Guarantor Networks, Financial Contagion, and Systemic Risk: Evidence from Kenya

Paul Brimble, Kevin Carney and Joshua W. Deutschmann

24 July 2026

Abstract

Social ties underpin much of formal and informal finance in low- and middle-income countries. One concern about the use of social ties to secure loans is the potential for coordinated or strategic behavior to cause defaults to cascade. In this paper, we provide a novel characterization of the network defined by loan guarantees in a financial cooperative in Kenya. The network is highly interconnected, with most members forming a single large connected component. We then study how production shocks propagate through this network. We find that production shocks increase the likelihood of loan nonperformance both directly, for the shocked borrower, and indirectly, for borrowers connected to them through guarantee ties. Because the network is densely connected, these indirect effects aggregate to a sizeable share of the total short-term repayment impact of a shock. We complement these causal estimates with a model of financial contagion that distinguishes correlated shocks from network cascades, and use it to quantify the SACCO's exposure to systemic risk. These findings highlight a trade-off between expanding credit access through social collateral and containing systemic risk.

JEL codes: G21, G41, G51, O16
Keywords: Loan Guarantees, Financial Contagion, Household Finance

Brimble: University of Michigan, pbrimble@umich.edu (corresponding author); Carney: University of Michigan, kecarney@umich.edu; Deutschmann: University of Chicago, jdeutschmann@uchicago.edu. Julia Eigner, Cynthia Ji, Rathan Sudheer, Antony Wainaina, and Shengning Zhang provided exemplary research assistance. We are thankful for thoughtful feedback from Cynthia Kinnan, Michael Kremer, and Tanya Rosenblat; and from audience members at CSAE, DevWorkshop (U. Michigan) and WEFIDEV. We acknowledge financial support from the Development Innovation Lab, the USAID Feed the Future Lab for Markets, Risk, and Resilience, and Wellspring.

1 Introduction

Networks underpin access to formal and informal financial services. There is a long history of using social ties as a substitute for physical collateral or to screen borrowers within the context of joint-liability microfinance (Besley and Coate, 1995; Ghatak and Guinnane, 1999; Ghatak, 2000; Karlan et al., 2009; Cai et al., 2025), although there has been an overall decline in the use of traditional joint-liability microfinance (De Quidt, Fetzer and Ghatak, 2018). Within villages, social networks facilitate insurance against idiosyncratic shocks (Townsend, 1994). It is less well understood how social ties generalize beyond villages in financial contexts, and what the implications are for the resilience of financial institutions to shocks. Do the networks created by these financial relationships expose institutions to systemic risk?

We provide novel insights into the structure of financial networks using anonymized administrative data from a savings and credit cooperative (SACCO) in Kenya. Financial cooperatives, SACCOs, and credit unions are an important source of credit in low- and middle-income countries, including in sub-Saharan Africa. As of 2023, financial cooperatives in 12 countries in Africa had a loan portfolio exceeding $10 billion USD (World Council of Credit Unions, 2023). Of these, Kenya represented the largest share, with more than $4.5 billion USD in active loans. The SACCO we study has about 25,000 members and primarily serves dairy farmers in central Kenya.

We focus on the use of social ties and cooperative shares to guarantee loans. In Kenya, SACCOs formalize the use of social collateral to expand credit access using a system of shares and loan guarantees. Individuals become members in the cooperative by purchasing a minimum number of shares, which pay an annual dividend and have a cash value. Members can then pledge their shares as collateral for loans taken by other members. The cooperative typically allows shares to be pledged up to some multiplier of their value, such that 3-5 times the value of the shares can be made as loans.

We first describe the nature of the network defined by these guarantor relationships. For the period we study, the network implied by active guarantor relationships is large and well connected. At the start of our period of study (January 2018), the network contains 4,368 members, and the average member has about four active connections defined by loan guarantees. Roughly seven in ten members in the network hold a guaranteed loan, with the remainder participating purely as guarantors, and many members simultaneously guarantee others and are guaranteed themselves.

We focus on a set of long-term loans (with a duration of at least 12 months). These loans have an average principal amount of about KSH 84,000 (roughly 650 USD), and, where guaranteed, typically have two to three guarantors. Only 1% of loans go into default during the period we study, but about 44% experience at least one month of nonperformance.

Next, we explore how the structure of the networks defined by guarantor relationships exposes the financial institution to contagion risk. The size and prevalence of shocks and features of the network could exacerbate or mitigate shock propagation (Acemoglu, Ozdaglar and Tahbaz-Salehi, 2015; Elliott, Golub and Jackson, 2014). The resilience of SACCOs is an area of popular concern in Kenya following several high profile SACCO failures.

To identify how shocks propagate within networks, we leverage plausibly exogenous variation in the timing of production shocks caused by reductions in agricultural income to study how shocks propagate through these financial networks. We follow the empirical strategy of Kinnan et al. (2024) and define a harmonized set of loans that ever experience similar shocks. We compare loans with relatively earlier shocks against loans with relatively later shocks.

We first establish that production shocks directly increase the likelihood that a borrower's own loan falls into nonperformance. Using administrative data on the quantity of milk sold to the cooperative in a month, we define a shock as the start of a spell of at least three consecutive months with no milk delivered during a loan's planned repayment term. Following a production shock, the shocked borrower is about 17 percentage points more likely to be in loan nonperformance, a large increase relative to the pre-shock average. The increase emerges immediately after the shock and is sustained through the post-shock window.

Next, we test whether these production shocks propagate to the loan performance of *other* borrowers connected through the guarantor network. We measure each non-shocked borrower's exposure to a shock by their proximity to the shocked member in the pre-shock network. Borrowers who are more closely connected to a shocked member experience a larger post-shock increase in nonperformance: a borrower linked by a single direct guarantee is about 1.4 percentage points more likely to enter nonperformance. This propagation operates through the financial guarantee channel rather than through correlated production risk — a connected borrower's own milk production is unaffected when their counterpart is shocked. Combining the direct and propagation estimates implies a per-link contagion probability of approximately 0.08.

[In progress.] *An earlier version of this analysis distinguished propagation upstream (to the shocked member's guarantors) from propagation downstream (to those the shocked member guarantees). The directed-network results are being re-incorporated; the current draft reports the undirected, proximity-based propagation effect above.*

Finally, we use the size and structure of the network to gauge the aggregate implications for the SACCO. Because the network is densely connected, the many small per-link propagation effects sum to substantially more than the single direct effect, so that the bulk of a shock's aggregate short-term repayment cost operates through contagion rather than the direct hit to the shocked borrower. This suggests that contagion of shocks through guarantor networks has costly implications for the SACCO.

We complement the causal analysis with two additional contributions. First, we use dyadic regressions to document that members who are closer in the guarantor network have more correlated milk production and more correlated loan repayment outcomes. These correlations are consistent with both correlated shocks — nearby members facing similar production risks — and contagion — defaults cascading through the network. Second, we develop a model of financial contagion and systemic risk that separates these two amplification channels. The model combines correlated shocks with a network cascade process, and allows us to quantify the probability that the SACCO reaches regulatory thresholds for financial vulnerability under different scenarios.

We contribute to the literature on the structure and fragility of financial networks and the transmission of shocks through those networks (Acemoglu, Ozdaglar and Tahbaz-Salehi, 2015; Baqaee, 2018; Cabrales, Gottardi and Vega-Redondo, 2017; Elliott, Golub and Jackson, 2014; Elliott and Golub, 2022; Glasserman and Young, 2016; Jackson and Pernoud, 2021), which typically look at networks of financial institutions or networks of firms, and typically in higher-income settings. Kinnan et al. (2024) is one notable exception, exploring health shock propagation through networks defined by labor and supply chain relationships. To our knowledge, this is the first paper to characterize the structure and propagation implications of financial networks defined by loan guarantees within a financial cooperative.

We additionally link to a related but largely distinct literature on social networks and finance in low- and middle-income countries (Jackson, Rogers and Zenou, 2017). Karlan et al. (2009) study social ties and informal borrowing, with a focus on the strength of ties serving as social collateral. We contribute by studying the use of social ties in a more formal structure, in which we observe the direction and magnitude of loan guarantees. Our administrative data provide novel insights into a network which would otherwise be prohibitively costly to measure or would require proxy measures (Breza et al., 2020).

2 Context and Data

We study the guarantor network of a mid-sized Savings and Credit Cooperative (SACCO) in central Kenya. In this section, we describe the institutional setting, our data sources, and the key features of loans, repayment, and milk production that motivate our analysis.

2.1 Savings and Credit Cooperative Organisations in Kenya

The context of our study is a mid-sized Savings and Credit Cooperative (SACCO) in central Kenya, which serves about 25,000 members. SACCOs are a form of credit union that are popular in many countries in Africa. Like credit unions in the United States, many SACCOs were founded with their membership based in a particular occupation. The SACCO we partner with was founded by dairy farmers. While there are several SACCOs that operate in central Kenya, our partner is the primary SACCO that historically focused on the dairy industry. The SACCO has an affiliated dairy cooperative that serves its members as they sell their milk into the national supply chain. Dairy farmers who are SACCO members and sell to the dairy cooperative have their milk income deposited directly into their SACCO account. As a result, the SACCO and dairy cooperative serve as a one-stop shop for the financial affairs of many of their members, handling their income, saving, and borrowing.

The SACCO sector in Kenya is financially important, particularly for agricultural households. In 2023, SACCOs held approximately one trillion Kenyan shillings (7B USD) in total assets, equivalent to approximately 6.5% of the country's GDP, and held 758 billion Kenyan shillings (5.4B USD) in active loans (SASRA, 2024). According to Kenya's most recent FinAccess Household Survey, SACCOs are among the top sources of financing for agricultural households, surpassing traditional banks, MFIs, and moneylenders in terms of the fraction of farmers who turn to them

as a primary source of credit (Central Bank of Kenya, 2021).

Like other forms of credit unions around the world, SACCOs are member owned. In order to use the full suite of SACCO financial services, including longer-duration loans, individuals must purchase shares in the SACCO to meet a minimum threshold. In the case of the SACCO we study, this threshold is equivalent to about 40 USD. In order to take a longer-duration (longer than 3 months) loan from the SACCO, members pledge some of their own shares as a deposit, and solicit loan guarantees from others to pledge to cover the remaining balance of the loan. For most loans, borrowers pledge 25–33 percent of the loan value themselves, although the SACCO does in some cases offer asset-collateralized loans, which do not require external guarantors and instead uses the asset being financed as collateral (Jack et al., 2023).

The SACCO's regulator, the SACCO Societies Regulatory Authority (SASRA), sets prudential requirements for SACCOs to remain in good standing. Among these, SASRA treats a non-performing loan (NPL) ratio of no more than 5% as the acceptable maximum; ratios above this threshold place a SACCO under heightened regulatory scrutiny. We describe the full set of SASRA concern levels, which we use to anchor our measure of institutional vulnerability, in Section 6.

2.2 Data

We use three main sources of data for our analysis. First, we use loans data which consists of important characteristics of all loans issued by the SACCO. This dataset includes start dates, the number of instalments and amount of each loan, as well as details about guarantors associated with each loan, including their identity and the amount being guaranteed. Second, we use ledger data which consists of all repayments and accrued interest or fees for all loans issued by the SACCO. This dataset allows us to track loan performance, including repayment behaviour and defaults throughout the duration of the loan. Third, we use milk production data which contains monthly quantities of milk delivered to the dairy cooperative, as well as milk income paid to the members. This data allows us to identify which SACCO members are primary dairy farmers, as well as to identify milk production shocks.

We have data on loans from 2012 to 2023, but the ledger is only available from 2018 onwards. The milk production data is available from 2013 to mid-2021. Therefore, we focus primarily on the period of time from 2018 to 2023 for our analysis. The SACCO issues a range of loans from 1-month advances to multi-year long-term loans. We focus throughout on long-term loans — those with monthly instalments of 12 months or more. We do this for two main reasons. First, long-term loans account for the majority of the lending from the SACCO; between 2018 and 2023, KSH 2.7 billion was lent to members, of which approximately KSH 1.9 billion was through long-term loans. Second, long-term loans are the ones that rely on guarantors, which is our object of interest.

Table 1 presents summary statistics for members, loans, repayment, and milk production.

2.3 Members

Panel A of Table 1 presents characteristics of the roughly 23,000 members who appear in our data (the SACCO has about 25,000 members in total). Over our study period (2018–2023), 38% of members take out at least one loan. Conditional on borrowing, the median member takes out 2

loans and borrows a total of KSH 225,000 on average.

The guarantee system creates a dense web of financial obligations across the membership. Throughout this period, 34% of members received a guarantee on their own loan, 31% provided a guarantee for another member's loan, and 21% had a reciprocal guarantee relationship — simultaneously guaranteeing and being guaranteed by the same member. This extensive overlap between borrowing and guaranteeing is central to the contagion mechanism we study: a member who experiences a shock may simultaneously affect loans they have guaranteed for others and trigger collection from their own guarantors. As this is a dairy SACCO, 79% of members are milk producers, providing the production data we use to identify income shocks.

2.4 Loans and Guarantors

Panel B of Table 1 presents characteristics of long-term loans issued between 2018 and 2023. These loans typically range from 12 to 48 months in duration, with a median of 24 months. The average loan amount is approximately KSH 84,000 (roughly 650 USD).

When taking out a loan, the amount is capped based on the borrower's own personal savings — i.e., their shares in the financial cooperative. Each borrower may take out a loan up to three times the amount of their shares, and must — in general — find other members of the SACCO to act as guarantors and pledge their own shares to cover the remaining amount of the loan. We find that 71% of long-term loans are guaranteed. Conditional on having a guarantor, borrowers typically have 2–3 guarantors who collectively cover on average 66% of the loan value. This is to be expected if borrowers take out the maximum amount to which they are entitled: their own shares cover only one-third of the loan, and thus the guarantors are responsible for covering the remaining two-thirds.

2.5 Loan Performance, Repayment and Defaults

Loans are designed to be paid off monthly, with a proportion of the repayment covering the interest, and the remaining going towards the principal. The SACCO's regulator, the SACCO Societies Regulatory Authority (SASRA), classifies loans by days in arrears: Performing (0 days), Watch (1–30 days), Substandard (31–180 days), Doubtful (181–360 days), and Loss (360+ days). A loan is considered performing as long as the borrower is behind by no more than half of a monthly repayment amount; completely missing a payment immediately moves the loan to Watch status. During the Watch period, the SACCO sends SMS notifications but the loan is still considered performing during the period. At the Substandard stage, the loan is now considered to be non-performing and the SACCO escalates to calls, visits, and freezing of guarantor accounts to encourage repayment. At the Doubtful stage, the SACCO begins recovering funds from the borrower's deposit shares and guarantor pledges. These regulatory classifications are relevant for our analysis of systemic risk in Sections 6 and 7, where we consider the share of non-performing loans as a measure of institutional vulnerability.

Panel C of Table 1 presents statistics on loan performance. Defaults are very rare in this setting, accounting for only 1% of the sample, in part because the SACCO is usually able to recover savings from both the borrower and their guarantors to prevent outright default. Of the remaining

loans, 66% were completed on time, 12% were completed late, and 22% are still ongoing.

However, temporary nonperformance is common. Up to 44% of loans are classified as non-performing at some point during their repayment period (this is a lower bound, as some of the ongoing loans may still be relatively new). Conditional on experiencing at least one month of non-performance, an average of 28% of the loan period is spent in non-performing status. Therefore, while defaults are a relatively minor concern in this setting, loan delinquencies in the form of nonperformance — resulting in delayed loan completion, administrative costs, and potential recovery actions against guarantors — are relatively common. Moreover, loan nonperformance is the key metric used by SASRA to evaluate the SACCO's viability.

2.6 Milk Production

Panel D in Table 1 presents statistics on monthly milk production from 2018 until mid-2021 for all milk-producing members using administrative data on milk quantities delivered to the dairy co-operative. When milk is delivered, members on average produce just under 200 litres per month. This is roughly the expected monthly output from one cow, consistent with the small-scale dairy farming that characterises our setting.

Members have positive milk production for only 59% of the months in our study period. Some of this reflects natural dry-off periods of 1–2 months when cows are not producing. However, we find that at least 73% of milk producers experience a production shock, defined as having no milk production for at least 3 consecutive months. Among farmers who experience such a shock, the typical member experiences only one, which lasts for approximately one year. We focus on quantities rather than revenue to avoid picking up price shocks, which are more likely to stem from aggregate market fluctuations than from individual idiosyncratic events.

3 Guarantor Networks

We now turn to the structure of the networks defined by these guarantor relationships. We first describe how we construct the network from our administrative data, then characterize the network's structure and evolution over time, and finally introduce the proximity measures that we will use in our empirical analysis.

3.1 Network Construction

We construct a guarantor networks from the set of active loans and their guarantor-borrower relationships. We define these networks at the monthly level — the time frame at which loans are paid — so that new links are added when a new loan is initiated and removed when a loan is paid off. We thus obtain 72 month-year snapshots from January 2018 to December 2023.[1]

Our primitive object is the *directed* network $\mathcal{D}_t$. For each ordered pair of members (i, j) at month-year t, the directed edge $d_{ijt} = 1$ if member i guarantees a loan held by member j, and $d_{ijt} = 0$ otherwise. The direction follows the money: $d_{ijt} = 1$ when i is the guarantor and j is the

[1]While the ledger data are available from 2018 onwards, they include pre-existing loans with data going back to 2013. The networks are therefore constructed using all loans active in a given month, regardless of when they originated.

borrower. We focus primarily on *unweighted* networks, in which edges take values in $\{0, 1\}$.[2] For simplicity, a member i appears as a node in $\mathcal{D}_t$ if and only if they are involved in at least one active guarantee relationship at month-year t. The network is time-varying $\mathcal{D}_t$ as members enter and exit guarantee relationships as new loans are initiated and existing loans are paid off.

From $\mathcal{D}_t$ we derive two further networks. The *undirected* network $\mathcal{U}_t$ has edge indicator $u_{ijt} = \max\{d_{ijt}, d_{jit}\}$, equal to 1 whenever members i and j are linked by at least one guarantee relationship in either direction. The *reciprocal* network $\mathcal{R}_t$ has edge indicator $r_{ijt} = d_{ijt} \cdot d_{jit}$, equal to 1 if and only if i and j guarantee each other. The reciprocal network captures the strongest form of pairwise commitment, in which both members guarantee one another. By construction, $u_{ijt} = u_{jit}$ and $r_{ijt} = r_{jit}$, while d_{ijt} is asymmetric in general (i.e., $d_{ijt} \neq d_{jit}$).

3.2 Network Structure

The guarantor network grows over our study period alongside the SACCO's expanding membership. Table 2 presents quarterly statistics on the directed network $\mathcal{D}_t$ and its undirected counterpart $\mathcal{U}_t$, covering network size, connectivity, and structure. In January 2018, $\mathcal{D}_t$ contains 4,368 nodes connected by 9,709 directed guarantee relationships, of whom 3,024 are active borrowers. The number of nodes exceeds the number of borrowers because some members participate only as guarantors — they pledge their shares to secure others' loans but do not hold active long-term loans themselves. At that point, 1,344 members serve exclusively as guarantors; by October 2021, this figure rises to 1,749. Overall, $\mathcal{D}_t$ grows to 5,755 nodes and 13,334 directed edges by October 2021, with 4,006 active borrowers. This growth reflects rising loan demand over the period, as new borrowers enter the network and recruit guarantors from among the existing membership.

Figure 1 visualizes $\mathcal{U}_t$ for January 2018. Each node represents a member, with larger nodes indicating higher degree (more guarantee relationships). Darker edges represent reciprocal guarantee ties (those in $\mathcal{R}_t$), while lighter edges represent non-reciprocal ties. The figure illustrates the dense, interconnected structure of the network: nearly all members are embedded in a single large cluster with no visible isolated subgroups.

Despite this growth, the network's key structural properties remain remarkably stable. The average degree in $\mathcal{U}_t$ — the number of distinct guarantee relationships per member — holds steady at approximately 4 throughout the period, ranging from 3.96 in January 2018 to 4.10 in October 2021. Reciprocity — the share of directed edges in $\mathcal{D}_t$ that are also in $\mathcal{R}_t$ — remains around 0.21–0.23, indicating that roughly one in five guarantee relationships is mutual. Clustering coefficients in $\mathcal{U}_t$ are low, averaging approximately 0.11. The clustering coefficient measures how often a member's guarantors also guarantee each other. A low clustering coefficient means that guarantee relationships do not form tight local groups — a member's guarantors are typically spread across different parts of the network rather than concentrated in a close-knit clique. This matters for contagion: because connections are dispersed, shocks are not contained within local clusters but

[2]Each guarantee relationship also carries a monetary value — the amount pledged by the guarantor — so that d_{ijt} can equivalently be defined as the guarantee amount, yielding a weighted version of the network; the corresponding weighted proximity measures are described in Appendix B.

can instead propagate to distant parts of the network. The stability of these structural properties suggests that as the network expands, new members form connections at a similar rate and pattern to existing members.

A defining feature of the network is that $\mathcal{U}_t$ forms essentially one large connected component. The largest connected component of $\mathcal{U}_t$ contains over 96% of all network members in every quarter, rising from 97% in January 2018 to 98% in October 2021. This means that nearly all members who participate in the guarantor network are reachable from one another through chains of guarantee relationships. The average shortest path length within this component is approximately 7–8, meaning that any two members are typically separated by around 6–7 intermediary guarantee connections. Combined with the near-universal membership in the giant component and the low clustering, these statistics indicate that a shock to any single member can, in principle, reach the vast majority of the network through a relatively small number of steps — a feature with important implications for the potential scope of contagion.

The structure of this network differs markedly from the village-level social networks that are more commonly studied in the development economics literature. Village networks are typically characterized by distinct, loosely connected clusters (Banerjee et al., 2013; Jackson, Rogers and Zenou, 2017). In contrast, the SACCO guarantor network forms a single, densely interconnected component spanning the entire membership, with low clustering and short path lengths connecting distant members. This high connectivity increases the potential scope for financial contagion: whereas shocks in a village network might remain localized within a cluster, shocks in this network can propagate to dispersed parts of the network through relatively short chains of guarantee relationships.

3.3 Network Proximity

A central object in our empirical analysis is the proximity between pairs of members in the guarantor network. Table 3 reports the distribution of dyad-level proximity in $\mathcal{U}_t$ at quarterly intervals. The number of borrowers grows from roughly 3,800 in 2018 to over 5,700 by 2023, generating between 7.4 and 16.4 million unique dyads. Of these pairs, between 29.8% and 36.6% are *unconnected* — they belong to separate components of $\mathcal{U}_t$ or one member is not in $\mathcal{U}_t$ at all (column 4).[3] For the remaining (connected) pairs, we focus primarily on proximity in the undirected network $\mathcal{U}_t$. An important feature of our setting is that network proximity is time-varying. As new loans are initiated and existing loans are repaid, the network and the corresponding distances evolve.

We use two complementary measures of proximity. The first, shortest-path *distance*, is the standard graph-theoretic measure of how many guarantee relationships separate two members. The second, effective *resistance*, is a construct from the electrical-network and random-walk literature that aggregates information across *all* paths connecting two members — not just the shortest — and is, to our knowledge, novel to the analysis of financial guarantor networks. We also intro-

[3]The bulk of these unconnected pairs involves at least one borrower who is not in the guarantor network (e.g., borrowers who did not rely on guarantors); a small residual reflects the few network members who lie outside the largest connected component.

duce their inverses respectively: *closeness* and *conductance*.

3.3.1 Distance and Closeness

Let $Distance_{ijt}$ denote the length of the shortest path from i to j in the undirected network $\mathcal{U}_t$, with $Distance_{ijt} = \infty$ if no such path exists.[4] This is the textbook graph-theoretic measure of proximity: members at $Distance_{ijt} = 1$ are linked by a single guarantee relationship, those at $Distance_{ijt} = 2$ are separated by one intermediary, and so on. We then define its inverse:

$$Closeness_{ijt} = \begin{cases} 1/Distance_{ijt} & \text{if } Distance_{ijt} < \infty, \\ 0 & \text{otherwise.} \end{cases}$$

Directly connected members have $Closeness_{ijt} = 1$, members separated by one intermediary have $Closeness_{ijt} = 0.5$, and unconnected members have $Closeness_{ijt} = 0$. We additionally define binary distance indicators $\mathbb{1}\{Distance_{ijt} = k\}$ for $k = 1, 2, 3, \ldots$, which equal one if members i and j are separated by exactly k steps in $\mathcal{U}_t$, and zero otherwise. These indicators allow us to estimate non-parametric effects at each distance, rather than imposing the functional form implied by the inverse-distance measure.

The distance bins in Table 3 give a sense of the network's spread. Direct guarantee links are strikingly rare: pairs at $Distance_{ijt} = 1$ account for only about 0.1% of dyads. The mass of connected pairs sits much further out — 8.4–10.8% at distances 2–5, 46.3–50.9% at distances 6–9 (the modal range), and 6.2–11.5% at distances of 10 or more. Appendix Table A1 reports the full per-distance breakdown for readers who want finer granularity than the coarse bins shown here.

3.3.2 Resistance and Conductance

Shortest-path distance has an important conceptual limitation in our setting: it uses only the single most direct route between i and j and discards information about all other paths through the network. Yet contagion in a guarantor network can plausibly spread through any of the (potentially many) chains of guarantee obligations connecting two members — not just the shortest one. Two pairs with identical $Distance_{ijt}$ can be exposed to very different cascade risks if one is linked through several parallel chains while the other is linked through only one.[5] We adopt the effective resistance distance (Klein and Randić, 1993), which connects our analysis directly to the electrical-network and random-walk literatures.

Treating each guarantee relationship in the undirected network $\mathcal{U}_t$ as a unit resistor and applying the series–parallel rules of circuit analysis, the effective resistance between i and j aggregates

[4] Analogous distances can be defined on the other two networks introduced in Section 3.1: a directed shortest-path distance $Distance^{\mathcal{D}}_{ijt}$ on $\mathcal{D}_t$ that follows the direction of guarantee edges from guarantor to borrower, and a reciprocal shortest-path distance $Distance^{\mathcal{R}}_{ijt}$ on $\mathcal{R}_t$. Of these, $Distance^{\mathcal{D}}_{ijt}$ is asymmetric in general ($Distance^{\mathcal{D}}_{ijt} \neq Distance^{\mathcal{D}}_{jit}$), and the pair $(Distance^{\mathcal{D}}_{ijt}, Distance^{\mathcal{D}}_{jit})$ allows us to distinguish downstream from upstream proximity — a distinction we exploit in the propagation analysis of Section 5.4.

[5] A family of network proximity measures captures exactly this multiplicity by aggregating information across edge-disjoint paths of different lengths between two nodes — examples include path-count and walk-count summaries, random-walk hitting times, harmonic distance (Acemoglu, Ozdaglar and Tahbaz-Salehi, 2015), and Katz-style discounted walk counts.

the conductance of every path connecting them.[6] By the series–parallel rules of circuit analysis, the inverse of the resistance admits a path-count decomposition. Let P_d^{ijt} denote the number of edge-disjoint paths of length d between i and j in $\mathcal{U}_t$. Then

$$\frac{1}{Resistance_{ijt}} \approx \sum_{d \geq 1} \frac{P_d^{ijt}}{d},$$

where each path contributes inversely to its length and parallel paths add. Two properties make this construction natural for our setting. First, $Resistance_{ijt} = 1$ corresponds exactly to a pair linked by a single direct guarantee edge with no other parallel paths — a single unit resistor — in which case the resistance distance and the shortest-path distance coincide. Second, by the parallel rule, two members linked by multiple paths receive a strictly lower resistance than those linked by a single path of the same length, capturing the intuition that path redundancy raises exposure to cascades. Like shortest-path distance, $Resistance_{ijt}$ is symmetric on $\mathcal{U}_t$ and equals ∞ for unconnected pairs — by construction, the same set of dyads that are unconnected under *Distance*. Bringing this electrical-network construct into the analysis of financial guarantor networks is, to our knowledge, novel.

We introduce the inverse of resistance and adopt the convention from the electrical-engineering literature and call it *conductance*:

$$Conductance_{ijt} = \begin{cases} 1/Resistance_{ijt} & \text{if } i \text{ and } j \text{ are in the same component of } \mathcal{U}_t, \\ 0 & \text{otherwise.} \end{cases}$$

For unconnected dyads, $Conductance_{ijt} = 0$ and it equals one for a pair linked by a single direct guarantee. The resistance bins in Table 3 give a sense of how path multiplicity shapes effective proximity. Among connected pairs, 3.9–6.6% have $Resistance_{ijt} \in [0, 0.5)$, 27.7–33.8% have $Resistance_{ijt} \in [0.5, 1)$, 12.4–16.4% have $Resistance_{ijt} \in [1, 1.5)$, 8.4–11.0% have $Resistance_{ijt} \in [1.5, 2)$, and 6.0–10.8% have $Resistance_{ijt} \geq 2$. The striking feature is that 32–40% of all dyads have $Resistance_{ijt} < 1$ even though only about 0.1% are linked by a direct guarantee edge. Because a single direct edge alone yields $Resistance_{ijt} = 1$, the only way a non-directly-linked pair can fall below this benchmark is through path multiplicity — multiple edge-disjoint chains of guarantee relationships running in parallel between i and j, each contributing to the conductance between them. The bulk of the network's effective proximity therefore comes not from direct ties but from the dense web of indirect routes that the resistance distance is designed to capture. The distribution is remarkably stable over time, consistent with the steady-state structural properties documented in Table 2.

4 Empirical Evidence on Correlated Outcomes in Guarantor Networks

Before turning to the causal analysis of financial contagion, we first document the relationship between network proximity and correlated outcomes. If guarantor networks are a channel through

[6] $Resistance_{ijt} = L_t^+(i,i) + L_t^+(j,j) - 2\,L_t^+(i,j)$, where $L_t = D_t - \mathcal{U}_t$ is the graph Laplacian of $\mathcal{U}_t$ (with D_t the diagonal degree matrix) and L_t^+ is its Moore–Penrose pseudoinverse.

which shocks propagate, we should expect members who are closer in the network to have more correlated production and repayment outcomes. In this section, we use dyadic regressions to test for these correlations, exploiting the large scale and time-varying structure of the guarantor network.

4.1 Empirical Framework for Correlated Outcomes

Our unit of observation is a borrower pair $\{i, j\}$ observed at month-year t. The sample consists of all dyads of members who hold an active long-term loan at any point during our study period. Because our undirected and reciprocal proximity measures are symmetric, we include each pair only once. Our main specification includes pair fixed effects and time fixed effects:

$$y_{ijt} = \alpha \, Closeness_{ijt} + \mu_{ij} + \lambda_t + \epsilon_{ijt} \tag{1}$$

where y_{ijt} is a symmetric time-varying outcome for the pair, $Closeness_{ijt}$ is the undirected network proximity between members i and j at month-year t as defined in Section 3.3, μ_{ij} are pair fixed effects, and λ_t are time fixed effects. The pair fixed effects absorb all time-invariant characteristics of the dyad, including geographic proximity, shared occupation, kinship ties, and any stable social relationships that may have led the two members to be close in the guarantor network. Identification of α therefore comes entirely from *within-pair* variation in network proximity over time. Such variation arises because the guarantor network evolves as loans are initiated and repaid. When a third member k takes out a new loan guaranteed by i, and k also has a guarantee relationship with j, the shortest-path distance between i and j may decrease — even though neither i nor j changed their own borrowing or guaranteeing behavior. Conversely, when intermediary loans are repaid and guarantee links dissolve, previously close pairs may become more distant.[7]

Even under pair fixed effects, the coefficient α admits two interpretations. First, it may reflect *correlated shocks*: when i and j become closer in the network through a new intermediary k, this intermediary may also expose i and j to more similar production risks — for instance, if k farms in the same area or faces the same market conditions as both i and j. The pair fixed effects rule out time-invariant sources of correlated shocks (such as i and j being in the same village or climate zone), but they cannot rule out time-varying common exposure introduced through network changes. Second, it may reflect *contagion*: when i and j become closer, a default by one member can more easily cascade to the other through the chain of guarantee relationships. Separating these two channels — correlated shocks and contagion — requires the identification strategy based on exogenous production shocks developed in Section 5.

As an alternative to the continuous $Closeness_{ijt}$, we also estimate equation 1 with binary dis-

[7] In Appendix C, we report an analogous specification with member fixed effects ϕ_i and ϕ_j in place of the pair fixed effects μ_{ij}. Member fixed effects absorb each individual's time-invariant characteristics — average default propensity, location, herd size, and farming ability — but leave several pair-level selection channels in ϵ_{ijt}, including: (i) geographic and kinship proximity between i and j specifically; (ii) shared occupation, common input suppliers, or correlated market exposure between the two; (iii) pre-existing social ties that led i and j to enter into a guarantee relationship in the first place; and (iv) any other time-invariant determinant of who chooses to guarantee whom. The gap between the member-fixed-effects coefficient and the pair-fixed-effects coefficient measures the contribution of these selection channels to the raw correlation between proximity and correlated outcomes.

tance indicators $\mathbb{1}\{Distance_{ijt} = k\}$ for $k = 1, 2, 3, 4\text{–}5, 6\text{–}7$, which allow for non-linear effects of network distance without imposing the inverse-distance functional form. Given the dyadic structure of the data, we cluster standard errors at the level of both members to account for the non-independence of observations that share a member (Cameron, Gelbach and Miller, 2011).

4.2 Joint Loan Performance and Network Proximity

We next examine whether members who are closer in the guarantor network are more likely to experience simultaneous loan nonperformance. We define the outcome y_{ijt} as an indicator for whether both members i and j are in nonperformance in month-year t, conditional on both having active loans. Because loan nonperformance status is well-defined at a month-year frequency, we can exploit the full month-year panel.

We estimate equation 1 with pair fixed effects, identifying α from within-pair changes in network proximity over time. A positive coefficient indicates that when two members become closer in the guarantor network (e.g., through a new intermediary guarantee relationship), they become more likely to experience simultaneous nonperformance. Because pair fixed effects absorb all time-invariant characteristics of the dyad, this coefficient cannot be explained by selection on geographic proximity, shared occupation, kinship ties, or other stable pair-level traits.

Column 1 of Table 4 reports the estimates. Pairs that are closer in the guarantor network are more likely to be in simultaneous nonperformance: the coefficient on *Closeness* is positive (Panel A), and the binary-distance specification (Panel B) shows that the association is concentrated among the most proximate pairs — directly linked pairs and those separated by one or two intermediaries — and attenuates toward zero at greater distances, relative to the unconnected omitted category.

The correlation between network proximity and joint nonperformance is consistent with two channels: correlated shocks (members who are close in the network face similar production risks) and contagion (the default of one member causally increases the likelihood of default for connected members). Separating these two channels requires the identification strategy developed in Section 5 and the model in Section 6.

4.3 Production Comovement and Network Proximity

We first examine whether members who are closer in the guarantor network have more correlated milk production. Correlated production could arise because geographically proximate farmers face common weather and disease shocks, because economically linked farmers share inputs or information, or simply because farmers with similar characteristics select into guarantee relationships with one another.

We measure production comovement using two approaches. First, we construct a binary indicator for whether both members i and j have positive milk production in the same month. Second, we compute the correlation of monthly milk quantities within each quarter for each pair, producing a quarterly measure of production comovement.

We estimate equation 1 with production comovement as the outcome variable. The coefficient α captures the extent to which closer members in the guarantor network have more correlated

production, net of any time-invariant pair-level selection. This is a necessary condition for correlated shocks to be a meaningful amplification channel, and provides empirical grounding for the correlated-shocks parameter ρ in the model developed in Section 6.

[In progress.] *Estimates of production comovement by network proximity (Table 4, columns 2–3) are being finalised.*

4.4 Robustness

As a robustness check, we re-estimate the closeness specification while successively dropping the most proximate dyads — first those at distance 1, then those at distances 1 and 2, and so on — so that identification comes only from changes in proximity among more distant pairs, where within-pair variation in the network is plausibly more exogenous.

[In progress.] *Results of this robustness exercise are being finalised.*

5 Empirical Evidence on Financial Contagion in Guarantor Networks

The correlated outcomes in Section 4 are consistent with both correlated shocks and contagion but cannot separate the two channels. To isolate contagion — the causal propagation of repayment difficulties through the guarantor network — we use plausibly exogenous variation in idiosyncratic milk-production shocks, closely following the identification strategy of Kinnan et al. (2024).

5.1 The Contagion Parameter

The central object of this section is the per-link contagion probability q: the probability that borrower i becomes nonperforming when a single guarantee counterpart j becomes nonperforming, in a configuration where i and j are connected only by the direct guarantee edge between them. Let $Y_i \in \{0, 1\}$ denote loan i's nonperformance status, and let $\Delta Y_i = 1$ indicate that loan i enters nonperformance (i.e., Y_i transitions from 0 to 1). Recall from Section 3.3 that P_d^{ij} is the number of edge-disjoint paths of length d between i and j in $\mathcal{U}_t$. The per-link contagion probability is then:

$$q \;\equiv\; \Pr\big(\,\Delta Y_i = 1 \,\big|\, \Delta Y_j = 1,\, u_{ij} = 1,\, P_d^{ij} = 0 \;\forall\, d \geq 2\,\big). \tag{2}$$

The conditioning event states that i and j share a direct guarantee edge ($u_{ij} = 1$) and that no path of length two or more connects them through any intermediary — the only path between i and j is the direct edge itself. This provides a very clean interpretation of the contagion, especially given the level of connectedness in the network.

Our strategy proceeds in two stages after outlining the empirical framework (Section 5.2). First, we estimate the direct effect of a production shock on the shocked borrower's own loan performance (Section 5.3). Second, we estimate the propagation effect of the same shock to non-shocked borrowers' loan performance which are differentially exposed due to different proximity (Section 5.4). Combining these two estimates, we are then able to estimate the per-link contagion parameter $\hat{q}$ (Section 5.5).

5.2 Empirical Framework for Financial Contagion

For each loan in the harmonised set — loans whose borrower ever experiences a milk-production shock during the study window — we define the shock period as beginning in the first month of a spell of at least three consecutive months with zero milk delivery during the loan's planned term (the first such spell, where a loan has more than one), consistent with the production-shock definition in Section 2.6. The sample is restricted to regular producers, defined as borrowers with positive milk deliveries in the three months before the loan starts. We then compare loans whose shock occurs early in the planned term (treated loans) against loans whose shock occurs later (control loans), before and after the event. For control loans, the event period is the placebo event corresponding to the borrower's true shock period minus a constant offset Δ equal to the midpoint of the planned remaining loan term.

We construct a balanced event-time panel covering nine months around each shock: three pre-shock months (one quarter) and six post-shock months (two quarters). The asymmetric window reflects the typical duration of a milk-production shock. Figure 2 (discussed below) shows that production falls sharply at the shock and remains depressed throughout the first post-shock quarter before recovering by the second quarter; the six-month post-window therefore covers both the production disruption itself and the subsequent period over which a shocked borrower's repayment difficulties may propagate to connected counterparts. The three-month pre-window is sufficient to verify parallel pre-trends. Throughout, we use $\tau \in \{-3, -2, -1, 0, 1, 2, 3, 4, 5\}$ to denote event time relative to the shock (in months) and t to denote calendar time. The omitted reference period in event-study specifications is $\tau = -1$.

5.2.1 Direct Effects

We first estimate the effect of a production shock on the shocked borrower's own loan performance, following the difference-in-differences design of Kinnan et al. (2024). We classify shocked loans by the timing of the shock within the loan's planned term: shocks in the first four months are excluded (too little pre-period to assess trends); "treated" loans are those shocked in the first half of the remaining term; and "control" loans are those shocked in the second half. We match each treated loan shocked at event time t to control loans whose placebo shock occurs Δ months later, and identify the direct effect from the before-versus-after change in nonperformance for treated loans relative to controls.

The dynamic specification for direct effects is

$$y_{it} \;=\; \sum_{\substack{k=-3 \\ k\neq -1}}^{5} \alpha_k \, \mathbb{1}\{\tau = k\} \times Shock_i \;+\; \sum_{\substack{k=-3 \\ k\neq -1}}^{5} \gamma_k \, \mathbb{1}\{\tau = k\} \;+\; \mu_i + \delta_t + \delta_t \times Degree_{is} + \epsilon_{it}, \qquad (3)$$

where y_{it} is the loan-performance outcome for loan i in calendar month t, $Shock_i = 1$ for treated loans and zero for controls, μ_i are loan fixed effects, δ_t are calendar-month fixed effects, and $\delta_t \times Degree_{is}$ flexibly absorbs differential time trends by the borrower's out-degree and in-degree at the start of loan i, where the subscript s throughout this section denotes a quantity evaluated at the start of loan i. The direct-effect coefficients of interest, α_k, trace out the dynamic effect of the shock

on the treated relative to the control group across event time. For a single-coefficient summary, we collapse the dynamic specification into a static post-shock indicator:

$$y_{it} \;=\; \alpha\, Post_{it} \times Shock_i + \gamma\, Post_{it} + \mu_i + \delta_t + \delta_t \times Degree_{is} + \epsilon_{it}, \tag{4}$$

where $Post_{it} = \mathbb{1}\{\tau \geq 0\}$. The direct-effect coefficient α is the average post-shock change in y for treated loans relative to controls. Standard errors are clustered at the loan level.

We apply three sample restrictions. We exclude (i) loans that were already completed in the month prior to the shock (no scope for further repayment behaviour); (ii) loans that were already in nonperformance in the month prior to the shock (the outcome is at the ceiling); and (iii) loans whose borrower has zero degree in the guarantor network at loan start. The third restriction aligns the direct sample with the propagation sample, ensuring comparability across the two regressions.

5.2.2 Propagation Effects

We measure non-shocked borrower i's exposure to j's shock by their proximity in the pre-shock guarantor network — the conductance $Conductance_{ijs}$ defined in Section 3.3, which equals zero for unconnected pairs and one for a pair joined by a single direct guarantee. Using the same difference-in-differences design, we identify propagation from the differential post-shock change in nonperformance between more- and less-exposed non-shocked loans, comparing equally well-connected loans by flexibly controlling for degree.

The propagation analysis asks whether shocks to a borrower j propagate to a non-shocked borrower i's loan performance via the guarantor network. We pair each shocked loan j with all active non-shocked loans i at the time of j's shock; the unit of observation is the (non-shocked-loan, shocked-loan, calendar-month) triple, indexed by (i, j, t). The non-shocked loans include both loans whose borrower experiences a shock later in the study window and loans from members for whom no shock data are available. Following the direct-effects design, we form a balanced nine-month event-time panel with $\tau \in \{-3, \ldots, +5\}$ around each shock event j.

We apply three sample restrictions to the non-shocked loans to ensure clean identification. A pair (i, j) is excluded if loan i: (i) is within four months of starting or finishing at the time of j's shock; (ii) is more than two-thirds of the way to its planned completion at the time of j's shock; or (iii) experiences its own shock within four months of j's. Restriction (i) avoids contamination from beginning- and end-of-loan dynamics; (ii) avoids loans with insufficient remaining time for propagation to materialise; (iii) avoids attributing i's own production shock to the cascade.

We measure i's exposure to j's shock using $Conductance_{ijs}$, the conductance between i and j on the undirected network evaluated at the start of loan i. Using the network at the loan's start avoids endogenous post-shock changes in network structure. By construction, $Conductance_{ijs} = 0$ for unconnected pairs and $Conductance_{ijs} = 1$ for pairs linked by a single direct guarantee edge with no other connecting paths. Coefficient estimates on $Post_{jt} \times Conductance_{ijs}$ are therefore interpreted as the change in i's loan-performance outcome from being unconnected to j to being connected by a single direct guarantee link — the contrast that identifies the per-link contagion parameter q defined in equation (2).

The dynamic specification is

$$y_{ijt} \;=\; \sum_{\substack{k=-3 \\ k\neq -1}}^{5} \beta_k \mathbb{1}\{\tau = k\} \times Conductance_{ijs} \;+\; \gamma\, Conductance_{ijs} \;+\; \mu_i + \omega_j + \theta_\tau + \delta_t + \delta_t \times Degree_{is} + \epsilon_{ijt}, \tag{5}$$

with non-shocked-loan fixed effects μ_i, shocked-loan fixed effects ω_j, event-time fixed effects θ_τ, calendar-month fixed effects δ_t, and time-varying degree controls $\delta_t \times Degree_{is}$. The propagation-effect coefficients of interest are β_k. The static counterpart is

$$y_{ijt} \;=\; \beta\, Post_{jt} \times Conductance_{ijs} + \gamma\, Conductance_{ijs} + \mu_i + \omega_j + \theta_\tau + \delta_t + \delta_t \times Degree_{is} + \epsilon_{ijt}. \tag{6}$$

Standard errors are two-way clustered at the non-shocked-loan (i) and shocked-loan (j) levels.

5.2.3 Coefficients of Interest

In our setting, α captures the direct effect of a shock on own loan nonperformance. The interpretation of β is more nuanced: it captures the effect on the probability of loan nonperformance of i from a one-unit increase in $Conductance_{ijs}$ between i and the shocked borrower j. Since a single guarantee edge between j and i corresponds to $Resistance_{ijs} = 1$, a one-unit increase corresponds to moving from being unconnected to the shocked borrower j to being connected by a single guarantee edge—so β is precisely the reduced-form per-link effect on i's nonperformance. Finally, because a shock does not cause nonperformance in all directly shocked borrowers, we recover the structural per-link contagion parameter—the probability of default propagation conditional on the counterpart's nonperformance—as $q = \beta/\alpha$.

5.3 Direct Effects of Shocks

We focus on *loan nonperformance* (an indicator that the loan is in Substandard status, the SACCO's measure of repayment distress) as our primary outcome, with *milk production* (litres delivered in a given month) serving as a first-stage check on the shock measure.

Figure 2 plots the event study with milk production as the outcome. Pre-shock coefficients are close to zero, providing no evidence of differential pre-trends between treatment and control loans. At $\tau = 0$, milk production falls sharply for treated farmers, and remains depressed throughout the first post-shock quarter ($\tau \in \{0, 1, 2\}$) before recovering by the second post-shock quarter ($\tau \in \{3, 4, 5\}$). This recovery profile motivates the asymmetric event window introduced in Section 5.2: a one-quarter pre-shock window is sufficient for parallel-trends inspection, while a two-quarter post-shock window covers both the production disruption and the subsequent period in which downstream repayment effects can materialise. Panel A of Appendix Table A2 reports the average effect: milk production falls by 73 litres per month over the six post-shock months ($p < 0.01$), off a pre-shock mean of 145 litres.

Panel A of Table 5 reports average direct effects on loan nonperformance, estimated from equation (4). The four columns progressively tighten the sample to align with the propagation specification. Column 1 includes all shocked loans regardless of pre-shock status; column 2 excludes loans already completed in the month prior to the shock; column 3 additionally excludes

loans already in nonperformance prior to the shock; and column 4 — our preferred sample — additionally restricts attention to loans whose borrowers also enter the propagation sample (non-zero degree in the guarantor network at loan start). The estimated effect of a production shock on the borrower's own loan nonperformance ranges from 16.6 percentage points in column 1 to 17.4 percentage points in column 4 (off a pre-shock mean of 7.7%), and is statistically significant at the 1% level in every column.

Figure 3 plots the corresponding event study using the column-4 sample. Pre-shock coefficients are close to zero and statistically insignificant. Following the shock, the probability of nonperformance rises sharply and remains elevated throughout the post-shock window: in contrast to milk production, repayment difficulties do not unwind by the second post-shock quarter, consistent with a shock that triggers persistent loan-classification changes.

Panel A of Appendix Table A2 reports analogous estimates for two further repayment outcomes. Production shocks raise the probability that a loan enters *Loan Watch* — the SACCO's pre-nonperformance early-arrears category — by 24.8 percentage points, and reduce the probability that a loan is *Finished* (paid off) over the post-shock window by 8.2 percentage points (both significant at the 1% level). Both estimates are consistent with the headline nonperformance result and confirm that production shocks delay repayment across the SACCO's full classification of loan health.

5.4 Propagation Effects through Guarantor Networks

We now ask whether shocks to a borrower j propagate to a non-shocked borrower i's loan performance via the guarantor network. The outcome is again loan nonperformance for borrower i; the treatment intensity is i's conductance to the shocked borrower j, computed at the start of loan i (Section 3.3). As a placebo for the correlated-shock channel, we re-estimate the specification with i's own milk production as the outcome.

Panel B of Table 5 reports estimates from equation (6) on loan nonperformance. The four columns vary the inclusion rule for connected pairs by maximum shortest-path distance, focusing inference on progressively more local neighbourhoods of the shocked member. Column 1 is the most permissive specification, retaining all connected pairs; column 2 excludes pairs at shortest-path distance 10 or more; column 3 additionally excludes pairs at distance 8 or more; and column 4 — our preferred sample — additionally excludes pairs at distance 6 or more, restricting inference to the local neighbourhood within which contagion is most plausible. The coefficient on $Post_{jt} \times Conductance_{ijs}$ ranges from 1.5 percentage points (column 1) to 1.4 percentage points (column 4), significant at the 5% level in every column. The column-4 estimate is the average increase in nonperformance probability for a non-shocked borrower moving from being unconnected to the shock to being connected to it by a single direct guarantee link.

Figure 4 plots the corresponding event study using the column-4 sample. Pre-shock coefficients are close to zero and statistically insignificant. Following the shock, nonperformance rises sharply and the effect is approximately constant throughout the post-shock window, consistent with a one-shot cascade rather than one that accumulates gradually.

The placebo test rules out correlated production risk as the channel. Panel B of Appendix Table A2 reports the propagation effect on the non-shocked member's own milk production, and Figure A1 plots the corresponding event study. If propagation operated through correlated production shocks — for instance, if i and j farm in the same area and face the same weather — we would expect i's own milk production to fall when j is shocked. Instead, the post-shock coefficients are economically small and statistically indistinguishable from zero, exactly as we would expect if propagation operates through the financial channel of guarantee obligations rather than through correlated income shocks.

Panel B of Appendix Table A2 also reports propagation effects on two further repayment outcomes. Production shocks raise the probability that a connected non-shocked borrower's loan enters *Loan Watch* by 1.5 percentage points (significant at the 5% level), mirroring the direct-effect pattern. The propagation effect on *Loan Finished* is negative but statistically insignificant, suggesting that connected borrowers may delay repayment without yet falling out of the SACCO's classification of loan health. Together, these results provide direct evidence that production shocks propagate through guarantor networks via the financial channel: connected borrowers experience repayment difficulties even though their own production is unaffected, ruling out correlated production risk as the operative channel and pointing instead to the cascade of guarantee obligations.

5.5 Estimates of the Contagion Parameter

The direct and propagation estimates combine into an estimate of the per-link contagion probability q defined in equation (2). From Panel A column 4 of Table 5, the direct effect of a production shock on the shocked borrower's own nonperformance is $\hat{\alpha} = 17.4$ percentage points; from Panel B column 4, the propagation effect on a directly-connected non-shocked borrower's nonperformance is $\hat{\beta} = 1.4$ percentage points. These two coefficients map to q via $\hat{q} = \hat{\beta} \,/\, \hat{\alpha} \approx 0.078$.

Conditional on a shocked member experiencing nonperformance, each directly connected non-shocked counterpart experiences nonperformance at $\hat{q}$ times that rate through the guarantee channel alone. We estimate $\hat{q}$ jointly on the stacked direct- and propagation-effect samples, which delivers a delta-method standard error on $\hat{q}$ that incorporates the joint covariance of $\hat{\alpha}$ and $\hat{\beta}$ under two-way clustering at the shocked- and non-shocked-loan levels.

Although $\hat{q}$ is well below one, the breadth of the network means that the aggregate cost of a shock is dominated by propagation rather than by the direct effect. Because each shocked borrower is connected — directly or indirectly — to many other borrowers, the many small per-link propagation effects sum to substantially more than the single direct effect.

[In progress.] *The back-of-the-envelope aggregate decomposition — the expected number of nonperformance events per shock, and the share attributable to the direct versus the propagation channel — is being finalised.*

In short, even with a per-link contagion rate well below one, the dense web of guarantee relationships in this SACCO means that propagation accounts for the overwhelming bulk of the aggregate repayment cost of a production shock. Network contagion through guarantor ties is, in

this setting, a meaningful source of aggregate repayment risk for the financial institution — even though no single shocked borrower passes her distress on with certainty.

6 Modelling Financial Contagion and Systemic Risk

The empirical evidence in the preceding sections documents two facts: loan outcomes are correlated across the guarantor network (Section 4), and production shocks propagate causally through guarantee ties (Section 5). These facts are consistent with two distinct amplification mechanisms: correlated shocks that simultaneously distress nearby borrowers, and contagion that cascades defaults through the network. In this section, we develop a model that incorporates both mechanisms and allows us to quantify their respective contributions to systemic risk.

The model has two key parameters: a correlated-shocks parameter ρ that governs the clustering of initial defaults, and a contagion parameter β that governs how those defaults spread through the network. Taking the observed network and the distribution of non-performing loans as given, we use the model to simulate the distribution of aggregate default outcomes and assess the probability that the SACCO reaches regulatory vulnerability thresholds.

6.1 Model Setup

Consider N homogenous borrowers, indexed by $i \in \{1, \ldots, N\}$. Borrowers are connected via an undirected guarantor network $\mathcal{U}$, where $u_{ij} = 1$ if borrowers i and j are connected by a guarantor relationship, and $u_{ij} = 0$ otherwise.

Each borrower draws an initial latent repayment shock x_i from a continuous distribution F, and their loan becomes nonperforming whenever x_i exceeds a threshold $\bar{x}$ defined by $\Pr(x_i > \bar{x}) = p$ — equivalently, $\bar{x} = F^{-1}(1-p)$. The parameter p is the unconditional probability that any given loan becomes nonperforming due to a repayment shock, so $\bar{x}$ is calibrated to deliver that target rate. Absent contagion, borrower i nonperforms whenever their latent financial stress exceeds this threshold:

$$Y_i^0 = \mathbf{1}\{x_i > \bar{x}\}.$$

We interpret Y^0 as the vector of *fundamental nonperformance*, driven only by repayment shocks. To capture contagion, we allow the nonperformance of connected borrowers to increase the likelihood that borrower i also nonperforms. We model contagion as an *independent cascade* with per-link contagion probability $q \in [0, 1]$. The parameter q has a transparent interpretation: it is the probability that a single nonperforming borrower would, on her own, cause a directly-connected performing counterpart to also nonperform. To formalise this, we associate with each guarantee edge $\{i, j\}$ in $\mathcal{U}$ an independent Bernoulli(q) transmission draw τ_{ij}, with $\tau_{ij} = 1$ indicating that nonperformance would propagate across this edge if either endpoint were to nonperform. Starting from the fundamental defaults $Y^{(0)}$, the cascade then proceeds in discrete rounds: borrower i enters nonperformance in round $m+1$ if any already-nonperforming borrower is connected to her through a transmitting edge:

$$Y_i^{m+1} \;=\; \max\Big\{ Y_i^{(m)},\; \mathbf{1}\Big\{ \exists\, j : u_{ij}\, \tau_{ij}\, Y_j^m = 1 \Big\} \Big\}. \tag{7}$$

The $\max\{\cdot\}$ operator makes nonperforming absorbing: once a borrower nonperforms, they remain in nonperformance for the rest of the cascade. The model has the clean property that a performing borrower exposed to k nonperforming neighbours is likely to nonperform from contagion with probability $1-(1-q)^k$, reflecting independent transmission across her k exposed guarantee links. The cascade converges to the limiting nonperformance vector:

$$Y_i^{(\infty)} = \lim_{m\to\infty} Y_i^{(m)}.$$

We denote by d the share of nonperforming loans:

$$d = \frac{\sum_{i=1}^{N} Y_i^{\infty}}{N}$$

6.2 The Separate Roles of Correlated Shocks and Contagion

This framework separates two distinct amplification mechanisms. First, correlated shocks increase the probability that borrowers who are close in the guarantor network become simultaneously distressed. Second, contagion generates additional defaults endogenously through the network, as the default of one borrower raises the likelihood of default among connected borrowers. In the empirical analysis below, we use this distinction to separate the role of correlated shocks from the role of contagion in generating SACCO-level vulnerability.

The two mechanisms have different implications for how network structure matters. Correlated shocks primarily reflect who is connected — borrowers linked by guarantees may face similar production risks due to geographic proximity, shared inputs, or common market exposure. The correlated outcomes documented in Section 4 are consistent with this channel. Contagion, by contrast, reflects the structure of the network itself — how defaults spread depends on the density, connectivity, and topology of the guarantee relationships. The causal propagation effects estimated in Section 5 provide evidence for this channel.

Disentangling the two is inherently difficult without a structural model, because the same observed patterns of clustered defaults are consistent with either mechanism. The model allows us to hold one channel fixed while varying the other, providing a decomposition of their respective contributions to aggregate vulnerability.

To accommodate correlated shocks, we generalise the iid latent structure of the previous subsection by drawing the latent shock vector $\mathbf{x} = (x_1, \ldots, x_N)$ from a multivariate normal distribution whose covariance is non-zero only on connected pairs in the guarantor network:

$$\mathbf{x} \;\sim\; \mathcal{N}(\mathbf{0},\, \Sigma(\rho))\,, \tag{8}$$

with $\Sigma_{ii}(\rho) = 1$ and, for $i \neq j$, $\Sigma_{ij}(\rho) = \rho \cdot A_{ij}$, where A is the 0/1 adjacency matrix of $\mathcal{U}$ and $\rho \in [0,1]$ governs the strength of network-based correlation. Each x_i retains its standard-normal marginal, so $\Pr(x_i > \bar{x}) = p$ regardless of ρ. Connected borrowers have pairwise correlation $\mathrm{Corr}(x_i, x_j) = \rho$, while unconnected pairs remain independent. The parameter $\rho = 0$ recovers the iid latent shocks of the previous subsection, while higher values generate shocks that cluster on the guarantor network even in the absence of contagion through loan repayment. Fundamental

nonperformance is then $Y_i^0 = \mathbf{1}\{x_i > \bar{x}\}$, exactly as in the previous subsection.

6.3 Characterising Systemic Risk

We define systemic risk as the probability that the SACCO's aggregate loan portfolio deteriorates beyond a threshold at which the institution faces regulatory intervention or financial instability. We anchor this threshold in the regulatory framework of SASRA, which classifies SACCOs into concern levels based on their share of non-performing loans (NPLs). Under SASRA guidelines, an NPL ratio below 2% indicates "best loan quality"; a ratio between 2% and 5% remains within the prudential threshold and is considered acceptable; a ratio between 5% and 10% is above the threshold and described as concerning; and a ratio above 10% indicates "serious deterioration in loan quality…a threat to [the SACCO's] compliance, performance, and sustainability." These thresholds provide a natural, policy-relevant definition of institutional vulnerability.

Formally, we say that the SACCO is *vulnerable* whenever the share of nonperforming borrowers exceeds a threshold $\bar{d}$. The threshold $\bar{d}$ corresponds to the SASRA regulatory levels described above. The central object of interest is the probability of vulnerability, $\Pr(d > \bar{d})$, which we compute via simulation in Section 7.

6.4 Calibration of Structural Parameters

The model requires estimates of three key parameters: the unconditional shock default parameter p, the contagion parameter q and the correlated-shocks parameter ρ.

We take the per-link contagion probability q from Section 5.5 and treat it as a draw from its estimated sampling distribution rather than as a single point, so that the estimation uncertainty in the contagion channel propagates through to the model's predictions. Specifically, we draw $q \sim \mathcal{N}\big(\hat{q},\, \widehat{\mathrm{se}}(\hat{q})^2\big)$, centred on the point estimate $\hat{q} \approx 0.078$ with standard deviation equal to its delta-method standard error ($\widehat{\mathrm{se}}(\hat{q}) \approx 0.035$), truncated to $[0, 1]$.

Holding the contagion channel fixed in this way, we back out the unconditional nonperformance probability p separately for each calendar month. The logic is that p is the only remaining free parameter with which to reconcile the model to the data: given the guarantor network observed in month t and the contagion parameter q, there is a unique level of fundamental risk p at which the model reproduces the nonperformance actually observed that month. Formally, for each month t from April 2018 to December 2023 we construct the borrower network $\mathcal{U}_t$ and solve for the value $\hat{p}_t$ that equates the model-implied expected nonperformance share to its empirical counterpart,

$$\mathbb{E}\big[\, d \,\big|\, \mathcal{U}_t,\, p,\, q \,\big] \;=\; d_t^{\mathrm{obs}}, \tag{9}$$

where d_t^{obs} is the observed share of nonperforming loans in month t. We evaluate the left-hand side by simulation: at each candidate p we draw S independent worlds — each consisting of a vector of fundamental defaults $Y^{(0)}$ with $\Pr(Y_i^0 = 1) = p$ and a set of per-link transmission indicators $\tau_{ij} \sim \mathrm{Bernoulli}(q)$ — iterate the cascade in equation (7) to convergence, and average the realised nonperformance share across worlds. Because $\mathbb{E}[\, d \mid \mathcal{U}_t,\, p,\, q \,]$ is strictly increasing in p, the solution $\hat{p}_t$ is unique and is recovered by a simple search.

This procedure yields one calibrated value $\hat{p}_t$ for each of the 69 months in the window. Figure A2 plots the resulting series against the observed nonperformance share: the two track each other closely, and the calibrated $\hat{p}_t$ lies everywhere below d_t^{obs}, the gap reflecting the contagion amplification that separates fundamental from realised nonperformance. The backed-out fundamental risk is stable across most of the sample but rises sharply in the second half of 2022, when a severe production downturn pushed the observed nonperformance share above 30%.

Pooling the monthly estimates characterises the distribution of fundamental risk over the sample. Excluding three outliers — the months August through October 2022, at the peak of the 2022 stress episode — the remaining 66 monthly values are well described by a normal distribution, $\hat{p}_t \sim \mathcal{N}(0.127, 0.028^2)$: a Shapiro–Wilk test fails to reject normality ($W = 0.99$, $p = 0.73$), in contrast to the full sample, for which normality is rejected ($p = 0.01$) owing to the right skew introduced by the 2022 outliers. We therefore treat fundamental risk as a draw from this fitted normal distribution when integrating over shock realisations.

Taken together, the calibration delivers a distribution over the structural inputs rather than a single point: a month drawn uniformly from the sample pins down the network, p is drawn from the fitted normal distribution of fundamental risk, and q is drawn from its estimated sampling distribution. Feeding these draws through the cascade model traces out the distribution of the aggregate nonperformance share d reported in Figure 5 and discussed in Section 7.

[In progress.] *The correlated-shocks parameter ρ is currently set to zero (independent fundamental shocks); estimating ρ by matching simulated to empirical moments — such as joint nonperformance rates and the clustering of nonperformance by network distance — is in progress.*

7 Quantifying Systemic Risk and Network Resilience

Using the calibrated model from Section 6, we now quantify the level of systemic risk in the SACCO's guarantor network. We first describe the simulation approach, then apply it to the observed guarantor networks, and finally conduct counterfactual exercises to assess how changes in network structure or policy would affect institutional resilience.

7.1 Quantifying Systemic Risk with Simulations

The model in Section 6 implies that the realized share of nonperforming borrowers, d, is a random variable. For a fixed network $\mathcal{U}$ and fixed parameter values (p, q, ρ), different realizations of the underlying repayment shocks and the per-link transmission draws generate different patterns of fundamental nonperformance and, in turn, different cascade outcomes. We therefore study the distribution of d by simulating the model across many realizations.

Formally, in each simulated world $s \in \{1, \ldots, S\}$, we draw a latent shock vector $\mathbf{x}^{(s)} = (x_1^{(s)}, \ldots, x_N^{(s)})$ from the correlated multivariate normal distribution defined in equation (8):

$$\mathbf{x}^{(s)} \sim \mathcal{N}(\mathbf{0}, \Sigma(\rho)).$$

The set of fundamental nonperforming loans is then

$$Y_i^{(0,s)} \;=\; \mathbf{1}\{x_i^{(s)} > \bar{x}\},$$

with threshold $\bar{x} = \Phi^{-1}(1-p)$ chosen to deliver the calibrated unconditional nonperformance rate p. Independently, we draw transmission indicators $\tau_{ij}^{(s)} \sim \text{Bernoulli}(q)$ for every edge $\{i,j\} \in \mathcal{U}$. Starting from the fundamental nonperformance vector $Y^{(0,s)}$, contagion propagates through the network according to the independent-cascade rule in equation (7):

$$Y_i^{(m+1,s)} \;=\; \max\Big\{ Y_i^{(m,s)},\; \mathbf{1}\Big\{ \exists\, j : u_{ij}\, \tau_{ij}^{(s)}\, Y_j^{(m,s)} = 1 \Big\} \Big\},$$

for rounds $m = 0, 1, 2, \ldots$, until the process converges to a limiting nonperformance vector $Y_i^{(\infty,s)}$.

The realized share of nonperforming borrowers in simulated world s is then

$$d^{(s)} \;=\; \frac{1}{N} \sum_{i=1}^{N} Y_i^{(\infty,s)}.$$

This object varies across simulated worlds because the underlying shocks and transmission draws vary. In some simulated worlds, shocks are mild and few borrowers fundamentally nonperform, so contagion never starts. In other worlds, fundamental nonperformance occurs in sparsely connected regions of the network, so propagation remains limited. In yet others, shocks fall on highly connected borrowers and trigger large cascades that drive a high realized nonperformance share.

By repeating this procedure for a large number of simulations S, we obtain the distribution of d, the central object for quantifying systemic risk in the institution. In particular, this distribution allows us to compute the probability that the SACCO becomes vulnerable,

$$\Pr(d > \bar{d}),$$

where $\bar{d}$ is the SASRA threshold above which the institution is considered financially vulnerable, as well as the expected nonperformance share $\mathbb{E}[d]$ and the probability of extreme outcomes in the upper tail.

7.2 Systemic Risk in the Observed Network

We apply the simulation framework to the observed SACCO guarantor network under the calibrated parameter values. Table 6 reports, for each quarter, the observed nonperformance share alongside the mean and standard deviation of the simulated nonperformance share d and the simulated probability that d exceeds 20%, holding the per-period shock intensity fixed. Table 7 reports the analogous quantities when the shock intensity $\hat{p}$ is recalibrated period-by-period to match the observed nonperformance rate. Integrating over the calibrated distributions of fundamental and contagion risk — drawing a month uniformly, p from its fitted normal distribution of fundamental risk, and q from its estimated sampling distribution (Section 6) — yields the distribution of the aggregate nonperformance share d shown in Figure 5. The simulated nonperformance share is centred at 0.185, exceeds the 10% SASRA vulnerability threshold in 97% of draws, and exceeds 20% in 35%, indicating that the institution operates with a substantial and persistent probability

of breaching regulatory concern levels.

[In progress.] *A full discussion of these simulation results — including the decomposition across scenarios (no amplification; correlated shocks only; contagion only; both) — is in progress.*

7.3 Quantifying the Role of Correlated Shocks

Show that if we allow lots of correlated shocks, with and without contagion, how does it change?

7.4 Counterfactual Network Resilience

A key advantage of the simulation approach is that it allows us to evaluate how changes in network structure would affect systemic risk. We consider several counterfactual exercises. First, we examine what would happen if the SACCO limited the number of loans that any single member could guarantee, reducing the maximum degree in the network. Second, we consider random rewiring of the network to assess whether the observed structure is more or less vulnerable than a comparable random network with the same degree distribution. Third, we explore the effects of increasing self-collateralization requirements, which would reduce the reliance on guarantors and thus the density of the network.

[In progress.] *Counterfactual simulation results are in progress.*

8 Conclusion

This paper provides novel evidence on the structure and consequences of guarantor networks within a Kenyan financial cooperative. Using rich administrative data on loans, repayments, and dairy production, we document that the SACCO's network is highly interconnected, with most members linked through multiple overlapping guarantees. This institutional arrangement broadens credit access but also embeds loans in a complex network of mutual responsibility, which may create pathways through which shocks can spread. Our empirical analysis shows that adverse production shocks among dairy farmers increase the likelihood of loan nonperformance, both directly for the affected borrower and indirectly for others connected through guarantee ties. Because the network is densely connected, the bulk of a shock's aggregate short-term repayment impact arises from contagion through guarantee ties rather than from the direct hit to the shocked borrower alone.

These findings speak to long-standing debates about the role of social collateral in financial markets in low- and middle-income countries. Joint-liability microfinance once highlighted the promise of peer monitoring and risk-sharing (Besley and Coate, 1995; Ghatak, 1999), but subsequent evidence underscored the potential for these same mechanisms to amplify stress (De Quidt, Fetzer and Ghatak, 2018). Our results extend this insight to a new institutional setting: credit cooperatives where guarantees are bilateral, long-term, and embedded in a cooperative shareholding structure. Unlike village insurance networks, where shocks may be pooled informally (Townsend, 1994), the guarantor network formalizes obligations in a way that allows repayment difficulties to cascade across otherwise unrelated borrowers. This helps explain why Kenyan poli-

cymakers and regulators have expressed concerns about SACCO resilience, especially following several high-profile cooperative failures in recent years.

From a policy perspective, the results suggest a tension between breadth of access and systemic stability. Dense networks of guarantees enable more members to borrow, but they also increase the exposure of the institution to correlated repayment problems. Because most of the aggregate effect of shocks stems from propagation, interventions that reduce contagion—such as limiting the number of loans one member can guarantee, weighting guarantees by their monetary size, or requiring more self-collateralization—could increase institutional resilience. At the same time, such restrictions may reduce access to credit for borrowers with limited assets. Policymakers therefore face a balancing act between risk mitigation and inclusivity. Mediating this trade-off is an area for future research.

Our findings also point to the importance of network position for individual risk exposure. Borrowers who rely heavily on guarantees from others may be especially vulnerable to network contagion, even if their own income streams are stable, while members who act primarily as guarantors may face less direct repayment risk but still shoulder long-term exposure through their pledged shares.

[In progress.] *Earlier drafts documented an asymmetry whereby repayment problems propagate more strongly downstream (to borrowers the shocked member guarantees) than upstream (to the shocked member's own guarantors); we are re-incorporating the directed-network results that support this claim.*

Understanding how members perceive these risks, whether they adjust their guarantee behavior accordingly, and how guarantor policy should respond, is an additional direction for future work.

Finally, our study underscores the value of bringing network analysis to household finance in developing country contexts. The dense web of guarantees we document resembles the interbank networks that have been central to research on systemic risk in developed financial systems (Acemoglu, Ozdaglar and Tahbaz-Salehi, 2015; Elliott, Golub and Jackson, 2014). Yet the household finance setting differs in crucial ways: the networks are defined not by wholesale financial flows but by household-level obligations, and the propagation of shocks is driven by idiosyncratic production risks rather than market-wide fluctuations. This makes individual borrowers a unique sample for studying how micro-level shocks aggregate into macro-level vulnerabilities. Further research to study heterogeneity across different types of members and examine heterogeneity of our results in terms of network structure will help assess the policy implications of these types of financial arrangements.

9 Tables and Figures

Table 1
Summary Statistics of Key Variables

			Percentiles					
	Mean	Std. Dev.	10^{th}	25^{th}	50^{th}	75^{th}	90^{th}	Observations
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Member Characteristics								
Any loan	0.38	0.49						22,294
Number of loans $\mid > 0$	2.68	2.00	1.00	1.00	2.00	3.00	5.00	8,449
Amount borrowed (000s) $\mid > 0$	225.3	843.6	18.0	36.0	90.0	216.9	461.0	8,449
Guarantee provided	0.34	0.47						22,294
Guarantee received	0.31	0.46						22,294
Guarantee reciprocated	0.21	0.41						22,294
Milk producer	0.79	0.40						22,294
Panel B. Loan Characteristics								
Monthly instalments	27.0	13.9	12.0	12.0	24.0	36.0	48.0	22,640
Amount (000s)	84.1	342.7	11.7	21.5	38.3	76.0	157.0	22,640
Any guarantor	0.71	0.46						22,640
Number of guarantors $\mid > 0$	2.7	2.1	1.0	1.0	2.0	3.0	5.0	15,972
Guarantor share $\mid > 0$	0.66	0.18	0.40	0.61	0.67	0.73	0.89	15,971
Panel C. Loan Performance Statistics								
Defaulted	0.01	0.11						22,640
Completed late	0.12	0.32						22,640
Completed ontime	0.66	0.47						22,640
Ongoing	0.22	0.41						22,640
Ever non-performing	0.44	0.50						22,640
Share of months non-performing $\mid > 0$	0.28	0.25	0.04	0.08	0.17	0.42	0.68	9,879
Panel D. Milk Production								
Monthly milk production (litres) $\mid > 0$	183	253	42	72	127	224	366	374,997
Share of months with milk production	0.59	0.32	0.12	0.31	0.64	0.88	1.00	17,705
Any shock	0.73	0.44						17,705
Number of shocks $\mid > 0$	1.47	0.67	1.00	1.00	1.00	2.00	2.00	12,959
Average length of shock $\mid > 0$	12.75	10.88	3.00	4.50	8.00	17.00	32.00	12,959

Notes: This table presents summary statistics of key variables for the period 2018–2023. Panel A presents member-level characteristics for the roughly 23,000 SACCO members. *Any loan* indicates having at least one long-term loan (12 months or more) during the period. *Guarantee provided* and *Guarantee received* indicate whether the member ever served as a guarantor or had their loan guaranteed, respectively. Panel B presents loan-level characteristics for long-term loans. *Guarantor share* is the fraction of the loan amount covered by guarantors. Panel C presents loan performance statistics at the loan level. *Ever non-performing* indicates the loan was classified as Substandard or worse by SASRA at any point during its term. Panel D presents monthly milk production statistics for milk-producing members from 2018 to mid-2021. *Any shock* indicates at least 3 consecutive months of zero milk production. Conditional statistics (denoted $\mid > 0$) are computed for the subsample with positive values.

Table 2

Summary Statistics of Guarantor Networks

	Borrowers		Full Network					Largest Component	
Time	All	Network	Nodes	Edges	Degree	Reciprocity	Clustering	Nodes	Length
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
2018-01	3,908	3,024	4,368	9,709	3.96	0.218	0.113	4,218	7.71
2018-04	3,835	2,932	4,353	9,377	3.86	0.206	0.112	4,196	7.90
2018-07	3,866	2,940	4,401	9,477	3.87	0.203	0.113	4,252	7.80
2018-10	3,950	3,023	4,482	9,775	3.92	0.205	0.112	4,368	7.74
2019-01	4,103	3,102	4,569	10,010	3.93	0.207	0.110	4,452	7.74
2019-04	4,157	3,176	4,657	10,326	3.98	0.206	0.115	4,540	7.65
2019-07	4,314	3,254	4,726	10,563	3.99	0.214	0.116	4,621	7.68
2019-10	4,472	3,314	4,784	10,929	4.08	0.216	0.121	4,682	7.64
2020-01	4,654	3,512	5,001	11,597	4.13	0.220	0.121	4,865	7.57
2020-04	4,778	3,603	5,141	11,992	4.15	0.219	0.118	4,983	7.59
2020-07	4,915	3,617	5,228	12,142	4.14	0.217	0.118	5,069	7.60
2020-10	4,906	3,601	5,228	12,099	4.11	0.224	0.116	5,120	7.63
2021-01	4,932	3,672	5,307	12,279	4.10	0.229	0.119	5,209	7.53
2021-04	5,022	3,731	5,404	12,417	4.06	0.231	0.122	5,321	7.65
2021-07	5,158	3,817	5,573	12,774	4.07	0.226	0.123	5,469	7.57
2021-10	5,289	4,006	5,755	13,334	4.10	0.230	0.122	5,632	7.60
2022-01	5,372	4,150	5,900	13,765	4.13	0.230	0.123	5,772	7.69
2022-04	5,521	4,239	6,011	14,036	4.12	0.234	0.123	5,881	7.79
2022-07	5,640	4,372	6,111	14,490	4.17	0.240	0.121	5,983	7.83
2022-10	5,645	4,359	6,136	14,420	4.14	0.236	0.119	6,044	7.93
2023-01	5,643	4,430	6,212	14,689	4.17	0.238	0.117	6,117	8.00
2023-04	5,708	4,518	6,328	14,989	4.17	0.240	0.117	6,252	8.03
2023-07	5,720	4,490	6,310	14,904	4.16	0.238	0.119	6,235	8.03
2023-10	5,665	4,495	6,412	15,125	4.15	0.241	0.117	6,339	8.06

Notes: This table presents monthly statistics of the guarantor network from 2018 to 2023. The network is constructed using all long-term loans (12 months or more) that are active in the first month of each period. A directed edge flows from the guarantor to the borrower. *Borrowers (All)* counts all members with active long-term loans. *Borrowers (Network)* counts borrowers who appear in the guarantor network (i.e., have at least one guarantor). *Nodes* counts all members in the network, including non-borrowing guarantors. *Edges* counts directed guarantee relationships. *Degree* is the average number of undirected connections per member. *Reciprocity* is the fraction of directed edges for which the reverse edge also exists. *Clustering* is the average local clustering coefficient of the undirected network, measuring how often a member's connections are also connected to each other. *Nodes (Largest Component)* counts members in the largest connected component of the undirected network. *Path Length* is the average shortest-path distance between all pairs of members within the largest component.

Table 3

Summary Statistics of Network Proximity

			Dyads by Distance (%)					Dyads by Resistance (%)				
Time	Borrowers	Dyads	∞	1	2–5	6–9	10+	$[0, 0.5)$	$[0.5, 1)$	$[1, 1.5)$	$[1.5, 2)$	2+
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
2018-01	3,908	7,634,278	29.8	0.09	10.6	50.9	8.7	4.7	29.0	15.5	10.6	10.4
2018-04	3,835	7,351,695	30.3	0.09	9.8	49.1	10.8	3.9	27.7	16.4	10.9	10.8
2018-07	3,866	7,471,045	30.7	0.09	10.3	49.6	9.3	3.9	28.6	15.9	10.4	10.5
2018-10	3,950	7,799,275	29.8	0.09	10.7	50.5	8.9	4.4	28.7	15.5	11.0	10.5
2019-01	4,103	8,415,253	31.5	0.08	10.5	49.1	8.9	4.5	27.9	15.4	10.8	10.0
2019-04	4,157	8,638,246	31.9	0.08	10.8	49.5	7.8	5.2	28.7	15.4	10.2	8.6
2019-07	4,314	9,303,141	31.9	0.08	10.5	49.2	8.4	5.1	28.9	15.3	10.5	8.5
2019-10	4,472	9,997,156	34.3	0.08	10.4	47.3	7.9	5.6	28.7	14.2	9.7	7.4
2020-01	4,654	10,827,531	33.3	0.08	10.7	48.6	7.4	6.2	30.2	13.8	9.7	6.9
2020-04	4,778	11,412,253	34.6	0.07	10.4	47.7	7.2	6.4	30.6	13.3	9.0	6.0
2020-07	4,915	12,076,155	36.6	0.07	9.8	46.3	7.2	6.0	29.5	12.7	8.6	6.6
2020-10	4,906	12,031,965	36.6	0.07	9.6	46.5	7.2	5.6	29.5	13.2	8.5	6.6
2021-01	4,932	12,159,846	35.6	0.07	10.1	48.0	6.2	5.4	30.4	12.9	8.9	6.8
2021-04	5,022	12,607,731	35.8	0.07	9.5	47.3	7.3	5.0	29.5	13.1	9.2	7.3
2021-07	5,158	13,299,903	35.5	0.07	9.7	48.4	6.3	5.3	29.3	12.9	9.4	7.6
2021-10	5,289	13,984,116	34.8	0.07	9.6	49.1	6.5	5.8	29.9	13.3	9.1	7.2
2022-01	5,372	14,426,506	33.5	0.07	9.4	49.5	7.5	6.1	31.1	13.7	9.1	6.5
2022-04	5,521	15,237,960	33.9	0.06	9.0	48.5	8.5	6.0	30.9	13.2	9.0	7.1
2022-07	5,640	15,901,980	34.9	0.06	8.8	47.4	8.9	6.4	30.4	12.7	8.7	6.9
2022-10	5,645	15,930,190	34.5	0.06	8.6	46.9	9.9	6.0	30.5	12.4	8.9	7.6
2023-01	5,643	15,918,903	33.6	0.06	8.5	47.5	10.3	6.5	31.9	12.6	8.4	7.1
2023-04	5,708	16,287,778	32.7	0.06	8.4	48.1	10.7	6.5	33.2	12.7	8.4	6.5
2023-07	5,720	16,356,340	32.9	0.06	8.4	47.9	10.7	6.5	32.4	13.1	8.5	6.6
2023-10	5,665	16,043,280	31.3	0.06	8.4	48.8	11.5	6.6	33.8	13.0	8.7	6.6

Notes: This table presents the distribution of pairwise shortest-path distances $Distance_{ijt}$ in the undirected network $\mathcal{U}_t$ among all borrower pairs at the start of each quarter from 2018 to 2023. *Borrowers* counts the total number of members with active long-term loans. *Dyads* is the total number of unique borrower pairs. Distance is measured as the shortest-path length in $\mathcal{U}_t$. Columns 4–11 report the share of dyads at each distance (1 through 7 individually; 8+ includes all pairs at distance 8 or greater). ∞ indicates pairs in separate components of $\mathcal{U}_t$.

Table 4
Joint Outcomes and Network Proximity

	Joint Loan Nonperformance (1)	Joint Milk Production Shock (2)	Correlated Milk Production (3)
Panel A. Closeness			
Closeness	0.0132*	·	·
	(0.0067)	(·)	(·)
Panel B. Distance Indicators			
$\mathbb{1}\{Distance = 1\}$	0.0106***	−0.0214***	·
	(0.0023)	(0.0033)	(·)
$\mathbb{1}\{Distance = 2\}$	0.0051**	−0.0176***	·
	(0.0020)	(0.0029)	(·)
$\mathbb{1}\{Distance = 3\}$	0.0036**	−0.0156***	·
	(0.0018)	(0.0027)	(·)
$\mathbb{1}\{Distance = 4\}$	0.0027	−0.0132***	·
	(0.0017)	(0.0025)	(·)
$\mathbb{1}\{Distance = 5\}$	0.0022	−0.0101***	·
	(0.0016)	(0.0024)	(·)
$\mathbb{1}\{Distance = 6\}$	0.0018	−0.0070***	·
	(0.0015)	(0.0023)	(·)
$\mathbb{1}\{Distance = 7\}$	0.0010	−0.0037*	·
	(0.0014)	(0.0022)	(·)
$\mathbb{1}\{Distance \geq 8\}$	−0.0007	0.0015	·
	(0.0014)	(0.0022)	(·)
Unconnected mean	0.0826	0.0884	·
Members	8,775	6,501	·
Pairs	32,684,663	18,987,120	·
Observations	881,638,212	437,965,198	·

Notes: This table presents coefficient estimates from dyadic regressions of equation 1 relating pair outcomes to network proximity. The unit of observation is a borrower pair $\{i, j\}$ at month-year t, with each pair included once. Three outcomes are studied: column (1) *Joint Loan Nonperformance*, an indicator that both i and j are simultaneously in nonperforming status; column (2) *Joint Production Shock*, an indicator that both members experience a milk production shock in the same window; and column (3) *Correlated Milk Production*, the within-quarter correlation of monthly milk quantities for the pair. Panel A reports the coefficient on continuous *Closeness* (the inverse of the undirected shortest-path distance between i and j). Panel B reports coefficients on binary distance indicators $\mathbb{1}\{Distance = k\}$ for $k = 1, 2, \ldots, 7$ and $\mathbb{1}\{Distance \geq 8\}$, with unconnected pairs as the omitted category. All specifications include pair fixed effects and time fixed effects. Standard errors in parentheses are two-way clustered at the level of both pair members. 1%, 5% and 10% statistical significance are indicated with ***, **, and * respectively.

Table 5

Direct and Propagation Effects of Shocks of Loan Nonperformance

	Loan Nonperformance			
	(1)	(2)	(3)	(4)
Panel A. Direct Effects				
Post × *Shock*	0.1659***	0.1738***	0.1727***	0.1739***
	(0.0201)	(0.0212)	(0.0200)	(0.0203)
Pre-shock mean	0.1458	0.1540	0.0759	0.0770
Shocks	379	366	329	321
Loans	716	678	606	598
Observations	6,443	6,101	5,453	5,381
Loans completed prior to shock	Yes	No	No	No
Loans nonperforming prior to shock	Yes	Yes	No	No
Loans not in propagation sample	Yes	Yes	Yes	No
Panel B. Propagation Effects				
Post × *Conductance*	0.0146**	0.0144**	0.0155**	0.0136**
	(0.0063)	(0.0062)	(0.0061)	(0.0058)
Pre-shock mean	0.0561	0.0562	0.0539	0.0507
Shocks	321	321	321	321
Loans	285	284	276	263
Pairs	25,855	23,186	16,971	11,527
Observations	232,695	208,674	152,739	103,743
Loans completed prior to shock	No	No	No	No
Loans nonperforming prior to shock	No	No	No	No
Connected pairs with distance ≥ 10	Yes	No	No	No
Connected pairs with distance ≥ 8	Yes	Yes	No	No
Connected pairs with distance ≥ 6	Yes	Yes	Yes	No
Panel C. Contagion Parameter				
q				0.0781**
				(0.0350)

Notes: This table presents coefficient estimates from differences-in-differences regressions estimating both the direct and propagation effects of milk production shocks. The two columns correspond to two outcomes: column (1) *Loan Nonperformance*, a binary indicator that the loan is in nonperforming (Substandard) status, and column (2) *Milk Production*, litres of milk delivered to the dairy cooperative in a given month. Panel A reports direct effects from equation 4: the regressor of interest is *Post* × *Shock*, where the unit of observation is a loan-month for the shocked borrower's own loan. Panel B reports propagation effects from equation 6: the regressor of interest is *Post* × *Conductance*, where the unit of observation is a (shocked-loan, non-shocked-loan, month) triple and *Conductance* is the inverse of the effective resistance between the shocked and non-shocked loan members in the undirected network at the start of the non-shocked loan, equal to 1 for pairs linked by a single direct guarantee edge and 0 for unconnected pairs. *Shock* is a binary variable equal to 1 if the loan experienced a shock in the first half of the remaining loan term (treated loans); for control loans, placebo shocks taking place Δ months prior to their actual shock are constructed, with Δ the midpoint of the remaining loan term. *Post* is a binary variable equal to 1 from the month of the shock onwards (event months 0 through 5). For each loan, three months of pre-shock and six months of post-shock observations form a balanced panel. The sample is restricted to loans where the borrower has non-zero degree at the start of the loan and the loan was not already paid off in the month prior to the shock. All specifications include loan and time fixed effects; Panel B additionally includes shocked-loan and time-to-shock fixed effects. Controls interacted with time fixed effects include the out-degree and in-degree of the loan member at the start of the loan. Standard errors in parentheses are clustered at the loan level for direct effects and two-way clustered at the shocked- and non-shocked-loan levels for propagation effects. 1%, 5% and 10% statistical significance are indicated with ***, **, and * respectively.

Table 6

Simulations Effects of Shocks Fixed Parameters

Time (1)	Borrowers (2)	Actual (3)	Expected (4)	SD (5)	$\Pr(d > 0.20)$ (6)
2018m03	3,214	0.157	0.137	0.009	0.000
2018m06	3,256	0.174	0.138	0.009	0.000
2018m09	3,304	0.188	0.138	0.009	0.000
2018m12	3,354	0.194	0.137	0.009	0.000
2019m03	3,429	0.167	0.139	0.009	0.000
2019m06	3,514	0.233	0.139	0.009	0.000
2019m09	3,620	0.187	0.140	0.009	0.000
2019m12	3,751	0.156	0.141	0.009	0.000
2020m03	3,854	0.146	0.141	0.009	0.000
2020m06	3,919	0.220	0.141	0.009	0.000
2020m09	3,914	0.191	0.141	0.009	0.000
2020m12	3,893	0.133	0.140	0.009	0.000
2021m03	3,990	0.113	0.140	0.009	0.000
2021m06	4,085	0.146	0.140	0.009	0.000
2021m09	4,232	0.130	0.141	0.008	0.000
2021m12	4,310	0.127	0.141	0.008	0.000
2022m03	4,447	0.118	0.141	0.008	0.000
2022m06	4,558	0.225	0.142	0.008	0.000
2022m09	4,568	0.315	0.142	0.008	0.000
2022m12	4,528	0.210	0.141	0.008	0.000
2023m03	4,619	0.180	0.142	0.008	0.000
2023m06	4,636	0.232	0.141	0.008	0.000
2023m09	4,651	0.206	0.141	0.008	0.000
2023m12	4,658	0.176	0.140	0.008	0.000

Notes: This table reports simulated systemic-risk statistics for the observed guarantor network at the start of each quarter, under the calibrated independent-cascade model of Section 6. Column (2) reports the number of borrowers; column (3) the observed share of nonperforming loans; columns (4)–(5) the mean and standard deviation of the simulated nonperformance share d across simulation draws, holding the per-period shock intensity fixed; and column (6) the simulated probability that d exceeds 0.20. The per-link contagion probability is set to the estimate $\hat{q}$ from Table 5.

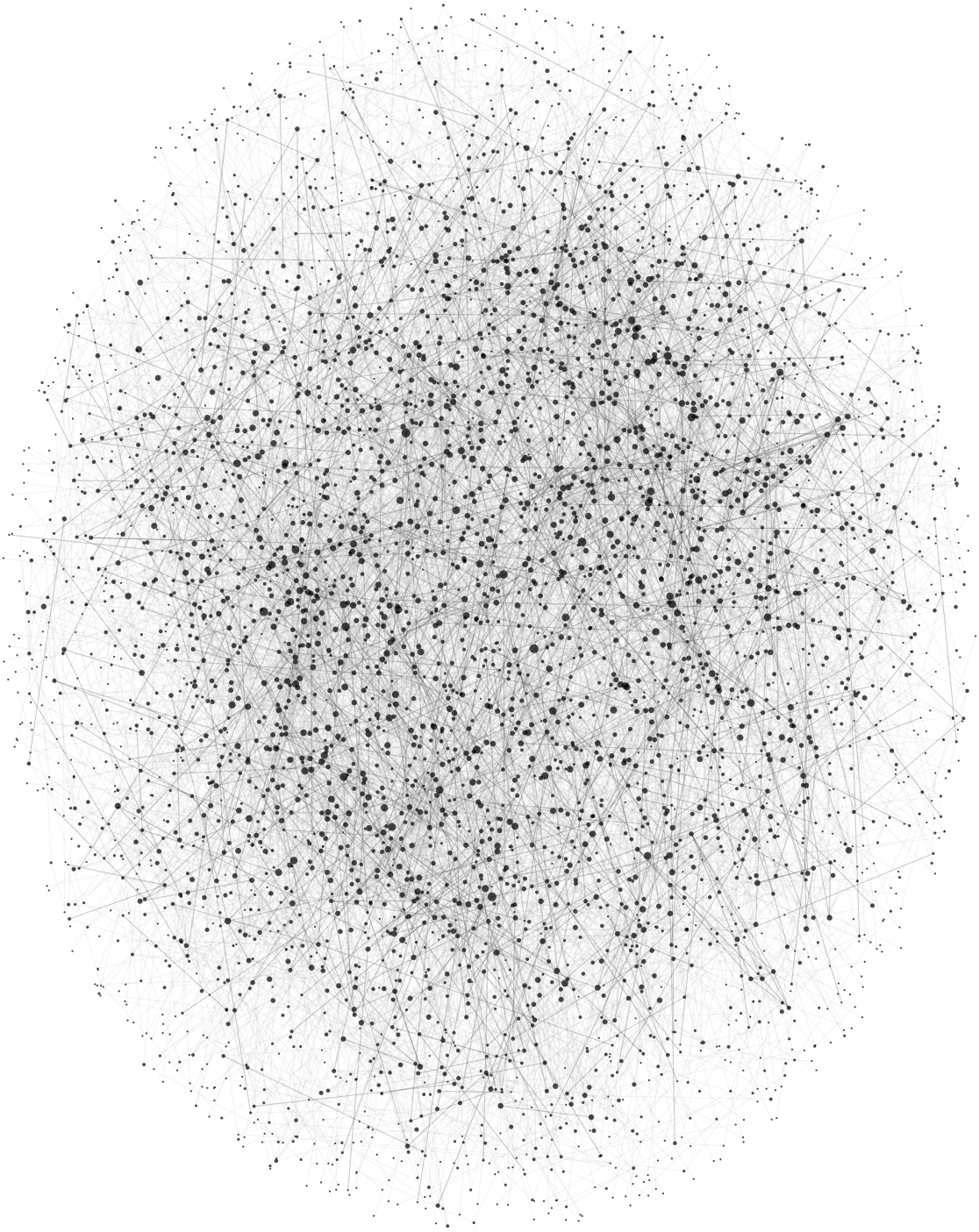
Table 7

Simulations Effects of Shocks with Calibration

Time	Borrowers	$\hat{p}$	Actual	Expected	SD	$\Pr(d > 0.20)$
(1)	(2)	(3)	(4)	(5)	(6)	(7)
2018m03	3,214	0.035	0.157	0.051	0.006	0.000
2018m06	3,256	0.125	0.174	0.173	0.009	0.003
2018m09	3,304	0.130	0.188	0.180	0.009	0.018
2018m12	3,354	0.140	0.194	0.192	0.009	0.194
2019m03	3,429	0.145	0.167	0.200	0.009	0.514
2019m06	3,514	0.125	0.233	0.175	0.009	0.003
2019m09	3,620	0.150	0.187	0.209	0.009	0.820
2019m12	3,751	0.125	0.156	0.177	0.009	0.004
2020m03	3,854	0.120	0.146	0.171	0.009	0.000
2020m06	3,919	0.120	0.220	0.171	0.008	0.001
2020m09	3,914	0.145	0.191	0.203	0.009	0.639
2020m12	3,893	0.115	0.133	0.163	0.008	0.000
2021m03	3,990	0.095	0.113	0.137	0.008	0.000
2021m06	4,085	0.075	0.146	0.109	0.007	0.000
2021m09	4,232	0.110	0.130	0.157	0.008	0.000
2021m12	4,310	0.095	0.127	0.138	0.008	0.000
2022m03	4,447	0.095	0.118	0.138	0.008	0.000
2022m06	4,558	0.125	0.225	0.178	0.008	0.003
2022m09	4,568	0.205	0.315	0.279	0.009	1.000
2022m12	4,528	0.210	0.210	0.284	0.009	1.000
2023m03	4,619	0.145	0.180	0.204	0.008	0.703
2023m06	4,636	0.165	0.232	0.229	0.008	1.000
2023m09	4,651	0.165	0.206	0.229	0.008	1.000
2023m12	4,658	0.145	0.176	0.202	0.008	0.611

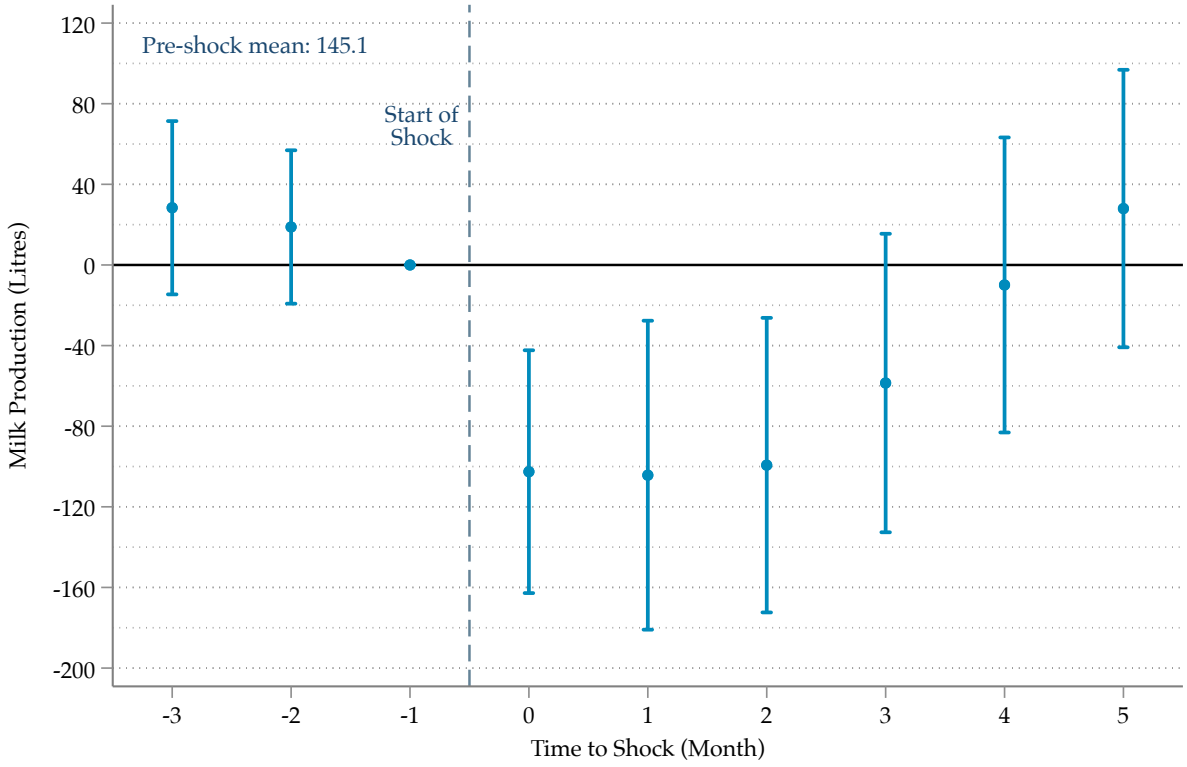
Notes: This table repeats the simulation in Table 6, recalibrating the per-period shock intensity $\hat{p}$ (column 3) so that the simulated nonperformance share matches the observed share in each quarter. Column (2) reports the number of borrowers; column (4) the observed nonperformance share; columns (5)–(6) the mean and standard deviation of the simulated share d; and column (7) the simulated probability that d exceeds 0.20. The per-link contagion probability is set to the estimate $\hat{q}$ from Table 5.

Figure 1
Guarantor Networks



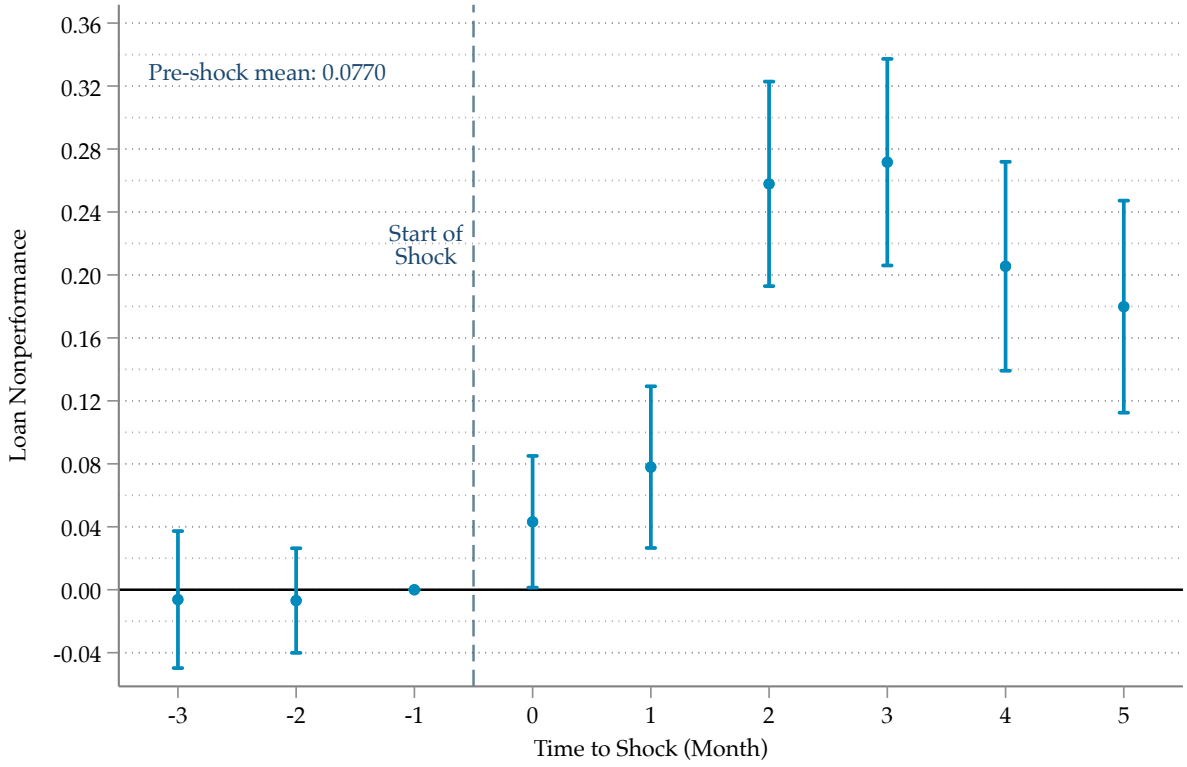
Notes: This figure visualizes the undirected guarantor network for long-term loans in January 2018. Each node represents a SACCO member who is either a borrower or a guarantor. Node size is proportional to the member's undirected degree (number of guarantee relationships). Edges represent guarantee relationships: darker edges indicate reciprocal ties (both members guarantee each other), while lighter edges indicate non-reciprocal ties (only one member guarantees the other). The network contains 4,368 members connected by 8,649 undirected edges, of which 1,060 are reciprocal. Node positions are determined by a force-directed layout algorithm.

Figure 2
Direct Effects of Shocks on Milk Production



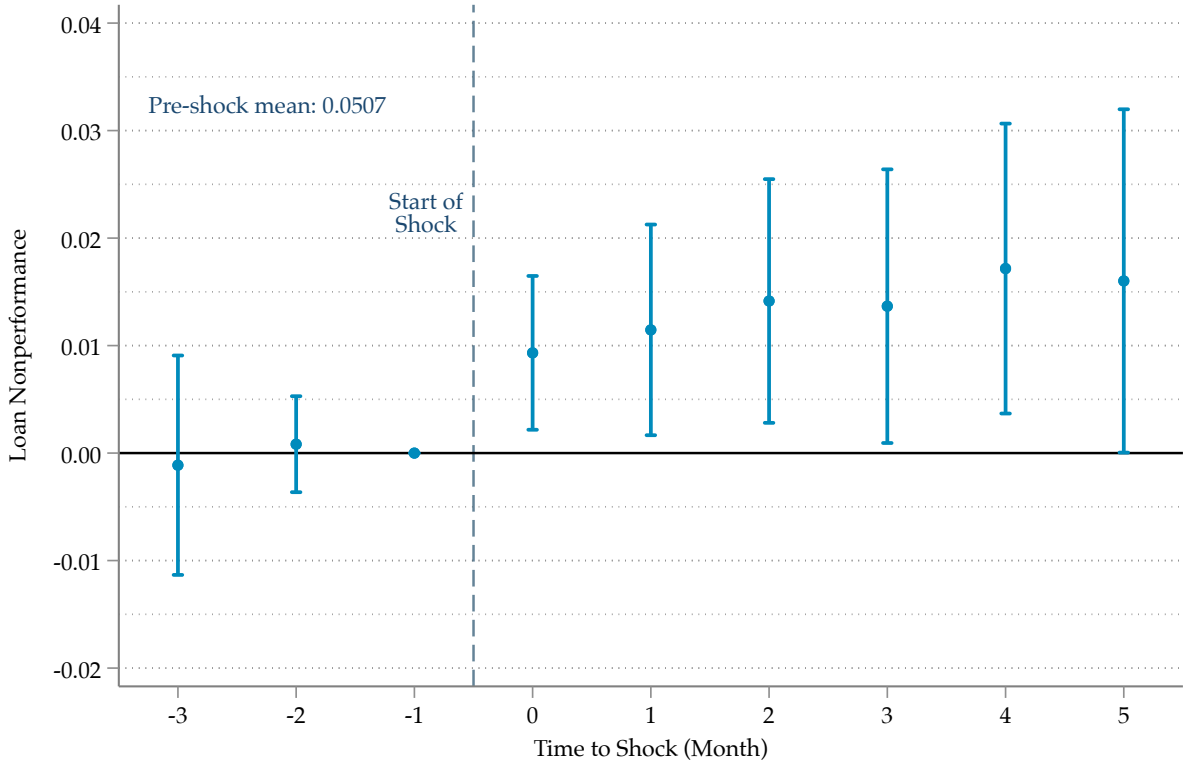
Notes: This figure plots coefficients from an event-study regression of equation 3 estimating the direct effects of milk production shocks on the shocked borrower's own milk production. The dependent variable is *Milk Production*, the litres of milk delivered to the dairy cooperative in a given month. Sample, treatment definition, fixed effects, controls, and standard errors are identical to Figure 3. Confidence intervals are shown at the 95% level, with horizontal caps denoting 90% intervals. The negative coefficients in the post-shock periods confirm that shocked farmers experience a real disruption to milk production, concentrated in the first post-shock quarter, providing a first-stage check on the shock measure.

Figure 3
Direct Effects of Shocks on Loan Nonperformance



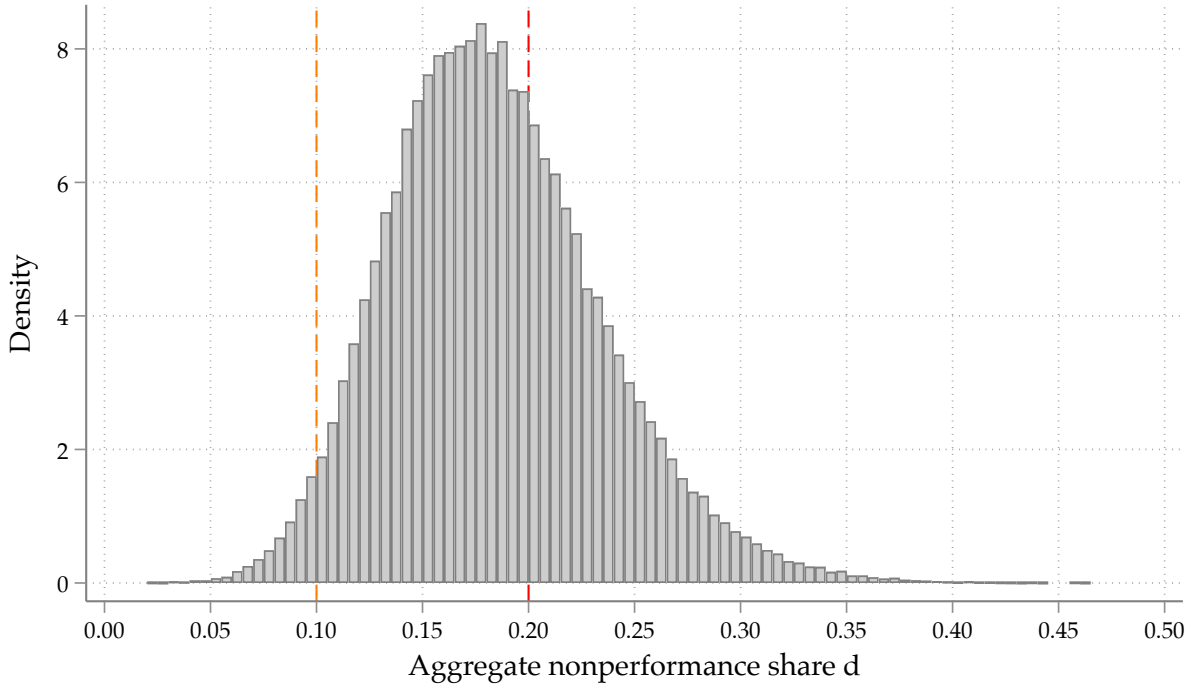
Notes: This figure plots coefficients from an event-study regression of equation 3 estimating the direct effects of milk production shocks on loan nonperformance. The dependent variable is *Loan Nonperformance*, a binary indicator that the loan is in nonperforming (Substandard) status. The treatment variable is *Shock*, a binary indicator equal to 1 if the loan experienced a shock in the first half of the remaining loan term (treated loans); for control loans, placebo shocks Δ months prior to their actual shock are constructed, with Δ the midpoint of the remaining loan term. Each loan contributes three months of pre-shock and six months of post-shock observations to form a balanced panel. The sample is restricted to loans where (i) the borrower has non-zero degree in the network at the start of the loan, and (ii) the loan was not already paid off in the month prior to the shock. The specification includes loan and time fixed effects, with controls (out-degree and in-degree at the start of the loan) interacted with time fixed effects. Standard errors are clustered at the loan level. Confidence intervals are shown at the 95% level, with horizontal caps denoting 90% intervals. The dashed annotation reports the pre-shock mean of the dependent variable.

Figure 4
Propagation Effects of Shocks on Loan Nonperformance



Notes: This figure plots coefficients from an event-study regression of equation 5 estimating the propagation effects of milk production shocks on loan nonperformance for non-shocked borrowers. The dependent variable is *Loan Nonperformance* for non-shocked borrower i, a binary indicator that loan i is in nonperforming status. The treatment variable is $Conductance_{ij}$, the inverse of the effective resistance between the shocked loan member j and the non-shocked loan member i in the undirected network at the start of the non-shocked loan; coefficients are interpreted as the differential probability of nonperformance between a non-shocked loan linked to the shocked member by a single direct guarantee edge ($Conductance = 1$) and an unconnected one ($Conductance = 0$). For each shocked loan, three months of pre-shock and six months of post-shock observations form a balanced panel of non-shocked-loan-by-month observations. Sample restrictions exclude pairs where the non-shocked loan is within four months of starting or finishing, is already two-thirds complete at the time of the shock, or experiences its own shock within four months of j's. The specification includes shocked-loan, non-shocked-loan, time, and time-to-shock fixed effects, with controls (out-degree and in-degree of the non-shocked loan member at the start of the non-shocked loan) interacted with time fixed effects. Standard errors are two-way clustered at the shocked- and non-shocked-loan levels. Confidence intervals are shown at the 95% level, with horizontal caps denoting 90% intervals.

Figure 5
Distribution of Aggregate Nonperformance



100000 draws. Month ~ U{2018m4..2023m12}; p ~ N(0.1273, 0.0277); q ~ N(0.0781, 0.0350); rho = 0.
Mean = 0.1846, SD = 0.0505. Pr(d > 0.10) = 0.9702, Pr(d > 0.20) = 0.3495.
Dashed lines: SASRA supervisory thresholds (10% orange, 20% red).

Notes: This figure plots the simulated distribution of the aggregate nonperformance share d under the calibrated independent-cascade model of Section 6, integrating over the three sources of uncertainty described in Section 6. Each of the 100,000 simulation draws (i) selects a calendar month uniformly from April 2018 to December 2023, which fixes the observed guarantor network for that month; (ii) draws the unconditional shock probability p from its calibrated distribution of fundamental risk, $\mathcal{N}(0.127,\ 0.028^2)$ (see Section 6 and Figure A2); and (iii) draws the per-link contagion probability q from its estimated sampling distribution, $\mathcal{N}(0.078,\ 0.035^2)$, with $\rho = 0$. For each draw, fundamental defaults and per-link transmission indicators $\tau_{ij} \sim \text{Bernoulli}(q)$ are realised, the cascade in equation (7) is iterated to convergence, and the resulting share d is recorded. Across draws the mean simulated nonperformance share is 0.185 (standard deviation 0.051); the simulated probability that d exceeds the 10% SASRA vulnerability threshold is 0.97, and the probability that it exceeds 20% is 0.35. Dashed vertical lines mark the SASRA supervisory thresholds at 10% (orange) and 20% (red).

References

Acemoglu, Daron, Asuman Ozdaglar, and Alireza Tahbaz-Salehi. 2015. "Systemic Risk and Stability in Financial Networks." *American Economic Review*, 105(2): 564–608.

Banerjee, Abhijit, Arun G. Chandrasekhar, Esther Duflo, and Matthew O. Jackson. 2013. "The Diffusion of Microfinance." *Science*, 341(6144): 1236498.

Baqaee, David Rezza. 2018. "Cascading Failures in Production Networks." *Econometrica*, 86(5): 1819–1838.

Besley, Timothy, and Stephen Coate. 1995. "Group Lending, Repayment Incentives and Social Collateral." *Journal of Development Economics*, 46(1): 1–18.

Breza, Emily, Arun G. Chandrasekhar, Tyler H. McCormick, and Mengjie Pan. 2020. "Using Aggregated Relational Data to Feasibly Identify Network Structure without Network Data." *American Economic Review*, 110(8): 2454–2484.

Cabrales, Antonio, Piero Gottardi, and Fernando Vega-Redondo. 2017. "Risk Sharing and Contagion in Networks." *The Review of Financial Studies*, 30(9): 3086–3127.

Cai, Jing, Muhammad Meki, Simon Quinn, Erica Field, Cynthia Kinnan, Jonathan Morduch, Jonathan De Quidt, and Farah Said. 2025. "Microfinance." VoxDev VoxDevLit.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2011. "Robust Inference with Multiway Clustering." *Journal of Business & Economic Statistics*, 29(2): 238–249.

Central Bank of Kenya. 2021. "2021 FinAccess Household Survey." Central Bank of Kenya, Kenya National Bureau of Statistics, and Financial Sector Deepening Kenya, Nairobi, Kenya.

De Quidt, Jonathan, Thiemo Fetzer, and Maitreesh Ghatak. 2018. "Commercialization and the Decline of Joint Liability Microcredit." *Journal of Development Economics*, 134: 209–225.

Elliott, Matthew, and Benjamin Golub. 2022. "Networks and Economic Fragility." *Annual Review of Economics*, 14(Volume 14, 2022): 665–696.

Elliott, Matthew, Benjamin Golub, and Matthew O. Jackson. 2014. "Financial Networks and Contagion." *American Economic Review*, 104(10): 3115–3153.

Ghatak, Maitreesh. 1999. "Group Lending, Local Information and Peer Selection." *Journal of Development Economics*, 60(1): 27–50.

Ghatak, Maitreesh. 2000. "Screening by the Company You Keep: Joint Liability Lending and the Peer Selection Effect." *The Economic Journal*, 110(465): 601–631.

Ghatak, Maitreesh, and Timothy W. Guinnane. 1999. "The Economics of Lending with Joint Liability: Theory and Practice." *Journal of Development Economics*, 60(1): 195–228.

Glasserman, Paul, and H. Peyton Young. 2016. "Contagion in Financial Networks." *Journal of Economic Literature*, 54(3): 779–831.

Jackson, Matthew O., and Agathe Pernoud. 2021. "Systemic Risk in Financial Networks: A Survey." *Annual Review of Economics*, 13(Volume 13, 2021): 171–202.

Jackson, Matthew O., Brian W. Rogers, and Yves Zenou. 2017. "The Economic Consequences of Social-Network Structure." *Journal of Economic Literature*, 55(1): 49–95.

Jack, William, Michael Kremer, Joost de Laat, and Tavneet Suri. 2023. "Credit Access, Selection, and Incentives in a Market for Asset-Collateralized Loans: Evidence From Kenya." *The Review of Economic Studies*, 90(6): 3153–3185.

Karlan, Dean, Markus Mobius, Tanya Rosenblat, and Adam Szeidl. 2009. "Trust and Social Collateral." *The Quarterly Journal of Economics*, 124(3): 1307–1361.

Kinnan, Cynthia, Krislert Samphantharak, Robert Townsend, and Diego Vera-Cossio. 2024. "Propagation and Insurance in Village Networks." *American Economic Review*, 114(1): 252–284.

Klein, D. J., and M. Randić. 1993. "Resistance distance." *Journal of Mathematical Chemistry*, 12(1): 81–95.

SASRA. 2024. "The SACCO Supervision Annual Report 2023." Sacco Societies Regulatory Authority (SASRA), Nairobi, Kenya.

Townsend, Robert M. 1994. "Risk and Insurance in Village India." *Econometrica*, 62(3): 539.

World Council of Credit Unions. 2023. "Statistical Report." World Council of Credit Unions.

A Appendix Tables and Figures

Table A1

Summary Statistics of Network Distance

			Dyads by Distance (%)										
Time	Borrowers	Dyads	∞	1	2	3	4	5	6	7	8	9	10+
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
2018-01	3,908	7,634,278	29.8	0.09	0.3	1.1	2.9	6.2	10.6	14.5	14.9	10.8	8.7
2018-04	3,835	7,351,695	30.3	0.09	0.3	1.0	2.7	5.7	9.7	13.5	14.5	11.4	10.8
2018-07	3,866	7,471,045	30.7	0.09	0.3	1.1	2.8	6.1	10.3	14.1	14.5	10.8	9.3
2018-10	3,950	7,799,275	29.8	0.09	0.3	1.1	2.9	6.4	10.9	14.5	14.5	10.5	8.9
2019-01	4,103	8,415,253	31.5	0.08	0.3	1.1	2.9	6.2	10.5	13.9	14.1	10.6	8.9
2019-04	4,157	8,638,246	31.9	0.08	0.3	1.1	3.0	6.4	10.9	14.3	14.1	10.2	7.8
2019-07	4,314	9,303,141	31.9	0.08	0.3	1.1	2.9	6.2	10.6	14.0	14.1	10.5	8.4
2019-10	4,472	9,997,156	34.3	0.08	0.3	1.0	2.8	6.2	10.5	13.6	13.4	9.8	7.9
2020-01	4,654	10,827,531	33.3	0.08	0.3	1.1	2.9	6.4	10.9	14.2	13.8	9.8	7.4
2020-04	4,778	11,412,253	34.6	0.07	0.3	1.0	2.9	6.2	10.6	13.9	13.7	9.6	7.2
2020-07	4,915	12,076,155	36.6	0.07	0.3	1.0	2.7	5.9	10.1	13.4	13.3	9.4	7.2
2020-10	4,906	12,031,965	36.6	0.07	0.3	0.9	2.6	5.8	9.9	13.3	13.5	9.7	7.2
2021-01	4,932	12,159,846	35.6	0.07	0.3	1.0	2.7	6.1	10.6	14.1	13.9	9.4	6.2
2021-04	5,022	12,607,731	35.8	0.07	0.3	0.9	2.6	5.8	10.1	13.7	13.8	9.7	7.3
2021-07	5,158	13,299,903	35.5	0.07	0.3	0.9	2.6	6.0	10.6	14.5	14.2	9.2	6.3
2021-10	5,289	13,984,116	34.8	0.07	0.3	0.9	2.6	5.9	10.5	14.6	14.5	9.4	6.5
2022-01	5,372	14,426,506	33.5	0.07	0.3	0.9	2.5	5.7	10.2	14.3	14.8	10.2	7.5
2022-04	5,521	15,237,960	33.9	0.06	0.3	0.9	2.4	5.5	9.7	13.8	14.5	10.4	8.5
2022-07	5,640	15,901,980	34.9	0.06	0.3	0.8	2.4	5.3	9.4	13.4	14.2	10.3	8.9
2022-10	5,645	15,930,190	34.5	0.06	0.2	0.8	2.3	5.2	9.2	13.1	14.1	10.5	9.9
2023-01	5,643	15,918,903	33.6	0.06	0.3	0.8	2.3	5.1	9.1	13.1	14.4	11.0	10.3
2023-04	5,708	16,287,778	32.7	0.06	0.3	0.8	2.3	5.0	8.9	13.0	14.6	11.6	10.7
2023-07	5,720	16,356,340	32.9	0.06	0.2	0.8	2.3	5.0	9.0	13.0	14.5	11.4	10.7
2023-10	5,665	16,043,280	31.3	0.06	0.3	0.8	2.3	5.0	9.0	13.1	14.8	11.9	11.5

Notes: This table presents the full distance-bin breakdown of pairwise shortest-path distances $Distance_{ijt}$ in the undirected network $\mathcal{U}_t$ among all borrower pairs at the start of each quarter from 2018 to 2023. *Borrowers* counts the total number of members with active long-term loans. *Dyads* is the total number of unique borrower pairs. Cell entries are percentages of dyads. Column 4 reports the share of *unconnected* pairs ($Distance_{ijt} = \infty$); columns 5–13 report the share at each individual distance from 1 to 9; column 14 reports the share at distance 10 or greater.

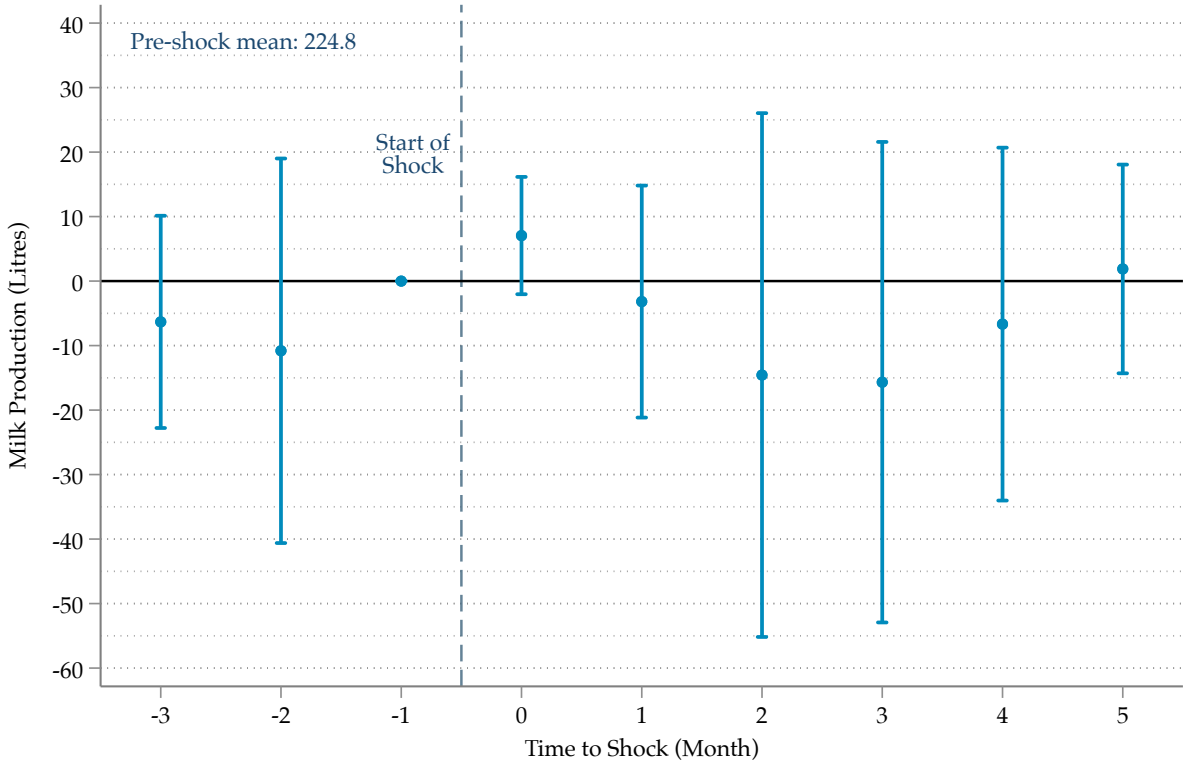
Table A2

Direct and Propagation Effects of Shocks of Milk Production and Loan Outcomes

	Milk Production (1)	Loan Watch (2)	Loan Finish (3)
Panel A. Direct Effects			
Post × *Shock*	-73.4***	0.2479***	-0.0815***
	(24.0)	(0.0255)	(0.0241)
Pre-shock mean	145.1	0.1283	0.0318
Shocks	321	321	321
Loans	598	598	598
Observations	5,381	5,381	5,381
Loans completed prior to shock	No	No	No
Loans nonperforming prior to shock	No	No	No
Loans not in propagation sample	No	No	No
Panel B. Propagation Effects			
Post × *Conductance*	0.6	0.0149**	-0.0092
	(17.6)	(0.0070)	(0.0079)
Pre-shock mean	224.8	0.0933	0.0527
Shocks	321	321	321
Loans	263	263	263
Pairs	11,527	11,527	11,527
Observations	103,743	103,743	103,743
Loans completed prior to shock	No	No	No
Loans nonperforming prior to shock	No	No	No
Connected pairs with distance ≥ 6	No	No	No

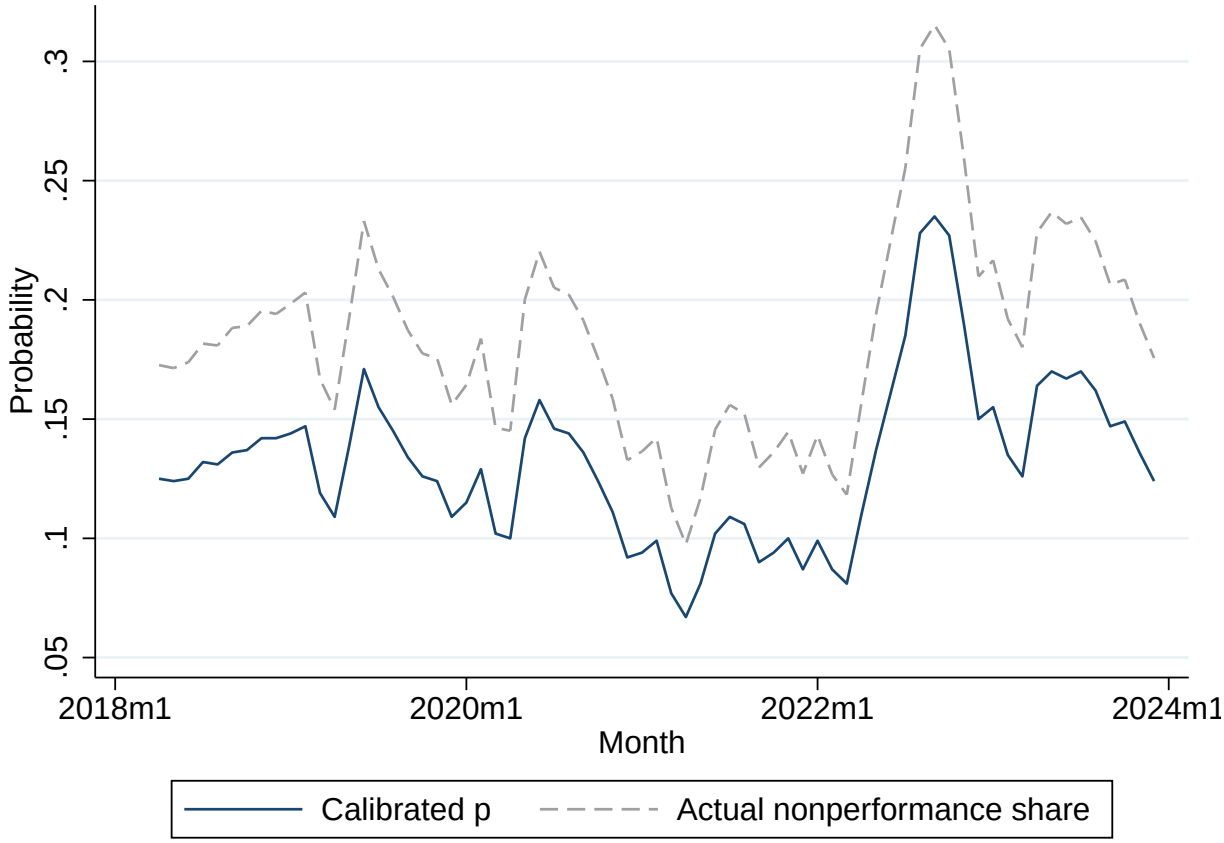
Notes: This table presents coefficient estimates from differences-in-differences regressions estimating both the direct and propagation effects of milk production shocks. The two columns correspond to two outcomes: column (1) *Loan Nonperformance*, a binary indicator that the loan is in nonperforming (Substandard) status, and column (2) *Milk Production*, litres of milk delivered to the dairy cooperative in a given month. Panel A reports direct effects from equation 4: the regressor of interest is *Post* × *Shock*, where the unit of observation is a loan-month for the shocked borrower's own loan. Panel B reports propagation effects from equation 6: the regressor of interest is *Post* × *Conductance*, where the unit of observation is a (shocked-loan, non-shocked-loan, month) triple and *Conductance* is the inverse of the effective resistance between the shocked and non-shocked loan members in the undirected network at the start of the non-shocked loan, equal to 1 for pairs linked by a single direct guarantee edge and 0 for unconnected pairs. *Shock* is a binary variable equal to 1 if the loan experienced a shock in the first half of the remaining loan term (treated loans); for control loans, placebo shocks taking place Δ months prior to their actual shock are constructed, with Δ the midpoint of the remaining loan term. *Post* is a binary variable equal to 1 from the month of the shock onwards (event months 0 through 5). For each loan, three months of pre-shock and six months of post-shock observations form a balanced panel. The sample is restricted to loans where the borrower has non-zero degree at the start of the loan and the loan was not already paid off in the month prior to the shock. All specifications include loan and time fixed effects; Panel B additionally includes shocked-loan and time-to-shock fixed effects. Controls interacted with time fixed effects include the out-degree and in-degree of the loan member at the start of the loan. Standard errors in parentheses are clustered at the loan level for direct effects and two-way clustered at the shocked- and non-shocked-loan levels for propagation effects. 1%, 5% and 10% statistical significance are indicated with ***, **, and * respectively.

Figure A1
Propagation Effects of Shocks on Milk Production



Notes: Placebo test analogous to Figure 4, but with *Milk Production* (litres delivered by the non-shocked loan member in a given month) as the dependent variable. The treatment variable, sample restrictions, fixed effects, controls, and standard errors are identical to Figure 4. Coefficients close to zero across the post-shock window indicate that propagation does not operate via correlated production shocks — *i*'s own milk production is not affected when a connected member *j* is shocked — and instead point to the financial channel of guarantee obligations as the mechanism behind the propagation effect on loan nonperformance documented in Figure 4. Confidence intervals are shown at the 95% level, with horizontal caps denoting 90% intervals.

Figure A2
Calibrated Fundamental Risk and Observed Nonperformance by Month



Notes: This figure plots, for each month from April 2018 to December 2023, the calibrated unconditional nonperformance probability $\hat{p}_t$ (solid navy) against the observed share of nonperforming loans d_t^{obs} (dashed grey). For each month, $\hat{p}_t$ is the value of the fundamental shock probability at which the calibrated independent-cascade model of Section 6 reproduces the observed nonperformance share on that month's guarantor network, holding the per-link contagion probability at its estimate $\hat{q} \approx 0.078$ and setting $\rho = 0$; it solves $\mathbb{E}[\, d \mid \mathcal{U}_t,\, p,\, q\,] = d_t^{\text{obs}}$, with the expectation evaluated by simulating S independent worlds at each candidate p (equation (9)). The calibrated series lies everywhere below the observed share, the gap reflecting contagion amplification through the guarantor network. Both series rise sharply in the second half of 2022; the three months August–October 2022 are the outliers excluded when fitting the normal distribution of fundamental risk reported in Section 6.

B Alternative Proximity Measures

This appendix documents the alternative proximity measures referenced in Section 3.3 that we considered but did not adopt as primary measures.

Weighted shortest-path distance. We construct a weighted analog of the shortest-path distance using the inverse of the guarantee amount as the edge cost. Stronger guarantees (larger g_{ij}) correspond to shorter weighted edges, so the weighted distance treats high-value guarantees as "closer" relationships. Formally, for the directed and undirected weighted networks $\mathcal{D}_t^W$ and $\mathcal{U}_t^W$, the weighted distance is

$$Distance_{ijt}^W = \min_{\pi: i \to j} \sum_{(u,v) \in \pi} \frac{1}{g_{uv,t}},$$

computed via Dijkstra's algorithm on the weighted graph.

Harmonic distance. Following Acemoglu, Ozdaglar and Tahbaz-Salehi (2015), we construct the harmonic distance h_{ij} as the expected hitting time of a random walk on the guarantee-amount-weighted directed network. The harmonic distance captures both directionality and path multiplicity simultaneously and is the natural distance measure for studying contagion in directed financial networks. It is defined only within the largest strongly connected component of $\mathcal{D}_t^W$.

Empirical performance. **[In progress.]** *Regression results using these alternative closeness measures* ($1/Distance^W$ *and* $1/h$) *— showing that they add no explanatory power beyond the shortest-path closeness* $1/Distance$ *and the conductance* $1/Resistance$ *used in the main text — are being finalised.*

C Dyadic Regressions with Member Fixed Effects

This appendix reports the alternative dyadic specification referenced in Section 4. We replace the pair fixed effects μ_{ij} in equation 1 with member fixed effects ϕ_i and ϕ_j:

$$y_{ijt} \;=\; \alpha\, Closeness_{ijt} \;+\; \phi_i + \phi_j \;+\; \lambda_t \;+\; \epsilon_{ijt}.$$

Member fixed effects absorb each individual's time-invariant characteristics (average default propensity, location, herd size, farming ability), but they leave several pair-level selection channels in the residual:

- Geographic and kinship proximity between i and j specifically;
- Shared occupation, common input suppliers, or correlated market exposure between the two;
- Pre-existing social ties that led i and j to enter into a guarantee relationship in the first place;
- Any other time-invariant determinant of who chooses to guarantee whom.

Interpretation. The gap between the member-fixed-effects coefficient and the pair-fixed-effects coefficient (Table 4) measures the contribution of pair-level selection to the raw correlation between proximity and outcomes. A larger member-FE coefficient indicates that selection accounts for some of the observed correlation; a similar coefficient under the two specifications suggests that selection effects are small.

Empirical results. **[In progress.]** *The table reporting member-fixed-effects coefficients alongside the pair-fixed-effects coefficients from the main text (for joint nonperformance, joint production shocks, and correlated milk production) is being finalised.*