1
Credit records per borrower, mean	4.32	2.32	2.74
Authorized credit per borrower (thousand dollars)			
10th percentile	5.9	2.2	2.6
Median	47.9	14.7	17.2
90th percentile	631.8	184.2	263.2
Interest rate (percent)			
10th percentile	9.3	10.0	9.9
Median	15.5	16.2	16.0
90th percentile	24.4	29.0	28.0
Mean	16.1	18.5	17.9
Relationships without collateral (percent, 2010-2024)	80.2	88.1	86.2
Collateral to authorized credit, collateralized relationships			
10th percentile	0.05	0.06	0.05
Median	0.45	0.48	0.47
90th percentile	1.45	1.31	1.36

Notes: The source for Panel A is the 2024 Economic Census, where informal is defined as an establishment that reports zero social security contributions and positive income. Firms that borrow under both statuses in different years (56,141) are counted in both columns of the unique borrower row. Panel B is the January snapshot and the source is the administrative social security records matched with the universe of loans at the firm level. In Panel B, a firm is formal if it appears in the social security records (i.e., it pays social security contributions). Percentiles are computed within each column. Authorized credit is the sum of authorized lines across the borrower's banks, in nominal pesos converted at 19 pesos per dollar. Interest rates are annualized gross rates at the bank relationship level, winsorized at the 1st and 99th percentiles within year. Collateral information exists from 2010. The ratio divides collateral value by the authorized line among relationships with positive collateral.

Table A.7: Estimation sample, census outcomes: establishments by formality and credit

	Mean	SD	p10	Median	p90	N
Panel A. Pooled levels or flows at the start of the window, per market						
All firms	8,874.6	29,241.7	415	2,319	20,316	1,394
Formal	1,094.3	3,872.7	13	138.5	2,695	1,394
Informal	7,780.3	25,628.2	393	2,159.5	17,436	1,394
Firms with bank credit	572.8	1,793.1	3	119	1,331	1,394
Formal	136.6	547.7	0	12	280	1,394
Informal	436.2	1,277.0	1	95	1,035	1,394
Entrants	4,139.1	12,961.3	164	984	9,540	996
Formal	449.9	1,672.9	5	50	973	996
Informal	3,689.1	11,400.7	156	923.5	8,454	996
Exits	3,092.8	10,563.8	104	693	6,861	996
Formal	337.7	1,364.8	3	37	774	996
Informal	2,755.1	9,258.3	99	654.5	5,934	996
Switchers, informal-to-formal	253.4	879.8	3	38.5	589	996
Switchers, formal-to-informal	229.0	814.4	3	37	537	996
Stayers, formal	390.0	1,427.8	2	42	938	996
Stayers, informal	3,820.8	13,773.9	190	1,097	8,542	996
Panel B. Five-year changes, log points						
All firms	0.144	0.134	0.029	0.146	0.270	1,394
Formal	0.127	0.335	−0.229	0.125	0.490	1,389
Informal	0.144	0.138	0.025	0.147	0.275	1,394
Firms with bank credit	0.907	1.718	−0.386	0.087	3.846	1,394
Formal	0.541	1.029	−0.439	0.211	2.095	1,394
Informal	1.089	2.080	−0.424	0.089	4.874	1,394
Entrants	0.076	0.255	−0.187	0.076	0.356	996
Formal	−0.029	0.520	−0.606	−0.016	0.511	979
Informal	0.082	0.259	−0.181	0.081	0.363	996
Exits	0.036	0.256	−0.211	0.044	0.300	996
Formal	−0.051	0.543	−0.693	−0.045	0.655	965
Informal	0.040	0.264	−0.211	0.046	0.306	996
Switchers, informal-to-formal	0.050	0.589	−0.693	0.061	0.690	957
Switchers, formal-to-informal	0.228	0.584	−0.434	0.222	0.916	958
Stayers, formal	0.250	0.407	−0.127	0.221	0.693	970
Stayers, informal	0.203	0.158	0.075	0.196	0.346	996

Notes: The establishments from the Economic Censuses are aggregated to local labor markets. All statistics pool the 1,394 market window observations (556 markets, windows 2009-2014, 2014-2019, 2019-2024), unweighted. N smaller where undefined. Informal: zero social security contributions and positive income, formal otherwise. Entrants, exits, switchers and stayers are measured between consecutive censuses as in Table 3.

Table A.8: Estimation sample, registry and social security records: outcomes and instrument

	Mean	SD	p10	Median	p90	N
Panel A. Pooled levels or flows at the start of the window, per market						
Borrowers	556.8	3,714.0	4	50.5	919	1,394
Formal	107.9	623.2	0	7	177	1,394
Informal	448.9	3,151.9	3	39	747	1,394
Number of loans	1,281.0	9,216.0	4	70	1,517	1,394
Total amount of authorized credit†	431.5	5,578.9	0.1	2.5	195.6	1,394
Average loan size‡	86.1	340.6	10.8	38.3	169.4	1,394
Loans per borrower	2.27	26.64	1.00	1.38	2.22	1,394
Collateral ratio	2.21	55.35	0.00	0.04	0.31	1,394
Interest rate (percent)	14.86	4.56	8.88	14.91	20.30	1,394
Banks in the market	10.1	9.0	2	7	21	1,394
IMSS workers	26,655	142,866	37	1,316	47,088	1,394
IMSS wage bill, daily‡	415.2	2,706.8	0.3	12.8	546.2	1,394
Panel B. Five-year changes, log points						
Borrowers	0.298	0.888	−0.676	0.295	1.424	1,394
Formal	0.558	0.830	−0.340	0.693	1.542	1,130
Informal	0.211	0.928	−0.807	0.182	1.409	1,379
Number of loans	0.468	1.103	−0.646	0.224	2.015	1,394
Total amount of authorized credit	1.032	1.478	−0.509	0.901	2.774	1,394
Average loan size	0.565	1.045	−0.459	0.520	1.663	1,394
Loans per borrower	0.169	0.536	−0.337	0.111	0.761	1,394
Collateral ratio	0.702	2.455	−1.600	0.597	3.128	820
Interest rate§	0.150	5.072	−5.980	0.476	5.971	1,394
IMSS workers	0.216	0.448	−0.095	0.149	0.555	1,386
IMSS wage bill	0.582	0.542	0.180	0.490	1.016	1,386
Panel C. Instrument						
Shift share instrument (standardized)	0.248	1.197	−0.893	−0.261	2.063	1,394
Coverage	0.921	0.110	0.762	0.967	1.000	1,394

Notes: Same sample and layout as Table A.7. Credit registry: as delivered, borrowers are firms with a positive authorized line at a regulated bank in January, formal if IMSS employers that year. Loans are outstanding credit. Collateral ratio and interest rate are credit-weighted market averages. Banks are those lending in that market. IMSS: workers and daily wage bill of IMSS employers in the market. † million, ‡ thousand dollars at 19 pesos per dollar. § percentage points. Instrument standardized over all market windows with coverage. Coverage is the share of baseline credit held by banks with an estimated shock.

Table A.9: Credit registry by year (January snapshot) and borrower status

	Formal borrowers					Informal borrowers				
Year	N	Credit†	Rate	No coll.	2+ banks	N	Credit†	Rate	No coll.	2+ banks
2005	2,516	31.6	16.00	-	5.0	21,428	21.1	16.62	-	2.0
2006	4,000	26.3	17.00	-	6.3	33,379	15.8	18.18	-	2.5
2007	5,717	21.1	16.30	-	6.8	50,164	13.7	18.00	-	3.0
2008	7,975	21.1	16.91	-	8.1	73,658	13.7	19.31	-	3.4
2009	9,568	18.8	18.10	-	8.7	78,573	13.2	19.50	-	3.8
2010	25,513	26.3	13.93	74.6	28.9	226,512	10.5	14.95	89.7	12.8
2011	28,888	26.3	14.07	75.4	29.3	228,728	10.9	15.38	88.9	13.0
2012	32,797	26.3	14.46	74.7	29.7	239,564	11.1	15.80	87.9	13.6
2013	38,581	26.3	14.79	77.4	28.3	259,762	11.6	16.24	88.9	14.0
2014	42,355	29.5	13.76	77.6	28.0	252,567	11.6	15.17	87.9	14.3
2015	46,402	32.6	13.24	80.6	27.8	252,332	12.8	14.59	88.9	14.3
2016	53,007	34.6	13.06	82.3	26.9	249,152	13.2	14.54	88.5	14.3
2017	80,833	48.5	16.00	79.6	21.3	279,487	15.2	17.55	83.7	7.5
2018	94,356	43.7	16.58	85.2	23.8	310,486	15.3	17.63	89.9	8.1
2019	99,452	46.7	17.21	81.8	23.7	299,635	15.7	18.34	88.2	8.2
2020	96,694	52.4	17.00	78.0	22.3	282,031	14.8	18.14	86.7	7.2
2021	90,647	52.6	14.96	76.8	21.2	246,308	15.8	15.85	86.0	6.8
2022	86,710	54.3	14.71	76.3	22.7	256,537	14.5	15.29	89.2	7.0
2023	77,402	67.1	15.69	92.0	23.3	154,974	22.1	15.69	92.6	9.9
2024	83,334	74.7	16.45	79.4	24.1	171,635	22.7	16.65	85.6	10.3
2025	83,187	84.2	16.20	-	25.8	165,867	26.6	15.95	-	12.4

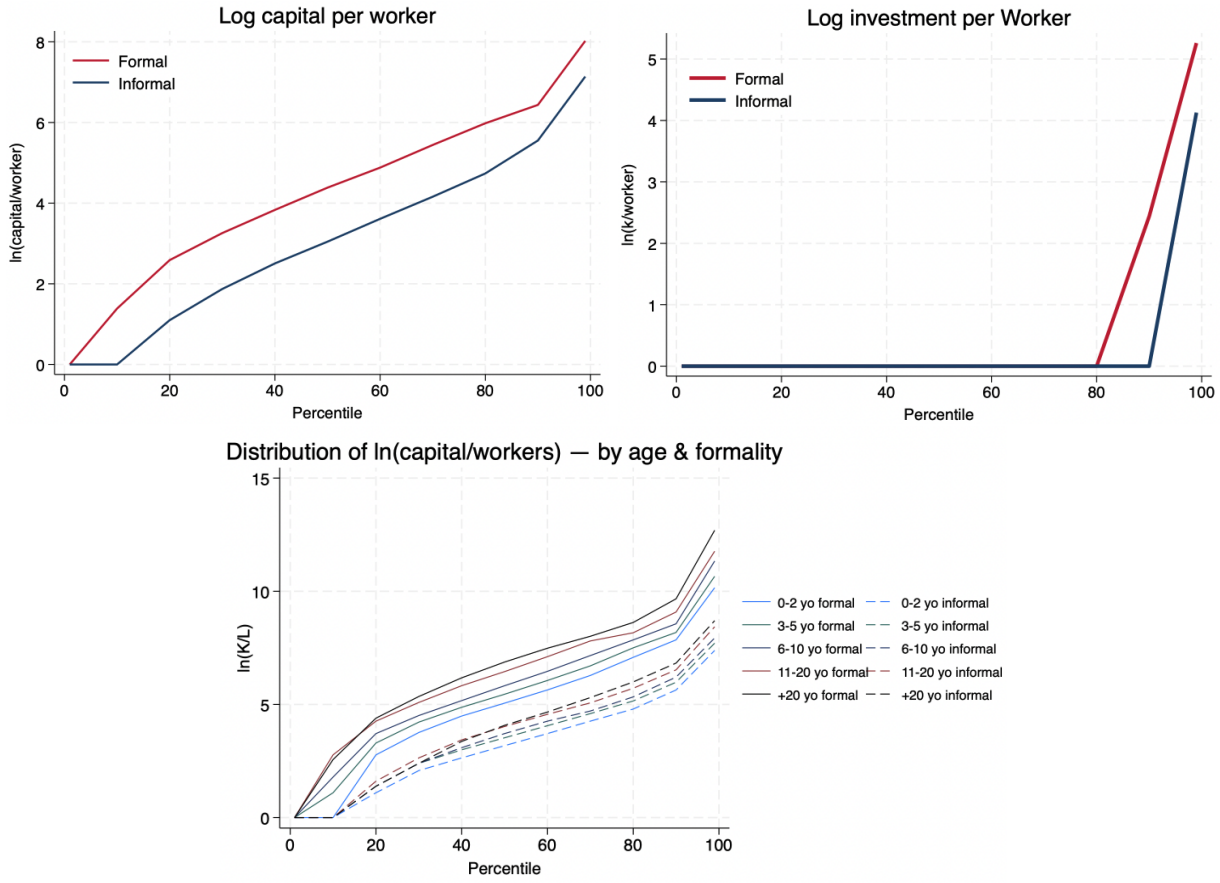
Notes: Credit registry matched to IMSS employer records. A borrower is formal in a given year if it appears as an IMSS employer that year. Borrowers: firms with a positive authorized line. Credit: median borrower authorized credit, †in nominal thousand dollars, at 19 pesos per dollar. Rate: median annualized gross rate at the bank relationship level, percent, winsorized at the 1st and 99th percentiles within year. No coll.: share of relationships without collateral, percent, reported for the years in which the registry carries the collateral field. 2+ banks: share of borrowers with two or more banks, percent.

Table A.10: Estimation sample and all market windows with an instrument

	All market windows			Estimation sample, by window					
				2009-2014		2014-2019		2019-2024	
	Mean	Median	N	Mean	Median	Mean	Median	Mean	Median
Panel A. Pooled levels or flows at the start of the window, per market									
All firms	7,244.4	1,699	1,748	8,980.7	2,675	8,564.4	2,113	9,078.4	2,271
Formal	882.9	88	1,748	1,096.9	159	1,072.9	139	1,112.1	120
Informal	6,361.6	1,605.5	1,748	7,883.8	2,453.5	7,491.4	1,989	7,966.3	2,128
With a loan (census)	465.5	73	1,748	98.4	10.5	787.6	208	739.1	199
Informal→Formal	212.4	25	1,204	-	-	236.1	44	269.2	34
Formal→Informal	191.9	23	1,204	-	-	207.7	30	248.5	42
Borrowers (registry)	470.5	33	1,662	218.7	28	610.9	58	766.4	66
Total credit†	345.3	1.3	1,748	28.9	0.7	663.2	2.2	526.3	6.0
Collateral ratio	1.87	0.03	1,662	-	-	0.11	0.09	5.84	0.10
Interest rate (%)	15.01	14.91	1,662	15.89	15.87	12.36	12.57	16.36	16.25
IMSS workers	22,360	837	1,674	25,687	1,642	25,289	1,195	28,654	1,168
Panel B. Instrument									
Shift share IV (std.)	0.000	−0.449	2,356	−0.225	−0.211	1.788	1.927	−0.804	−0.809
Coverage	0.702	0.938	2,356	0.937	0.991	0.925	0.975	0.905	0.945

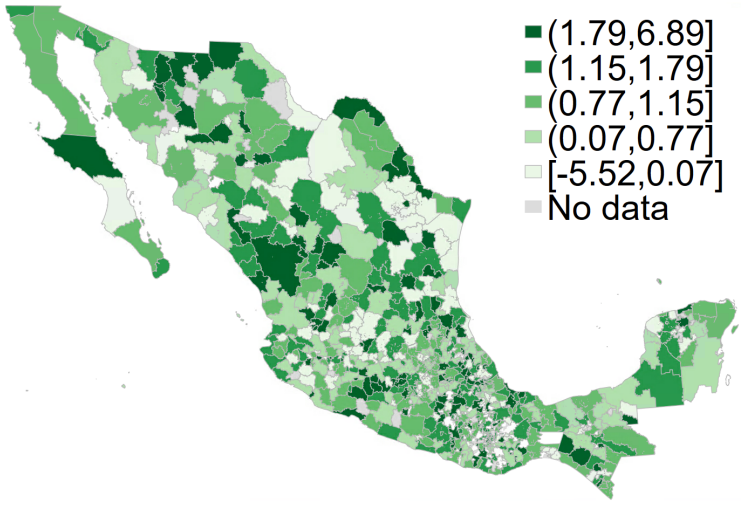
Notes: First three columns pool, unweighted, every market year with a defined instrument (Panel B: 2,356 market years at the four census years, 2009 included, where no window can end) or, in Panel A, every market window whose five-year change is defined, that is, the three windows ending in 2014, 2019 and 2024 with the market in the panel at the baseline (1,748). N per row, smaller where a variable is not observed. Switcher levels at the 2009 baseline would need the 2004 census, so those rows cover two windows. Last six columns are the estimation sample of Table 4 by window, with 398, 477 and 519 markets per window, respectively. Variables and sources as in that table. † Million dollars at 19 pesos per dollar.

Figure A.2: Capital, investment, and debt by formality



Notes: The source is the 2019 Economic Census. Formal versus informal firms.

Figure A.3: Five-year credit growth across markets



Notes: Realized five-year log change in bank credit by local labor market, the endogenous variable in the first stage.

B Model

This appendix collects the formal material behind the two-period model of the main text: the static and dynamic value functions and their envelope properties, the cost ordering behind the threshold structure and the cash-rich region, the master threshold and its four instances, the comparative statics of the thresholds and the transition semi-elasticities, the competition in the goods market, and the closed form of the model-implied estimand. Statements keep the notation of the main text, and proofs are in Appendix B.13.

B.1 Static problem: labor demand and the profit function

Let $\tau_{y,s}$ and $\tau_{\ell,s}$ denote the taxes effectively paid in sector s: $\tau_{y,f} = \tau_y$, $\tau_{\ell,f} = \tau_\ell$, $\tau_{y,i} = \tau_{\ell,i} = 0$. Conditional on operating with state (θ, k), the firm solves $\max_{\ell \geq 0} (1-\tau_{y,s})\theta k^\alpha \ell^\nu - (1+\tau_{\ell,s}) w\, \ell$; substituting the optimal ℓ back into the objective yields the operating-profit function

$$\pi_s(\theta, k) \;=\; (1-\nu)\,(1-\tau_{y,s})\,\theta k^\alpha \left[\frac{\nu\,(1-\tau_{y,s})\,\theta k^\alpha}{(1+\tau_{\ell,s})\,w}\right]^{\frac{\nu}{1-\nu}},$$

which the next lemma rewrites in the separable form used throughout.

Lemma B.1 (Labor demand and profits)**.** *Optimal labor and the operating-profit function are*

$$\ell_s^*(\theta, k) = \left[\frac{\nu(1-\tau_{y,s})\,\theta k^\alpha}{(1+\tau_{\ell,s})w}\right]^{\frac{1}{1-\nu}}, \qquad \pi_s(\theta, k) \;=\; \Omega_s\, x\, k^\gamma, \tag{B.1}$$

where

$$x \equiv \theta^{\frac{1}{1-\nu}}, \qquad \gamma \equiv \frac{\alpha}{1-\nu} \in (0,1), \qquad \Omega_s \equiv (1-\nu)\,(1-\tau_{y,s})^{\frac{1}{1-\nu}} \left[\frac{\nu}{(1+\tau_{\ell,s})w}\right]^{\frac{\nu}{1-\nu}}. \tag{B.2}$$

Labor demand is strictly increasing in θ and k ($\partial\ell^*/\partial\theta > 0$, $\partial\ell^*/\partial k > 0$) and strictly decreasing in the wage and in own-sector taxes ($\partial\ell^*/\partial w < 0$, $\partial\ell_f^*/\partial\tau_y < 0$, $\partial\ell_f^*/\partial\tau_\ell < 0$). The profit function is strictly increasing and strictly concave in k ($\gamma < 1$), and the tax wedge

$$T \;\equiv\; \frac{\Omega_f}{\Omega_i} \;=\; (1-\tau_y)^{\frac{1}{1-\nu}}(1+\tau_\ell)^{-\frac{\nu}{1-\nu}} \;<\; 1 \tag{B.3}$$

summarizes the fiscal cost of formality per unit of the profitability index x.

All proofs are collected in Appendix B.13. The index $x = \theta^{1/(1-\nu)}$ is the natural unit of firm heterogeneity: profits are linear in x, so all expectations below reduce to means of a log-normal random variable. Conditional on the transition-stage information set, define the expected profitability index as

$$\hat{x} \;\equiv\; \mathbb{E}[x_2 \mid \mathcal{I}_1] = \exp\left\{\frac{\rho \ln\theta_1 + (1-\rho)\ln\omega}{1-\nu} + \frac{\sigma_\varepsilon^2}{2(1-\nu)^2}\right\} \quad \text{(incumbents)}, \tag{B.4}$$

and $\hat{x}_e(\omega) = \exp\{\ln\omega/(1-\nu) + \sigma_\varepsilon^2/(2(1-\nu)^2)\}$ for entrants (whose period-2 productivity is $\ln\theta_2 = \ln\omega + \varepsilon_2$). $\hat{x}$ is strictly increasing in θ_1 and in ω; it is the single sufficient statistic for expected productivity in every decision below.

B.2 Investment and financing: formal statements

Lemma B.2 (Capital and borrowing policy)**.** *The solution to Equation* (8) *is*

$$k_s^*(\hat{x}, n) = \begin{cases} \bar{k}_s(\hat{x}), & n \geq \bar{k}_s(\hat{x}) & \textit{(R1: unconstrained)} \\ n, & \tilde{k}_s(\hat{x}) \leq n < \bar{k}_s(\hat{x}) & \textit{(R2: self-financed at the kink)} \\ \tilde{k}_s(\hat{x}), & (1-b_s)\tilde{k}_s(\hat{x}) \leq n < \tilde{k}_s(\hat{x}) & \textit{(R3: interior borrowing)} \\ \dfrac{n}{1-b_s}, & 0 \leq n < (1-b_s)\tilde{k}_s(\hat{x}) & \textit{(R4: collateral-constrained)} \end{cases} \tag{B.5}$$

with borrowing $a_s^* = \max\{0, k_s^* - n\}$*, strictly positive exactly in R3-R4. This policy is continuous, weakly increasing in* $\hat{x}$ *and in* n*, and satisfies the following derivatives*

$$\begin{aligned} &\frac{\partial k_s^*}{\partial \phi_s} = -\frac{\tilde{k}_s}{(1-\gamma)R_s} < 0 \textit{ in R3}, = 0 \textit{ otherwise}; \qquad \frac{\partial k_s^*}{\partial b_s} = \frac{n}{(1-b_s)^2} > 0 \textit{ in R4}, = 0 \textit{ otherwise}; \\ &\frac{\partial k_s^*}{\partial n} = 1 \textit{ in R2}, \frac{1}{1-b_s} > 1 \textit{ in R4}, = 0 \textit{ otherwise}; \quad \frac{\partial k_s^*}{\partial \hat{x}} > 0 \textit{ in R1, R3}, = 0 \textit{ otherwise}. \end{aligned} \tag{B.6}$$

Four observations connect Lemma B.2 to the data. First, investment-cash-flow sensitivity ($\partial k^*/\partial n > 0$) is positive only for constrained firms (R2, R4), and is largest exactly where the collateral constraint binds (R4). This is the firm-level signature of the credit channel that the Shift-Share design picks up at the market-level. Second, the marginal cost of external funds ϕ_s acts on the intensive margin of constrained borrowers (R3), while collateral capacity b_s acts on the most severely constrained firms (R4), typically young, low-n entrants. Therefore, the two instruments of a credit expansion reach different parts of the firm distribution. Third, borrowing incidence $\mathbb{P}(a^* > 0)$ rises with expected productivity and falls with internal funds. Finally, notice that this four-region finance-restriction policy is a generalization of the hierarchy of finance with three endogenous regimes of Bond and Meghir (1994).

Lemma B.3 (Value of operating)**.** *Define* $M_s \equiv \beta(1-\lambda_s)\Omega_s$ *and then let* $\Psi_s(\hat{x}; R) \equiv (1-\gamma)\,\gamma^{\frac{\gamma}{1-\gamma}}\left(M_s\hat{x}\right)^{\frac{1}{1-\gamma}} R^{-\frac{\gamma}{1-\gamma}}$*, which is strictly increasing and strictly convex in* $\hat{x}$ *and strictly*

decreasing in R. Then

$$G_s(\hat{x}, n) = \begin{cases} \Psi_s(\hat{x}; R^0), & R1 \\ M_s \hat{x}\, n^\gamma - R^0 n, & R2 \\ \phi_s\, n + \Psi_s(\hat{x}; R_s), & R3 \\ \phi_s\, n + M_s \hat{x} \left(\frac{n}{1-b_s}\right)^\gamma - R_s \frac{n}{1-b_s}, & R4, \end{cases} \tag{B.7}$$

where G_s is continuous, strictly increasing and convex in $\hat{x}$ (strictly on R1 and R3), and concave and non-decreasing in n. Therefore, the marginal value of internal funds, $q_s \equiv 1 + \partial G_s/\partial n$, satisfies

$$q_s = \begin{cases} 1, & R1 \\ 1 + \gamma M_s \hat{x} n^{\gamma-1} - R^0 \in (1, 1+\phi_s], & R2 \\ 1 + \phi_s, & R3 \\ 1 + \phi_s + \frac{\gamma M_s \hat{x}\, k_s^{*\,\gamma-1} - R_s}{1-b_s} \;>\; 1 + \phi_s, & R4, \end{cases} \tag{B.8}$$

The function q_s is the shadow price of a peso inside the firm. It exceeds one for constrained firms and jumps above $1 + \phi_s$ when collateral binds.

Remark B.4 (Borrowing participation and the fixed cost of borrowing)**.** *Define $G_s(\hat{x}, \tilde{n}) = m$ and the borrowing participation premium as the surplus gain from capital expansion using external funds over staying at the kink in capital without external funding,*

$$\pi_s^B(\hat{x}, n) \;\equiv\; G_s^{\mathrm{R3/R4}}(\hat{x}, n) - \left[M_s \hat{x}\, n^\gamma - R^0 n\right],$$

which equals zero at the R2/R3 boundary and is strictly increasing in $\hat{x}$ beyond it (its $\hat{x}$-derivative is $M_s(k_s^{*\gamma} - n^\gamma) > 0$ whenever $k_s^* > n$). With $\phi_{0s} > 0$ the firm borrows if and only if $\pi_s^B \geq \phi_{0s}$, so the R2/R3 boundary shifts to the unique root of $\pi_s^B = \phi_{0s}$, which is a borrowing-participation cutoff in $\hat{x}$ given n. The population mass beyond the cutoff, i.e. the borrowing incidence, is strictly decreasing in ϕ_{0s} given (ϕ_s, b_s).

The next lemma is the analytical workhorse of the paper. By the envelope theorem, the value of choosing a sector responds (i) to its spread, in proportion to equilibrium borrowing, and (ii) to its collateral capacity, in proportion to the shadow value of pledgeability.

Lemma B.5 (Value of financial conditions)**.** *Since G_s, and hence U_s, are differentiable with respect to (ϕ_s, b_s), except at region boundaries, the derivatives exist and are given by Equation (10) of the main text.*

B.3 Formality threshold

Assumption 1 (Finance-driven formality). *Define* $T_\lambda \equiv T \cdot \dfrac{1-\lambda_f}{1-\lambda_i}$ *as the tax wedge adjusted by the survival advantage of formality. Then*

$$\left(\frac{R_f}{R_i}\right)^{\gamma} \;<\; T_\lambda. \tag{B.9}$$

Assumption 1 states that, for a firm that has to borrow, the financing advantage of formality ($\left(\frac{R_f}{R_i}\right)^{\gamma} < 1$) outweighs its fiscal cost net of the survival advantage ($T_\lambda \lesseqgtr 1$). This assumption is equivalent to $A > 0$, and it is the exact sense in which formality is, in this model, access to capital. Absent a financing or survival advantage ($R_f = R_i$, $\lambda_f = \lambda_i$), $T < 1$ implies $A < 0$ and no borrower would ever formalize.

Assumption 2 (Interior thresholds and cost ordering). *(a)* $C_o + B\, n_o > 0$, *(b) The switching and entry fees satisfy* $0 < c_{if}, c_{fi}$, *the ordering* $0 \le c_i^e \le c_f^e$, *and*

$$(1+\phi_f)\, c_f^e - (1+\phi_i)\, c_i^e \;<\; (1+\phi_f)\, c_{if}, \tag{B.10}$$

$$(1+\phi_f)\, c_f^e + (1+\phi_i)\, c_{fi} \;>\; (1+\phi_i)\, c_i^e. \tag{B.11}$$

Part (a) of Assumption 2 rules out the case in which formality is preferred to informality at every productivity level, and thus guarantees an interior formality threshold. Part (b) states that entering sector d and switching to d' at once costs more than entering d' directly. Notice that conditions (B.10)-(B.11) are exactly equivalent (by Lemma B.8) to the cost ordering $C_{f\text{-inc}} < C_{\text{e}} < C_{i\text{-inc}}$.

Assumption 3 (Viable informal continuation). *For each origin* o *and internal funds* n_o, *the informal destination is strictly preferred to exit or to staying out at the formality cutoff, that is,* $\hat{x}_o^{ei}(n_o) < \hat{x}_o^{if}(n_o)$. *Equivalently,* $G_i(\hat{x}_o^{if}, n_o - \kappa_{oi}) > \kappa_{oi} + \beta(1-\lambda_i)c_i$, *so that the quasi-rents of informal operation at the cutoff cover the informal fee and operating cost.*

Assumption 3 means that a firm productive enough to be at the formal-informal indifference threshold is productive enough to prefer operating informally over quitting. If it failed, firms would skip the informal option altogether.

Proposition B.6 (Formality threshold). *Under Assumptions 1-3 and* $c_i > 0$, *for each origin* o *and internal funds* n_o, *the R3×R3 difference* (12) *is strictly increasing and strictly convex in* $\hat{x}$ *and has a unique zero at*

$$\hat{x}_o^F(n_o) \;=\; \left[\frac{B\, n_o + C_o}{A}\right]^{1-\gamma}. \tag{B.12}$$

This is the operative formality threshold whenever both destinations are in R3 at $\hat{x}_o^F(n_o)$, *the empirically relevant case. Appendix B.12 gives the other branches and the self-consistency*

rule that selects among them. Thus, the firm chooses the formal sector if and only if $\hat{x} \geq \hat{x}_o^{\mathrm{F}}(n_o)$. *The threshold is strictly increasing in internal funds, so*

$$\frac{\partial \hat{x}_o^F}{\partial n_o} = (1-\gamma)\,\frac{B}{A}\left[\frac{Bn_o + C_o}{A}\right]^{-\gamma} > 0, \tag{B.13}$$

that is, conditional on expected productivity, firms with more internal funds are less likely to formalize, because self-finance substitutes for the financial advantage of formality.

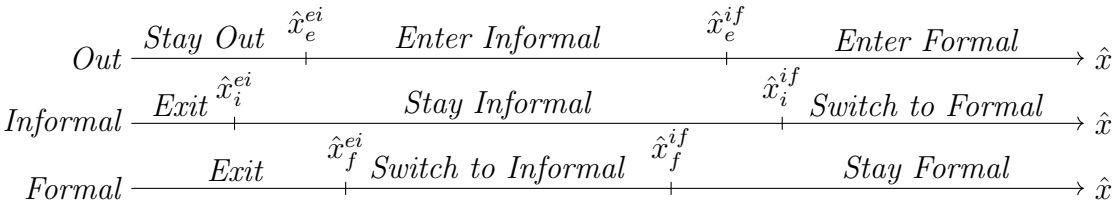
Proposition B.6 formalizes the sorting logic of the model. The firms that select into formality are those whose growth needs are large relative to their internal resources, i.e. high $\hat{x}$, low n. These are precisely the firms for which a bank credit supply shock is decisive, which is why the Shift-Share elasticities of Section 6 are informative about (ϕ_s, b_s).

For unconstrained, cash-rich firms the finance advantage is irrelevant, and with taxes as the only non-financial difference a sufficiently wealthy firm prefers informality; a strong enough survival advantage of formality ($T_\lambda \geq 1$) or the size-dependent audit cost of the general model removes this artifact, and the sizable mass of formal-to-informal transitions in the data suggests that for many firms the net non-financial advantage of formality is indeed weak (Proposition B.11).

B.4 The full transition map and its comparative statics

Each exit and entry margin reduces to one comparison, continuation in the informal sector (the relevant alternative at low $\hat{x}$, by the ordering below) against consuming the outside value, and each is characterized by a condition of the form $G_s(\hat{x}, \tilde{n}) = m$, where $\tilde{n}$ are the funds carried into the continuation and m collects the operating and transition costs. Solving branch by branch of the financing regions yields a single master threshold whose four instances are exit from either sector and entry into either sector. Appendix B.12 develops the complete taxonomy. Two of its properties matter below: a firm without internal funds exits or fails to enter at any productivity (financial exit), and a constrained firm whose quasi-rents on internal funds cover the relevant costs never exits. The thresholds are ordered as follows.

Proposition B.7 (Ordered thresholds)**.** *Under Assumptions 1-3,* $c_i > 0$ *and* $0 \leq c_{ei} < c_{fi}$, *and at a common level of internal funds across origins, the policy of every origin state is a partition of the* $\hat{x}$ *line, such that* $\hat{x}_i^{ei} \leq \hat{x}_e^{ei} < \hat{x}_f^{ei} < \hat{x}_f^{if} < \hat{x}_e^{if} < \hat{x}_i^{if}$, *with the first inequality strict when* $c_{ei} > 0$.



B.5 Cost ordering and the cash-rich region

Lemma B.8 (Cost ordering). *Evaluate (14) at each origin's fee pair* $(\kappa_{of}, \kappa_{oi})$[33] *and define* $K \equiv \beta[(1-\lambda_f)c_f - (1-\lambda_i)c_i]$. *Then the costs are*

$$C_{i\text{-}inc} = (1+\phi_f)c_{if} + K, \quad C_{f\text{-}inc} = -(1+\phi_i)c_{fi} + K, \quad C_e = (1+\phi_f)c_f^e - (1+\phi_i)c_i^e + K, \tag{B.14}$$

and $C_e < C_{i\text{-}inc} \iff$ (B.10), *while* $C_{f\text{-}inc} < C_e \iff$ (B.11). *Under* $c_{if}, c_{fi} > 0$ *alone,* $C_{f\text{-}inc} < C_{i\text{-}inc}$ *always holds;* (B.10)-(B.11) *are what place the entrant threshold strictly between the two incumbent thresholds.*

Remark B.9 (Cash-rich firms and the role of the audit margin). *For unconstrained firms (R1 × R1) the finance advantage is irrelevant and* $sgn\{\Delta_o\} = sgn\{\Psi_f(\hat{x}; R^0) - \Psi_i(\hat{x}; R^0) - costs\}$, *which under* $T_\lambda < 1$ *is negative for large* $\hat{x}$. *This means that with taxes as the only non-financial difference, a sufficiently wealthy firm prefers informality. Two observations discipline this artifact. First, with* $\lambda_i - \lambda_f$ *large enough,* $T_\lambda \geq 1$ *and formality remains absorbing even for the cash-rich. Second, the sizable mass of formal-to-informal transitions in the data, on average* 290 *per market and window against* 314 *formalizations (Section 6), suggests that for many firms the net non-financial advantage of formality is weak, that is,* T_λ *close to or below one.*

B.6 The master threshold and its four instances

Combining the exit option with Proposition B.6 yields the complete policy map of each origin state. Define

$$A_s \equiv (1-\gamma)\Big(\frac{\gamma}{R_s}\Big)^{\frac{\gamma}{1-\gamma}} M_s^{\frac{1}{1-\gamma}}, \quad A_s^0 \equiv (1-\gamma)\Big(\frac{\gamma}{R^0}\Big)^{\frac{\gamma}{1-\gamma}} M_s^{\frac{1}{1-\gamma}}, \quad \mathcal{K}_s \equiv \beta(1-\lambda_s)c_s, \tag{B.15}$$

so that $\Psi_s(\hat{x}; R_s) = A_s \hat{x}^{1/(1-\gamma)}$, and define the identities $A = A_f - A_i$ for the formality-margin slope in (13) and $A^0 = A_f^0 - A_i^0$ for the cash-rich slope of Remark B.9. Each margin is characterized by a condition of the form $G_s(\hat{x}, \tilde{n}) = m$, where $\tilde{n}$ are the funds carried into the continuation and m collects the operating and transition costs. Solving branch by branch of (B.7) defines the following master threshold

$$\mathcal{X}_s(\tilde{n}; m) = \begin{cases} \big(m/A_s^0\big)^{1-\gamma}, & \text{R1} \\ \dfrac{m + R^0\tilde{n}}{M_s\,\tilde{n}^\gamma}, & \text{R2} \\ \left[\dfrac{m - \phi_s\tilde{n}}{A_s}\right]^{1-\gamma}, & \text{R3, if } m > \phi_s\tilde{n} \\ (1-b_s)^\gamma \dfrac{m + \tilde{n}\,\dfrac{R^0 + \phi_s b_s}{1-b_s}}{M_s\,\tilde{n}^\gamma}, & \text{R4,} \end{cases} \tag{B.16}$$

[33] $(c_{if}, 0)$ for informal incumbents, $(0, c_{fi})$ for formal incumbents, (c_f^e, c_i^e) for entrants

where $\mathcal{X}_s$ is continuous across the R1/R2, R2/R3 and R3/R4 boundaries, weakly decreasing in $\tilde{n}$, independent of financial conditions on R1, and satisfies $\mathcal{X}_s \to \infty$ as $\tilde{n} \to 0^+$ on R4. Notice that firms without internal funds exit or fail to enter at any productivity (financial exit). Moreover, on R2, G_s is linear in $\hat{x}$, hence the absence of a power form. If $m \leq \phi_s \tilde{n}$ on R3 (or the analogous condition on R4), the constrained firm never exits because the quasi-rents $\phi_s \tilde{n}$ earned by internal funds keep it alive at any productivity.

Hence, the four thresholds are instances of (B.16):

$$\text{Exit as informal:} \quad \hat{x}_i^{xi}(n) = \mathcal{X}_i\big(n;\ \mathcal{K}_i\big), \qquad n = n^{\text{exit}}, \tag{B.17}$$

$$\text{Exit as formal:} \quad \hat{x}_f^{xi}(n) = \mathcal{X}_i\big(n - c_{fi};\ \mathcal{K}_i + c_{fi}\big), \qquad n = n^{\text{exit}}, \tag{B.18}$$

$$\text{Entry as informal:} \quad \hat{x}_x^{xi}(k_0) = \mathcal{X}_i\big(k_0 - c_i^e;\ \mathcal{K}_i + c_i^e\big), \tag{B.19}$$

$$\text{Entry as formal:} \quad \hat{x}_x^{if}(k_0) = \hat{x}_e^F(k_0) = \left[\frac{B\,k_0 + C_e}{A}\right]^{1-\gamma} \qquad \text{(Prop. B.6).} \tag{B.20}$$

A complete taxonomy of these thresholds is developed in Appendix B.12.

B.7 Comparative statics and transition semi-elasticities

Proposition B.10 (Comparative statics of all thresholds)**.** *On the borrowing regions, with $D_o \equiv Bn_o + C_o > 0$ and a_s^* evaluated at the relevant margin, the thresholds respond to financial conditions as in Table 6.*

The thresholds translate into transition probabilities, the objects the Shift-Share design estimates, by integrating over the period-1 population. Consider informal incumbents with permanent component ω and internal funds n (the same algebra applies to each origin state). Since $\ln \theta_1 = \ln \omega + \varepsilon_1$ and $\hat{x}$ is monotone in θ_1, the probability of an informal-to-formal switch is given by Equation (15) of the main text, where the productivity threshold inverts (B.4):

$$\ln \bar{\theta}_{if}(\omega, n) = \frac{1}{\rho}\Big[(1-\nu)\ln \hat{x}_i^F(n) - (1-\rho)\ln \omega\Big] - \frac{\sigma_\varepsilon^2}{2\rho(1-\nu)}. \tag{B.21}$$

Expression (B.21) is common to all regions. Its derivation uses only the log-linear map (B.4) and the existence of a cutoff in $\hat{x}$-units, so it converts any threshold, whichever branch of (B.16) delivers it, closed-form or implicit, into a cutoff for θ_1. The financing region enters only through the value of the threshold and through its derivatives in the chain rule below. Analogous expressions hold for $\mathbb{P}(\text{exit} \mid i) = \Phi(z_{iX})$, $\mathbb{P}(f \to i) = \Phi(z_{fi}) - \Phi(z_{fX})$, entry probabilities (with $\hat{x}_e(\omega)$ and thresholds in ω), and all the other transitions.

Proposition B.11 (Transition semi-elasticities). *Define $p = \mathbb{P}(i \to f \mid \omega, n)$ from (15). Then, when both destinations borrow at an interior optimum (R3) at the margin,*

$$\frac{\partial p}{\partial \phi_f} = -\,\phi(z_{if})\,\frac{(1-\nu)(1-\gamma)}{\rho\,\sigma_\varepsilon}\,\frac{a_f^*}{D_i} \;<\; 0, \qquad \frac{\partial p}{\partial \phi_i} = \phi(z_{if})\,\frac{(1-\nu)(1-\gamma)}{\rho\,\sigma_\varepsilon}\,\frac{a_i^*}{D_i} \;>\; 0, \tag{B.22}$$

while, when the relevant destination is collateral-constrained (R4) at the margin, the collateral responses are

$$\frac{\partial p}{\partial b_f} = \phi(z_{if})\,\frac{1-\nu}{\rho\,\sigma_\varepsilon}\,\frac{\mu_f k_f^*}{N_i} \;\geq\; 0, \qquad \frac{\partial p}{\partial b_i} = -\,\phi(z_{if})\,\frac{1-\nu}{\rho\,\sigma_\varepsilon}\,\frac{\mu_i k_i^*}{N_i} \;\leq\; 0, \tag{B.23}$$

where N_i is the numerator of the operative branch of $\hat{x}_i^F$ at the crossing[34] and $\mu_s k_s^*$ is the shadow value of collateral of Lemma B.5. The same structure, with the appropriate threshold and sign from Proposition B.10, applies to every other transition probability. The factor $(1-\gamma)$ is specific to the power branch (R3) and drops on the linear branches; all responses are scaled by $\phi(z)/\sigma_\varepsilon$, which is the density of firms at the margin, so aggregate responses are largest where the firm distribution is thick near thresholds.

Two aggregation implications are worth highlighting. First, because (B.21) is region-free, the chain rule composes with whichever region's log-derivative is operative at the margin for the given (ω, n). Thus, the ϕ-formulas displayed are the R3 instances, the collateral formulas the R4 instances, and R1/R2 self-financed firms contribute exactly zero. Second, aggregate transition responses are therefore incidence-weighted integrals. This is $\phi(z)$ times the region-specific term, integrated over the joint distribution of (ω, n), that concentrates predicted responses among small, young, low-net-worth firms.

Remark B.12 (Internal funds of incumbents). *For an incumbent, internal funds depend on θ_1 through period-one profits, $n_o = \Omega_o x_1 k_1^\gamma + \bar{n}_o$ with $x_1 \equiv \theta_1^{1/(1-\nu)}$ and $\bar{n}_o \equiv (1-\delta)k_1 - c_o$, so the crossing is solved in θ_1 at the predetermined state (ω, k_1) rather than at fixed n_o. On R3×R3 the gap $\ln \hat{x} - \ln \hat{x}_o^F(n_o(\theta_1))$ has derivative ρJ with respect to $\ln x_1$, where $J \equiv 1 - \frac{1-\gamma}{\rho}\frac{B\,\Omega_o x_1 k_1^\gamma}{B\,n_o + C_o}$. When $\rho > 1-\gamma$ and $B\bar{n}_o + C_o \geq 0$, $J \in (0, 1]$, the crossing is unique, and Equations (15)-(B.21) hold with n_o evaluated at the crossing. The derivatives of Proposition B.11 are then divided by J: the feedback from productivity to internal funds raises their magnitude and leaves their signs unchanged. For entrants, $n_e = k_0$ does not depend on productivity and the formulas apply as stated.*

B.8 From thresholds to transition rates

Corollary B.13 (Credit expansion and the switching margins). *Consider a credit supply expansion that improves both sectors' borrowing terms: the spread of sector s falls by the*

[34] It equals $D_i \hat{x}^*$ on R4×R4. See Appendix B.12 for more details.

proportion $\chi_s^\phi \geq 0$*, that is, by the absolute amount* $\varepsilon_s \equiv \chi_s^\phi \phi_s$*,* $\mathrm{d}\phi_s = -\varepsilon_s$*, and/or collateral capacities rise,* $\mathrm{d}b_s \geq 0$*. The conditions below are stated for the spread components; a collateral improvement adds* $\mu_s k_s^* \mathrm{d}b_s$ *to its own sector's side of each inequality. Then:*

(i) Entry and exit. *Under Proposition B.7, the binding entry margin and both exit margins,* $\hat{x}_i^{\mathrm{E}}$*,* $\hat{x}_i^{\mathrm{X}}$*,* $\hat{x}_f^{\mathrm{X}}$*, are informal-continuation comparisons, so they respond only to informal-side terms: all three fall when* $\varepsilon_i > 0$ *or* $\mathrm{d}b_i > 0$*, and are unchanged under a purely formal-side expansion. Hence total entry rises and exit falls at both incumbent margins whenever the expansion reaches the informal side; the formal/informal* composition *of entry is governed by* $\hat{x}_e^{\mathrm{F}}$*, which obeys the same condition as the switching margins of (ii). These are credit effects at a fixed output price. Proposition B.16 adds the price effect, which raises every cutoff. In the final effect (credit and price effects combined), exit rises in both sectors and, of the two entry margins, only formal entry rises.*

(ii) Switching up. $\mathbb{P}(i \to f)$ *rises if and only if*

$$\varepsilon_f \, a_f^* \;>\; \varepsilon_i \, a_i^* \qquad \text{at the } i \to f \text{ margin.} \tag{B.24}$$

Because the switch fee travels with the formal option, $a_f^* - a_i^* = \tilde{k}_f - \tilde{k}_i + c_{if} > 0$ *under Assumption 1; hence equal absolute improvements* ($\varepsilon_f = \varepsilon_i$) *always raise informal-to-formal switching. Under equal* proportional *improvements* ($\chi_f^\phi = \chi_i^\phi$)*, (B.24) reads* $\phi_f a_f^* > \phi_i a_i^*$*: the borrowing-scale gap must outweigh the spread gap. The same condition, with the* a^**'s evaluated at the entrant's funds* ($a_s^* = \tilde{k}_s - (k_0 - c_s^e)$)*, governs the entrant margin* $\hat{x}_e^{\mathrm{F}}$*: there* $a_f^* - a_i^* = \tilde{k}_f - \tilde{k}_i + c_f^e - c_i^e > 0$*, so equal absolute improvements also raise the formal share of entrants.*

(iii) Switching down. $\mathbb{P}(f \to i)$ *rises if and only if*

$$\varepsilon_i \, a_i^* \;>\; \varepsilon_f \, a_f^* \qquad \text{at the } f \to i \text{ margin,} \tag{B.25}$$

where now the fee travels with the informal option, $a_i^* - a_f^* = \tilde{k}_i - \tilde{k}_f + c_{fi}$*. Both switching margins can therefore rise together only if the expansion has a sufficiently informal-side component* (ε_i*, or* b_i $\uparrow$)*. Condition (B.25) concerns the switching threshold* z_{fi} *alone. Since* $\mathbb{P}(f \to i) = \Phi(z_{fi}) - \Phi(z_{fX})$ *and, by (i), the formal exit boundary* z_{fX} *falls with any informal-side improvement, an expansion with an informal-side component also raises* $\mathbb{P}(f \to i)$ *through the exit boundary: the formal firms it keeps alive are low-productivity firms that switch down. Under a purely formal-side expansion the exit boundary is unchanged and (B.25) is necessary and sufficient. In general it is sufficient, and a fall in* $\mathbb{P}(f \to i)$ *requires the formal-side dominance at the switching margin to outweigh the exit-boundary term.*

This incidence algebra is the model's central payoff for the empirical strategy, and the estimates of Section 6 discipline it. Formalization rises and formal-to-informal transitions fall; by (ii) and (iii) both movements require the same inequality, that the formal-side improvement dominates at each switching margin, $\varepsilon_f a_f^* > \varepsilon_i a_i^*$, while the size of the formal-to-informal transitions decline is informative about the informal side, because a larger informal-side improvement attenuates it. The entry and exit block points the same way: informal entry is flat and exit rises, so the informal-side threshold effects of (i) are absent from the data, and the exit increase instead reflects the turnover of marginal cohorts, the subject of the next subsection, and the price effect of Appendix B.10. The estimated switching pattern is thereby converted into a statement about which sector's borrowing terms the shock reached.

B.9 Composition: marginal entrants, turnover, and average productivity

The last set of results concerns the composition of the response. This is where the model meets the value-added estimates of Section 6: informal value added falls, formal value added does not move, and the average across firms is at most marginally positive. Allow period-2 firms a shutdown option: after observing θ_2, an active firm operates if and only if variable profit covers the fixed cost, $\Omega_s x_2 k_2^{\gamma} \geq c_s$; otherwise it stays idle for the period (it still resells its capital). Shutdown is the two-period analogue of endogenous exit for a young cohort. For an entering cohort into sector i (the same algebra applies to a formal entering cohort, the margin the estimated shock moves most), entry occurs if and only if $\omega \geq \bar{\omega}(k_0)$, where $\bar{\omega}$ solves $\hat{x}_e(\bar{\omega}) = \hat{x}_i^{\mathrm{E}}(k_0)$, and:

$$\mathbb{P}(\text{shutdown of an entrant with}(\omega, k_2)) : \pi^X(\omega) = \Phi\left(\frac{(1-\nu)\ln\big(c_i/(\Omega_i k_2^{\gamma})\big) - \ln\omega}{\sigma_\varepsilon}\right), \tag{B.26}$$

which is strictly decreasing in ω (directly, and indirectly through k_2).

Proposition B.14 (Composition of a credit expansion)**.** *Let financial conditions improve for the informal sector ($\phi_i \downarrow$ or $b_i \uparrow$), so that $\bar{\omega}$ falls for every constrained k_0. Then, for the entering cohort:*

(i) Entry rises*: the mass of entrants $\int \mathbb{P}(\omega \geq \bar{\omega}(k_0))\,\mathrm{d}F(k_0)$ increases.*

(ii) Average entrant productivity falls*: for $\ln\omega \sim N(0, \sigma_\omega^2)$, $\mathbb{E}[\omega \mid \omega \geq \bar{\omega}]$ is strictly increasing in $\bar{\omega}$; hence a lower threshold mechanically lowers the mean (and every quantile) of entrant productivity, and with it average value added per entrant.*

(iii) Cohort turnover rises*: the cohort shutdown rate $\mathbb{E}[\pi^X(\omega) \mid \omega \geq \bar{\omega}]$ is strictly decreasing in $\bar{\omega}$; the marginal entrants admitted by the expansion fail at higher rates*

than infra-marginal ones, so measured exit rises with a lag even at a fixed output price, where incumbent exit thresholds fall. Moreover, with the price response of Appendix B.10, they rise as well.

Part (iii) of the proposition isolates the composition channel: for the collateral-constrained entrants where the margin lives, capital is pinned by wealth and collateral, so a pure spread reduction leaves their capital unchanged and the composition effect is the entire response, while a collateral relaxation deepens their capital and partially offsets it. The proposition thus reproduces, inside the two-period model, the composition signature of Section 6: the expansion admits marginal, lower-ω entrants and raises their subsequent turnover, a model counterpart of higher exit, with a lag, while on the informal side the formality threshold falls, the most productive informal firms leave the pool, and informal average value added declines. The two forces lower average productivity within the informal sector, the response that is robust in the data, and leave the economy-wide average roughly flat, in line with the estimates: the model reads the value-added results as reallocation plus selection, not as productivity gains within infra-marginal firms.

B.10 Competition in the goods market

This subsection formalizes the extension of Subsection 4.4. Output is a homogeneous good sold at price $P > 0$ in units of the good in which capital, loans, wages and fixed costs are paid, which is the numéraire; in the baseline the two goods coincide and $P = 1$. The wage w and the financial parameters remain exogenous. Inverse demand is $P = \mathcal{P}(Q)$, where Q is the aggregate output after the transition stage. There is a continuum of incumbents, with a given measure over $(o, \omega, \theta_1, k_1)$, and of potential entrants, with a given measure over (ω, k_0). Both measures are finite, atomless, and have positive densities. By the law of large numbers, Q and P are deterministic, so at the transition stage every firm takes P as given. Period one is predetermined, so the internal funds n_o of Equation (6) do not depend on P.

Lemma B.15 (Price neutrality in index units)**.** *With revenue $P\theta k^{\alpha}\ell^{\nu}$, optimal labor and profits are those of Lemma B.1 evaluated at $P\theta$ in place of θ,*

$$\ell_s^*(\theta, k; P) = \ell_s^*(P\theta, k), \qquad \pi_s(\theta, k; P) = \Omega_s \, P^{\frac{1}{1-\nu}} \, x \, k^{\gamma}, \tag{B.27}$$

and the expected output after the transition stage of a firm that installs k in sector s is $\mathbb{E}[y_2 \mid \mathcal{I}_1] = \upsilon_s \, P^{\frac{\nu}{1-\nu}} \, \hat{x} \, k^{\gamma}$, with $\upsilon_s \equiv [\nu(1-\tau_{y,s})/((1+\tau_{\ell,s})w)]^{\frac{\nu}{1-\nu}}$. Consequently, Lemmas B.2–B.5, Propositions B.6–B.10, Corollary B.13 and the master threshold (B.16) still hold with $P^{\frac{1}{1-\nu}}\hat{x}$ in place of $\hat{x}$. So the cutoff values $\hat{x}_j$ of Propositions B.6–B.10 do not involve P. A firm compares them with its price-adjusted index $P^{\frac{1}{1-\nu}}\hat{x}$, so in terms of its unadjusted

index the cutoff of every margin j becomes $\hat{x}_j\, P^{-\frac{1}{1-\nu}}$. In productivity units,

$$\begin{aligned} \ln \bar{\theta}_{1,j}(P) &= \ln \bar{\theta}_{1,j}(1) - \frac{1}{\rho} \ln P \quad \textit{(incumbents)}, \\ \ln \bar{\omega}_j(P) &= \ln \bar{\omega}_j(1) - \ln P \quad \textit{(entrants)}, \end{aligned} \tag{B.28}$$

where $\bar{\theta}_{1,j}$ is the period-one productivity cutoff of (B.21) *and $\bar{\omega}_j = \hat{x}_j^{1-\nu} \exp\{-\sigma_\varepsilon^2/(2(1-\nu))\}$ the entrant cutoff in ω, which inverts the entrant index $\hat{x}_e(\omega)$ defined after* (B.4).

Let $\mathcal{A}(P;\Xi)$ denote the set of active firms at price P under financial conditions $\Xi \equiv (\phi_i, \phi_f, b_i, b_f)$, i.e., the incumbents and entrants whose price-adjusted index exceeds the exit or entry threshold of their origin (Proposition B.7). Define aggregate output as the expected output of active firms that survive the death shock,

$$Q(P;\Xi) \;\equiv\; \int_{\mathcal{A}(P;\Xi)} (1-\lambda_d)\, \mathbb{E}[y_2 \mid \mathcal{I}_1]\, \mathrm{d}\mu, \tag{B.29}$$

with d the destination chosen by each firm and μ the measure of firms.

Assumption 4 (Demand and supply)**.** *(a) $\mathcal{P} : (0,\infty) \to (0,\infty)$ is continuous, strictly decreasing, and $\mathcal{P}(Q) \to 0$ as $Q \to \infty$. (b) $Q(\cdot\,;\Xi)$ is continuous and strictly increasing, with $Q(P;\Xi) \to 0$ as $P \to 0$ and $Q(P;\Xi) \to \infty$ as $P \to \infty$. (c) For the comparative statics, $\mathcal{P}$ and Q are continuously differentiable and the credit supply expansion of intensity ζ in* (B.30) *below raises output at a fixed price, $Q_\zeta \equiv \partial Q/\partial \zeta > 0$.*

Part (b) rules out the only case in which supply could slope down, which is when the output of formalizing firms falls because the payroll tax lowers their labor demand per unit of capital. All other components of Q rise with the price: the active set expands, since all exit and entry cutoffs fall by (B.28); capital rises, since k_s^* is weakly increasing in the price-adjusted index (Lemma B.2); and labor rises, since ℓ_s^* is increasing in $P\theta$ (Lemma B.1). Part (c) requires that the credit effect raises output at a fixed price. Its components (more entry, less exit of borrowers, and deeper capital of borrowers in R3 and R4) all do. The only one that can lower output is again the output change of the firms that formalize, which is assumed not to dominate.

Proposition B.16 (Competition in the goods market)**.** *Let the assumptions of Propositions B.6 and B.7 and Assumption 4 hold. Call* price thresholds *the informal entry, formal exit and informal exit thresholds,* (B.19)*,* (B.18) *and* (B.17)*, and* credit thresholds *the formal entry threshold* (B.20) *and the two switching thresholds $\hat{x}_i^F$ and $\hat{x}_f^F$ of* (B.12)*. Then:*

(i) Equilibrium. *There is a unique price $P^*(\Xi)$ such that $P^* = \mathcal{P}\big(Q(P^*;\Xi)\big)$.*

(ii) Price and output responses. *Let a credit supply shock shift financial conditions by*

$$\mathrm{d}\Xi = (-\chi_i^\phi \phi_i,\, -\chi_f^\phi \phi_f,\, \chi_i^b b_i,\, \chi_f^b b_f)'\, \zeta, \tag{B.30}$$

with ζ its intensity and $(\chi_i^\phi, \chi_f^\phi, \chi_i^b, \chi_f^b) \geq 0$ its direction (per unit of intensity, the spread of sector s falls by the proportion χ_s^ϕ and its collateral capacity rises by the proportion χ_s^b), and let $Q_P \equiv \partial Q/\partial P$ at a fixed price. Then

$$\psi \;\equiv\; -\frac{\mathrm{d}\ln P^*}{\mathrm{d}\zeta} = \frac{|\mathcal{P}'(Q)|\, Q_\zeta}{P^*\big[1 + |\mathcal{P}'(Q)|\, Q_P\big]} \;>\; 0, \tag{B.31}$$

$$\frac{\mathrm{d}Q}{\mathrm{d}\zeta} = \frac{Q_\zeta}{1 + |\mathcal{P}'(Q)|\, Q_P} \;\in\; (0, Q_\zeta). \tag{B.32}$$

The price falls, and the equilibrium output response keeps the sign of the credit effect and is smaller. In elasticity form, $\psi = (Q_\zeta/Q)/(e_D + e_S)$, where $e_D \equiv P/(Q|\mathcal{P}'(Q)|)$ and $e_S \equiv PQ_P/Q$ are the price elasticities of demand and supply.

(iii) Thresholds. *For every margin j, let s_j be its credit effect, which is the direct log shift in index units per unit of ζ from Proposition B.10. The equilibrium shift of its productivity cutoff is the sum of the credit and price effects,*

$$\begin{aligned} \frac{\mathrm{d}\ln \bar\theta_{1,j}}{\mathrm{d}\zeta} &= \frac{1}{\rho}\Big[\underbrace{(1-\nu)\, s_j}_{\textit{credit effect}} + \underbrace{\psi}_{\textit{price effect}}\Big] \quad \textit{(incumbents)}, \\ \frac{\mathrm{d}\ln \bar\omega_j}{\mathrm{d}\zeta} &= \underbrace{(1-\nu)\, s_j}_{\textit{credit effect}} + \underbrace{\psi}_{\textit{price effect}} \qquad \textit{(entrants)}. \end{aligned} \tag{B.33}$$

(iv) Price thresholds. *At a price threshold (informal entry, formal exit or informal exit), the credit effect s_j depends only on the informal loan terms: $s_j = 0$ when the informal continuation does not borrow at the margin (R1 or R2) or when the expansion reaches the formal side alone ($\chi_i^\phi = \chi_i^b = 0$), and $s_j = -\big[a_i^* \chi_i^\phi \phi_i + \mu_i k_i^* \chi_i^b b_i\big]/(\Psi_i' \hat{x}_j) \leq 0$ otherwise. The price effect dominates the credit effect, which raises the cutoff so that exit rises in both sectors and informal entry falls, if and only if $\psi > (1-\nu)|s_j|$, and in particular whenever $s_j = 0$, in which case the three price cutoffs move only by the price effect.*

(v) Credit thresholds. *Under the formal-side dominance condition of Corollary B.13, the credit effect is negative ($s_j < 0$) at the credit thresholds (formal entry and the two switching thresholds). It dominates the price effect, so that its sign survives in equilibrium, if and only if $\psi < (1-\nu)|s_j|$. This is the case when demand is elastic enough: by (ii), $\psi = (Q_\zeta/Q)/(e_D + e_S)$, so the condition holds whenever $e_D + e_S > (Q_\zeta/Q)/\big((1-\nu)|s_j|\big)$, and it holds at all three credit thresholds at once whenever the sum of the two elasticities exceeds $(Q_\zeta/Q)/\big((1-\nu)\min_j |s_j|\big)$.*

Remark B.17 (Reasons for the dominance of the price effect at the price thresholds)**.** *Parts (iv) and (v) compare the price effect, ψ, which is common to all margins, with the credit effect, s_j, of each threshold. Two features of the model make $|s_j|$ smaller at the price thresholds than at the credit thresholds. First, by Proposition B.10 the credit effect of a price threshold contains only the informal terms, χ_i^ϕ and χ_i^b, whereas that of a credit threshold is a difference in which the formal terms carry the larger weight, $a_f^*\chi_f^\phi\phi_f - a_i^*\chi_i^\phi\phi_i$ and its collateral analogue. When the formal terms improve more than the informal ones, as in Section 6, the informal component is small and the difference large. Second, every credit effect is proportional to the borrowing of the marginal firm (Lemma B.5). At a common level of internal funds, informal borrowing, $a_i^* = \tilde{k}_i(\hat{x}) - n$, is increasing in productivity, $\hat{x}$, on R3 (Lemma B.2). Since the price thresholds lie below the credit thresholds of the same origin on the productivity line (Proposition B.7), the marginal firm at a price threshold borrows less. Hence a demand elasticity that keeps ψ below the credit effects of the credit thresholds can at the same time leave it above those of the price thresholds, which is the configuration described in Section 4.4. Since ψ is the direct output response divided by the sum of the demand and supply elasticities, the condition in (v) holds whenever demand is elastic enough. The baseline results are recovered as the slope of inverse demand vanishes, $\mathcal{P}'(Q) \to 0$, that is, as demand becomes perfectly elastic.*

Notice that the shutdown option of Appendix B.9 is also subject to the price effect, with the condition $\Omega_s P^{\frac{1}{1-\nu}} x_2 k_2^\gamma \geq c_s$. Thus, the shutdown probability (B.26) rises as P falls, and the turnover of every cohort, not only of the marginal entrants of Proposition B.14, rises with the expansion.

B.11 The model-implied estimand in closed form

The model makes precise what the IV coefficient identifies. Let ζ_m, the (exogenous) bank-supply shock in market m and the model counterpart of the standardized instrument $\hat{Z}_{mt}^{LOO}$, shift financial conditions by $\mathrm{d}\Xi_m = (-\chi_i^\phi\phi_i, -\chi_f^\phi\phi_f, \chi_i^b b_i, \chi_f^b b_f)'\,\zeta_m$. Let the direction $(\chi_i^\phi, \chi_f^\phi, \chi_i^b, \chi_f^b)$ of Equation (B.30) be the same in every market. Aggregate borrowing in a market, C_{mt} in the data, is $C = \int a^*\,\mathrm{d}\Lambda$ in the model, where Λ is the population of active firms. Its log change per unit of ζ_m is the model counterpart of the first-stage coefficient π of Equation (21):

$$\pi \equiv \frac{\mathrm{d}\ln C}{\mathrm{d}\zeta} = \frac{1}{C}\Bigg[\underbrace{\sum_{d\in\{i,f\}}\left(\int_{\mathrm{R3}_d}\frac{\chi_d^\phi\phi_d\,\tilde{k}_d(\hat{x})}{(1-\gamma)R_d}\,\mathrm{d}\Lambda + \int_{\mathrm{R4}_d}\frac{\chi_d^b b_d\,n}{(1-b_d)^2}\,\mathrm{d}\Lambda\right)}_{\text{intensive: loan deepening}} - \underbrace{\sum_j \int w_j\,s_j\,\Delta a_j^*\,\mathrm{d}\Lambda_{o(j)}}_{\text{extensive: loans created, eliminated, or repriced}}\Bigg]. \tag{B.34}$$

Each firm-count outcome Y in the data (stocks, entry, exit, transitions by sector) aggregates the threshold rules of the model. Its log change per unit of ζ_m, which adds up the threshold shifts, is the model counterpart of the reduced-form coefficient a of Equation (22):

$$a \equiv \frac{\mathrm{d}\ln Y}{\mathrm{d}\zeta} = -\frac{1}{Y}\sum_{j\in\mathcal{J}_Y} \iota_j^Y \int w_j\, s_j\, \mathrm{d}\Lambda_{o(j)}. \tag{B.35}$$

Therefore, to first order, the IV coefficient b of Equation (20) is their ratio:

$$b = \frac{a}{\pi}, \tag{B.36}$$

where $\mathcal{J}_Y$ is the set of thresholds that delimit Y and, for each threshold j, s_j is its log shift per unit of intensity, obtained from Proposition B.10 along the direction of the shock (the credit effect of Proposition B.16): a difference of borrowing-weighted spread gains at the formality thresholds and a one-sided informal-continuation term at the exit and entry thresholds, with collateral terms on R4-financed thresholds. The two intensive integrals run over continuing borrowers to d in R3 and in R4. The weight w_j is the density of firms at the threshold, $\phi(z_j)(1-\nu)/(\rho\sigma_\varepsilon)$ for incumbents and the type density for entrants; $\iota_j^Y = \pm 1$ records whether an upward crossing of j moves a firm into or out of Y; Δa_j^* is the borrowing gained on an upward crossing, a new loan at the entry and exit thresholds and the change in loan size at the formality thresholds; and $\Lambda_{o(j)}$ is the population of the origin state $o(j)$ whose firms face threshold j.

Three lessons follow. First, the numerator is a density-weighted sum of threshold shifts: the IV coefficient is large where many firms sit near their thresholds, young, small, constrained firms, which is consistent with responses that concentrate on the formal side, the sector whose borrowing terms the shock improves. Second, the denominator mixes intensive responses (continuing borrowers deepen loans by $\chi_s^\phi \phi_s \tilde{k}_s/[(1-\gamma)R_s]$ each where the spread binds and by $\chi_s^b b_s n/(1-b_s)^2$ each where collateral binds) with extensive ones (loans created, eliminated, or repriced at the margins), and both are built from the same primitives: spread levels, participation costs, and collateral caps. Third, because ζ enters numerator and denominator symmetrically, b is invariant to the scale of the shock for small ζ. Adding the price response of Appendix B.10 turns both responses, Equations (B.34) and (B.35), into total derivatives: the price effect adds the common term ψ of Proposition B.16(iii) to every threshold shift, and the ratio keeps its form. We use the ratio to interpret the estimates, their signs and relative magnitudes, and not to compute them: doing so would require the incidence of the shock on each loan term, (χ_s^ϕ, χ_s^b), and the density of firms at each margin, which are not identified without estimating the full model.

B.12 Complete threshold taxonomy

This appendix reports the fully substituted closed form of every exit and entry threshold on every financing region. Throughout, $\mathcal{K}_s \equiv \beta(1-\lambda_s)c_s$ (so that $K = \mathcal{K}_f - \mathcal{K}_i$), $A_s = (1-$

$\gamma)(\gamma/R_s)^{\frac{\gamma}{1-\gamma}} M_s^{\frac{1}{1-\gamma}}$ and $A_s^0 = (1-\gamma)(\gamma/R^0)^{\frac{\gamma}{1-\gamma}} M_s^{\frac{1}{1-\gamma}}$, with $A = A_f - A_i$ and $A^0 = A_f^0 - A_i^0$. Each one-sided margin (exit or entry against the outside option, informal continuation being the relevant alternative by Proposition B.7) is the solution of $G_i(\hat{x}, \tilde{n}) = m$ with the margin-specific pair $(\tilde{n}, m)$:

margin	$\tilde{n}$	m	comparison
exit as informal	n	$\mathcal{K}_i$	$U_i(\hat{x}, n) \gtrless n$
exit as formal	$n - c_{fi}$	$\mathcal{K}_i + c_{fi}$	$U_i(\hat{x}, n - c_{fi}) \gtrless n$
entry as informal	$k_0 - c_i^e$	$\mathcal{K}_i + c_i^e$	$U_i(\hat{x}, k_0 - c_i^e) \gtrless k_0$

with $n = n^{\text{exit}}$ of Equation (6). Substituting the four branches of (B.7):

Exit as informal.

$$\hat{x}_i^{\mathrm{X}}(n) = \begin{cases} \left(\mathcal{K}_i / A_i^0\right)^{1-\gamma}, & \text{R1} \\ \dfrac{\mathcal{K}_i + R^0 n}{M_i\, n^\gamma}, & \text{R2} \\ \left[\dfrac{\mathcal{K}_i - \phi_i n}{A_i}\right]^{1-\gamma}, & \text{R3, if } \mathcal{K}_i > \phi_i n \\ (1-b_i)^{\gamma-1}\, \dfrac{\mathcal{K}_i(1-b_i) + n\left(R^0 + \phi_i b_i\right)}{M_i\, n^\gamma}, & \text{R4.} \end{cases} \tag{B.37}$$

Exit as formal (against switching down).

$$\hat{x}_f^{\mathrm{X}}(n) = \begin{cases} \left[(\mathcal{K}_i + c_{fi}) / A_i^0\right]^{1-\gamma}, & \text{R1} \\ \dfrac{\mathcal{K}_i + c_{fi} + R^0(n - c_{fi})}{M_i\,(n - c_{fi})^\gamma}, & \text{R2} \\ \left[\dfrac{\mathcal{K}_i + (1+\phi_i)c_{fi} - \phi_i n}{A_i}\right]^{1-\gamma}, & \text{R3, if numerator} > 0 \\ (1-b_i)^{\gamma-1}\, \dfrac{(\mathcal{K}_i + c_{fi})(1-b_i) + (n - c_{fi})\left(R^0 + \phi_i b_i\right)}{M_i\,(n - c_{fi})^\gamma}, & \text{R4.} \end{cases} \tag{B.38}$$

Entry as informal. Replace (n, c_{fi}) by (k_0, c_i^e) in (B.38); in particular on R3,

$$\hat{x}_i^{\mathrm{E}}(k_0) = \left[\frac{\mathcal{K}_i + (1+\phi_i)c_i^e - \phi_i k_0}{A_i}\right]^{1-\gamma}, \tag{B.39}$$

and in type units $\bar{\omega}(k_0) = \hat{x}_i^{\mathrm{E}}(k_0)^{\,1-\nu} \exp\{-\sigma_\varepsilon^2/(2(1-\nu))\}$.

Entry as formal (the entrant formality margin). This margin compares two continuation values, each on its own region, with funds $\tilde{n}_f = k_0 - c_f^e$ and $\tilde{n}_i = k_0 - c_i^e$. Because $G_s(\cdot, \tilde{n})$ is a power function of $\hat{x}$ on R1/R3 but *linear* in $\hat{x}$ on R2/R4, closed forms

obtain on the diagonal region pairs:

$$\text{R3} \times \text{R3}: \quad \hat{x}_e^{\mathrm{F}}(k_0) = \left[\frac{B\,k_0 + C_e}{A}\right]^{1-\gamma}, \qquad C_e = (1+\phi_f)c_f^e - (1+\phi_i)c_i^e + K; \tag{B.40}$$

$$\text{R1} \times \text{R1}: \quad \hat{x}_e^{\mathrm{F}} = \left(C_e^0/A^0\right)^{1-\gamma}, \qquad C_e^0 = c_f^e - c_i^e + K, \tag{B.41}$$

(direction of the inequality per Remark B.9);

$$\text{R2} \times \text{R2}: \quad \hat{x}_e^{\mathrm{F}} = \frac{K + \beta(1-\delta)\left(c_f^e - c_i^e\right)}{M_f \tilde{n}_f^{\gamma} - M_i \tilde{n}_i^{\gamma}}, \qquad \text{formal above iff the denominator} > 0 \tag{B.42}$$

(at kink, only cost is forgone resale value, hence $\beta(1-\delta)$, not R^0);

$$\text{R4} \times \text{R4}: \quad \hat{x}_e^{\mathrm{F}} = \frac{K + R_f h_f - R_i h_i + (1+\phi_i)\tilde{n}_i - (1+\phi_f)\tilde{n}_f}{M_f h_f^{\gamma} - M_i h_i^{\gamma}}, \tag{B.43}$$

with $h_s \equiv \tilde{n}_s/(1-b_s)$ the capital installed at the binding cap.

Off-diagonal pairs (one destination borrowing while the other self-finances, e.g. R3×R2) mix $\hat{x}^{1/(1-\gamma)}$ and $\hat{x}$ terms: the indifference condition remains strictly monotone in $\hat{x}$ at the crossing and has a unique root, but no closed form; it is solved implicitly. The same diagonal/off-diagonal taxonomy applies verbatim to the incumbent formality margins $\hat{x}_i^{\mathrm{F}}, \hat{x}_f^{\mathrm{F}}$ with the fee pairs of Lemma B.8.

Branch selection. Which branch applies is a self-consistency requirement: compute the candidate threshold from each branch and retain the one whose funds fall in that branch's bracket *evaluated at the candidate*,

$$\begin{aligned}
&\text{R1: } \tilde{n} \geq \bar{k}_s(\hat{x}^*);\\
&\text{R2: } \tilde{k}_s(\hat{x}^*) \leq \tilde{n} < \bar{k}_s(\hat{x}^*);\\
&\text{R3: } (1-b_s)\tilde{k}_s(\hat{x}^*) \leq \tilde{n} < \tilde{k}_s(\hat{x}^*);\\
&\text{R4: } \tilde{n} < (1-b_s)\tilde{k}_s(\hat{x}^*)
\end{aligned}$$

Exactly one branch is self-consistent: G_s is continuous and strictly increasing in $\hat{x}$ ($\partial G_s/\partial \hat{x} = M_s k^{*\gamma} > 0$ by the envelope theorem), and at fixed $\tilde{n}$ the firm moves monotonically through R1→R2→R3→R4 as $\hat{x}$ rises, so $G_s(\cdot, \tilde{n})$ crosses any constant m exactly once.

B.13 Proofs

Proof of Lemma B.1. The objective $(1-\tau_{y,s})\theta k^{\alpha}\ell^{\nu} - (1+\tau_{\ell,s})w\ell$ is strictly concave in ℓ ($\nu < 1$); the first-order condition $\nu(1-\tau_{y,s})\theta k^{\alpha}\ell^{\nu-1} = (1+\tau_{\ell,s})w$ gives ℓ_s^* in (B.1). Substituting, using $(1+\tau_{\ell,s})w\,\ell^* = \nu(1-\tau_{y,s})\theta k^{\alpha}\ell^{*\nu}$, yields $\pi_s = (1-\nu)(1-\tau_{y,s})\theta k^{\alpha}\ell^{*\nu}$ and, after substituting $\ell^{*\nu}$, expression (B.2). Monotonicity claims follow by inspection of the closed forms: $\partial \ln \ell^*/\partial \ln \theta = 1/(1-\nu) > 0$, $\partial \ln \ell^*/\partial \ln k = \alpha/(1-\nu) > 0$, $\partial \ln \ell^*/\partial \ln w =$

$-1/(1-\nu) < 0$, and both tax terms enter with negative sign. Strict concavity of π_s in k follows from $\gamma = \alpha/(1-\nu) < 1$, itself implied by $\alpha + \nu < 1$. $T < 1$ is immediate from $\tau_y, \tau_\ell \geq 0$, with at least one strictly positive. □

Proof of Lemma B.2. The objective in Equation (8) is continuous and, on each branch of $a = \max\{0, k-n\}$, strictly concave in k (the return term has curvature $\gamma < 1$; the cost terms are linear). Marginal return to capital is $g'(k) = \gamma M_s \hat{x} k^{\gamma-1} + \beta(1-\delta)$, strictly decreasing, with range $(\beta(1-\delta), \infty)$. Marginal cost is 1 for $k < n$ and $1+\phi_s$ for $k > n$. If $g'(n) \leq 1$, i.e. $n \geq \bar{k}_s$, the self-financed optimum $\bar{k}_s$ (where $g' = 1$) is attainable without borrowing: region R1. If $1 < g'(n) \leq 1+\phi_s$, i.e. $\tilde{k}_s \leq n < \bar{k}_s$, the kink at $k = n$ is optimal: region R2. If $g'(n) > 1+\phi_s$, i.e. $n < \tilde{k}_s$, the firm borrows to the interior optimum $\tilde{k}_s$ (where $g' = 1+\phi_s$) provided the collateral constraint $k - n \leq b_s k \iff k \leq n/(1-b_s)$ permits, i.e. $\tilde{k}_s \leq n/(1-b_s)$: region R3. Otherwise the constraint binds and $k = n/(1-b_s)$: region R4. The boundaries in (B.5) restate these conditions. The derivative signs follow from $\tilde{k}_s = (\gamma M_s \hat{x}/R_s)^{1/(1-\gamma)}$ and $\bar{k}_s = (\gamma M_s \hat{x}/R^0)^{1/(1-\gamma)}$, so that $\partial \tilde{k}_s/\partial \phi_s = -\tilde{k}_s/[(1-\gamma)R_s]$, $\partial \tilde{k}_s/\partial \hat{x} = \tilde{k}_s/[(1-\gamma)\hat{x}]$ and $\partial \bar{k}_s/\partial \hat{x} = \bar{k}_s/[(1-\gamma)\hat{x}]$; from $k = n/(1-b_s)$ in R4, whose b_s-derivative is $n/(1-b_s)^2$; and by inspection elsewhere. Continuity at boundaries is direct substitution. □

Proof of Lemma B.3. Substitute the policy (B.5) into Equation (8). R1: $G = M_s \hat{x} \bar{k}^\gamma - R^0 \bar{k}$; using $\gamma M_s \hat{x} \bar{k}^{\gamma-1} = R^0$, $G = R^0 \bar{k}(1-\gamma)/\gamma = \Psi_s(\hat{x}; R^0)$ after substituting $\bar{k}$. R2: $k = n$, $a = 0$: $G = M_s \hat{x} n^\gamma - R^0 n$. R3: $G = -\tilde{k} - \phi_s(\tilde{k} - n) + M_s \hat{x} \tilde{k}^\gamma + \beta(1-\delta)\tilde{k} = \phi_s n + M_s \hat{x} \tilde{k}^\gamma - R_s \tilde{k}$, and the same substitution with R_s gives $\phi_s n + \Psi_s(\hat{x}; R_s)$. R4: substitute $k = n/(1-b_s)$ directly. Monotonicity and convexity of Ψ_s in $\hat{x}$ follow from the exponent $1/(1-\gamma) > 1$; G inherits them. The formulas for q_s are the n-derivatives of (B.7); the R2 bounds because $g'(n) \in (1, 1+\phi_s]$ there, and the R4 inequality because the bracketed marginal-return gap is strictly positive when the collateral constraint binds. Concavity of G in n holds piecewise and at kinks because q_s is (weakly) decreasing in n region by region as n rises through R4→R3→R2→R1. □

Proof of Lemma B.5. On R1-R2, $a^* = 0$ and (ϕ_s, b_s) do not enter the realized objective: derivatives are zero, consistent with (10). On R3-R4 apply the envelope theorem to Equation (8): the derivative of the value with respect to ϕ_s at the optimum is the partial derivative of the objective, $-a_s^*$. For b_s, attach the multiplier $\mu_s \geq 0$ to the constraint written as $b_s k - a \geq 0$, so that the envelope derivative is $\mu_s k_s^*$: zero when the constraint is slack (R3) and, in R4, where $a = k - n$ and the first-order condition in k reads $\gamma M_s \hat{x} k^{\gamma-1} - R_s - \mu_s(1-b_s) = 0$, equal to

$$\mu_s k_s^* = \left[\gamma M_s \hat{x} k_s^{*\gamma-1} - R_s\right] \frac{n}{(1-b_s)^2} > 0,$$

using $k_s^* = n/(1-b_s)$; differentiating (B.7) in R4 with respect to b_s directly gives the same expression. (Writing the constraint as $k \le n/(1-b_s)$ instead rescales the multiplier to the wedge $\gamma M_s \hat{x} k_s^{*\gamma-1} - R_s$ and the derivative of the bound to $n/(1-b_s)^2$; the product is unchanged.) Cross-sector derivatives are zero because s'-parameters do not enter sector-s payoffs. □

Proof of Proposition B.6. On R3×R3, (12) follows by subtracting the R3 branch of (B.7) for each destination and collecting the $\hat{x}^{1/(1-\gamma)}$, n_o, and constant terms, using Equation (6). $A > 0$ iff $M_f^{1/(1-\gamma)} R_f^{-\gamma/(1-\gamma)} > M_i^{1/(1-\gamma)} R_i^{-\gamma/(1-\gamma)}$, i.e. iff $M_f/M_i > (R_f/R_i)^\gamma$; since $M_f/M_i = T_\lambda$ by definition, this is Assumption 1. Strict monotonicity and convexity in $\hat{x}$ follow from $A > 0$ and the exponent $1/(1-\gamma) > 1$; since $\Delta_o \to -(Bn_o + C_o) < 0$ as $\hat{x} \to 0$ (by Assumption 2) and $\Delta_o \to \infty$ as $\hat{x} \to \infty$, a unique zero exists and solves (B.12). Differentiating (B.12) in n_o gives (B.13). Off R3×R3, the envelope theorem gives $\partial \Delta_o / \partial \hat{x} = M_f k_f^{*\gamma} - M_i k_i^{*\gamma}$, so Δ_o is strictly increasing in $\hat{x}$ wherever $M_f k_f^{*\gamma} > M_i k_i^{*\gamma}$. Under Assumption 1 this holds whenever the formal destination is not collateral-constrained, since then $k_f^* = \tilde{k}_f$ and $k_i^* \le \tilde{k}_i$. When the formal destination is in R4, $k_f^* = h_f \equiv (n_o - \kappa_{of})/(1-b_f) < \tilde{k}_f$ and the inequality requires in addition $M_f h_f^\gamma > M_i k_i^{*\gamma}$, which fails when the formal fee leaves little net internal funds. Under these conditions Δ_o is continuous and strictly increasing in $\hat{x}$ on the borrowing range, hence the crossing is unique globally on the borrowing range. Otherwise uniqueness is branch specific and the selection rule of Appendix B.12 applies. □

Proof of Lemma B.8. Substituting each origin's fee pair into (14) gives (B.14) directly; the common constant K cancels from every pairwise difference. Upper comparison:

$$C_{i\text{-inc}} - C_e = (1+\phi_f)c_{if} - \left[(1+\phi_f)c_f^e - (1+\phi_i)c_i^e\right] \;>\; 0 \iff \text{(B.10)}.$$

Lower comparison:

$$C_e - C_{f\text{-inc}} = (1+\phi_f)c_f^e - (1+\phi_i)c_i^e + (1+\phi_i)c_{fi} \;>\; 0 \iff \text{(B.11)}.$$

Note that $c_i^e \le c_f^e$ alone does *not* imply (B.11), because $\phi_i > \phi_f$ can make $(1+\phi_i)c_i^e$ exceed $(1+\phi_f)c_f^e$. If $c_{fi} \ge c_i^e$, however, $(1+\phi_i)(c_{fi} - c_i^e) \ge 0$, and the sum is strictly positive in either case: if $c_i^e > 0$, then $c_f^e \ge c_i^e > 0$ and $(1+\phi_f)c_f^e > 0$; if $c_i^e = 0$, then $c_{fi} - c_i^e = c_{fi} > 0$. Hence (B.11) holds. □

Proof of Proposition B.7. The three-interval structure for each origin follows from: (i) values of continuation are continuous and strictly increasing in $\hat{x}$ while exit/out values are constant in $\hat{x}$, giving unique exit/entry cutoffs; (ii) Proposition B.6 gives the unique formality cutoff above them whenever $\hat{x}^{\mathrm{X}} < \hat{x}^{\mathrm{F}}$, which is Assumption 3. The formal incumbent's middle interval is switching down rather than staying informal by definition of its choice set. The three one-sided thresholds solve the same informal-continuation comparison with

different fees, $G_i(\hat{x}, n-\kappa) = \kappa + \beta(1-\lambda_i)c_i$ with $\kappa = 0$, c_i^e and c_{fi} for the informal incumbent, the entrant and the formal incumbent; since G_i is strictly increasing in $\hat{x}$ and non-decreasing in its second argument (Lemma B.3), the root is increasing in κ, which gives $\hat{x}_i^{\mathrm{X}} \le \hat{x}^{\mathrm{E}} < \hat{x}_f^{\mathrm{X}}$ under $0 \le c_i^e < c_{fi}$. The ordering $\hat{x}_f^{\mathrm{F}} < \hat{x}_e^{\mathrm{F}} < \hat{x}_i^{\mathrm{F}}$ follows from (B.12): thresholds are increasing in C_o, and $C_{f\text{-inc}} < C_{\mathrm{e}} < C_{i\text{-inc}}$ by Assumption 2. □

Proof of Proposition B.10. Each threshold is defined by an indifference condition $\Delta(\hat{x}^*, \cdot) = 0$ between two options. By the implicit function theorem, $\partial\hat{x}^*/\partial\xi = -(\partial\Delta/\partial\xi)/(\partial\Delta/\partial\hat{x})$, and $\partial\Delta/\partial\hat{x} > 0$ at an upward crossing. The numerators are given by Lemma B.5: for the formality margins, $\partial\Delta/\partial\phi_f = -a_f^*$, $\partial\Delta/\partial\phi_i = +a_i^*$, $\partial\Delta/\partial b_f = \mu_f k_f^* \ge 0$, $\partial\Delta/\partial b_i = -\mu_i k_i^* \le 0$. Using $\partial\Delta/\partial\hat{x} = A\hat{x}^{\gamma/(1-\gamma)}/(1-\gamma) = D_o/[(1-\gamma)\hat{x}^*]$ at the threshold (substituting $A\hat{x}^{*1/(1-\gamma)} = D_o$) delivers the log-derivatives in the first row. For the informal entry/exit margins the indifference is $G_i = \text{const}$, so $\partial\Delta/\partial\phi_i = -a_i^*$ and $\partial\Delta/\partial\hat{x} = \Psi_i' > 0$, giving the stated positive derivative (thresholds rise as informal finance becomes dearer); formal-side parameters do not enter, hence the zeros. For formal exit, the relevant comparison at low $\hat{x}$ under the ordering of Proposition B.7 is exit versus switching down, whose value is $U_i(\hat{x}, n - c_{fi})$; its derivatives in informal financial parameters are those of U_i (Lemma B.5), delivering the third row's informal-side entries, and formal-side parameters do not enter the comparison at all, hence its zeros. Off that ordering - if staying formal were the binding alternative at the exit margin - the cutoff would be $\mathcal{X}_f(n; \mathcal{K}_f)$ and easier formal terms would weakly lower it (≥ 0 in ϕ_f, ≤ 0 in b_f), which is the caveat recorded in the table note. □

Proof of Proposition B.11. $p = 1 - \Phi(z)$ with $z = (\ln\bar{\theta} - \ln\omega)/\sigma_\varepsilon$ and, from (B.21), $\partial \ln\bar{\theta}/\partial\xi = \frac{1-\nu}{\rho}\,\partial \ln\hat{x}_i^{\mathrm{F}}/\partial\xi$ for any financial parameter ξ. Chain-differentiate and substitute the log-derivatives from Proposition B.10; $\phi(\cdot) > 0$ delivers the signs. For (B.23): on the linear branches the threshold is the ratio of a linear numerator to a slope (Appendix B.12); the quotient rule with the root identity $N_i = D_i\hat{x}^*$ gives $\partial \ln\hat{x}_i^{\mathrm{F}}/\partial b_f = -\mu_f k_f^*/N_i$ and $\partial \ln\hat{x}_i^{\mathrm{F}}/\partial b_i = +\mu_i k_i^*/N_i$, and the same chain rule applies; the $(1-\gamma)$ factor is absent because the operative branch is linear in $\hat{x}$. □

Proof of Corollary B.13. *(i)* Under Proposition B.7 the binding entry margin and both exit margins compare an *informal* continuation with the outside value, so by Lemma B.5 they respond only to informal-side terms (second and third rows of Proposition B.10): all three thresholds fall when $\varepsilon_i > 0$ or $\mathrm{d}b_i > 0$, and total entry, the mass of entrants above $\hat{x}_i^{\mathrm{E}}$, rises.

(ii)-(iii) By Lemma B.5, total differentiation of the formality margin of origin o gives

$$\mathrm{d}\Delta_o = -a_f^*\,\mathrm{d}\phi_f + a_i^*\,\mathrm{d}\phi_i = \varepsilon_f a_f^* - \varepsilon_i a_i^*,$$

so $\hat{x}_o^{\mathrm{F}}$ falls, formality gains ground, if and only if $\varepsilon_f a_f^* > \varepsilon_i a_i^*$. Since $p_{i\to f}$ falls in $\hat{x}_i^{\mathrm{F}}$ while $p_{f\to i}$ rises in $\hat{x}_f^{\mathrm{F}}$, this is (B.24) at the $i \to f$ margin and, reversed, (B.25) at the $f \to i$

margin. For the borrowing gaps: at the $i \to f$ margin the formal option carries funds $n - c_{if}$ and the informal option n, so

$$a_f^* - a_i^* = \big[\tilde{k}_f - (n - c_{if})\big] - \big[\tilde{k}_i - n\big] = \tilde{k}_f - \tilde{k}_i + c_{if}.$$

For the entrant margin the fee pair (c_f^e, c_i^e) gives $a_f^* - a_i^* = \tilde{k}_f - \tilde{k}_i + c_f^e - c_i^e$, so the same argument applies to $\hat{x}_e^{\mathrm{F}}$.

Under Assumption 1, $T_\lambda > (R_f/R_i)^\gamma$ implies $T_\lambda R_i/R_f > (R_i/R_f)^{1-\gamma} > 1$, hence $\tilde{k}_f/\tilde{k}_i = \big(T_\lambda R_i/R_f\big)^{1/(1-\gamma)} > 1$: the gap is strictly positive, and $\varepsilon_f = \varepsilon_i$ delivers (B.24). At the $f \to i$ margin the fee switches sides, $a_i^* - a_f^* = \tilde{k}_i - \tilde{k}_f + c_{fi}$, giving the statement in (iii). □

Proof of Proposition B.14. (i) The entry threshold $\bar{\omega}$ falls by Proposition B.10 (second row), so the entry probability $1-F_\omega(\bar{\omega})$ rises for every k_0 with a constrained margin. (ii) For a log-normal ω the truncated mean $m(\bar{\omega}) \equiv \mathbb{E}[\omega \mid \omega \geq \bar{\omega}]$ satisfies $m'(\bar{\omega}) = h(\bar{\omega})\,[m(\bar{\omega}) - \bar{\omega}] > 0$, where $h = f_\omega/(1 - F_\omega)$ is the hazard and $m(\bar{\omega}) > \bar{\omega}$; hence a lower threshold strictly lowers the truncated mean, and first-order stochastic dominance of the truncated distributions delivers the same for every quantile and for average value added (an increasing transform of ω). (iii) Write the cohort shutdown rate as $S(\bar{\omega}) = \mathbb{E}[\pi^X(\omega) \mid \omega \geq \bar{\omega}]$ with π^X strictly decreasing (by (B.26) and $\partial k_2/\partial\omega \geq 0$). Then $S'(\bar{\omega}) = h(\bar{\omega})\big[S(\bar{\omega}) - \pi^X(\bar{\omega})\big] < 0$ because $\pi^X(\bar{\omega}) > \pi^X(\omega)$ for all $\omega > \bar{\omega}$ implies $\pi^X(\bar{\omega}) > S(\bar{\omega})$. A lower $\bar{\omega}$ therefore raises the cohort shutdown rate. □

Proof of Lemma B.15. With revenue $P\theta k^\alpha \ell^\nu$ the static objective is $(1-\tau_{y,s})P\theta k^\alpha \ell^\nu - (1+\tau_{\ell,s})w\ell$, which is the objective of Lemma B.1 with $P\theta$ in place of θ; its first-order condition and the substitution in the proof of Lemma B.1 go through unchanged, giving $\ell_s^*(P\theta, k)$ and $\pi_s = \Omega_s (P\theta)^{\frac{1}{1-\nu}} k^\gamma = \Omega_s P^{\frac{1}{1-\nu}} x k^\gamma$. Output is $y_2 = \theta_2 k^\alpha \ell_s^{*\nu}$; substituting $\ell_s^*(P\theta_2, k)$ gives $y_2 = \upsilon_s P^{\frac{\nu}{1-\nu}} \theta_2^{\frac{1}{1-\nu}} k^\gamma$, and $\mathbb{E}[\theta_2^{1/(1-\nu)} \mid \mathcal{I}_1] = \hat{x}$ by (B.4). The price enters the transition-stage problem (8), the values U_s and the differences Δ_o only through the product $P^{\frac{1}{1-\nu}}\hat{x}$ that multiplies $M_s k^\gamma$, so every statement that holds as a function of $\hat{x}$ holds as the same function of $P^{\frac{1}{1-\nu}}\hat{x}$, including the roots that define the cutoff values $\hat{x}_j$; a cutoff $\hat{x}_j$ is therefore reached at $\hat{x} = \hat{x}_j P^{-\frac{1}{1-\nu}}$ in unadjusted units. For an incumbent, with $\hat{x}$ from (B.4), the condition $P^{\frac{1}{1-\nu}}\hat{x} \geq \hat{x}_j$ reads $\rho \ln \theta_1 \geq (1-\nu)\ln \hat{x}_j - \ln P - (1-\rho)\ln \omega - \sigma_\varepsilon^2/(2(1-\nu))$, which compared with (B.21) gives the first line of (B.28). For an entrant, $\hat{x}_e(\omega) = \exp\{\ln \omega/(1-\nu) + \sigma_\varepsilon^2/(2(1-\nu)^2)\}$ turns the same condition into $\ln \omega \geq (1-\nu)\ln \hat{x}_j - \ln P - \sigma_\varepsilon^2/(2(1-\nu))$, which at $P = 1$ is the cutoff $\bar{\omega}_j$ of the statement and in general gives the second line of (B.28). □

Proof of Proposition B.16. (i) Define $H(P) \equiv \mathcal{P}(Q(P;\Xi)) - P$ on $(0,\infty)$. By Assumption 4(a)–(b), H is continuous and strictly decreasing, being the sum of a decreasing and a strictly decreasing function. As $P \to 0$, $Q \to 0$ and $\mathcal{P}(Q) \to \lim_{Q\to 0}\mathcal{P}(Q) \in (0,\infty]$, so

$H > 0$ for small P; as $P \to \infty$, $Q \to \infty$ and $\mathcal{P}(Q) \to 0$, so $H \to -\infty$. Hence H has exactly one zero, $P^*(\Xi)$.

(ii) Write the equilibrium condition as $F(P, \zeta) \equiv P - \mathcal{P}(Q(P; \Xi(\zeta))) = 0$. By Assumption 4(c), F is continuously differentiable with $F_P = 1 - \mathcal{P}'(Q)Q_P = 1 + |\mathcal{P}'(Q)|Q_P > 0$ and $F_\zeta = -\mathcal{P}'(Q)Q_\zeta = |\mathcal{P}'(Q)|Q_\zeta > 0$. The implicit function theorem gives $\mathrm{d}P^*/\mathrm{d}\zeta = -F_\zeta/F_P = -|\mathcal{P}'(Q)|Q_\zeta/(1 + |\mathcal{P}'(Q)|Q_P) < 0$; dividing by P^* gives (B.31). The equilibrium output response is $\mathrm{d}Q/\mathrm{d}\zeta = Q_\zeta + Q_P\,\mathrm{d}P^*/\mathrm{d}\zeta = Q_\zeta\big[1 - |\mathcal{P}'|Q_P/(1 + |\mathcal{P}'|Q_P)\big] = Q_\zeta/(1 + |\mathcal{P}'|Q_P)$, which is (B.32) and lies in $(0, Q_\zeta)$. For the elasticity form, with e_D and e_S as defined in (ii), substituting $|\mathcal{P}'| = P/(Qe_D)$ in (B.31) gives $\psi = (Q_\zeta/Q)/(e_D + e_S)$.

(iii) Differentiate (B.28) along the equilibrium path. The cutoff value in index units moves by the credit effect s_j per unit of ζ and $-\mathrm{d}\ln P^*/\mathrm{d}\zeta = \psi$ is the price effect, so $\mathrm{d}\ln\bar{\theta}_{1,j}/\mathrm{d}\zeta = \rho^{-1}[(1-\nu)s_j + \psi]$ for an incumbent and $\mathrm{d}\ln\bar{\omega}_j/\mathrm{d}\zeta = (1-\nu)s_j + \psi$ for an entrant.

(iv) Under the ordering of Proposition B.7, the three price thresholds are instances of the master threshold with the informal continuation as the alternative, (B.17), (B.18) and (B.19), so by Proposition B.10 their derivatives with respect to (ϕ_f, b_f) are zero and those with respect to (ϕ_i, b_i) are $a_i^*/\Psi_i' \geq 0$ and ≤ 0, both vanishing when the informal continuation does not borrow ($a_i^* = 0$ and $\mu_i = 0$ on R1 and R2). Along (B.30), the chain rule and Lemma B.5 give $s_j = -[a_i^*\chi_i^\phi\phi_i + \mu_i k_i^*\chi_i^b b_i]/(\Psi_i'\hat{x}_j) \leq 0$, which is zero in the two cases stated. By (iii) the cutoff rises if and only if $(1-\nu)s_j + \psi > 0$, that is, $\psi > (1-\nu)|s_j|$; when $s_j = 0$ only the price effect remains, ψ/ρ for the two exit cutoffs and ψ for the informal entry cutoff in ω units. Since the measures are atomless with positive densities, a higher exit cutoff raises the mass of incumbents below it, the exits, in each sector, and a higher informal entry cutoff lowers the mass of entrants above it.

(v) Under the dominance condition, the first row of Proposition B.10 and Corollary B.13(ii)–(iii) give $s_j < 0$ at the three credit thresholds, so by (iii) the equilibrium shift $(1-\nu)s_j + \psi$ is negative if and only if $\psi < (1-\nu)|s_j|$. Substituting the elasticity form of (ii), $\psi = (Q_\zeta/Q)/(e_D + e_S)$, the condition reads $e_D + e_S > (Q_\zeta/Q)/((1-\nu)|s_j|)$; it holds at the three credit thresholds at once when the sum of the elasticities exceeds the largest of the three right-hand sides, which is the one with $\min_j |s_j|$. □

C Shift-Share design

Table C.1: Firm-level responses to bank credit supply shocks

	Coefficient (Std. error)	Mean	N
Within-firm credit growth on bank shock†	0.131*** (0.021)	0.038	1,020,447
weighted by lagged credit	0.454*** (0.052)	−0.425	1,020,447
Firm credit growth on its banks' shocks	0.016*** (0.004)	0.035	3,704,022
two or more banks	0.092*** (0.013)	0.067	430,148
single bank	0.011*** (0.004)	0.030	3,272,491

Khwaja and Mian (2008) estimates on the annual firm-bank panel. Rows 1 and 2: log credit growth of a firm-bank pair on the bank's credit supply shock (log points), with firm-by-year fixed effects, among firms with two or more banks. The second row weights pairs by lagged credit. Rows 3 to 5: a firm's log credit growth on the standardized lagged-share-weighted average of its banks' shocks, with market-by-year fixed effects, among firm-years whose banks with a shock hold at least half of its lagged credit. The bank count classifies 3,702,639 of the 3,704,022 pooled firm years; the remaining 1,383 (0.04 percent) have no bank counted with positive lagged credit because the count and the exposure shares lag credit differently and enter the pooled row only. Standard errors clustered at the market level in parentheses, instead of the exposure-robust (AKM) standard errors of Adão et al. (2019), as the conservative choice in this design: for the IV coefficients the AKM standard errors are the smaller of the two. †Clustered by bank, the standard error is 0.045. Means of the dependent variable in each sample, weighted where the regression is. N: firm-bank-years in rows 1 and 2, firm-years in rows 3 to 5. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.2: Annual event study coefficients by horizon

		Credit		Interest rate (pp)		Collateral ratio	
h	Credit	Formal	Informal	Formal	Informal	Formal	Informal
−4	−0.082**	−0.097*	−0.077	0.168	−0.046	0.151***	0.071*
	(0.041)	(0.056)	(0.049)	(0.111)	(0.098)	(0.050)	(0.040)
−3	0.004	−0.039	0.016	−0.013	0.118**	0.090	0.074
	(0.027)	(0.024)	(0.026)	(0.042)	(0.052)	(0.058)	(0.045)
−2	−0.127***	−0.064	−0.132***	0.185**	−0.080	−0.031	0.101***
	(0.033)	(0.049)	(0.041)	(0.089)	(0.069)	(0.059)	(0.037)
−1	−0.043	−0.054	−0.006	−0.064	−0.103***	0.021	0.021
	(0.027)	(0.072)	(0.055)	(0.075)	(0.035)	(0.060)	(0.053)
0	0.140***	0.192***	0.076***	−0.274***	−0.021	−0.075	−0.075
	(0.031)	(0.043)	(0.030)	(0.096)	(0.110)	(0.131)	(0.084)
1	0.128***	0.200***	0.057	−0.172**	0.118*	−0.267***	−0.158
	(0.033)	(0.035)	(0.037)	(0.087)	(0.066)	(0.076)	(0.115)
2	0.220***	0.223***	0.158**	−0.262***	0.073	−0.208*	−0.193***
	(0.053)	(0.044)	(0.065)	(0.083)	(0.060)	(0.115)	(0.071)
3	0.227***	0.215***	0.162***	−0.133	0.029	−0.238**	−0.198***
	(0.051)	(0.044)	(0.061)	(0.129)	(0.066)	(0.118)	(0.068)
4	0.201***	0.176***	0.146**	−0.519***	−0.120	−0.289**	−0.138
	(0.052)	(0.044)	(0.066)	(0.130)	(0.121)	(0.124)	(0.089)

Coefficients plotted in Figures 5, 6, 7 and 8: reduced-form estimates on the annual market panel of the outcome h years after the shock on the IV, one regression per horizon and column, with the annual analogs of the main specification's controls, weighted by the market's lagged credit, on market-years with instrument coverage of at least 0.5. Standard errors clustered at the market level in parentheses, instead of the exposure-robust (AKM) standard errors of Adão et al. (2019), as the conservative choice in this design: for the IV coefficients the AKM standard errors are the smaller of the two. Stars from the corresponding two-sided tests. Credit columns in logs, interest rates in percentage points, the collateral ratio in logs. Market-years per regression range from 7,697 to 10,002 in the credit column, 5,859 to 7,840 in the formal credit and rate columns, 7,622 to 9,915 in the informal ones, 2,805 to 4,314 and 3,780 to 5,689 in the formal and informal collateral columns. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.3: Annual event study coefficients by horizon: additional outcomes

	Years since shock, h								
	−4	−3	−2	−1	0	1	2	3	4
Number of loans (log)									
Formal	0.36	1.45	−3.41	−1.65*	−0.80	−0.73	−0.07	0.01	0.52
	(0.82)	(1.18)	(2.50)	(0.95)	(0.83)	(0.96)	(0.99)	(1.00)	(1.04)
Informal	0.52	1.91	−4.18	−2.07	−3.22*	−3.28	−0.62	−0.12	0.06
	(1.00)	(1.54)	(2.73)	(1.38)	(1.95)	(2.07)	(3.26)	(2.81)	(3.27)
Number of borrowers (log)									
Formal	0.11	1.25	−3.26	−1.66*	−0.69	−0.93	−0.49	−0.31	0.17
	(0.87)	(1.10)	(2.50)	(0.98)	(0.74)	(0.81)	(0.82)	(0.82)	(0.88)
Informal	0.41	1.86	−4.20	−2.00	−3.23*	−3.22	−0.88	−0.50	−0.33
	(1.02)	(1.52)	(2.70)	(1.44)	(1.87)	(1.97)	(3.10)	(2.61)	(3.07)
Loans per borrower (log)									
Formal	−0.25	0.20	−0.15	0.02	−0.10	0.20	0.42*	0.31	0.35
	(0.18)	(0.17)	(0.10)	(0.14)	(0.18)	(0.22)	(0.25)	(0.26)	(0.30)
Informal	0.12	0.05	0.02	−0.07	0.02	−0.05	0.26	0.38	0.39
	(0.09)	(0.10)	(0.08)	(0.09)	(0.13)	(0.17)	(0.20)	(0.24)	(0.25)
Credit per loan (log)									
Formal	−10.06*	−5.31**	−3.03	−3.79	19.96***	20.74***	22.35***	21.48***	17.07***
	(5.58)	(2.55)	(4.84)	(7.70)	(3.74)	(2.85)	(3.88)	(3.80)	(3.70)
Informal	−8.25*	−0.28	−8.99***	1.44	10.86***	8.98***	16.38***	16.29***	14.53***
	(4.47)	(2.42)	(2.98)	(5.03)	(2.95)	(3.40)	(4.75)	(4.84)	(5.37)
Share of loans without collateral									
Formal	−0.06	0.13	−0.20	0.07	−0.79***	−0.50	−0.83***	−0.70**	−0.76**
	(0.10)	(0.19)	(0.13)	(0.14)	(0.30)	(0.30)	(0.30)	(0.33)	(0.34)
Informal	−0.14**	0.14	−0.05	0.07	−0.58**	−0.18	−0.38***	−0.26	−0.36*
	(0.07)	(0.12)	(0.06)	(0.07)	(0.23)	(0.17)	(0.14)	(0.17)	(0.20)
Formal sector, IMSS (log)									
Employment	0.19	0.33*	0.25	−0.07	0.10	0.09	0.53*	0.59*	0.22
	(0.16)	(0.19)	(0.16)	(0.17)	(0.18)	(0.21)	(0.32)	(0.32)	(0.43)
Mean wage	0.22**	0.24**	0.06	0.22***	0.01	−0.02	0.16	0.16	0.04
	(0.09)	(0.12)	(0.11)	(0.07)	(0.08)	(0.16)	(0.17)	(0.20)	(0.22)

Same specification as Table C.2, one regression per horizon and row. Coefficients and standard errors are multiplied by 100, so log outcomes read as percent and the share of loans without collateral in percentage points. Loans, borrowers, loans per borrower and credit per loan are market totals or ratios for loans to formal and to informal borrowers. The share without collateral is the share of loans with no posted collateral. Employment and mean wage are the market's IMSS-registered workers and their mean wage. Standard errors clustered at the market level in parentheses, instead of the exposure-robust (AKM) standard errors of Adão et al. (2019), as the conservative choice in this design: for the IV coefficients the AKM standard errors are the smaller of the two. Market years per regression range from 5,859 to 7,840 for the formal borrower series, 7,622 to 9,915 for the informal ones and 7,738 to 9,969 for employment and wages. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.4: Annual event study coefficients by horizon: informal minus formal differences

	Informal minus formal		
Years since shock, h	Credit	Interest rate (pp)	Collateral ratio
−4	0.018	−0.211	−0.077
	(0.082)	(0.166)	(0.081)
−3	0.055*	0.130**	−0.021
	(0.029)	(0.062)	(0.046)
−2	−0.069	−0.258***	0.128
	(0.071)	(0.091)	(0.080)
−1	0.051	−0.043	−0.009
	(0.118)	(0.087)	(0.088)
0	−0.117***	0.257*	0.009
	(0.045)	(0.141)	(0.156)
1	−0.147***	0.294**	0.111
	(0.047)	(0.125)	(0.109)
2	−0.069	0.344***	0.021
	(0.081)	(0.120)	(0.106)
3	−0.057	0.171	0.042
	(0.074)	(0.149)	(0.112)
4	−0.034	0.409**	0.155
	(0.075)	(0.166)	(0.107)

Coefficients plotted in the right panels of Figures 6, 7 and 8: reduced-form estimates on the annual market panel of the difference between the informal and the formal series of the market h years after the shock on the instrument, one regression per horizon and column, with the same specification as Table C.2. Standard errors clustered at the market level in parentheses, instead of the exposure-robust (AKM) standard errors of Adão et al. (2019), as the conservative choice in this design: for the IV coefficients the AKM standard errors are the smaller of the two. Stars from the corresponding two-sided tests. Credit in logs, interest rate in percentage points, the collateral ratio in logs. Market-years per regression range from 5,801 to 7,760 in the credit and interest rate columns and from 2,651 to 4,040 in the collateral column. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.5: Shift-Share IV estimates of the effect of bank shocks on the number of firms with credit

Dep. var.: Five-year log change, No. firms with credit			
	Total	Formal	Informal
$\widehat{\Delta_5 \ln(C)}$	0.035***	0.042**	0.034**
	(0.011)	(0.018)	(0.014)
	[0.486]	[0.262]	[0.748]
FS Coeff	1.113***	1.113***	1.113***
	(0.068)	(0.068)	(0.068)
FS F	265.81	265.81	265.81
Mean	776	185	591
Obs.	1,394	1,394	1,394
Markets	556	556	556

Notes: See Table 9 for specification and inference details. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.6: Census worker counts

	$\Delta_5 \ln$(Census employment)		
	Total	Formal	Informal
$\widehat{\Delta_5 \ln(C)}$	0.037	−0.071	0.074
	(0.082)	(0.193)	(0.204)
FS Coeff	−2.184	−2.261	−2.184
	(1.589)	(1.454)	(1.589)
FS F	1.89	2.42	1.89
Mean	51,646	5,050	46,819
Obs.	519	496	519
Markets	519	496	519

Notes: Census worker counts exist only in the 2019 and 2024 waves, leaving a single five-year window (a cross-section of markets, so clusters equal observations). The first stage on this subsample is weak (first-stage F between 1.9 and 2.4), so these estimates are reported for completeness only and are not informative. Exposure-robust (AKM) standard errors following Adão et al. (2019) in parentheses for the IV coefficients.

Tables C.7 to C.16 examine the robustness of the five-year results. We proceed in three steps. First, we test the identifying assumption directly. We check that the instrument is balanced on predetermined market characteristics, we control for lagged outcome changes, and we run placebo regressions on pre-period outcomes. Second, we vary the construction of the instrument. We drop the Leave-One-Out correction, renormalize by coverage, move the coverage threshold from zero to 0.8, and rebuild the instrument without the largest lender. Third, we vary the population and the estimator. We exclude the three metropolitan areas, compare the IV estimates with their OLS and reduced-form counterparts, and re-estimate everything at the bank level with re-identified shocks. The corresponding tables follow. One result survives every exercise: credit supply shifts the composition of firms toward formality without changing their number.

A central identifying assumption of our design is that the bank credit supply shocks aggregated by the Shift-Share instrument are as-good-as-randomly assigned conditional on the controls in our main specification. Following the balance tests suggested by Borusyak et al. (2022) and Goldsmith-Pinkham et al. (2020), we regress each of the seven predetermined market controls of the main specification on the Leave-One-Out Shift-Share instrument, conditional on year fixed effects and the coverage control and weighted by baseline market credit. Table C.7 reports the results with each control measured five and ten years before the baseline period, that is, ten and fifteen years before the outcome wave. None of the fourteen coefficients is statistically distinguishable from zero at the ten percent level (the smallest p-value is 0.12) and the point estimates are economically small. As a complementary dynamic check, Table C.8 re-estimates the main specifications adding the dependent variable's own lagged five-year change as a control. The core results are essentially unchanged, with the formal-firm response at 0.048 against 0.050 in the baseline; the exceptions are the aggregate and formal value-added responses, which reverse sign under this control and on which we therefore do not lean, while the value-added result we emphasize, the reallocation of production away from the informal sector, strengthens under it.

As a direct test of the identifying assumption that our Shift-Share instrument is uncorrelated with pre-existing dynamics in local outcomes, Table C.9 reports placebo IV estimates in which the dependent variable is the five-year log change of each outcome before the shock window, $\Delta_5 Y_{t-5} = \ln Y_{t-5} - \ln Y_{t-10}$, with the endogenous regressor, instrument, controls, weights, and sample condition unchanged. The answer is margin-specific. The stock of existing firms displays a positive pre-period correlation that is common to the two sectors rather than differential (formal 0.016, informal 0.019), whereas the main specification's estimated effects load entirely on formal firms. Moreover, the pre-period correlation is not what drives the estimates: adding the lagged outcome change to the main specification leaves the contemporaneous formal-firm effect essentially unchanged, 0.048 against 0.050.

Wages and value added attract significant negative placebo coefficients (-0.164 for average value added), which we read as mean reversion in bank lending: shocks were directed disproportionately toward markets whose financial and real conditions had been deteriorating, so the contemporaneous wage and value-added estimates partly capture recoveries and we do not lean on them. The entry, exit, and formality-transition placebos are uninformative, because the required further lag exists in a single census window and the first stage collapses (a first-stage F of about 1.9), and we report them for transparency rather than as evidence in either direction. The stock pre-period correlation does bear on the transition estimates of Table 9, which are log changes in the number of switching firms between consecutive windows and therefore include the growth of the origin stock over the previous window. That term is common to every destination from the same origin and cannot produce the pattern in the table: from a formal origin stayers rise and switchers fall, and from an informal origin switchers rise and stayers do not move. Nor can the rise in exit (Table 10) explain it by itself, since the exit of would-be switchers alone would leave the number of stayers unchanged.

A natural question is whether our findings depend on the Leave-One-Out construction of the Shift-Share instrument. Table C.10 re-estimates every outcome with the conventional construction, in which the bank shocks $\hat{g}_{bt}$ are estimated on the full sample including each market's own loans, so that $\hat{Z}_{mt} = \sum_b s_{bm(t-5)}\, \hat{g}_{bt}$ assigns every market the same shock value for a given bank; the sample, controls, weights, and inference are otherwise identical to the main specification. As Borusyak et al. (2022) emphasize, this construction suffers from own-observation overlap: each market's own credit dynamics enter its instrument, which pulls the IV estimate toward its OLS counterpart. Qualitatively nothing changes, and the modest deviations move exactly as the overlap predicts, toward the attenuated OLS counterparts: the formal-firm elasticity falls from 0.050 to 0.035, with a first stage of similar strength (F of 245 against 266). We therefore retain the Leave-One-Out instrument as our preferred specification: the comparison shows that the refinement, far from generating our results, if anything strengthens the composition margins.

A potential concern with our Leave-One-Out instrument is that its magnitude scales mechanically with coverage: in a market where only half of the base-period banks have an identifiable shock, the instrument is a coverage-weighted sum that is mechanically half as large as it would otherwise be. Table C.11 re-estimates the main specification with the renormalized instrument $\tilde{z}_{mt} = z_{mt}/\text{coverage}_{mt}$, which converts the coverage-weighted sum into a coverage-weighted average, standardized as in the main specification and estimated on the identical sample. The estimates are nearly indistinguishable from the main ones, with the formal-firm elasticity at 0.052 against 0.050 and a very strong first stage throughout (an F above 286 in every family); the coverage-driven scaling of the instrument drives no result in the paper.

Our main analysis restricts the estimation sample to market-windows for which the Leave-One-Out instrument is constructed from at least half of the local base-period banking exposure. Table C.12 re-estimates every outcome at the two extremes: a threshold of zero, which keeps every market-window with a constructible instrument (1,520 observations in 568 markets), and a threshold of 0.8, which keeps only market-windows where the instrument pools the great majority of local banking exposure (1,218 observations in 541 markets). The coverage requirement behaves like the measurement restriction it is: the first-stage F rises with the threshold, from about 269 to between 288 and 325, while the substantive conclusions do not move, with the formal-firm elasticity in a narrow band around the main estimate (0.053 at zero, 0.050 in the main specification, 0.045 at 0.8) and formal entry at 0.068 and 0.059 at the two extremes; the changes are confined to margins we already treat cautiously. The threshold matters for the precision of the instrument, not for any substantive conclusion of the paper.

A remaining concentration concern is that the identifying variation might come disproportionately from a single dominant lender: the shock-level diagnostics show an effective sample of roughly seventeen independent bank-window shocks, with the largest carrying 17 percent of the weight. Table C.13 therefore rebuilds the instrument from scratch excluding the bank with the largest share of aggregate baseline credit, removing both its shocks and its exposure shares, and re-estimates the main specification on the identical sample. The cost appears exactly where it should, in the first-stage F, which falls to about 35, while no conclusion changes: the formal stock rises by a similar magnitude as in the main specification (0.046 against 0.050, with standard errors of 0.010 and 0.013), turnover still runs through total and informal exits, and the formality transitions keep their signs with formalization and formal-to-informal if anything larger (0.088 and -0.074). What the largest lender contributes is first-stage strength, not the result.

The credit-weighted regression places disproportionate weight on the three largest metropolitan areas (Mexico City, Monterrey, and Guadalajara). To address the concern that the main estimates reflect dynamics specific to these urban banking centers, Table C.14 re-estimates every specification excluding them, nine of the 1,394 market-window observations. The composition core is essentially unchanged, with the formal-firm response at 0.049 against 0.050 and formal entry identical at 0.067, and several margins strengthen outside the metros (the employment response rises to 0.028 and the informal value-added decline deepens to -0.115). Two results soften and are disclosed as such: the informal-to-formal switching flow loses significance, with the composition shift operating instead through formal retention (0.055), fewer switches into informality (-0.049), and a declining count of informal stayers (-0.025), and the mean-wage decline disappears, indicating that the wage-composition effect of marginal hires is a metropolitan phenomenon. Dropping the largest markets therefore leaves the formalization core intact.

As a model-free check on our identification, Table C.15 reports, for every outcome, the naive OLS regression on credit growth and the reduced-form regression on the standardized Leave-One-Out instrument, with the same controls, weights, and sample restrictions as the main IV specification and standard errors clustered at the market level. The reduced form confirms that the identifying correlation is present in the raw data, with significance concentrated in the same margins the paper emphasizes (formal firms 0.056, s.e. 0.016; formal entry 0.075; informal exit 0.062), and the OLS-IV comparison shows why instrumenting matters: OLS is attenuated relative to IV on the composition margins (the formal stock at 0.025 against 0.050, formalization at 0.031 against 0.052), the pattern expected when demand shocks dominate observational credit variation.

As a complementary identification check, Table C.16 moves the analysis to the shock level. The credit supply shocks are re-identified with the cell-fixed-effects approach of Degryse et al. (2019), an extension of the within-firm framework of Khwaja and Mian (2008) in which credit demand is absorbed at the market-sector (LLM $\times$ industry) cell rather than within multi-bank firms. Both multi-bank and single-bank firms now contribute to the shock estimates, a consequential expansion in Mexico, where single-bank firms dominate local credit markets; the unit of observation is the bank-year, outcomes and market credit growth are aggregated across the markets each bank serves using baseline exposure shares, and the shocks, units, aggregation, controls, weights, and samples all differ from the market-level design, so nothing mechanical links the two sets of estimates: if our results reflected a particular scheme for purging demand, the own-market overlap, or fragile specification choices, this regression would be free to disagree.

It reproduces the composition of the response outcome by outcome: the total stock of firms in a bank's markets does not respond while the formal stock rises (0.036 against 0.050 at the market level), turnover again runs through informal exit, the formality transitions match sign by sign, and the fragile value-added responses are absent, consistent with our decision not to lean on them. That a design with different shocks, different units, different aggregation, and different inference reproduces the composition of the response, and not merely its direction, supports the credit supply interpretation of our results.

The formality composition result at the center of the paper is not the product of any single methodological choice. It survives the balance tests, lagged-outcome controls, and placebo scrutiny, in which the only pre-trends are common to the two sectors or attached to level outcomes we already treat cautiously. It survives replacing the Leave-One-Out construction with the conventional one, renormalizing by coverage, moving the coverage threshold from zero to 0.8, and rebuilding the instrument without the largest lender. It is reproduced outcome by outcome in a bank-level design with re-identified shocks. Sample composition matters only for precision, since the identifying power of the data resides where the credit is. In contrast, the value-added aggregates are the margins sensitive

to the specification, and the paper reads them, together with the smaller, though more stable, employment and wage responses, with caution, as secondary evidence. The main result stands: bank credit supply changes the composition of firm activity in Mexican local economies.

Table C.7: Balance test: Shift-Share instrument and predetermined market characteristics

	Lagged 5 years		**Lagged 10 years**	
Dependent variable	$\widehat{\beta}$	Std. error	$\widehat{\beta}$	Std. error
Employment share — Manufacturing	0.040	(0.026)	0.039	(0.027)
Employment share — Commerce	−0.008	(0.011)	−0.012	(0.012)
Bank credit share — Manufacturing	−0.007	(0.031)	0.027	(0.053)
Bank credit share — Commerce	0.010	(0.034)	−0.057	(0.061)
Credit-weighted average interest rate	0.323	(0.256)	0.756	(0.804)
Log number of banks	−0.039	(0.072)	−0.072	(0.069)
Log number of loans	0.205	(0.314)	−0.034	(0.315)

Notes: Each entry reports the coefficient from a separate regression of the indicated predetermined market characteristic on the standardized instrument, conditional on year fixed effects and the coverage share interacted with year fixed effects, weighted by baseline market credit. Standard errors clustered at the market level in parentheses, instead of the exposure-robust (AKM) standard errors of Adão et al. (2019), as the conservative choice in this design: for the IV coefficients the AKM standard errors are the smaller of the two. Lags are relative to the baseline period, so the two blocks measure each characteristic ten and fifteen years before the outcome wave. Sector shares are the manufacturing and commerce shares of market employment and of market bank credit. Services and other activities are the omitted group. Sample restricted to market-windows with instrument coverage ≥ 0.5. The five-year lag columns use 1,204–1,360 observations in 521–545 markets. The ten-year lag columns use 770–938 observations in 459–517 markets. None of the fourteen coefficients is statistically distinguishable from zero at the 10 percent level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.8: Pre-trend test: main estimates with the lagged outcome change as a control

	Main specification		**With pre-trend control**	
Dependent variable	$\widehat{\beta}$	Std. error	$\widehat{\beta}$	Std. error
Log existing firms — Total	0.004	(0.005)	0.003	(0.005)
Log existing firms — Formal	0.050***	(0.013)	0.048***	(0.013)
Log existing firms — Informal	−0.002	(0.006)	−0.002	(0.006)
Formal workers (IMSS)	0.020***	(0.005)	0.019***	(0.006)
Formal wage bill (IMSS)	0.024**	(0.009)	0.020**	(0.010)
Mean daily wage (IMSS)	−0.012***	(0.004)	−0.008**	(0.003)
Avg. value added	0.029*	(0.015)	−0.031*	(0.018)
Total value added	0.032*	(0.017)	−0.045**	(0.019)
Formal value added	0.028	(0.031)	−0.109***	(0.029)
Informal value added	−0.049**	(0.021)	−0.071***	(0.014)

Notes: The pre-trend columns re-estimate the main specification adding the dependent variable's own lagged five-year change as a control. The sample shrinks because the lagged change requires data two windows back. The main-specification columns use 1,386–1,394 observations in 552–556 markets. The pre-trend columns use 933–983 observations in 527–543 markets. Exposure-robust (AKM) standard errors following Adão et al. (2019) in parentheses. The main-specification columns repeat the corresponding entries of the results tables for comparison. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.9: Placebo tests: Shift-Share IV on pre-period (five-year) outcome changes

Pre-period outcome ($\Delta_5 Y_{t-5}$, log)	$\widehat{\Delta_5 \ln(C)}$	Std. error	First-stage F	N
Log existing firms — Total	0.016***	(0.005)	290.11	983
Log existing firms — Formal	0.016**	(0.008)	290.06	979
Log existing firms — Informal	0.019***	(0.006)	290.11	983
Log firm entry — Total	0.034	(0.114)	1.89	514
Log firm entry — Formal	−0.078	(0.296)	1.89	508
Log firm entry — Informal	0.048	(0.095)	1.89	514
Log firm exit — Total	−0.043	(0.120)	1.89	514
Log firm exit — Formal	0.146	(0.368)	1.89	495
Log firm exit — Informal	−0.081	(0.108)	1.89	514
Formal–Formal	−0.017	(0.077)	1.59	488
Informal–Formal	−0.112	(0.204)	1.89	496
Formal–Informal	0.075	(0.197)	1.89	485
Informal–Informal	−0.047	(0.127)	1.89	514
Firms with credit — Total	0.079	(0.058)	290.11	983
Firms with credit — Formal	0.091**	(0.037)	290.11	983
Firms with credit — Informal	0.006	(0.120)	290.11	983
Formal workers (IMSS)	−0.011*	(0.006)	290.13	970
Formal wage bill (IMSS)	−0.024***	(0.005)	290.13	970
Mean daily wage (IMSS)	−0.015***	(0.006)	290.13	970
Avg. value added	−0.164***	(0.014)	290.11	983
Total value added	−0.148***	(0.013)	290.11	983
Formal value added	−0.174***	(0.021)	290.11	983
Informal value added	−0.036*	(0.019)	290.11	983

Notes: Each row reports a placebo IV estimate in which the dependent variable is the five-year log change in the indicated outcome before the credit-supply shock window, $\Delta_5 Y_{t-5} = \ln Y_{t-5} - \ln Y_{t-10}$. The endogenous regressor, instrument, controls, baseline-credit weights, and sample condition follow the main specification (Table 10). The sample additionally requires a non-missing outcome ten years prior. Exposure-robust (AKM) standard errors following Adão et al. (2019) in parentheses. The first-stage F statistic is computed on each restricted sample. The entry, exit, and transition outcomes require lagged flows that exist in a single census window, so their samples halve and the first stage collapses. We report these rows for transparency only. The census-based worker outcomes cannot be estimated in the pre-period at all and are omitted. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.10: Shift-Share IV estimates using the conventional (Non-Leave-One-Out) instrument

Dependent variable	$\widehat{\Delta_5 \ln(C)}$	Std. error	N
Log existing firms — Total	0.002	(0.005)	1,394
Log existing firms — Formal	0.035***	(0.007)	1,389
Log existing firms — Informal	−0.002	(0.006)	1,394
Log firm entry — Total	0.007	(0.009)	996
Log firm entry — Formal	0.043***	(0.009)	979
Log firm entry — Informal	0.007	(0.010)	996
Log firm exit — Total	0.049***	(0.009)	996
Log firm exit — Formal	0.021	(0.018)	965
Log firm exit — Informal	0.059***	(0.008)	996
Formal–Formal	0.026***	(0.005)	970
Informal–Formal	0.049***	(0.019)	957
Formal–Informal	−0.043***	(0.012)	958
Informal–Informal	−0.003	(0.003)	996
Firms with credit — Total	0.024***	(0.007)	1,394
Firms with credit — Formal	0.018*	(0.011)	1,394
Firms with credit — Informal	0.030***	(0.009)	1,394
Formal workers (IMSS)	0.015***	(0.004)	1,386
Formal wage bill (IMSS)	0.014**	(0.007)	1,386
Mean daily wage (IMSS)	−0.009***	(0.003)	1,386
Avg. value added	0.016	(0.016)	1,394
Total value added	0.017	(0.016)	1,394
Formal value added	−0.001	(0.028)	1,394
Informal value added	−0.027**	(0.011)	1,394

Notes: Two-stage least squares estimates using the conventional (non Leave-One-Out) instrument, in which the bank shocks are estimated on the full sample including each market's own loans, standardized to unit variance. Specification, controls, baseline-credit weights, and sample condition follow Table 10. The estimation samples are identical to the corresponding leave-one-out regressions. Exposure-robust (AKM) standard errors following Adão et al. (2019) in parentheses. The first-stage F statistic ranges from 245 to 258. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.11: Robustness to instrument renormalization

Dependent variable	$\widehat{\Delta_5 \ln(C)}$	Std. error	N
Log existing firms — Total	0.004	(0.006)	1,394
Log existing firms — Formal	0.052***	(0.016)	1,389
Log existing firms — Informal	−0.002	(0.007)	1,394
Log firm entry — Total	0.016	(0.011)	996
Log firm entry — Formal	0.070***	(0.021)	979
Log firm entry — Informal	0.013	(0.011)	996
Log firm exit — Total	0.048***	(0.014)	996
Log firm exit — Formal	0.041**	(0.018)	965
Log firm exit — Informal	0.054***	(0.014)	996
Formal–Formal	0.042***	(0.011)	970
Informal–Formal	0.049**	(0.021)	957
Formal–Informal	−0.039**	(0.018)	958
Informal–Informal	−0.008	(0.005)	996
Firms with credit — Total	0.041***	(0.014)	1,394
Firms with credit — Formal	0.046**	(0.022)	1,394
Firms with credit — Informal	0.042**	(0.017)	1,394
Formal workers (IMSS)	0.021***	(0.007)	1,386
Formal wage bill (IMSS)	0.026**	(0.012)	1,386
Mean daily wage (IMSS)	−0.012**	(0.005)	1,386
Avg. value added	0.032*	(0.018)	1,394
Total value added	0.035*	(0.020)	1,394
Formal value added	0.033	(0.039)	1,394
Informal value added	−0.058**	(0.027)	1,394

Notes: Two-stage least squares estimates of the main specification with the renormalized instrument $\tilde{z}_{mt} = z_{mt}/\text{coverage}_{mt}$, which converts the instrument from a coverage-weighted sum of bank shocks into a coverage-weighted average, standardized as in the main specification. The estimation sample, controls, weights, and inference are identical to Table 10. Exposure-robust (AKM) standard errors following Adão et al. (2019) in parentheses. The first-stage F statistic ranges from 287 to 311. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.12: Sensitivity to instrument coverage threshold

	Coverage ≥ 0			**Coverage** ≥ 0.8		
Dependent variable	$\hat{\beta}$ (Std. error)	N	M	$\hat{\beta}$ (Std. error)	N	M
Log existing firms — Total	0.006 (0.005)	1,520	568	0.003 (0.005)	1,218	541
Log existing firms — Formal	0.053*** (0.013)	1,514	564	0.045*** (0.012)	1,213	537
Log existing firms — Informal	−0.000 (0.005)	1,520	568	−0.001 (0.006)	1,218	541
Log firm entry — Total	0.016* (0.009)	1,092	568	0.012 (0.009)	862	527
Log firm entry — Formal	0.068*** (0.017)	1,071	555	0.059*** (0.016)	846	516
Log firm entry — Informal	0.013 (0.009)	1,091	568	0.010 (0.010)	862	527
Log firm exit — Total	0.051*** (0.011)	1,092	568	0.047*** (0.012)	862	527
Log firm exit — Formal	0.037*** (0.014)	1,049	549	0.031** (0.013)	834	510
Log firm exit — Informal	0.057*** (0.012)	1,092	568	0.056*** (0.013)	862	527
Formal–Formal	0.038*** (0.008)	1,051	553	0.034*** (0.007)	837	513
Informal–Formal	0.064*** (0.018)	1,042	540	0.051*** (0.017)	827	500
Formal–Informal	−0.047*** (0.015)	1,040	541	−0.035** (0.014)	829	505
Informal–Informal	−0.004 (0.004)	1,091	568	−0.004 (0.004)	862	527
Firms with credit — Total	0.029*** (0.011)	1,520	568	0.026** (0.011)	1,218	541
Firms with credit — Formal	0.042** (0.018)	1,520	568	0.031* (0.017)	1,218	541
Firms with credit — Informal	0.022 (0.014)	1,520	568	0.023* (0.014)	1,218	541
Formal workers (IMSS)	0.020*** (0.005)	1,511	563	0.015*** (0.004)	1,211	538
Formal wage bill (IMSS)	0.025*** (0.009)	1,511	563	0.016** (0.007)	1,211	538
Mean daily wage (IMSS)	−0.013*** (0.004)	1,511	563	−0.011*** (0.004)	1,211	538
Avg. value added	0.025* (0.014)	1,520	568	0.025* (0.014)	1,218	541
Total value added	0.030* (0.016)	1,520	568	0.027* (0.016)	1,218	541
Formal value added	0.020 (0.030)	1,520	568	0.014 (0.028)	1,218	541
Informal value added	−0.040** (0.020)	1,520	568	−0.031* (0.018)	1,218	541

Notes: Two-stage least squares estimates of the main specification at alternative thresholds on instrument coverage. The left block includes all market-years with a non-missing instrument (`cov_z_5` ≥ 0). The right block restricts to market-years for which the instrument pools at least 80% of the local base-period banking exposure (`cov_z_5` ≥ 0.8). The headline specification in Tables 10–11 uses the intermediate threshold `cov_z_5` ≥ 0.5. All other specification details are unchanged. N denotes observations and M markets. The first-stage F statistic ranges from 269 to 293 in the left block and from 288 to 325 in the right block. Exposure-robust (AKM) standard errors following Adão et al. (2019) in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.13: Shift-Share IV estimates excluding the biggest bank

Dependent variable	$\widehat{\Delta_5 \ln(C)}$	Std. error	N
Log existing firms — Total	0.001	(0.005)	1,394
Log existing firms — Formal	0.046***	(0.010)	1,389
Log existing firms — Informal	−0.005	(0.007)	1,394
Log firm entry — Total	0.006	(0.011)	996
Log firm entry — Formal	0.056***	(0.012)	979
Log firm entry — Informal	0.003	(0.012)	996
Log firm exit — Total	0.044***	(0.015)	996
Log firm exit — Formal	−0.001	(0.016)	965
Log firm exit — Informal	0.059***	(0.015)	996
Formal–Formal	0.028***	(0.009)	970
Informal–Formal	0.088***	(0.018)	957
Formal–Informal	−0.074***	(0.022)	958
Informal–Informal	0.000	(0.004)	996
Firms with credit — Total	0.001	(0.012)	1,394
Firms with credit — Formal	0.005	(0.021)	1,394
Firms with credit — Informal	0.017	(0.016)	1,394
Formal workers (IMSS)	0.013*	(0.007)	1,386
Formal wage bill (IMSS)	0.009	(0.011)	1,386
Mean daily wage (IMSS)	−0.009*	(0.005)	1,386
Avg. value added	−0.002	(0.021)	1,394
Total value added	−0.002	(0.024)	1,394
Formal value added	−0.035	(0.043)	1,394
Informal value added	−0.026	(0.022)	1,394

Notes: Two-stage least squares estimates of the main specification with the instrument reconstructed to exclude the bank holding the largest share of aggregate baseline credit: both its shocks and its exposure shares are removed, and the resulting instrument is standardized as in the main specification. The estimation sample, controls, weights, and sample condition are identical to Table 10. Exposure-robust (AKM) standard errors following Adão et al. (2019) in parentheses. The first-stage F statistic ranges from 34.95 to 35.13, and its decline relative to the main specification reflects the excluded bank's substantial contribution to the identifying variation. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.14: Shift-Share IV estimates excluding the three largest metropolitan areas

Dependent variable	$\widehat{\Delta_5 \ln(C)}$	Std. error	N
Log existing firms — Total	−0.003	(0.006)	1,385
Log existing firms — Formal	0.049***	(0.014)	1,380
Log existing firms — Informal	−0.009	(0.007)	1,385
Log firm entry — Total	0.018	(0.015)	990
Log firm entry — Formal	0.067***	(0.019)	973
Log firm entry — Informal	0.014	(0.015)	990
Log firm exit — Total	0.041**	(0.018)	990
Log firm exit — Formal	0.067***	(0.018)	959
Log firm exit — Informal	0.037**	(0.018)	990
Formal–Formal	0.055***	(0.009)	964
Informal–Formal	0.014	(0.020)	951
Formal–Informal	−0.049**	(0.023)	952
Informal–Informal	−0.025***	(0.006)	990
Firms with credit — Total	0.114***	(0.020)	1,385
Firms with credit — Formal	0.065***	(0.021)	1,385
Firms with credit — Informal	0.143***	(0.024)	1,385
Formal workers (IMSS)	0.028***	(0.008)	1,377
Formal wage bill (IMSS)	0.036***	(0.014)	1,377
Mean daily wage (IMSS)	0.001	(0.004)	1,377
Avg. value added	0.022	(0.021)	1,385
Total value added	0.018	(0.019)	1,385
Formal value added	0.029	(0.038)	1,385
Informal value added	−0.115***	(0.031)	1,385

Notes: Same specification and inference as the main tables. Exposure-robust (AKM) standard errors following Adão et al. (2019) in parentheses. This sample excludes Mexico City, Monterrey, and Guadalajara. These markets account for nine of the 1,394 market-window observations of the main sample. The first-stage F statistic ranges from 82 to 102. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.15: Reduced-form: Shift-Share instrument and local labor market outcomes

	OLS on credit growth		**Reduced form on $\hat{Z}$**		
Dependent variable	$\widehat{\beta}$	Std. error	$\widehat{\beta}$	Std. error	N
Existing firms — Total	0.007	(0.005)	0.004	(0.012)	1,394
Existing firms — Formal	0.025**	(0.012)	0.056***	(0.016)	1,389
Existing firms — Informal	0.006	(0.006)	−0.002	(0.013)	1,394
Entry — Total	0.017*	(0.010)	0.017	(0.022)	996
Entry — Formal	0.033**	(0.016)	0.075**	(0.030)	979
Entry — Informal	0.017	(0.011)	0.013	(0.023)	996
Exit — Total	0.031***	(0.008)	0.055***	(0.011)	996
Exit — Formal	0.025**	(0.011)	0.043*	(0.025)	965
Exit — Informal	0.034***	(0.009)	0.062***	(0.011)	996
Informal→Formal	0.031*	(0.017)	0.059**	(0.027)	957
Formal→Informal	−0.010	(0.013)	−0.044*	(0.026)	958
Formal→Formal	0.027**	(0.014)	0.045***	(0.016)	970
Informal→Informal	0.004	(0.006)	−0.007	(0.013)	996
Total	0.008	(0.018)	0.039	(0.037)	1,394
Formal	0.025	(0.023)	0.046	(0.031)	1,394
Informal	0.016	(0.021)	0.038	(0.051)	1,394
Workers (IMSS)	0.014**	(0.006)	0.022*	(0.012)	1,386
Total wage bill	0.012	(0.010)	0.026	(0.019)	1,386
Mean daily wage	−0.010***	(0.003)	−0.013*	(0.007)	1,386
Average VA	0.034**	(0.014)	0.032	(0.029)	1,394
Total VA	0.041***	(0.016)	0.036	(0.030)	1,394
VA formal	0.032	(0.025)	0.031	(0.046)	1,394
VA informal	0.007	(0.028)	−0.055	(0.063)	1,394

Notes: Each row reports two separate regressions with the same controls, weights, and sample as the main specification: OLS of the outcome on the five-year log change in market credit, and the reduced form of the outcome on the standardized leave-one-market-out instrument. Standard errors clustered at the market level in parentheses, instead of the exposure-robust (AKM) standard errors of Adão et al. (2019), as the conservative choice in this design: for the IV coefficients the AKM standard errors are the smaller of the two. Each reduced-form coefficient equals the product of the corresponding first-stage and IV coefficients. Census worker counts (single-window subsample) are reported in Table C.6. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.16: Bank-level IV estimates: Khwaja-Mian-style identification at the bank-year

Dependent variable	$\widehat{\Delta_5 \ln(C)}$ (Std. error)	N
Log existing firms — Total	−0.004 (0.005)	131
Log existing firms — Formal	0.036*** (0.008)	131
Log existing firms — Informal	−0.009* (0.005)	131
Log firm entry — Total	−0.001 (0.008)	105
Log firm entry — Formal	0.045*** (0.011)	105
Log firm entry — Informal	−0.004 (0.009)	105
Log firm exit — Total	0.058*** (0.011)	105
Log firm exit — Formal	0.020 (0.016)	105
Log firm exit — Informal	0.070*** (0.010)	105
Formal–Formal	0.021*** (0.008)	105
Informal–Formal	0.068*** (0.013)	105
Formal–Informal	−0.077*** (0.015)	105
Informal–Informal	−0.003 (0.004)	105
Firms with credit — Total	−0.003 (0.011)	131
Firms with credit — Formal	0.003 (0.016)	131
Firms with credit — Informal	0.017 (0.010)	131
Formal workers (IMSS)	0.009 (0.007)	131
Formal wage bill (IMSS)	0.008 (0.012)	131
Mean daily wage (IMSS)	−0.015*** (0.004)	131
Avg. value added	0.022 (0.015)	131
Total value added	0.018 (0.017)	131
Formal value added	−0.006 (0.037)	131
Informal value added	−0.013 (0.017)	131

Notes: Two-stage least squares estimates at the bank-census-year level. The unit of observation is a bank-year. Market outcomes are aggregated to the bank level using the bank's baseline market exposure shares. The endogenous regressor is the bank-level aggregate of five-year market credit growth, instrumented by the bank's credit-supply shock identified from market-sector cell fixed effects following Degryse et al. (2019), an extension of Khwaja and Mian (2008). Controls are census-year fixed effects and the bank-level analogs of the two sector-share controls. Estimation is weighted by the bank's baseline credit. The sample keeps bank-years in which the aggregated outcome covers at least half of the bank's baseline credit, so the sample varies across outcomes. Standard errors clustered at the bank level in parentheses, since at the shock level conventional clustered standard errors are valid (Borusyak et al., 2022). The first-stage F statistic ranges from 30.40 to 34.91. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.