BANK OF ENGLAND

Staff Working Paper No. 833

A structural model of interbank network formation and contagion

Patrick Coen and Jamie Coen

October 2019

BANK OF ENGLAND

Staff Working Paper No. 833

A structural model of interbank network formation and contagion

Patrick Coen[(1)] and Jamie Coen[(2)]

Abstract

The interbank network, in which banks compete with each other to supply and demand differentiated financial products, fulfils an important function but may also result in risk propagation. We examine this trade-off by setting out a model in which banks form interbank network links endogenously, taking into account the effect of links on default risk. We estimate this model based on novel, granular data on aggregate exposures between banks. We find that the decentralised interbank market is not efficient: a social planner would be able to increase surplus on the interbank market by 13% without increasing mean bank default risk or decrease mean bank default risk by 4% without decreasing interbank surplus. We then propose two novel regulatory interventions (caps on aggregate exposures and pairwise capital requirements) that result in efficiency gains.

Key words: Contagion, systemic risk, interbank network, network formation.

JEL classification: L13, L51, G28, G18.

(1) London School of Economics. Email: p.coen@lse.ac.uk
(2) London School of Economics and Bank of England. Email: w.j.coen@lse.ac.uk

The views expressed in this paper are those of the authors, and not necessarily those of the Bank of England or its committees. We are particularly grateful to Alessandro Gavazza and Christian Julliard for many helpful discussions. We are also grateful for comments by Áureo de Paula, Péter Kondor, Caterina Lepore, Jack McKeown, Martin Oehmke, Amar Radia, Wolfgang Ridinger, Claudia Robles-Garcia, Caspar Siegert, Rhiannon Sowerbutts, John Sutton, Matthew Willison, Kathy Yuan, Shengxing Zhang and Jean-Pierre Zigrand, as well as seminar participants at the London School of Economics and the Bank of England and conference participants at the RES Junior Symposium 2019 and EARIE 2019. We are grateful to the Bank of England for providing the data. Both authors acknowledge the financial support of the Economic and Social Research Council.

The Bank's working paper series can be found at www.bankofengland.co.uk/working-paper/staff-working-papers

Bank of England, Threadneedle Street, London, EC2R 8AH
Email publications@bankofengland.co.uk

ISSN 1749-9135 (on-line)

1 Introduction

Direct interconnections between banks are important in two ways. First, these interconnections fulfill a function, in that there are gains to trade. The interconnection could, for example, involve providing liquidity or acting as the other party in a hedging transaction, which may result in surplus on both sides of the trade. Second, interconnections can open up at least one side of the transaction to counterparty risk: a lender, for example, runs the risk that the borrowing bank will not pay it back. Both effects were important during the crisis and remain important today, and consequently there is significant debate about optimal regulation in this context (Yellen, 2013).

We consider the following two fundamental economic questions. How do banks form the interbank network,[1] given the effect of such exposures on their default risk? What inefficiencies exist in network formation? The answers to these economic questions then lead us to two questions about regulation. Given equilibrium responses by banks, is regulation effective in reducing default risk? If it does reduce default risk, does it do so efficiently in a way that preserves interbank surplus?

We answer these questions by estimating a structural model of interbank network formation and network outcome, which in this case is the propagation of default risk along the interbank network. In the network formation part of our model, banks compete with each other to supply and demand differentiated financial products on the interbank market. A bank that supplies a financial product to another bank receives a return, but also acquires an exposure that makes it riskier, via the default risk part of the model. We estimate this model based on novel, rich Bank of England data on interbank exposures, and show that the model fits the data well. We are the first, to our knowledge, to estimate a structural model of the trade-off between surplus on the interbank market and the causal effect of the interbank market on bank default risk. This allows us to draw novel conclusions about the efficiency of the interbank market, the systemic importance of banks within this network and optimal regulation in this context.

We use novel Bank of England data on interbank exposures. These data are collected by the Bank of England through periodic regulatory surveys of banks from 2012 to 2018, in which they report the exposures they have to their most important banking counterparties. The exposures reported cover a range of financial instruments, and the dataset contains

[1] The term “the interbank market” is often used to describe short-term (often overnight) lending between banks. Here we use it more generally to cover any form of direct interconnection between banks.

rich detail on which instruments make up each exposure. This combination of breadth and granularity enables us to study features of the interbank network that have previously been hidden from view. The data reveal certain trends including pairwise persistence in exposures, but also increasing concentration across the same period. The data also show significant variation across pairs in the mix of financial products being traded: different banks are supplying and demanding different things. This variation guides our modelling and estimation in that we emphasize heterogeneity across pairs wherever possible.

Our model consists of three parts: (1) the default risk process that relates the default risk of a bank to that of other banks and the exposures between them, (2) the demand for interbank financial products and (3) their supply.

We model the default risk process as being spatially[2] autocorrelated, but with a generalisation: the effect of exposures on default risk (in other words the network effect, or *contagion intensity*) is allowed to be heterogeneous across pairs. That is, some links are inherently more risky than others. The main reason why one might expect such variation is correlation in the underlying fundamentals of the banks involved. Banks demand interbank financial products to maximise profits from heterogeneous technologies that take these differentiated interbank products as inputs. Banks supplying financial products face a trade-off between the returns they get from supplying financial products and the effect of supplying on their default risk (via the default risk process described above). The key mechanism in our model is that the cost of capital a bank faces is a function of its default risk, meaning that a bank supplying financial products endogenously increases its cost of capital when it does so. This occurs in two ways: by increasing the price it pays to be supplied financial products on the interbank market, and by increasing its cost of raising capital outside the interbank market. In practice, the model is essentially a multi-product Cournot game with a non-linear cost function with cost linkages across products and cost externalities across competing firms.

Equilibrium trades and prices depend in an intuitive way on the key parameters of the model: (i) variation in contagion intensity is a key driver of link formation; links form where they are least risky, (ii) risky banks pay more to be supplied financial products because of contagion and (iii) risky banks supply less, as their cost of doing so is higher. The sources of market failure are market power, inefficient allocation of outputs given cost differences and network externalities. We show that our model is consistent with the key summary statistics

[2]Where the measure of "proximity" in this case is the size of the exposure between two nodes. That is, the effect of bank j's default risk on bank i depends on the exposure of i to j.

in our data, as well as some additional stylised facts from the financial crisis.

The central part of the network formation model considers how equilibrium exposures change as the riskiness of banks changes. We identify this relationship by exploiting variation in regional equity indices as an instrumental variable for bank risk: for example, we take a shock to a Japanese equity index as a shock that affects Japanese banks and European banks differently. We argue that this shock is plausibly exogenous when we include the rich set of fixed effects that our panel network data enables (that is, we can control for any unobserved variation that varies across banks but not across pairs): validity requires only that these indices are not correlated with unobserved *pairwise* variation in the interbank market.

The central part of the default risk process considers how default risk changes as exposures change. In contrast to large parts of the network econometric literature (see De Paula (2017) for a summary) when estimating this relationship we consider the endogeneity of the network directly, using insights from the network formation part of our model. The default risk process is, by assumption, linear in the fundamentals of banks, but our network formation game shows that equilibrium network links are *non-linear* functions of bank fundamentals. We therefore use non-linear variation in bank fundamentals as instruments for equilibrium links in the default risk process.

We estimate our model and show that it fits the data well. We find significant variation in pairwise contagion intensity: some links are inherently riskier than others. This heterogeneity in network effect has an important implication for the identification of systemically important banks within our network. Using standard measures, a systemically important bank has large links to other banks that are systemically important, loosely speaking. Heterogeneous network effects and endogenous network formation together show why this approach is likely to be flawed: *some links are large because they are inherently safe*. Banks with large links like these would be incorrectly characterised as systemically important using standard network centrality measures based on unweighted network data. We propose an alternative measure of systemic importance based on network data that is weighted by the heterogeneous network effect parameters (that is, an inherently risky (safe) link is scaled up (down)), and show that this implies different rankings among banks.

We then use our estimated results to answer the key questions set out above. First, we consider the efficiency of the decentralised interbank market, which we do by deriving an efficient frontier that shows the optimal trade-off between interbank surplus and bank default risk. We find that the decentralised interbank market is not on the frontier: a social planner would be able to increase interbank surplus by 13.2% without increasing mean bank default risk or decrease mean bank default risk by 4.3% without decreasing interbank surplus. This result is driven by the fact that our empirical results indicate that banks supplying exposures have market power and network externalities are significant. The social planner internalises the externality by considering the effect that a given link has on the risk of other banks, with the result that the social planner would (i) reduce aggregate exposures and (ii) reduce inherently risky exposures by relatively more than inherently safe exposures.

Second, we use our model to simulate various forms of regulation, including a cap on individual exposures (BCBS, 2014b, 2018b) and an increase in regulatory capital requirements (BCBS, 2018a). We find that a cap on individual links reduces interbank surplus with only a small effect on mean bank default risk, as in equilibrium banks shift their supply to uncapped links. We find that this effect can be mitigated by capping aggregate exposures held by each bank, rather than a cap on individual exposures: an aggregate cap prevents a bank moving capped supply to another bank, and so is more effective in reducing bank default risk.

We find that a general increase in capital requirements that applies equally across exposures to all banks decreases mean bank default risk, but at the cost of reduced interbank surplus. We find that this effect can be mitigated by a pairwise adjustment to capital requirements based on their heterogeneous contagion intensity: we give links that are relatively risky (less risky) greater (lower) capital requirements. In other words, we propose directly risk-weighting interbank exposures based on contagion intensity, as this targets regulatory intervention more closely at the network externalities that are the key driver of inefficiency in our model. Our results suggest that our proposed alternative produces better outcomes than a risk-insensitive capital requirement, such that it would be strictly preferred by a social planner.

We discuss related literature below. In Section 2, we introduce the institutional setting and describe our data. In Section 3, we set out our model. In Section 4, we describe our identification strategy. In Section 5, we set out our results. In Section 6, we undertake counterfactual analyses. In Section 7, we conclude and discuss further applications of our work.

1.1 Related literature

Our work is related to three strands of literature: (i) endogenous network formation in financial markets, (ii) the effects of network structure on outcomes in financial markets and (iii) optimal regulation in financial markets.

There is a growing theoretical literature on network formation in financial markets (Babus, 2016; Farboodi, 2017; Chang and Zhang, 2018), but little empirical work (Cohen-Cole et al., 2010; Craig and Ma, 2018; Blasques et al., 2018). This literature typically studies the drivers of the structure of financial markets, and does not consider the implications of this structure for systemic risk. Our contribution is that we are the first, to our knowledge, to structurally estimate a model of network formation in which banks trade off gains to interbank trade against contagion. Importantly, this allows us to quantify the extent of inefficiency in the market, and to study the implications of network structure for systemic risk.

There is an extensive literature on the effect of network structure on outcomes in financial markets, both theoretical (Acemoglu et al., 2015; Ballester et al., 2006; Elliott et al., 2014) and empirical (Denbee et al., 2017; Eisfeldt et al., 2018; Gofman, 2017; Iyer and Peydro, 2011) The empirical literature typically takes as given the financial network, and then studies how exogenous changes in the network's structure affect outcomes like risk and total surplus, as well as identifying systemically important banks within the network. By explicitly modelling the formation of the network, our contribution is to more clearly identify network effects and consider more realistic counterfactual scenarios. Regarding identification, by endogenising the network we are able to separately identify the effects of market structure from the effects of the determinants of market structure. This also enables us to better identify systemically important banks. Regarding counterfactuals, various papers (Eisfeldt et al. (2018) and Gofman (2017), for example) adjust the network arbitrarily (usually by simulating a failure) and show the impact on market outcomes holding network structure otherwise fixed. In our model, network structure responds endogenously to such a change.

There is a specialist literature regarding optimal regulation in financial markets (Duffie, 2017; Baker and Wurgler, 2015; Greenwood et al., 2017; Batiz-Zuk et al., 2016). Our primary contribution is that by considering bank default risk we are able to evaluate bank regulation comprehensively. Various papers consider the effect of bank regulation on outcomes in specific markets,[3] but without considering bank default risk (which was arguably the pri-

[3]Including Kashyap et al. (2010) on bank lending, Kotidis and Van Horen (2018) on the repo market

mary focus of much recent banking regulation) it is not possible to draw any conclusions about whether regulation is optimal. Furthermore, our network formation model allows us to specifically address how the network will respond endogenously to any change in regulation.

2 Institutional setting and data

2.1 Institutional setting

Direct connections between banks fulfill an important function: "*there is little doubt that some degree of interconnectedness is vital to the functioning of our financial system*" (Yellen, 2013). Debt and securities financing transactions between banks are an important part of liquidity management, and derivatives transactions play a role in hedging. There is, however, widespread consensus that direct connections can also increase counterparty risk, with implications for the risk of the system as a whole (see, for example, Acemoglu et al. (2015)). This can be thought of, in loose terms, as a classic risk/reward trade-off. The importance of both sides of this trade-off is such that direct interconnections between banks are the subject of extensive regulatory and policy-making scrutiny, whose aim is to: "*preserve the benefits of interconnectedness in financial markets while managing the potentially harmful side effects*" (Yellen, 2013).

After the 2008 financial crisis, a broad range of regulation was imposed on these markets. In this paper, we focus on two in particular: (1) caps on large exposures and (2) increases in capital requirements. We focus on these two because we think they are most relevant to our underlying economic research question, which is to examine the *efficiency* with which this risk/reward trade-off is balanced.

2.1.1 Large exposures cap

In 2014 the Basel Committee on Banking Supervision (BCBS) set out new standards for the regulatory treatment of banks' large exposures (BCBS, 2014b, 2018b). The new regulation, which came into force in January 2019, introduces a cap on banks' exposures: a bank can have no single bilateral exposure greater than 25% of its capital.[4] For exposures held between

and Bessembinder et al. (2018) and Adrian et al. (2017) on the bond market.

[4] Where the precise definition of capital, in this case "Tier 1 capital", is set out in the regulation (BCBS, 2014b, 2018b)

two "globally systemic institutions", as defined in the regulation, this cap is 15%.

These requirements represent a tightening of previous rules, where they existed. For example, in the EU exposures were previously measured relative to a more generous measure of capital and there was no special rule for systemically important banks (AFME, 2017; European Council, 2018).

2.1.2 Capital requirements

Banks are subject to capital requirements, which mandate that their equity (where the precise definition of capital, Common Equity Tier 1, is set out in the regulation) exceeds a given proportion of their risk-weighted assets. Additional equity in principle makes the bank more robust to a reduction in the value of its assets, and so less risky. The total amount of capital E_{ij} that bank i is required to raise to cover asset j is the product of the value of the asset A_j, its risk-weighting ρ_{ij} and the capital requirement per unit of risk-weighted asset λ_i:

$$E_{ij} = \rho_{ij}\lambda_i A_j$$

The risk-weights, ρ_{ij}, can be calculated using banks' internal models or based on a standardised approach set out by regulators. Whilst risk-weights from banks' internal models are likely to vary by counterparty, the standardised approach is based on the credit rating relevant to the asset, and for the significant majority of interbank transactions between major banks this will be AAA or AA, the highest credit rating. In other words, for interbank transactions the standardised approach involves very little variation across i or j.[5]

In 2013 all banks in our sample faced the same capital requirement per risk-weighted unit, λ_i, which was 3.5%.[6] Since then, regulators have changed capital requirements in three ways. First, and most importantly, the common minimum requirement that applies to all banks has increased significantly. Second, capital requirements vary across banks, as systemically important banks face slightly higher capital requirements than non-systemically important banks. Third, capital requirements vary countercylically, in that in times of financial distress they are slightly lower (BCBS, 2018a). The result of these changes is that mean capital requirements for the banks in our sample has increased significantly, from

[5] Banks are also subject to a leverage ratio requirement (BCBS, 2014a) which does not weight exposures according to risk.

[6] We use the minimum capital requirements as published by BCBS (2011) as the minimum requirements for banks. National supervisors can add discretionary buffers on top of these requirements, which we do not include in our empirical work.

3.5% to over 9% in 2019. There have also been changes to the definition of capital and the measurement of risk-weighted assets, with the general effect of making capital requirements more conservative.

2.2 Data

2.2.1 Exposures

We define in general terms the exposure of bank i to bank j at time t as the immediate loss that i would bear if j were to default, as estimated at time t. The way in which this is calculated varies from instrument to instrument, but in general terms this can be thought of as (1) the value of the instrument, (2) less collateral, (3) less any regulatory adjustments intended to represent counterfactual variations to value or collateral in the event of default (for example, regulation typically requires a "haircut" to collateral when calculating exposures, as in the event of default any financial instruments provided as collateral are likely to be worth less).

We use regulatory data on bilateral interbank exposures, collected by the Bank of England. The dataset offers a unique combination of breadth and detail in measuring exposures. Much of the existing literature (such as Denbee et al. (2017)) on empirical banking networks relies on data from payment systems. This is only a small portion of the activities that banks undertake with each other and is unlikely to adequately reflect the extent of interbank activity or the risk this entails.

18 of the largest global banks operating in the UK report their top 20 exposures to banks over the period 2011 to 2018. Banks in our sample report their exposures every six months from 2011 to 2014, and quarterly thereafter. They report exposures across debt instruments, securities financing transactions and derivative contracts. The data are censored: we only see each bank's top 20 exposures, and only if they exceed £5 million. The data include granular breakdowns of each of their exposures: by type (e.g. they break down derivatives into interest rate derivatives, credit derivatives etc.), currency, maturity and, where relevant, collateral type.

We use this dataset to construct a series of snapshots of the interbank market between these 18 banks. We calculate the total exposure of bank i to bank j at time t, which we denote C_{ijt}, as the sum of exposures across all types of instrument in our sample. We winsorize exposures at the 99th percentile. The result is a panel of $N = 18$ banks over $T = 21$ periods from 2011 to 2018 Q2, resulting in $N(N-1)T = 6,426$ observations. For each C_{ijt}, we use the granular breakdowns to calculate underlying "exposure characteristics" that summarise the type of financial instrument that make up the total exposure. These 8 characteristics, which we denote d_{ijt}, relate to exposure type, currency, maturity and collateral type.

Although the dataset includes most of the world's largest banks, it omits banks that do not have a subsidiary in the UK.[7] Furthermore, for the non-UK banks that are included in our dataset, we observe only the exposures of the local sub-unit, and not the group. For non-European banks, this sub-unit is typically the European trading business.

2.2.2 Default risk

We follow Hull et al. (2009) and Allen et al. (2011) in calculating the (risk-neutral) probability of bank default implied by the spreads on publicly traded credit default swaps (data obtained from Bloomberg). This represents the market's estimate of bank default risk, as well as wider effects that are unrelated to the default risk of an individual bank (notably variations in the risk premium):

$$Prob(Default_{itT}) = 100(1 - (1 + (CDS_{itT}/10000)(1/rr))^{-T}$$

where rr is the assumed recovery rate, T is the period covered by the swap and CDS_{itT} is the spread.

2.2.3 Other data

We supplement our core data with the following:

- Geographic source of revenues for each bank from Bloomberg. Bloomberg summarises information from banks' financial statements about the proportion of their revenues that come from particular geographies, typically by continent, but in some cases by country.
- Macro-economic variables from the World Bank Global Economic Monitor, a panel of 348 macro series from a range of countries.

[7] This is particularly relevant for some major European investment banks, who operate branches rather than subsidiaries in the UK, and hence do not appear in our dataset.

- Commodity prices from the World Bank "Pink Sheet", which is a panel of 74 commodity prices.
- S&P regional equity indices for US, Canada, UK, Europe, Japan, Asia, Latin America.

2.3 Summary statistics

The data reveal certain empirical observations about exposures and how they vary cross-sectionally and inter-temporally in our sample: (1) exposures in our data are large, (2) our observed network is dense and reciprocal, (3) network links are heterogeneous in intensity and characteristics and (4) the network has become more concentrated over our sample network. We discuss below how we use these empirical observations to guide our modelling.

Empirical fact 1: Exposures are large

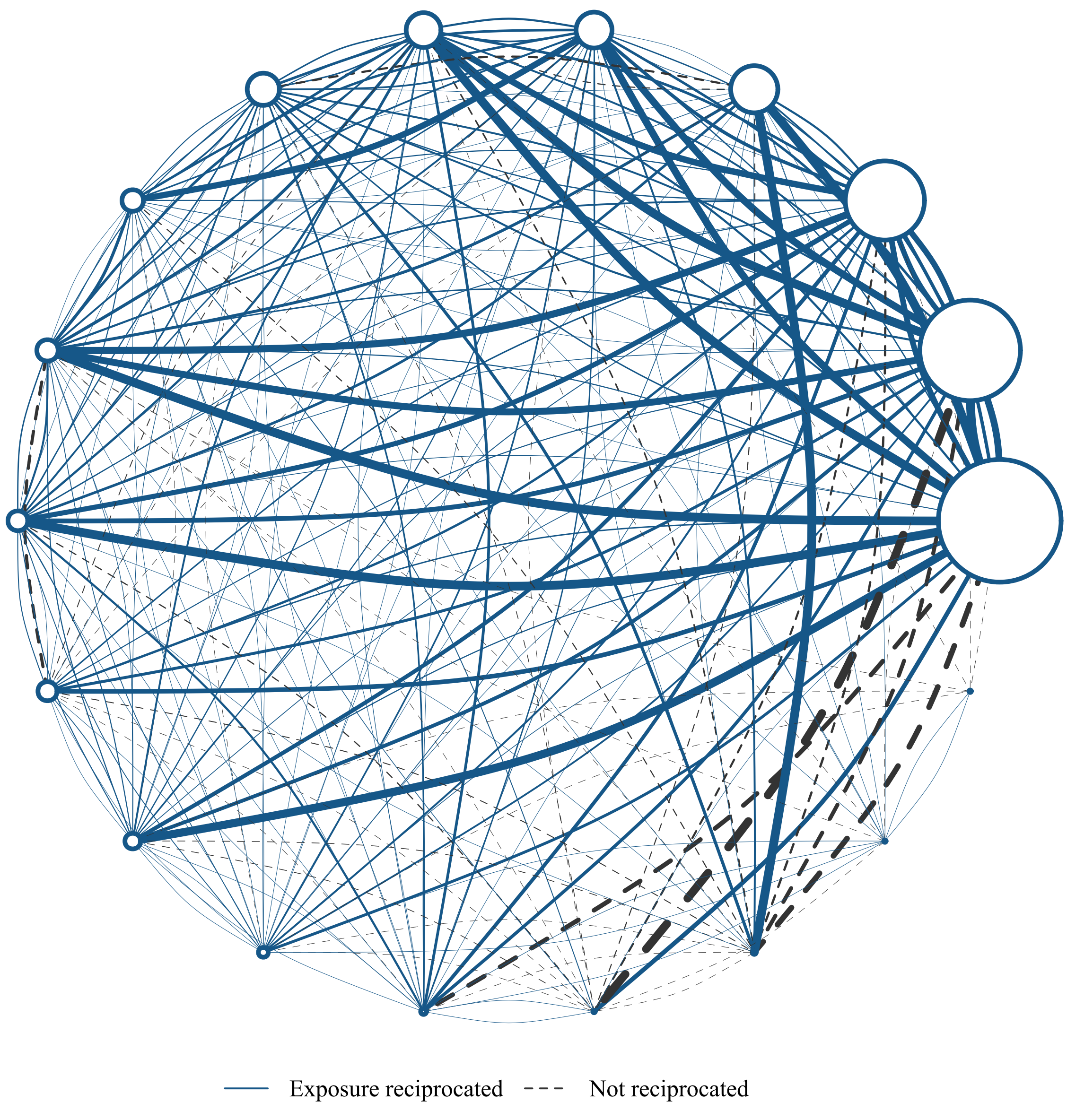
The primary advantage of our data, relative to others used in the literature, is that it is intended to capture a bank's *total* exposures. The largest single exposure in our sample is GBP $7,682$m, the largest total exposures to other banks in a given period is GBP $26,367$m. The mean exposure is GBP 285m and the mean total exposure to other banks in a given period is GBP $4,851$m.

In this respect, our data has two important advantages over many of the data used in the literature. First, our dataset is the closest available representation of *total* exposures, when most other empirical assessments of interbank connections rely on a single instrument, such as CDS (Eisfeldt et al., 2018) or overnight loans (Denbee et al., 2017). Second, our data are on exposures, rather than simply market value, in that when banks report their exposures they account for collateral and regulatory adjustments. Data based solely on market value are a representation of bank activity, rather than counterparty risk.

Empirical fact 2: The network is dense and reciprocal

Figure 1 shows the network of exposures between banks in 2015 Q2. Our sample is limited to the core of the banking network, and does not include its periphery. Our observed network is, therefore, dense: of the $N(N-1)T$ links we observe in total, only approximately 30% are

Figure 1: The aggregate network in H1 2015



Note: the solid line shows reciprocated links (each bank supplies the other) and the dashed lines shows unreciprocated exposures (that go in one direction only). The line width is proportional to the size of the exposure. The size of the node is proportional to its total outgoings.

0. One implication of the density of the network is that it is reciprocal: of the $N(N-1)T/2$ possible bilateral relationships in our sample, 55% are reciprocal, in that they involve a

strictly positive exposure in each direction (that is, bank i has an exposure to bank j and bank j has an exposure to bank i).

Empirical fact 3: The network is heterogeneous in intensity and characteristics

Although the network is dense and so not particularly heterogeneous in terms of the presence of links, it is heterogeneous in the intensity of those links (that is, the size of the exposure), as shown in Figure 1. We further demonstrate this in Table 1, which contains the results of a regression of our observed exposures C on fixed effects. The R^2 from a regression on it fixed effects is 0.43: if all of bank i's exposures in a given time period were the same, then this would be 1.00. In other words, the low R^2 indicates that there is significant variation in the size of exposures.

There is significant persistence in exposures, as set out in Table 1, in which we show that the R^2 for a regression of C_{ijt} on pairwise ij fixed effects is 0.67. In other words, a large proportion of the variation in exposures is between pairs rather than across time.

Table 1: Variation and persistence in network

	it	jt	ij
C_{ijt}	$R^2 = 0.43$	0.16	0.67
No. obs	6,426	6,426	6,426

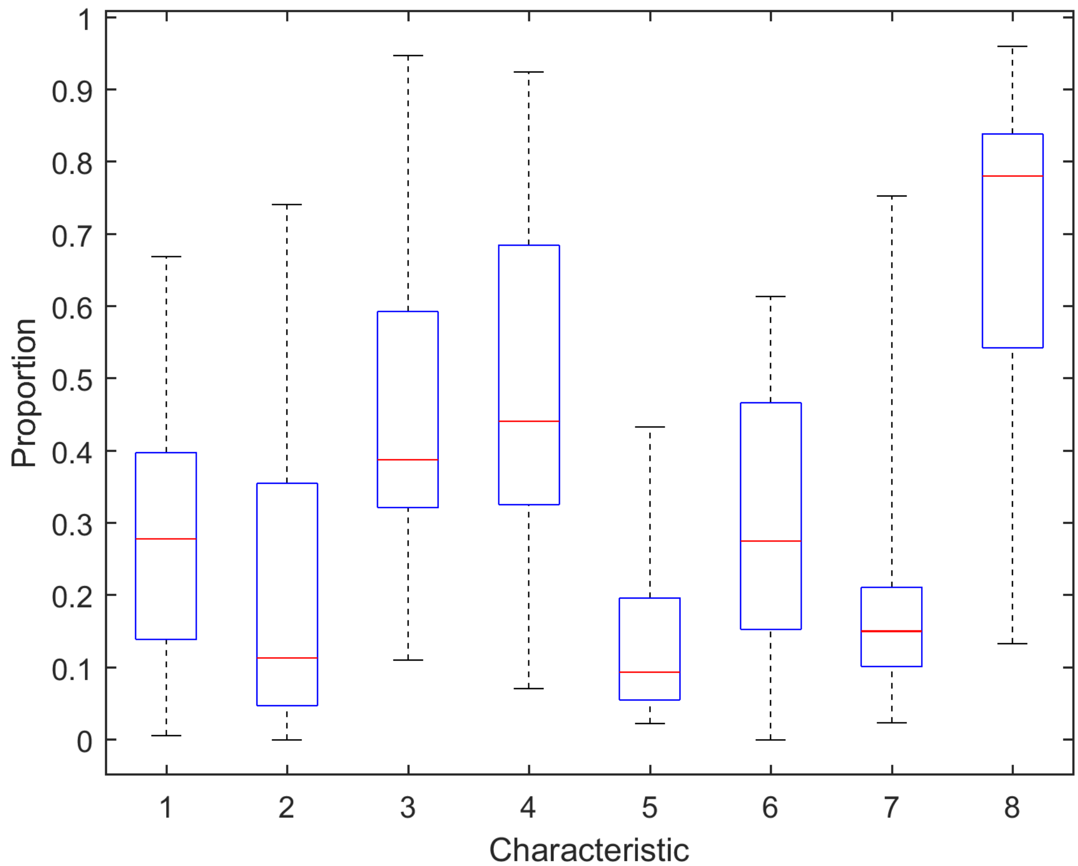
Note: This table shows the R^2 obtained from regressing observed network links on dummy variables. jt, for example, indicates that the regressors are dummy variables for each combination of j and t.

There is significant variation in product characteristics, as set out in Figure 2. This figure shows the minimum, interquartile range, median and maximum value for each of the exposure characteristics in the data.

Empirical fact 4: The network has increased in concentration over time

Even though the network is persistent, there is still inter-temporal variation. In particular, concentration in the interbank market has increased over time, in that the Herfindahl-

Figure 2: Variation in exposure characteristics



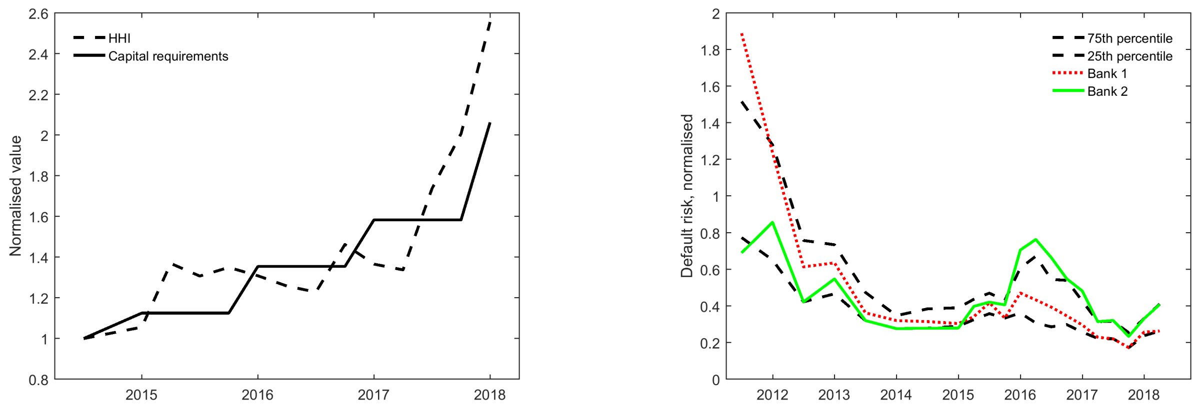
Characteristic: (1) proportion in EUR, (2) proportion in GBP, (3) proportion in USD, (4) maturity less than 3 months, (5) maturity 3 months to 1 year, (6) maturity open or overnight, (7) proportion that is repo, (8) proportion that is derivatives.

Hirshmann index[8] over exposure supply has increased, as set out in Figure 3. In Figure 3, we show that the HHI index and regulatory capital requirements are closely correlated. It is obviously not possible to draw any causal conclusions from such a graph, but the relationship between concentration and capital requirements will be an important part of our model and identification.

[8] $HHI_t = \frac{1}{N}\sum_j \sum_i s_{ij}^2$, where s_{ij} is the share of bank i in the total supply to bank j: $s_{ij} = \frac{C_{ij}}{\sum_i C_{ij}}$. Larger HHI indicates greater concentration. Because of the group-to-unit measurement issue we describe above, we weight exposures in our calculation of HHI by $(\frac{1}{NT}\sum_t \sum_j C_{ijt})^{-1}$. In this sense our measure of HHI is concentration within the i-bank.

Figure 3: Summary statistics

(a) Concentration (b) Default risk



Note: There was a change in the way our data was collected that mean comparing concentration before and after 2014 is not meaningful, so in the left-hand graph we restrict our sample to 2014 onwards.

Empirical fact 5: Bank default risk has decreased

Our sample runs from 2011 to 2018, and therefore earlier periods feature the end of the European debt crisis. Bank default risk has broadly reduced across all banks, as we set out in Figure 3. Importantly, though, there is cross-sectional variation across banks, and inter-temporal variation in that cross-sectional variation. We show this in Figure 3, in which we highlight the default risk of two specific banks. Bank 1 (Bank 2) was in the top (bottom) quartile by bank default risk in 2011, but the bottom (top) quartile by 2018.

2.4 Stylised facts

Our sample starts in 2011, so it does not feature the financial crisis that began in 2008. We note three features that were observed on the interbank market during the 2008 crisis, on the basis that a good model of interbank network formation should be able to replicate what happened during the crisis. First, risky banks were not supplied; in other words, they experienced *lockout* (Welfens, 2011). Second, risky banks did not supply, which we loosely term *liquidity hoarding* (Gale and Yorulmazer, 2013). Third, in the worst periods of the financial crisis there was effectively *market shutdown* in markets for certain instruments, in that very few banks were supplied anything on the interbank market (Allen et al., 2009).

3 Model

3.1 Setup and notation

There are N banks. At time t, the interbank market consists of an $N \times N$ directed adjacency matrix of total exposures, $\mathbf{C_t}$. C_{ijt} is the element in row i and column j of $\mathbf{C_t}$, and indicates the total exposure of bank i to bank j at time t. $\mathbf{C_t}$ is directed in that it is not symmetric: bank i can have an exposure to bank j, and bank j can have a (different) exposure to bank i. For each bank i, $\mathbf{d_i}$ is an $L \times 1$ vector of product characteristics for the exposures that it supplies.

$\mathbf{p_t}$ is an $N \times 1$ vector of bank default risks: the element in position i is the probability of default of bank i. $\mathbf{p_t}$ is a function of $\mathbf{C_t}$ and an $N \times K$ matrix of bank fundamentals, which we denote $\mathbf{X_t}$, and which update over time according to some exogenous process. This function is the default risk process, and the effect of $\mathbf{C_t}$ on $\mathbf{p_t}$ represents "contagion", as we will define more formally below.

C_{ijt} results in profits to bank i (we term this supply of exposures) and to bank j (demand for exposures). These profits depend on bank default risk, in a way we will formalise below. The equilibrium interbank network $\mathbf{C_t}$ is formed endogenously based on the supply- and demand-sides, such that markets clear. Banks choose their supply and demand decisions simultaneously. For simplicity, there is no friction between changes in bank fundamentals and the formation of the network: once fundamentals change, the equilibrium network changes immediately.[9]

3.2 Default risk process

Understanding the effect of exposures on default risk is a key part of our research question. In our approach to modelling this default risk process, we are guided by the summary statistics we set out above in three important ways:

- First, in our dataset the *exposures are large and complete* (empirical fact 1), which means that the exposures could reasonably have an impact on the default risk of the

[9] It is straightforward to introduce some friction in the timing, such that the network does not update immediately once fundamentals change. This would allow more detailed consideration of shock propagation in the *short-run*, which we define as the interval in which the network has not updated. We consider these short-run effects in further work, and consider in this paper only the *long-run* effects of changes in fundamentals.

banks that hold these exposures. In other words, the size of our observed exposures leads us to consider financial contagion of default risk through these exposures.

- Second, there is *cross-sectional variation in exposure characteristics* (empirical fact 3): in other words, firms are trading different financial products. Some financial products may not impact default risk in the same way as others: as a trivial example, holding GBP 100m of senior debt of bank j may have a smaller effect on the default risk of bank i than holding GBP 100m of junior debt. This empirical fact means that we need to take a flexible approach to modelling contagion that accounts for this heterogeneity.

- Third, there is *cross-sectional variation in bank default risk* (empirical fact 5). There is a broad theoretical literature on the importance of such cross-sectional variation for financial contagion: the effect of an exposure to bank j on bank i's default risk is likely to depend on the extent to which their underlying fundamentals are correlated (Glasserman and Young, 2015; Elliott et al., 2018). Our model of contagion, therefore, needs to be sufficiently flexible to account for this heterogeneity.

We model a bank's default risk process as the sum of two components: a set of fundamentals and a spatially autocorrelated component whereby bank i's default risk depends on its aggregate exposure to bank j, C_{ijt}, and bank j's default risk, p_{jt}. In matrix form:

$$\underbrace{\mathbf{p_t}}_{\text{Default risk}} = \underbrace{\boldsymbol{X_t\beta}}_{\text{Fundamentals}} + \underbrace{\tau_t(\boldsymbol{\Gamma} \circ \mathbf{C_t})\mathbf{p_t}}_{\text{Network spillovers}} + \mathbf{e_t^p}$$

where $\mathbf{p_t}$ is an $N \times 1$ vector of bank default risks, $\mathbf{C_t}$ is an $N \times N$ directed adjacency matrix of aggregate pairwise exposures, $\boldsymbol{\beta}$ is a $K \times 1$ vector that represents each bank's loadings on a $N \times K$ matrix of fundamentals $\mathbf{X}$, $\boldsymbol{\Gamma}$ is an $N \times N$ matrix of contagion intensities, τ_t is a scalar that allows for contagion intensities to vary across time and $\circ$ signifies the Hadamard product.

This is a spatially autocorrelated regression, as is commonly used in network econometrics (De Paula, 2017), with a generalisation: the parameter governing the size of the network effect, Γ_{ij}, is allowed to be heterogeneous across bank pairs. Before we explain the effect of this generalisation, we first define *contagion* from bank j to bank i as $\frac{\partial p_{it}}{\partial p_{jt}} > 0$: that is, the default risk of bank j has a causal impact on the default risk of bank i. In our model, $\frac{\partial p_{it}}{\partial p_{jt}} = \tau_t \Gamma_{ij} C_{ijt}$, such that the strength of contagion depends on the size of the exposure and this parameter Γ_{ij}.

Γ can be thought of as *contagion intensity* in that $\Gamma_{ik} > \Gamma_{im}$ implies that $\frac{\partial p_{it}}{\partial p_{kt}} > \frac{\partial p_{it}}{\partial p_{mt}}$ for any common $C_{ikt} = C_{imt}$. That is, bank i's default risk is more sensitive to exposures to bank k than to bank m, holding exposures and fundamentals constant. This heterogeneity could come from three sources. First, it could be a result of correlations in the underlying fundamentals, as described above, whereby if bank i and k (m) have fundamentals that are positively (negatively) correlated then exposure C_{ik} (C_{im}) is particularly harmful (benign). This implies a relationship between the fundamentals processes and Γ_{ij} which we leave open for now, but consider in our empirical analysis. Second, it could be a result of variations in product characteristics, as described above. This difference across products could be modelled using a richer default risk process that separately includes exposures matrices for each instrument type with differing contagion intensities, but this would introduce an infeasible number of parameters to take to data. Third, it could be a result of some other relevant pairwise variation that is unrelated to fundamentals or product, such as geographic location. It could be, for example, that recovery rates in the event of default are lower if bank i and bank j are headquartered in different jurisdictions, making cross-border exposures riskier than within-border exposures.

We allow for contagion intensity to vary across time via τ_t because there are, in principle, things that could affect contagion intensity. One of the purposes of the increase in capital requirement, for example, was to make holding a given exposure C_{ijt} safer, in the sense of Modigliani and Miller (1958) (because it means bank i has a greater equity buffer if bank j defaults). We do not make any assumptions about the relationship between τ_t and capital requirements λ at this stage, but consider it in estimation.

Subject to standard regularity conditions on $\mathbf{\Gamma}$ and $\mathbf{C}$ this spatially autocorrelated process can be inverted and expanded as a Neumann series as follows, which we term the Default Risk Process ("DRP"):

$$\mathbf{p_t} = (\mathbf{I} - \tau_t \mathbf{\Gamma} \circ \mathbf{C_t})^{-1}(\mathbf{X_t}\boldsymbol{\beta} + \mathbf{e^p_t}) = \sum_{s=0}^{\infty} (\tau_t \mathbf{\Gamma} \circ \mathbf{C_t})^s (\mathbf{X_t}\boldsymbol{\beta} + \mathbf{e^p_t})$$

3.3 Demand

In our approach to modelling demand we are guided by one important empirical fact: *product characteristics are heterogeneous* across banks (empirical fact 3). In other words, banks are supplying and demanding different financial products. This has two important implications:

- First, this heterogeneity has implications for the specificity with which we model the payoffs to demanding financial products. For example, if our empirical exposures were uniquely debt, then we would be able to include a standard model of liquidity management on the demand-side (as in Denbee et al. (2017)). If instead our empirical exposures were uniquely CDS contracts, then we would be able to include a model of credit risk management (as in Eisfeldt et al. (2018)). Instead, we need to model the demand-side in a general way that is applicable across the range of financial products that feature in our data.

- Second, this heterogeneity has implications for how we model competition between banks. In particular, this heterogeneity means we need to consider the extent to which exposures supplied by one bank are substitutable for those supplied by another bank (product differentiation, in other words).

Each j-bank has a technology that maps inputs into gross profit, from which the cost of inputs is subtracted to get net profits. Inputs are funding received from other banks $C_{ij}, \forall i \neq j$ and an outside option C_{0j} designed to capture funding from banks outside our sample and non-bank sources. Net profits are given by:

$$\Pi_{jt}^{D} = (\zeta_{ij} + \delta_{jt} + e_{ijt}^{D}) \sum_{i=0}^{N} C_{ijt}$$
$$-\frac{1}{2}\left(B \sum_{i=0}^{N} C_{ijt}^{2} + 2 \sum_{i=0}^{N} \sum_{k \neq i}^{N} \theta_{ik} C_{ijt} C_{kjt}\right)$$
$$-\sum_{i=0}^{N} r_{ijt} C_{ijt}$$

where ζ_{ij} and δ_{jt} represent heterogeneity in the sensitivity of the j-bank's technology to product i, B governs diminishing returns to scale and θ_{ik} governs the substitutability of product i and k. Before we motivate our choices about functional form in more detail, it is helpful to set out what this implies for the j-bank's optimal actions. Bank j chooses C_{ijt}^{D} to maximise net profit taking interest rates as given, resulting in optimal C_{ijt}^{D} such that inverse demand is as follows:

$$r^D_{ijt} = \underbrace{\zeta_{ij} + \delta_{jt}}_{\text{Technology}} - \underbrace{BC_{ijt}}_{\text{Own-effect}} - \underbrace{\sum_{k \neq i} \theta_{ik} C_{kjt}}_{\text{Cross-effect}} + \underbrace{\theta_0 C_{0jt}}_{\text{OOP}} + e^D_{ijt}$$

In other words, our functional form assumptions imply that the bank demanding exposures has linear inverse demand.

We assume that the j-bank has an increasing but concave objective function in the funding that it receives. We justify its concavity on the basis that the j-bank undertakes its most profitable projects first (or conversely, if its funding is restricted for whatever reason, it terminates its least profitable projects rather that its most profitable projects). Concavity also means that the returns to receiving funding decrease, in that the j-bank only has a limited number of opportunities for which it needs funding.

The intercept is comprised of three parts: δ_{jt}, ζ_{ij} and e^D_{ijt}. δ_{jt} ensures that the returns that the j-bank gets from funding are time-varying. This time variation is left general, although it could be related to the j-bank's fundamentals. It could be, for example, that when the j-bank's fundamentals are bad then the payoff to receiving funding is greater, in that the projects being funded are more important (if, for example, it needs this funding to undertake non-discretionary, essential projects or to meet margin calls on other funding). This is intended to allow for the importance of the interbank market in times of distress. The technologies possessed by each j-bank vary by ζ_{ij}, which governs the importance of the i-bank's product to the j-bank's technology. We allow this technology to be heterogeneous across pairs. e^D_{ijt} is an iid shock to the returns that bank j gets from receiving funding from bank i.

We also allow for product differentiation, in that the product supplied by bank i may not be a perfect substitute for the product supplied by bank k. We parameterise this product differentiation in parameters we denote θ_{ik}. General θ_{ik} cannot be reasonably estimated from our dataset; instead we parameterise it as being a logistic function of certain product characteristics, including maturity, currency and instrument-type.

$$\theta_{ik} = \frac{exp\Bigg(\tilde{\theta} - \sum_l^L \tilde{\theta}_l (d_{i,l} - d_{k,l})^2\Bigg)}{1 + exp\Bigg(\tilde{\theta} - \sum_l^L \tilde{\theta}_l (d_{i,l} - d_{k,l})^2\Bigg)} + e^\theta_{ik}$$

where $d_{i,l}$ denotes the value for characteristic l of bank i, e^{θ}_{ik} denotes random, unobserved variation in product differentiation and $\tilde{\theta}_l > 0$. For instrument type, for example, $d_{i,l=type}$ is the proportion of i's product that is derivatives. If banks i and k have very different product characteristics, then θ_{ik} is small and the two are not close substitutes. If, on the other hand, banks i and k have very similar product characteristics then θ_{ik} is large and the two are close substitutes. This parameterisation replaces θ_{ik} (which across all pairs has dimension N^2) with $\tilde{\theta}_l$ (which has dimension $L+1$).

3.4 Supply

In our approach to modelling the supply side, we are guided by the following empirical observations: the network we are seeking to model is *dense with heterogeneous intensities* (empirical facts 2 and 3). Much of the literature focuses on explaining sparse core-periphery structures, which are often rationalised by *fixed* costs to link formation (Craig and Ma (2018), for example, have a fixed cost of link formation relating to monitoring costs). Variation in fixed cost cannot explain heterogeneity in link intensity, however, so this empirical observations leads us to focus on heterogeneity in *marginal* cost instead.

Bank i has an endowment E_{it} that it can either supply to another bank or to an outside option. When it supplies its product to bank j it receives return r_{ijt} and incurs a per-unit cost puc_{ijt}. We model this per-unit cost as the cost of the equity that the bank has to raise to satisfy its capital requirements; that is, when bank i supplies bank j it pays a certain rate to raise the necessary equity. We parameterise the cost of equity as a linear function of the bank's default risk: $c^e_{it} = \phi p_{it}$. The riskier a bank is, the higher the cost of raising equity:

$$\underbrace{puc_{ijt}}_{\text{Per-unit cost}} = \underbrace{\lambda_{ijt}}_{\text{Reg'n}} \underbrace{c^e_{it}}_{\text{Cost of K}} = \lambda_{ijt}\ \phi p_{it}$$

where λ_{ijt} is the equity bank i needs to raise per-unit of exposure to bank j,[10] c^e_{it} is the cost of raising that equity, p_{it} is the default risk of bank i and ϕ is a parameter governing the relationship between default risk and cost of equity.

This simple parameterisation has three important implications. First, p_{it} is endogenously dependent on bank i's supply decisions, via the default risk process that we define above. In

[10] For ease of exposition we have collapsed the risk-weighting (ρ, using the notation from Section 2) and the capital required per risk-weighted assets (λ) into a single parameter, λ.

other words, when bank i supplies bank j, it takes into account the fact that doing so makes it riskier and so makes it costlier to raise capital. Second, p_{it} is endogenously dependent on the supply decisions of *other* banks, via the default risk process that we define above. In other words, there are network cost externalities. Third, p_{it} is endogenously dependent on regulation λ_{ijt} through the default risk process described above. In other words, in the spirit of Modigliani and Miller (1958), an increase in λ_{ijt} has two effects on the total cost of capital for firm i: it increases the amount of capital that the i bank needs to raise, but makes the bank safer and so makes the cost of a given unit of capital lower.

Bank i's problem in period t is to choose $\{C_{ijt}\}_j$ to maximise the following, taking $p_{k\neq i,t}$ as given:

$$\begin{aligned}\Pi_{it} &= \Pi_{it}^S + \Pi_{it}^D \\ &= \underbrace{\sum_j C_{ijt}[r_{ijt} - puc_{ijt} + e_{ijt}^S]}_{\text{Interbank supply}} + \underbrace{(E_{it} - \sum_j C_{ijt})r_{i0t}}_{\text{Supply to OOP}} + \Pi_{it}^D\end{aligned}$$

such that $C_{ijt} \geq 0$, $E_{it} - \sum_j C_{ijt} \geq 0$ and $puc_{ijt} = \lambda_{ijt}\phi p_{it}$.

For interior solutions the first order condition is as follows:

$$\underbrace{r_{ijt}^S + \frac{\partial r_{ijt}^S}{\partial C_{ijt}} C_{ijt} + e_{ijt}^S}_{\text{MB}} = puc_{ijt} + \underbrace{\sum_k \frac{\partial puc_{ikt}}{\partial C_{ijt}} C_{ikt}}_{\Delta \text{ Aggregate K cost}} - \underbrace{\frac{\partial \Pi_{it}^D}{\partial p_{it}} \frac{\partial p_{it}}{\partial C_{ijt}}}_{\Delta \text{ D-side cost}} - \underbrace{r_{i0t}}_{\text{OOP}}$$

The left-hand side is the marginal benefit to i of supplying bank j. The right-hand side is the marginal cost, which consists of four parts (i) the per-unit cost it pays, (ii) the marginal change in the per-unit cost, (ii) the marginal change in i's payoff from demanding interbank products and (iv) the outside option.

Bank i, when choosing to supply C_{ijt}, therefore balances the return it gets from supplying against the effect of its supply on its default risk, via the default risk process described above. Being riskier harms bank i by increasing the price it pays to access capital in two ways. First, it increases the marginal cost bank i pays when supplying interbank exposures (the third term in the equation above). Second, being riskier means that bank i pays higher interest rates when demanding exposures (the fourth term in the equation above).

3.5 Equilibrium

Before considering equilibrium, we summarise what our model implies for the definition of a bank. In our model, bank i is the following tuple: $(E_{it}, d_{i,l}, \beta_{\mathbf{i}}, \zeta_{\mathbf{i}}, \mathbf{\Gamma_i})$: respectively, an endowment, a set of product characteristics, a set of loadings on fundamentals, a technology and a set of contagion intensities. In other words, although the model is heavily parameterised, it allows for rich heterogeneity among banks.

Definition 1 *In this context we define a Nash equilibrium in each period t as: an $N \times N$ matrix of exposures $\mathbf{C_t^*}$ and $N \times 1$ vector of default risks $\mathbf{p_t^*}$ such that markets clear and every bank chooses its links optimally given the equilibrium actions of other banks.*

For interior solutions where $C_{ijt} > 0$, market clearing requires that supply and demand are equal, such that the following equilibrium condition holds, which we term the Equilibrium Condition ("EQC"):

$$\begin{aligned} 0 &= \delta_{jt} + \zeta_{ij} + e_{ijt}^{D} - 2BC_{ijt} - \sum_{k \neq i}^{N} \theta_{ik} C_{kjt} + e_{ijt}^{S} \\ &\quad -\lambda_{ijt}\phi_1 p_{it}(\mathbf{C_t}) - \phi_1 \tau_t \Gamma_{ij} p_{jt}(\mathbf{C_t}) \sum_{k \neq i}^{N} C_{ikt}\lambda_{ikt} - r_{i0t} \\ &\quad -\phi_1 \tau_t^2 \Gamma_{ij} p_{jt}(\mathbf{C_t}) \sum_{k} C_{kit}\Gamma_{ki} \sum_{m} C_{kmt}\lambda_{kmt} \end{aligned}$$

We show our calculations in Appendix B. Note that a bank's default risk is a function of $\mathbf{C_t}$, as we set out in the default risk process, which we repeat here for convenience:

$$\mathbf{p_t} = (\mathbf{I} - \tau_t \mathbf{\Gamma} \circ \mathbf{C_t})^{-1}(\mathbf{X_t}\boldsymbol{\beta} + \mathbf{e_t^p}) = \sum_{s=0}^{\infty} (\tau_t \mathbf{\Gamma} \circ \mathbf{C_t})^{s} (\mathbf{X_t}\boldsymbol{\beta} + \mathbf{e_t^p})$$

Substituting $\mathbf{p}$ out of EQC using DPR gives a system of equations in $\mathbf{C}^*$. The form of DPR is such that the EQC become a system of infinite-length series of polynomials, such that in general no analytical solution exists. Instead, we solve these equilibrium conditions numerically. We make no general claims about uniqueness or existence at this stage, but confirm numerically that our estimated results are an equilibrium that is, based on numerical simulations, unique.

3.6 Optimal networks

There are three immediate potential sources of inefficiency in our model (plus a fourth one we will define later):

1. Network externalities
2. Market power
3. Inefficient cost allocations

First, there are externalities within the interbank network, as bank k's default risk p_{kt} is affected by C_{ijt} provided that bank k has a chain of strictly positive exposures to i. If $C_{kit} > 0$ then this is trivially true, but it is also true if bank k has a strictly positive exposure to another bank that has a strictly positive exposure to i, and so on. Banks i and j do not fully account for the effect on p_{kt} when they transact, such that this negative externality implies that exposures are too large relative to the social optimum. Second, the banks supplying financial products may have market power, such that exposures are too small relative to the social optimum. Third, equilibrium allocations among suppliers may not be efficient, given differing marginal costs. In equilibrium high cost suppliers might supply positive quantities when it would be more efficient for low cost suppliers to increase their supply instead.

These inefficiencies mean that aggregate interbank surplus may not be maximised in equilibrium, where we define aggregate interbank surplus as the sum of aggregate surplus on the demand-side and aggregate surplus on the supply-side across all N banks. In other words, a social planner could specify an exposure network that increased aggregate interbank surplus.

In this context, however, it is insufficient to consider aggregate surplus within the interbank market. A bank's default risk can impact agents outside of the interbank market, such as its depositors, creditors, debtors and various other forms of counterparty. A crisis in the interbank market could, in principle, lead to a wider crisis with implications for the "real" economy. In other words, a social planner would not set exposures and default risk solely to maximise surplus in the interbank market, but instead to maximise total surplus in the economy, including aggregate interbank surplus and real surplus, which we define as follows.

Definition 2 : Real surplus : *We define "real surplus" as surplus outside of the interbank market, and denote it by R_t.*

The relationship between bank default risk and real surplus is important, as if there is such a relationship then it reveals a fourth possible inefficiency:

4. Real externalities: Banks do not take this into account the effect of their network formation decisions on real surplus.

Characterising the relationship between real surplus and default risk, or estimating it empirically, is not straightforward. We do not model or estimate this relationship, but only make the following directional assumption:

Assumption 1 *Suppose real surplus R_t is a function of the mean default risk of banks $\bar{p}_t$: $R_t = r(\bar{p}_t)$. We assume that R_t is strictly decreasing in $\bar{p}_t$.*

This assumption is clearly an approximation of what is likely to be a complex relationship between real surplus and bank default risk. It may not always hold; it may be, for example, that when bank default risk is very low, some additional bank default risk increases real surplus. It could also be that *mean* bank default risk is not the only thing that is important, but also some measure of dispersion or the minimum or maximum. Nevertheless, we think that this assumption reasonably represents the fundamental trade-off that regulators face when intervening in these markets: the trade-off between default risk and surplus in the market.

In particular, this assumption allows us to think about optimal default risk and interbank surplus in the sense of Pareto-optimality. That is, denote total surplus in the interbank market by TS_I (where the I subscript emphasises that this is total surplus in the interbank market only) and mean default probability by $\bar{p}$, and suppose $TS_I^H > TS_I^L$ and $\bar{p}^H > \bar{p}^L$. Assumption 1 implies that $(TS_I^H, \bar{p}^L) \succ^{SP} (TS_I^L, \bar{p}^H)$, where $\succ^{SP}$ denotes the social planner's preferences, but it does not allow us to rank $(TS_I^H, \bar{p}^H)$ and $(TS_I^L, \bar{p}^L)$, as we illustrate in Figure 4.

It is helpful to think about the trade-off between TS_I and $\bar{p}$ in terms of constrained maximisation of interbank surplus subject to a default risk constraint.

Definition 3 : Efficient frontier : *For an arbitrary, exogenous value of mean default risk, $\bar{p}^F$, define $TS_I^F = max_{\mathbf{C}}\ TS_I(\mathbf{C})\ st\ \bar{p}(\mathbf{C}) = \bar{p}^F$. We define the efficient frontier as the locus traced out in $(\bar{p}^F, TS_I^F)$ space as $\bar{p}^F$ is varied.*

In other words, the efficient frontier is agnostic about the scale of externalities outside of the interbank market. It requires only that there is no feasible alternative $(TS_I^A, \bar{p}^A)$ that is a Pareto-improvement in the sense that (i) $TS_I^A > TS_I^F$ and $\bar{p}^F \leq \bar{p}^A$ or (ii) $TS_I^A \geq TS_I^F$ and $\bar{p}^F < \bar{p}^A$. If such a Pareto-improvement existed, we can conclude from Assumption 1 that $(TS_I^A, \bar{p}^A) \succ^{SP} (TS_I^F, \bar{p}^F)$. The extent to which a given point is inefficient can then be loosely characterised by its vertical or horizontal distance from the frontier, as we set out in the definitions below. Figure 4 shows the frontier and illustrates what conclusions we can draw using this model about different outcomes.

Definition 4 : p inefficiency : *The default risk inefficiency of some allocation $(TS_I, \bar{p})$ is the percentage decrease in $\bar{p}$ that could be obtained without decreasing TS_I. In other words, it is the vertical distance in percentage terms from the frontier.*

Definition 5 : TS inefficiency : *The total surplus inefficiency of some allocation $(TS_I, \bar{p})$ is the percentage increase in TS_I that could be obtained without increasing $\bar{p}$. In other words, it is the horizontal distance in percentage terms from the frontier.*

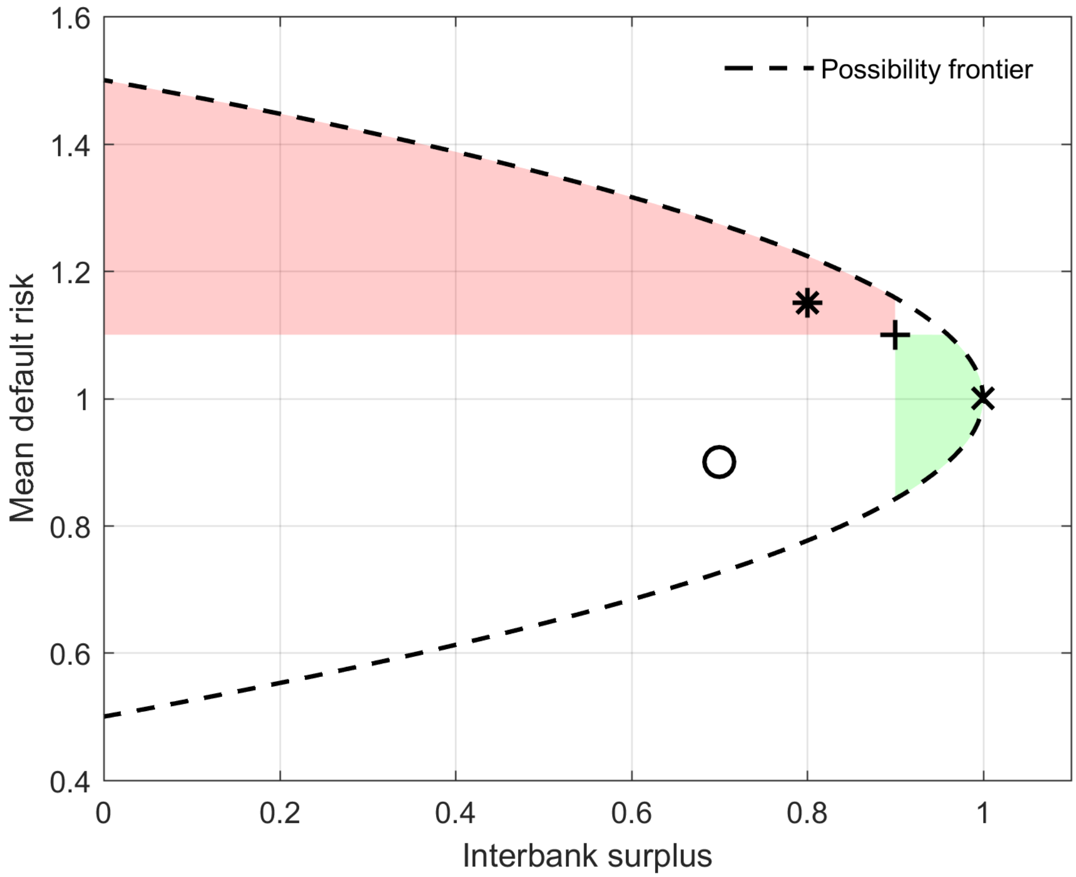
Finally, we note that although it is straightforward to consider efficient allocations, it is much more difficult to calculate optimal regulation (in our model, λ_{ijt}^{SP}, imposed by the social planner) that fully implements efficient allocations. We consider feasible regulations that are efficiency improvements over the perfectly decentralised market in the section below on counterfactual analysis.

3.7 Comparative statics

We demonstrate how the model works by arguing the following:

- Our model is consistent with the empirical facts we set out above.
- Our model is consistent with the stylised facts we set out above regarding how direct interbank connections behaved during the financial crisis.
- Our model is sufficiently flexible to be able to reasonably answer our fundamental economic question regarding the extent of inefficiency in the market.

Figure 4: Stylised example: Interbank surplus and default risk



Note: Point $+$ dominates any point in the red area but is dominated by any point in the green area. For example, $\times \succ^{SP} + \succ^{SP} *$, but we cannot rank $\circ$ relative to the other points. We cannot even rank $\circ$ relative to $\times$ despite $\times$ being on the efficient frontier: the social planner's preferences over $\times$ and $\circ$ depend on the scale of externalities outside of the interbank market, which we leave open. The extent of inefficiency of point $\circ$ can be expressed as the vertical distance south to the efficient frontier and the horizontal distance east to the frontier.

3.7.1 The model is consistent with our empirical facts

We set out certain empirical facts above that we used to guide our modelling. In this subsection, we explain in more detail how exactly the model is consistent with these empirical facts.

First, our empirical network is *heterogeneous* in the intensity of links. There are three main sources of such heterogeneity in our model: (i) firms have heterogeneous technologies

ζ_{ij} that require differing inputs from other firms, (ii) contagion intensity Γ_{ij} is heterogeneous, such that some links are intense because they are less risky and (iii) firms have heterogeneous fundamentals X_{it}, such that some links are intense because the banks involved have good fundamentals.

Second, our empirical network is *persistent* over time. Each of the sources of heterogeneity discussed above is also a source of persistence: ζ_{ij} and Γ_{ij} are by assumption fixed over time, and X_{it} vary over time but are persistent.

Third, we observe *increased concentration* in our data. In our model this results from the increase in capital requirements across our sample. Consider bank i's decision to supply bank j and/or bank k, where bank k's fundamentals are worse than bank j. For a given level of capital requirement λ, the fact that bank k is riskier means that ceteris paribus bank i supplies more to bank j than bank k. An increase in λ then makes supplying bank k relatively more costly compared to supplying bank j. In other words, an increase in capital requirements penalises risky links that are already likely to be small, resulting in an increase in concentration.

3.7.2 The model is consistent with our stylised facts

We also set out above three stylised facts from the crisis. Our model can match each of these stylised facts.

First, risky banks may choose to supply less total exposures, which we loosely term *liquidity hoarding*. All other things being equal, if a bank experiences a negative shock to its fundamentals it supplies less, as it is riskier and so its cost of capital is higher. This is not strictly liquidity hoarding in a structural sense, in that the bank is not lending less because it needs to preserve liquidity for the future, but the effect is the same. In that sense, this mechanism can be thought of as a reduced form for liquidity hoarding.

Second, risky banks may be supplied less, which we term *market lockout*. A shock to the fundamentals of bank j makes supplying it more risky and therefore more costly. This is true holding fixed δ_{jt}, which are fixed effects governing inter-temporal variation in demand. If this is related to X_{jt}, then the effect of variations in fundamentals is more complicated.

Third, when all banks are risky, liquidity hoarding and market lockout combine to result in *market shutdown*, where no bank is supplied anything at all. This follows in our model as the combination of the two previous effects.

3.7.3 The model can answer our question

Finally, we note that the model is flexible enough to answer the economic questions we are seeking to answer, in that the magnitude of each of the sources of inefficiency discussed above depends on specific parameter values (the identification of which we will discuss below). This is important for the robustness with which we answer these questions, as it shows that our answers to these questions are guided by the data rather than by our modelling assumptions.

We illustrate this by reference to a baseline set of parameters,[11] and then showing the effect of varying certain key parameters. We set out the results of these simulations in Table 2.

Table 2: Comparative statics

	[A] Baseline	[B] $\downarrow mean(\theta_{ij})$	[C] $\uparrow var(\Gamma_{ij})$	[D] $\uparrow mean(\Gamma_{ij})$
p inefficiency	4.3%	5.4%	6.0%	8.7%
TS inefficiency	13.2%	15.6%	14.6%	14.2%

Note: [A] is our baseline results set out below; [B] is the baseline, with every θ_{ij} multiplied by a factor of 0.8; [C] is the baseline, with a mean-preserving spread of Γ_{ij} such that its variance increases by a factor of 1.5; [D] is the baseline, with every Γ_{ij} multiplied by a factor of 1.5.

First, market power is determined by θ_{ij}, which governs the extent of product differentiation. If θ_{ij} is large (small), then products i and j are close substitutes and market power is low (high). We illustrate the impact of increased market power by multiplying every θ_{ij} by a factor of 0.8 (Column B in Table 2). As set out in Table 2, this increases the distance between the decentralised outcome and the efficient frontier.

Second, the efficiency of decentralised cost allocations is driven by the extent of variation in marginal cost across banks. If marginal cost is the same for all banks, then decentralised cost allocations are not inefficient. If marginal cost is highly variable, then the decentralised equilibrium will inefficiently involve some high cost links being positive. The extent of variation in marginal cost across banks is driven primarily by the extent of variation in contagion intensity Γ_{ij}. We illustrate this by applying a mean-preserving spread to Γ_{ij} such

[11] The baseline we choose is actually the parameter values we go on to estimate below. These comparative statics mostly hold for any baseline, although obviously some effects are switched off when certain parameters are zero.

that its variance increases by a factor of 1.5 (Column C in Table 2). This increases the distance between the decentralised outcome and the efficient frontier.

Third, the extent of externalities depends on the scale of network effects, which in our model is the size of Γ_{ij}. If these are large, then there are significant externalities and the decentralised equilibrium is more likely to be inefficient. We illustrate this by increasing every Γ_{ij} by a factor of 1.5. This also increases the distance between the decentralised outcome and the efficient frontier.

4 Identification

In describing our approach to identification, we first set out what data we use to model bank fundamentals. We then consider identification of the network formation game and of the default risk process. We then briefly describe our approach to controlling for the fact that we only partially observe some links. Finally, we return to our research question, and discuss in intuitive terms the key variation that we use to identify each of the key parameters that determine our answer to this research question.

4.1 Modelling fundamentals

To represent bank fundamentals $\mathbf{X}$ we use bank-specific and common data.

For bank-specific variation, we take the relevant equity index to be a bank-specific weighted average of global equity indices from S&P, where the weightings are the proportion of the bank's revenues that come from that geography (data provided annually by Bloomberg, based on corporate accounts). For example, suppose that at time t bank k obtained 70% of its revenues from the US and the remaining 30% from Japan. In this case, $Z^p_{kt} = 0.7 \times S\&P500_t + 0.3 \times S\&PJapan_t$. Absolute index values are not meaningful, so we normalise each S&P index by its value on 1 June 2019. Although this is clearly an imperfect measure of the bank's fundamentals, we argue it has informative value: this bank k would plausibly be more affected by a slowdown in Japan than some other bank with no Japanese revenues. The S&P indices we use are for the US, Canada, the UK, Europe, Japan, Asia and Latin America.

To capture common variation in bank fundamentals, we use a broad panel of macroeconomic and commodity data from the World Bank. We calculate the first three principal components of this panel, which collectively account for more than 99% of total variation, and include these three variables in $\mathbf{X}$. We also include the Chicago Board Options Exchange Volatility Index, more commonly known as "VIX", which represents expected variation in option prices, and the Morgan Stanley World Index.

4.2 Network formation

The EQC and DRP allow us to solve for equilibrium $\mathbf{C}$ and $\mathbf{p}$ as a function of λ, $\mathbf{X}$ and the jt and it fixed effects described above. In other words, identification is significantly easier when we solve for equilibrium exposures, because the endogenous exposures of other banks and endogenous default risks are substituted out of our empirical specification.

We assume bank fundamentals, as defined above, are exogenous. Treating this as exogenous assumes that a bank's revenue distribution and the equity indices themselves are independent of *pairwise* structural errors in the interbank network. We emphasise that the fact that we are able to include it and jt fixed effects means that the only remaining unobservable variation is pair-specific. We think it is a reasonable assumption that, for example, HSBC, which has deep roots in Asia, would not shift its geographic revenue base in response to pair-specific shocks in the interbank market. Similarly, we think it is a reasonable assumption that the equity indices that form the basis of our bank-specific fundamentals are independent of pair-specific shocks in the interbank market.

We treat product characteristics as exogenous, in keeping with the literature on demand estimation in characteristic space. We treat λ, regulatory capital requirements, as exogenous, in keeping with the literature on the empirical analysis of bank capital requirements (Robles-Garcia, 2018; Benetton, 2018). It is informative to consider how we are able to separately identify the effect of common time variation in capital requirements from the it and jt fixed effects. This relates to Figure 3, in which we show the correlation between concentration in the interbank market over time and changes in capital requirements. In our model the effect of the common increases in capital requirements on equilibrium exposures depends on the fundamentals of the banks supplying and demanding the exposures: in other words, although the changes in capital requirements are common across all banks, their effect on exposures is pair-specific.

B is not separately identifiable from the other parameters. We normalise $B = 1$ on the basis that in models of quantity competition what matters for market power is θ/B, not the absolute value of B.

4.3 Default risk process

We repeat DRP for convenience:

$$\mathbf{p_t} = (\mathbf{I} - \tau_t \mathbf{\Gamma} \circ \mathbf{C_t})^{-1}(\mathbf{X_t}\boldsymbol{\beta} + \mathbf{e_t^p}) = \sum_{s=0}^{\infty}(\tau_t \mathbf{\Gamma} \circ \mathbf{C_t})^s(\mathbf{X_t}\boldsymbol{\beta} + \mathbf{e_t^p})$$

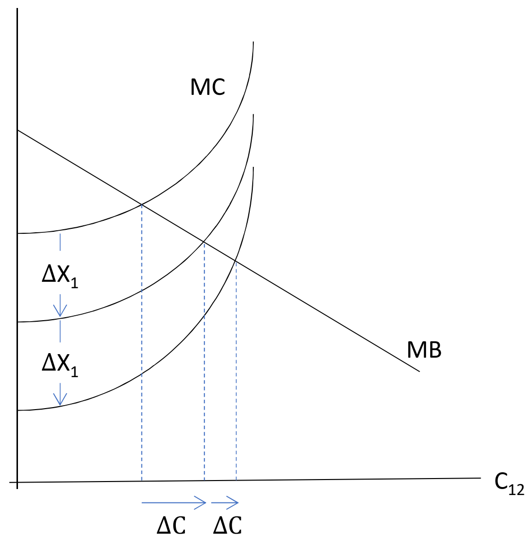
The advantage of explicitly considering network formation is that we can account for the endogeneity of the network in our spatial DRP model. The key insight to our identification strategy is that DRP is a *linear* function of bank fundamentals $\mathbf{X_t}$, but equilibrium exposures $\mathbf{C_t}$ are a *non-linear* function of $\mathbf{X_t}$. We therefore use non-linear variation in $\mathbf{X_t}$ as pair-specific, time-varying instruments for the network. We motivate this more clearly in three steps. First, we show that equilibrium exposures are indeed non-linear in bank fundamentals. Second, we show that this gives us the pair-specific variation that we need. Third, we set out exactly which variables we use as instruments.

The fact that equilibrium exposures are non-linear in bank fundamentals comes from the non-linearity of the cost function. The key intuition for this is that the cost function is convex in C_{ijt}, such that in equilibrium C_{ijt} would never grow linearly with fundamentals as that would lead to marginal cost becoming very large. Consider a simple example with three banks, 1, 2 and 3, and suppose, for the sake of simplicity, that in equilibrium every network link between those banks is strictly positive. In equilibrium C_{12}^* is such that the marginal cost of supplying exposures is equal to the marginal benefit. The marginal benefit is linear in C_{12}, whereas the marginal cost is convex in C_{12}, as set out in Figure 5. Suppose the fundamentals of banks 1, 2 and 3 improve, worsen and remain unchanged, respectively. In these circumstances, we show in Figure 5 that C_{12} changes non-linearly relative to the size of these. In Appendix B we show, for a simplified version of our model for which an analytical solution exists, that equilibrium C are a non-linear function of $\mathbf{X}$.

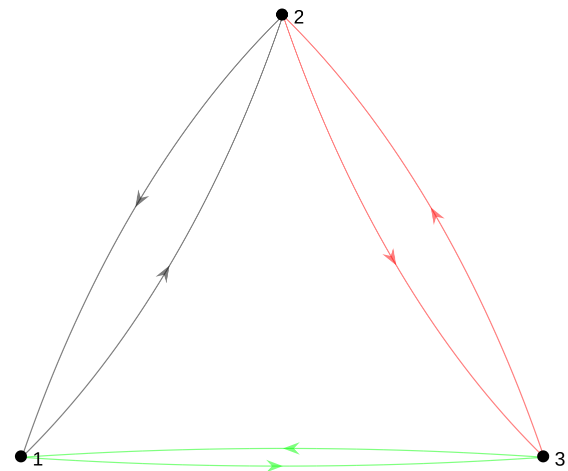
Having shown that exposures are non-linear in fundamentals, it is straightforward, using the same simple example, to show that changes in fundamentals then give us the pair-specific variation that we need for them to be instruments for C_{ijt}. Assume again that the fundamentals of banks 1, 2 and 3 improve, worsen and remain unchanged, respectively. This causes links between banks 1 and 3 to increase (because the improvement in bank 1's fundamentals mean that the marginal cost to bank 1 of supplying bank 3 has gone down, and the marginal cost to bank 3 of supplying bank 1 has gone down). For analogous, but opposite, reasons, links between bank 2 and bank 3 decrease. For links between banks 1 and

Figure 5: Non-linear bank fundamentals as instruments for C

(a) Non-linear effect of X on C



(b) Pairwise variation



Note to Figure 5: Suppose the fundamentals of bank 1, 2 and 3 improve, worse and do not change, respectively. In panel (a) we show that equilibrium exposures are non-linear with respect to this variation in fundamentals. In panel (b) we show that this this has differing pairwise effects on equilibrium link intensity, where link intensity between 1 and 2 increases, link intensity between 2 and 3 decreases and link intensity between 1 and 2 does not change.

2 it is not possible to sign the effect, as some elements of marginal cost have gone up and some have gone down. In summary, provided there is reasonable cross-sectional variation in bank fundamentals (which we show in Figure 3), then that variation has differing exogenous implications for each of the pairs.

We define $\tilde{X}_{ijt} = \frac{1}{N-2}\sum_{k\neq i,j} X_{kt}$ (that is, average fundamentals of other banks). As instruments for C_{ijt} we use $[X^2_{it},\ X^2_{jt},\ \tilde{X}^2_{ijt},\ X_{it}/X_{jt},\ X_{it}/\tilde{X}_{ijt}...]$, as well as these terms interacted with λ_{ijt} to leverage its time variation. We show the results of first stage regressions in the appendix. Assuming these bank fundamentals are orthogonal to unobserved shocks to bank default risk is more restrictive than in the case of the network formation data, as we have fewer fixed effects available to use. We assume that the equity indices on which we rely are independent of unobserved bank default risk. We justify this on the basis that, although the banks in our sample are large, none are a material proportion of these equity indices.

4.4 Incomplete data

As described in Section 2, for non-British banks we only observe local-unit-to-group exposures, under-estimating their total exposure.

We denote local-unit-to-group exposures by $\tilde{C}_{ijt}$ and group-to-group (that is, total) exposures by C_{ijt}. We assume that $C_{ijt} = (1 + a_i)\tilde{C}_{ij}$, where a_i are bank-specific parameters that we estimate. These parameters a_i are identified given that (i) some variables, such as X_{jt} and p_{it}, enter the EQC with non-bank-specific coefficients and (ii) for the British banks we know $a = 0$. In principle, a_{ijt} is identifiable in this way, but we restrict variation to a_i to preserve degrees of freedom.

4.5 Identification: Back to the research question

Having described our approach to identification, we summarise by considering how identification relates to our core research question regarding the inefficiency of the interbank market. We show in Table 2 that the extent of inefficiency in the market depends on three key parameters: (1) the size of θ_{ij}, which governs the extent of market power, (2) the size of average Γ_{ij}, which governs the extent of network externalities and (3) the cross-sectional dispersion of Γ_{ij}, which governs the extent to which equilibrium allocations go to high-cost banks. We summarise the key variation that identifies each of these parameters in Table 3. This is important for the robustness with which we answer our research question, as it shows that our answers to these questions are guided by the data rather than by our modelling assumptions.

Table 3: Key variation

	Key parameter	Key variation
[1]	$mean(\theta_{ik})$	$Cov(C_{ijt}, X_{kt} \mid d_i - d_k)$
[2]	$mean(\Gamma_{ij})$	$Cov(C_{ijt}, X_{jt})$, $Cov(p_{it}, X_{jt} \vert Z^C_{ijt})$
[3]	$var(\Gamma_{ij})$	$Cov(s_{ijt}, \lambda_t)$

Note: s_{ijt} denotes proportion of bank i's total supply that is to bank j. All other notation as previously defined.

θ_{ik} determines how closely banks i and banks k compete. We identify the size of θ_{ik}

by the covariance between C_{ijt} and X_{kt}, which is an exogenous measure of bank k's cost, conditional on the extent to which the two banks have similar product characteristics. If this covariance is high, then θ_{ik} is high.

Γ_{ij} determines the contagion intensity from j to i. There are two sources of empirical variation for this: from the network formation data and from the default risk data. On the network formation side, Γ_{ij} is identified by the covariance between C_{ijt} and X_{jt}. If C_{ijt} is sensitive to the fundamentals of bank j, then in the context of our model this means that Γ_{ij} is large. On the default risk side, Γ_{ij} is identified by the covariance between bank i's default risk and the fundamentals of bank j, conditional on the instruments we describe above for the size of C_{ijt}. If this conditional covariance is large, then this means that bank i's default risk is particularly sensitive to bank j's default risk, which in the context of our model means that Γ_{ij} is large.

Finally, we describe a further source of variation that helps identify the dispersal in Γ_{ij}. We set out above how a general increase in capital requirements leads to concentration, as it affects high and low marginal cost links differentially. Γ_{ij} is a key determinant of which links are high and low marginal cost. If, following an increase in capital requirements, bank i supplies relatively less to bank j, then this concentration indicates that Γ_{ij} is high.

5 Estimation

5.1 Approach to estimation

We first describe two parameterisations that we make when we take this model to data. We then describe the structure of our estimation approach.

The first parameterisation we make is with respect to Γ_{ij}. General symmetric Γ_{ij} consists of $N(N-1)/2 = 153$ elements. These are individually identifiable, as we show above, but because the length of our panel is limited we cannot estimate them with reasonable power. For this reason, our baseline estimation approach imposes the following structure on Γ_{ij}:

$$\Gamma_{ij} = \tilde{\Gamma}_i \tilde{\Gamma}_j$$

where $\tilde{\mathbf{\Gamma}}$ is an $N \times 1$ vector of parameters. This parameterisation is significantly more parsimonious but retains variation at the ij level. It does result in some loss of generality, in that loosely speaking it implies that if Γ_{12} and Γ_{23} are high, then Γ_{13} must also be high. This kind of structure is broadly consistent with each of the three motivations for heterogeneous Γ_{ij} that we introduce above. In particular, in this example, if the fundamentals of banks 1 and 2 are highly correlated (suggesting Γ_{12} is high) and the fundamentals of banks 2 and 3 are highly correlated, then it is likely that the fundamentals of banks 1 and 3 are also highly correlated.

The second parameterisation we make relates to τ_t. We include τ_t to allow for time-variation in contagion intensity because higher capital requirements are intended to make a given exposure safer. General τ_t, with a different multiplicative parameter for each time period, is in principle identifiable. In practice, we parameterise τ_t based on capital requirements:

$$\tau_t = e^{-\tau(\lambda_t - \lambda_1)}$$

where λ_t is the mean capital requirement at time t, λ_1 is the mean capital requirement in the first period of our sample, 2011, and τ is a scalar parameter. Thus $\tau_1 = 1$, but $\tau_{t>1}$ can be lower depending on the size of τ. If $\tau = 0$ then $\tau_t = 1$ for $\forall t$ and there is no time-variation in contagion intensity, if τ is large then there is significant time-variation. This is a more parsimonious approach that directly addresses the underlying reason why allowing for time-variation in contagion is important.

The parameters we seek to estimate are $\boldsymbol{\Theta} = (\tilde{\boldsymbol{\Gamma}}, \tau, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{\zeta}, \tilde{\boldsymbol{\theta}}, \phi)$; respectively, contagion intensities, time-variation in contagion intensities, fundamentals, demand intercept variation, pairwise technology importance, characteristic-based product differentiation, and the cost multiplier. Our estimation process involves two loops. In the inner loop, we solve our model numerically to calculate the network links and default risks implied by a given parameter vector; respectively, $\hat{C}(\Theta)$ and $\hat{p}(\Theta)$. In the outer loop, we search over parameter vectors Θ to minimise two sets of moments, where the relevant instruments are set out in the preceding section: (1) network formation: $\mathbb{E}[\mathbf{Z}'(\hat{\mathbf{C}}(\boldsymbol{\Theta}) - \mathbf{C})] = 0$ and (2) contagion: $\mathbb{E}[\mathbf{Z}'(\hat{\mathbf{p}}(\boldsymbol{\Theta}) - \mathbf{p})] = 0$. We express $\mathbf{p}$ in logs.

5.2 Results

We set out our results in Table 4.

We draw the following immediate implications for contagion intensity from our results:

- **Contagion is material**: on average 12.1% of mean bank default risk is due to inter-

Table 4: Results

		[1]	
ϕ		1.84**	
		(1.85)	
τ		9.26***	
		(6.03)	
	Min	*Median*	*Max*
Γ_i	0.23***	0.24	0.30***
	(4.60)	(1.60)	(10.70)
θ_k	0.24***	4.63	28.48***
	(2.59)	(0.40)	(5.08)
a_i	0.01	0.69	5.53**
	(0.06)	(1.01)	(2.03)
Network			
FE		ij, it, jt	
R^2		0.84	
No. obs		6,426	
Default risk			
FE		i	
Controls		Y	
R^2		0.84	
No. obs		378	

Notes: SEs clustered at bank level. Figures in parentheses are t-stats. ***, **, * indicate difference from 0 at 1%, 5% and 10% significance, respectively. For the heterogeneous parameters we report estimates and t-stats for the minimum, median and maximum, and plot the full distribution below. **Notation**: ϕ governs the sensitivity of cost of equity to default risk, τ governs the extent to which contagion intensity varies over time, Γ_i governs contagion intensity, θ_k governs product differentiation based on characteristics, a_i scales exposures for non-UK banks and the controls in the default risk equation are the fundamentals X.

bank contagion, with the remainder due to bank fundamentals.[12] This can be thought of as an aggregate representation of the network effect. We also re-run our estimation taking the network as exogenous in our estimation of the default risk process (that is, without using the instruments for the endogenous network that are implied by our network formation game). This results in parameter estimates that imply 9.8% of mean bank default risk is due to interbank contagion. In other words, incorrectly assuming that the network is exogenous biases our estimation of the network effect downwards.

- **Contagion is heterogeneous**: there is substantial pairwise variation in contagion intensity Γ_{ij}: some links are nearly twice as costly as others, in terms of their effect on default risk. We plot the estimated distribution of Γ_{ij} in Figure 6.

- **Contagion is related to risk sharing**: we set out above various motivations for why contagion intensity Γ_{ij} could be heterogeneous. One of these motivations is heterogeneity in the extent to which bank fundamentals are correlated; risk sharing, in other words. This implies a relationship between fundamentals, which we estimate as $X\beta$, and contagion intensity Γ_{ij}. We do not impose this relationship in estimation, but estimate general Γ_{ij} and test the existence of such a relationship post-estimation. These post-estimation tests, which we describe in Appendix C, support risk-sharing: Γ_{ij} is higher when the fundamentals of banks i and j are more closely positive correlated.

- **Contagion is time-varying**: there is evidence that contagion intensity has decreased across our sample, in line with increasing capital requirements. Estimated τ implies that mean contagion intensity decreased by 36% between 2011 and 2018, as we plot in Figure 6. This is consistent with a significant improvement in bank default risk in response to the banks becoming better capitalised.

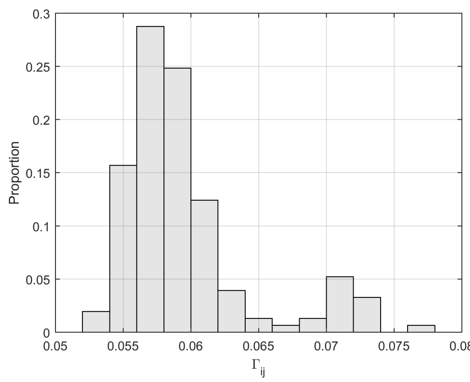
Our results also have implications for the form of competition between banks. We plot our estimated $\widehat{\theta}_{ij}$ in Figure 6, and show that there is significant product differentiation based on product characteristics. Generally, most $\widehat{\theta}_{ij}$ are close to zero, indicating that only pairs producing very similar products are substitutes. The most important product characteristics in determining substitutability are (i) the proportion of total exposures that is denominated in EUR and (ii) the proportion of exposures with maturity greater than 1 year.

Having described our results, we now discuss two important implications of our results

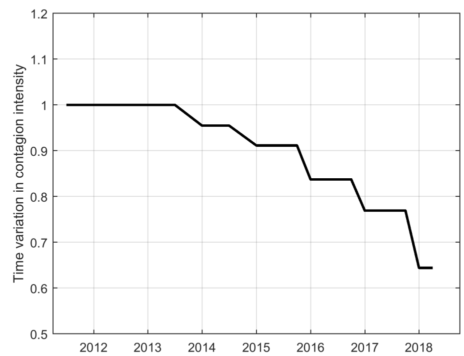
[12] We calculate this by calculating mean bank default based solely on fundamentals, $p_{it} = X_{it}\beta$, and comparing it to actual bank default risk.

Figure 6: Distributions of parameter estimates

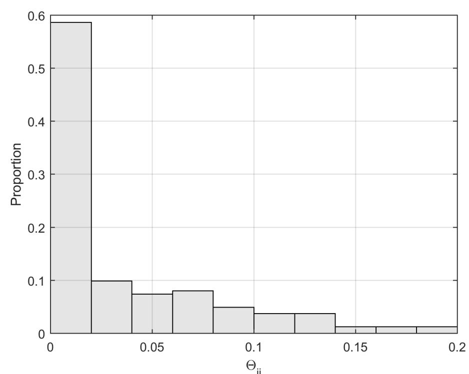
(a) Heterogeneous contagion intensity



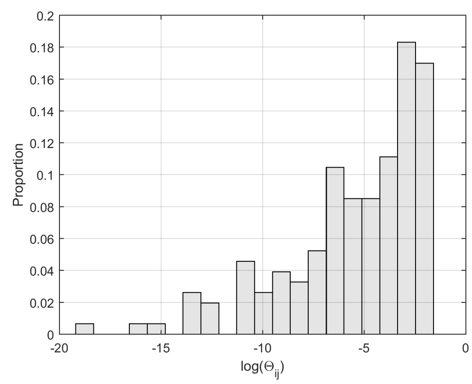
(b) Time variation in contagion intensity



(c) Product differentiation



(d) Product differentiation (in logs)



.

regarding (1) forward simulation of recessions using our model and (2) the identification of systemically important banks.

5.2.1 Forward simulation

In Figure 7 below we simulate the effect of a recession on the interbank network and default risk. We do this by simulating an arbitrary increase (deterioration) in bank fundamentals. As the shock increases in severity the network shrinks and, when the recession is sufficiently severe, dries up. This is an important cross-check of our work, as although we do not have network data from the financial crisis we know some elements of the interbank market shut down during the financial crisis. As Figure 7 shows, our model replicates this feature, at a similar level of bank fundamentals as were observed during the crisis. One implication of this is that bank default risk is convex with respect to bank fundamentals: as fundamentals deteriorate, the endogenously declining network dampens the effect of the change on fundamentals on default risk. There is, however, a zero lower bound, such that once the network has dried up then it cannot dampen the response to fundamentals. In other words, bank default risk is more sensitive to fundamentals in severe recessions.

This fact also has implications for forecasting. Suppose, for example, that when modelling the response of default risk p to fundamentals X the endogenous network was ignored, and instead p was simply regressed on X. Because severe recessions are very infrequently observed in our data, a regression of p on X in normal times would *understate* the extent to which p would respond to X in a severe recession. We show true simulated default risk (the black solid line) and such a naively estimated default risk (the red dashed line) in Figure 7.

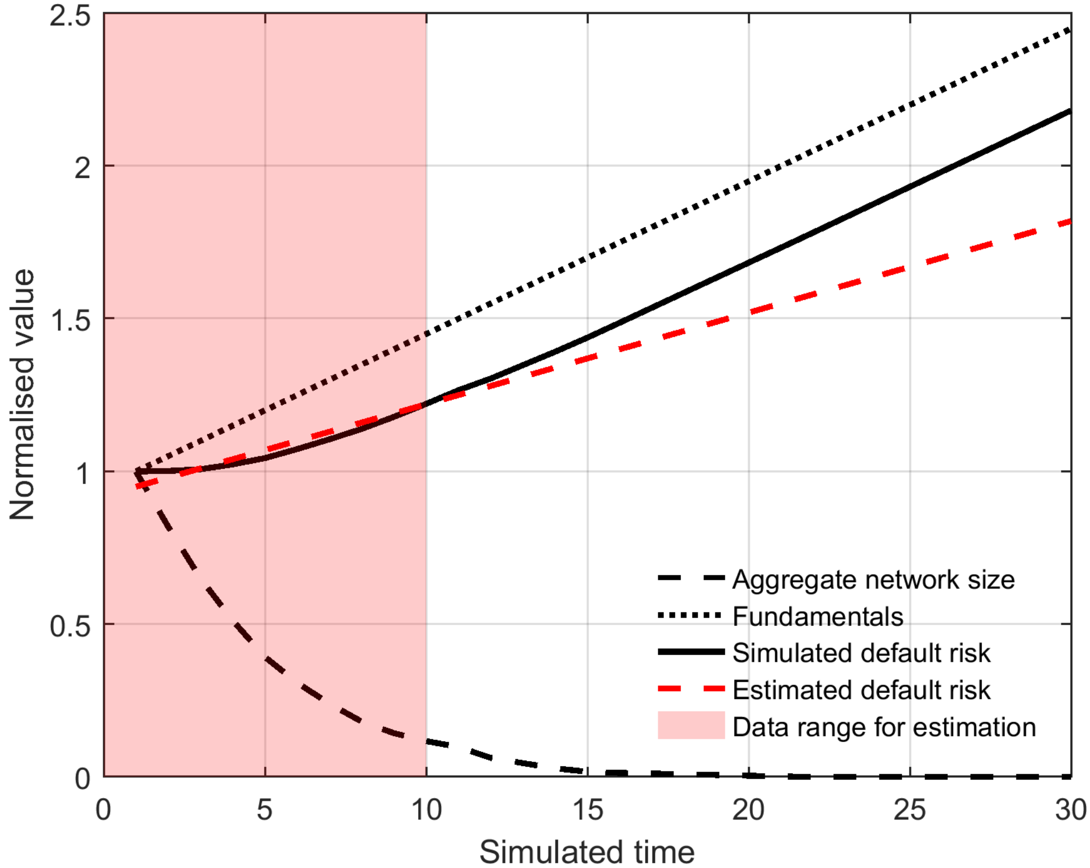
5.2.2 Systemic importance

A recurring issue in the network literature is the identification of "important" nodes. We have an equilibrium process that relates an outcome (bank default risk, in our case) to a network, and it is reasonable to ask which node in the network contributes most to the outcome in which we are interested. Understanding this communicates important information about this equilibrium process, but may also have implications for regulation (as we describe above, large parts of the banking regulatory framework are stricter for banks that are judged to be "systemically important" (BCBS, 2014b)). Various measures of systemic importance, or centrality, exist, where the most appropriate measure depends on the context and on the way in which nodes interact with each other (Bloch et al., 2017). Our contribution to this literature is not about the most appropriate measure, but instead about how any such measure should be calculated: it must account for the heterogeneity in contagion intensity Γ_{ij}.

We illustrate this by reference to one of the simplest measures of centrality: Eigenvector Centrality. Broadly speaking, node n's centrality score is the n'th entry in the eigenvector associated with the maximal eigenvalue of the adjacency matrix $\boldsymbol{C_t}$. A central node using this measure is close to other nodes that are central: this measure of centrality is in this sense self-referential. Nodes that have many large links to other nodes that have many large links are more central.

Applying this centrality measure to the network $\boldsymbol{C_t}$ therefore gives a ranking of which banks are most systemically important in driving bank default risk. If contagion intensity is

Figure 7: Simulated recession



Note: We simulate a recession by arbitrarily inflating (where an increase is a deterioration) bank fundamentals by an increasing factor (the dotted black line). As fundamentals deteriorate, the interbank network (the black dashed line) contracts and eventually dries up. Mean bank default risk (the solid black line) increases, but is convex because the network contraction dampens the effect of fundamentals. The red dashed line shows the results of observing a limited set of data (the red shared area) and fitting a linear regression of default risk on bank fundamentals: ignoring endogenous network formation understates how bank default risk changes with fundamentals in (infrequently observed) recessions.

homogenous, $\Gamma_{ij} = \Gamma$, then the level of Γ has no impact on this relative ranking. If, however, contagion intensity is heterogeneous, then accounting for this heterogeneity is important when assessing centrality: a more reasonable measure of centrality would be based on the weighted adjacency matrix $\boldsymbol{\Gamma} \circ \boldsymbol{C_t}$. Importantly, the effect of this weighting on the ranking of systemic importance is not random noise, because the equilibrium network depends on this weighting. More specifically, links C_{ij} where Γ_{ij} is low (high) are relatively safe (unsafe) and so are more likely to be large (small), all other things being equal. In other words, assessing centrality based on the raw, unweighted exposures matrix is likely to overstate the centrality of more central nodes and understate the centrality of less central nodes. This holds only when holding other things equal: in our model of network formation, links can be large even if they are not safe (if they are technologically important through ζ_{ij}, for example).

In Figure 8, we show that calculating Eigenvector Centrality based on unweighted $\boldsymbol{C_t}$ and weighted $\boldsymbol{\Gamma} \circ \boldsymbol{C_t}$ lead to quite different rankings of systemic importance. Bank 18, for example, would be identified as the most systemically important node based on the unweighted network. Based on the weighted network, however, 8 other banks are most systemically important than Bank 18: in other words, Bank 18's links are large because its links are relatively safe. Bank 5's centrality, on the other hand, is significantly understated when looking solely at the unweighted network: in other words, Bank 5's links are small because its links are relatively unsafe. We do this for Eigenvector Centrality, but the same point applies to other measures (including, for example, Katz-Bonacich centrality).

6 Counterfactual Analysis

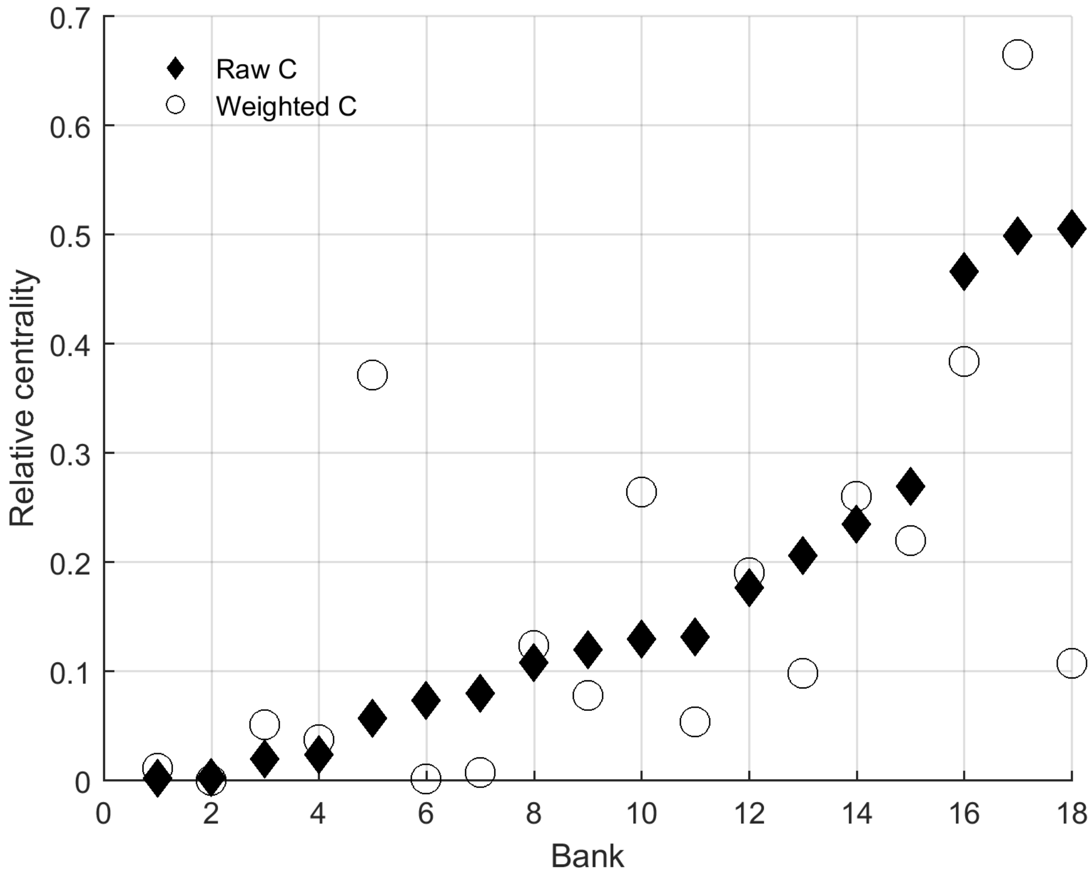
In our counterfactual analyses, we first consider the social planner's solution, and show what that implies for efficiency. We then consider two broad forms of regulation: caps on exposures and capital ratios.

Before we describe the counterfactual analyses in detail, we describe two uses of our model that play an important role in each of these counterfactual analyses. Our model, together with the parameters we have estimated, allow us to do two things. First, the estimated model provides a mapping from any arbitrary network of exposures $\mathbf{C_t}$ to (i) bank default risk and (ii) interbank surplus. Second, the estimated model provides a mapping from the exogenous parts of the model (fundamentals, regulation, etc) to decentralised equilibrium exposures $\mathbf{C_t}$. Together, these two uses of our model and results allow us to quantify surplus and default risk in counterfactual equilibria.

6.1 Efficiency

We describe above how our model implies a trade-off between mean bank default risk and interbank surplus, and how there is an efficient frontier on which this trade-off is optimised.

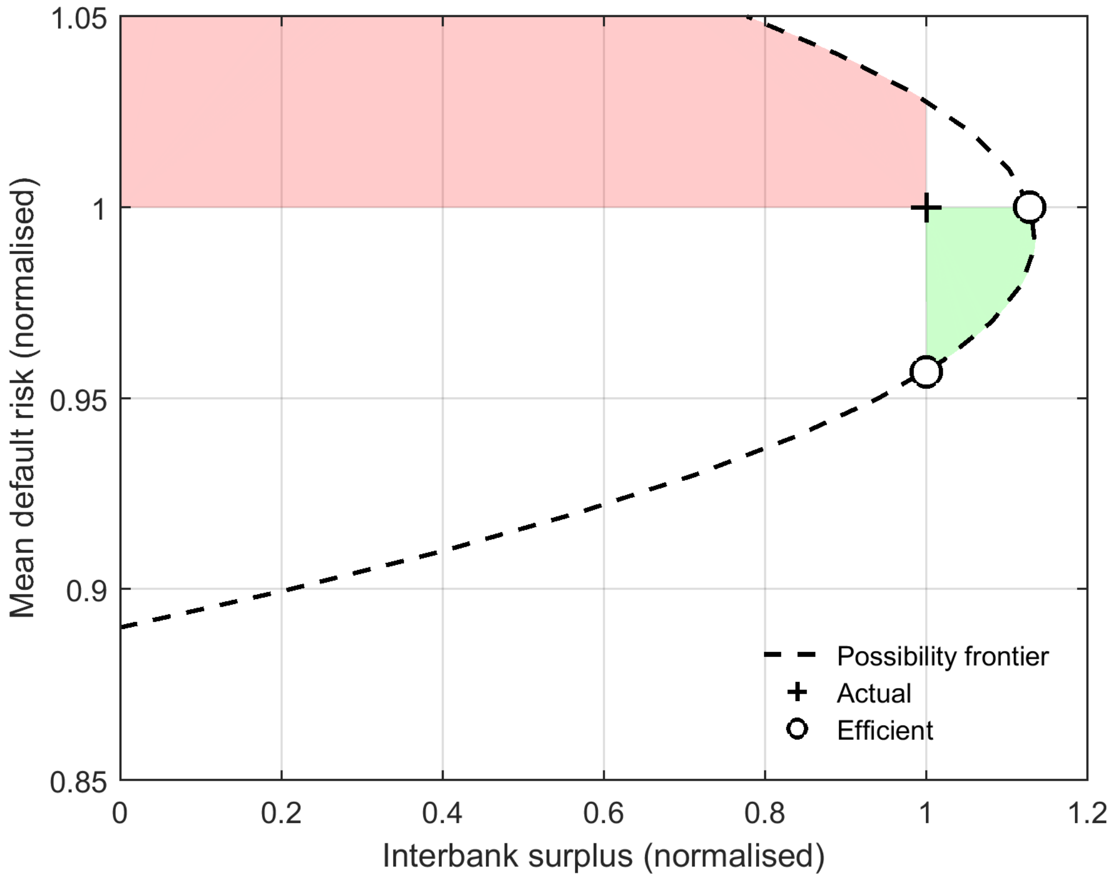
Figure 8: Identifying systemic nodes



We use our estimated model to derive this frontier, by choosing $\mathbf{C_t}$ to maximise interbank surplus, subject to mean bank default risk being less than some critical value. We then vary this critical value to trace out the efficient frontier. As described above, we do not know what allocations a social planner that was maximising aggregate surplus would choose, as we do not directly model the relationship between bank default risk and real surplus. We do know that this optimal allocation would be somewhere along the efficient frontier. The distance to the frontier in either direction is in this sense an estimate of inefficiency, as we describe above when we define p inefficiency and TS inefficiency.

We find that the decentralised interbank market is not on the efficient frontier: a social planner would be able to increase interbank surplus by 13.2% without increasing mean bank default risk or decrease mean bank default risk by 4.3% without decreasing interbank surplus, as set out in Figure 9. This result comes primarily from the fact that contagion (and thus network externalities) is significant.

Figure 9: Decentralised inefficiency



6.2 Caps on exposures

As discussed in Section 2, in 2019 a cap on individual exposures came into force: a bank can have no single bilateral exposure greater than 25% of its capital.[13] For exposures held between two "globally systemic institutions"[14] this cap is 15%.

We evaluate the effects of a cap on individual exposures by simulating new equilibrium exposures C_{ij}^{C} under a generic cap, using our estimated parameters and assuming that fun-

[13]Where the precise definition of capital, "Tier 1 capital", is set out in the regulation.
[14]As defined in the regulation.

damentals are unchanged. We consider a generic, binding cap at the i-bank level:

$$C_{ij}^{C} \leq 0.9 \cdot max_j\{C_{ij}\}$$

In other words, we assume that any exposure held by bank i has to be less than or equal to 90% of its largest exposure. This cap is stylised, in that it is defined relative to observed exposures, rather than relative to its capital. This avoids issues about measuring capital appropriately and measuring total exposures (our exposures do not include every possible financial instrument), while still showing the economic effect of a cap in general. We simulate the effect of this cap in Figure 10 below, and find that such a cap has a very small impact on default risk, for two reasons. First, a cap on individual exposures binds on the bank's largest exposures, which are more likely to be relatively safe (that is, they have low Γ_{ij}). Second, a cap on individual links creates excess supply and unmet demand that causes other uncapped links in the network to increase. That is, the network topology changes endogenously.

We propose an alternative form of regulation in which total exposures held by bank i are capped, rather than individual exposures:

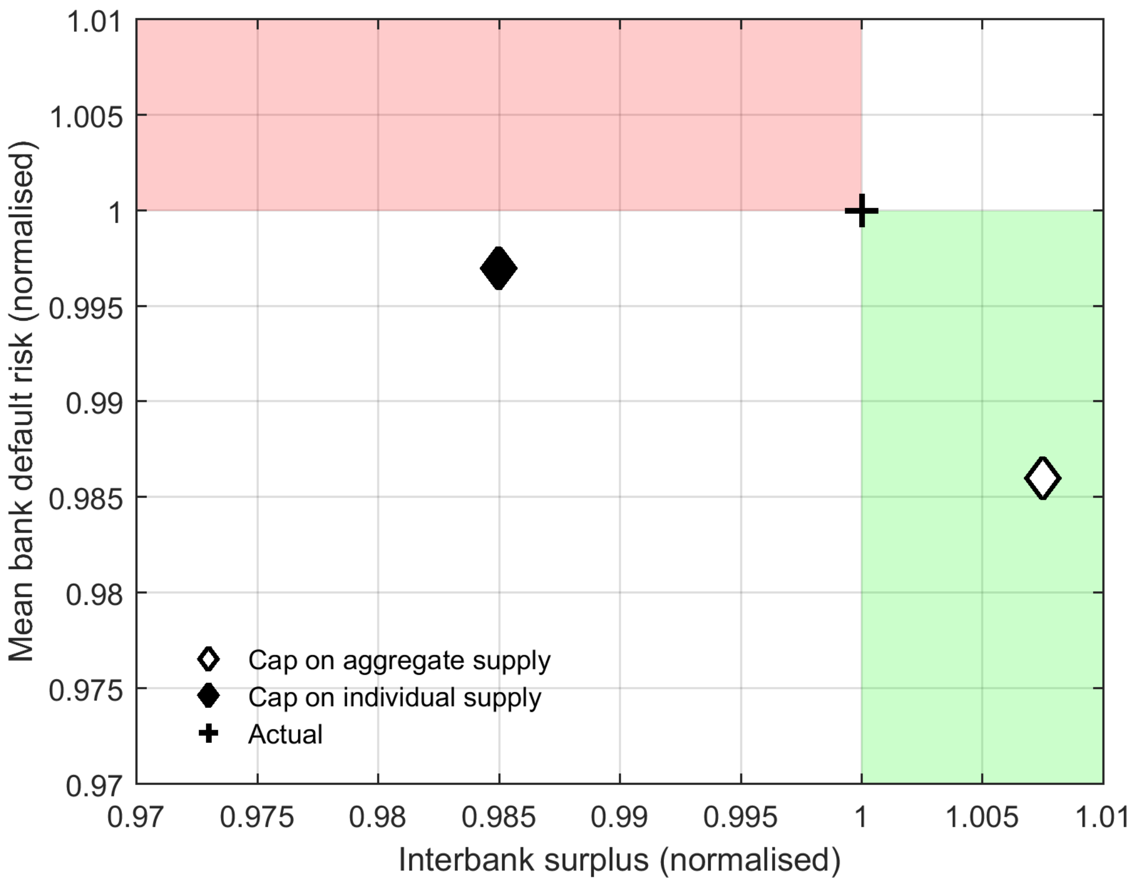
$$\sum_j C_{ijt}^{C} \leq 0.9 \sum_j C_{ijt}$$

A cap on total exposures held by bank i prevents other parts of the network from increasing in response to a capped link. A cap on total exposures also causes bank i to reduce risky (high Γ_{ij}) exposures by relatively more than safe (low Γ_{ij}) exposures. In other words, a cap on individual exposures targets relatively safe exposures, whereas a cap on total exposures targets relatively risky exposures. We simulate the effect of this cap in Figure 10, and find that it reduces mean default risk by significantly more than an individual cap and actually *increases* interbank surplus. Our results suggest a social planner therefore would strictly prefer a cap on total exposures to a cap on individual exposures.

6.3 Capital ratios

The second form of regulation we consider is a minimum capital requirement, as applied by regulators since the crisis. As described in Section 2, there is very little variation in risk-weights for exposures to banks under the standardised approach to risk-weighting. To assess the effect of a stylised risk-insensitive capital requirements, we simulate a further increase in

Figure 10: Counterfactual analysis of caps



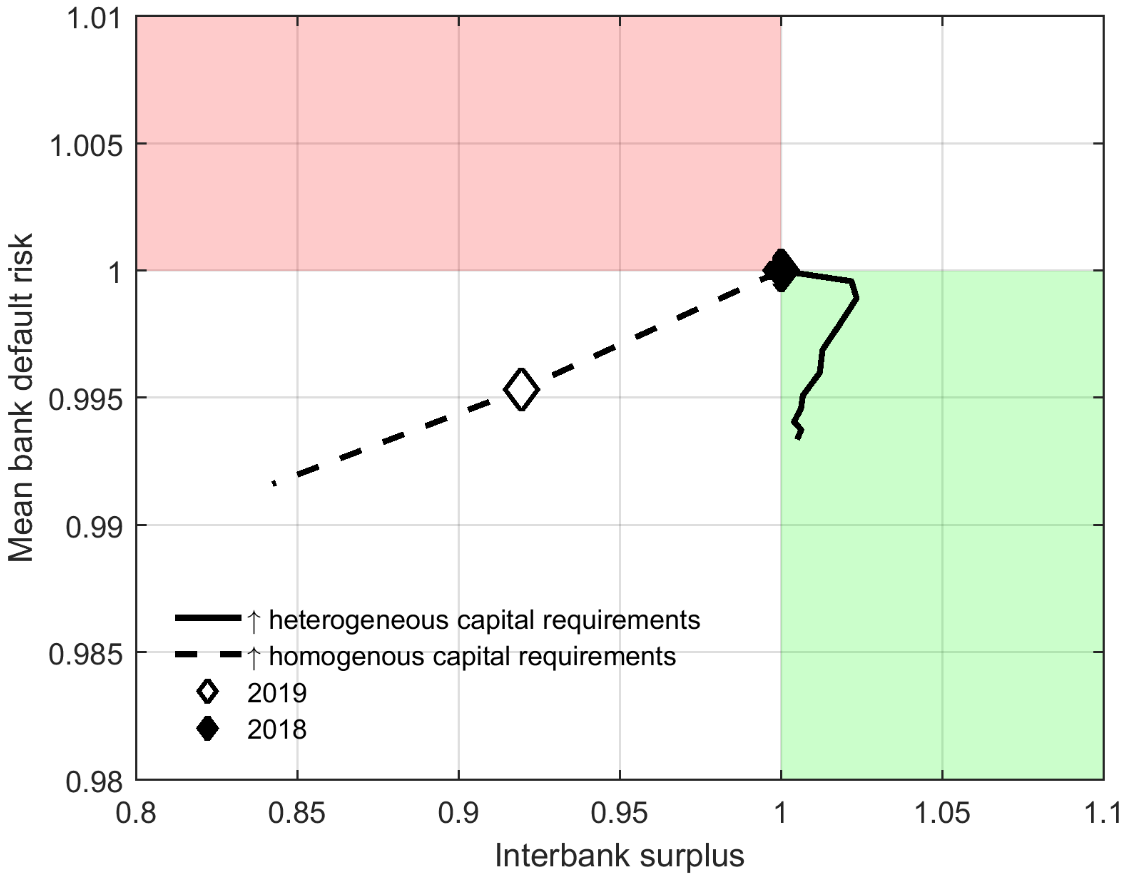
Note: The + sign indicates normalised default risk and interbank surplus in Q2 2018. The black diamond simulates a cap on individual exposures, C_{ij}. The white diamond simulates a cap on each bank's aggregate exposures, $\sum_j C_{ij}$.

λ_{it} by up to 2% holding bank fundamentals constant, as set out in Figure 11.

We propose a pairwise adjustment (that is, we allow λ_{ijt} to vary at the pair level) to capital ratios that is more closely targeted at network externalities. The key parameter in our model is Γ_{ij}, contagion intensity: links where this is high are particularly costly, in terms of their effect on default risk. We propose increasing the capital requirements for any link with $\Gamma_{ij} > median(\mathbf{\Gamma})$ ("high risk links") by some value b (where we increase the value b from 0% to 10% in Figure 11). For any link where Γ_{ij} is less than the 20th percentile of the distribution ("low risk links"), we propose *decreasing* the associated capital requirements by $b + 1.5\%$.[15] Our results suggest a social planner would strictly prefer this targeted change in capital ratios to a risk-insensitive increase in capital ratios.

[15] Any spread like this is an improvement over homogeneous capital requirements, this particular spread is one we have chosen arbitrarily as one that produces good results.

Figure 11: Counterfactual analysis of capital requirements



Note: This figure starts with normalised default risk and interbank surplus in Q2 2018 (the black diamond). We then plot the effect of (i) homogenous increases in capital requirements for all banks up to an additional 2% (the dashed line) and (ii) heterogeneous adjustments to capital requirements, as we describe in the text (the solid line). Heterogeneous capital requirements can reduce bank default risk by the same amount as homogeneous capital requirements, whilst increasing interbank surplus. For example, a homogenous increase in capital requirements was implemented in 2019 (the white diamond), when a heterogeneous adjustment to capital requirements could have reduced bank default risk to the same extent but with interbank surplus around 9% higher.

7 Conclusion

In this paper we set out a structural model of network formation and contagion. In contrast to much of the literature on financial networks, our model of network formation is in the spirit of the wider industrial organisation literature on demand estimation, in two ways in particular. First, we specify network formation as the interaction of demand (with a focus on the role of product characteristics in determining substitutability) and supply (with a focus on identifying the relevant underlying cost function). Second, in specifying our model and taking it to data we pay particular attention to the role of unobserved firm- and pair-level heterogeneity. In particular, the core of this paper is heterogeneity in contagion intensity, including (i) why one might reasonably expect contagion intensity to be heterogeneous, (ii) how this heterogeneity can be identified empirically and (iii) what implications this heterogeneity has for strategic interactions between firms and their regulation. The primary message of this paper is that this heterogeneity in contagion intensity has material implications for efficiency, centrality and optimal regulation.

References

Acemoglu, D., Ozdaglar, A., and Tahbaz-Salehi, A. (2015). Systemic risk and stability in financial networks. *American Economic Review*, 105(2):564–608.

Adrian, T., Boyarchenko, N., and Shachar, O. (2017). Dealer balance sheets and bond liquidity provision. *Journal of Monetary Economics*, 89:92–109.

AFME (2017). Crd 5: The new large exposures framework. Technical report, Association for Financial Markets in Europe.

Allen, F., Babus, A., and Carletti, E. (2009). Financial crises: theory and evidence. *Annu. Rev. Financ. Econ.*, 1(1):97–116.

Allen, J., Hortaçsu, A., and Kastl, J. (2011). Analyzing default risk and liquidity demand during a financial crisis: The case of Canada. Technical report, Bank of Canada Working Paper.

Babus, A. (2016). The formation of financial networks. *The RAND Journal of Economics*, 47(2):239–272.

Baker, M. and Wurgler, J. (2015). Do strict capital requirements raise the cost of capital? bank regulation, capital structure, and the low-risk anomaly. *American Economic Review*, 105(5):315–20.

Ballester, C., Calvó-Armengol, A., and Zenou, Y. (2006). Who's who in networks. wanted: The key player. *Econometrica*, 74(5):1403–1417.

Batiz-Zuk, E., López-Gallo, F., Martínez-Jaramillo, S., and Solórzano-Margain, J. P. (2016). Calibrating limits for large interbank exposures from a system-wide perspective. *Journal of Financial Stability*, 27:198–216.

BCBS (2011). Basel III: a global regulatory framework for more resilient banks and banking systems. Technical report, Bank for International Settlements.

BCBS (2014a). Basel III leverage ratio framework and disclosure requirements. Technical report, Bank for International Settlements.

BCBS (2014b). Supervisory framework for measuring and controlling large exposures. Technical report, Bank for International Settlements.

BCBS (2018a). Countercyclical capital buffer. Technical report, Bank for International Settlements.

BCBS (2018b). The treatment of large exposures in the Basel capital standards. Technical report, Bank for International Settlements.

Benetton, M. (2018). Leverage regulation and market structure: An empirical model of the UK mortgage market.

Bessembinder, H., Jacobsen, S., Maxwell, W., and Venkataraman, K. (2018). Capital commitment and illiquidity in corporate bonds. *The Journal of Finance*, 73(4):1615–1661.

Blasques, F., Bräuning, F., and Van Lelyveld, I. (2018). A dynamic network model of the unsecured interbank lending market. *Journal of Economic Dynamics and Control*, 90:310–342.

Bloch, F., Jackson, M. O., and Tebaldi, P. (2017). Centrality measures in networks.

Chang, B. and Zhang, S. (2018). Endogenous market making and network formation.

Cohen-Cole, E., Patacchini, E., and Zenou, Y. (2010). Systemic risk and network formation in the interbank market.

Craig, B. and Ma, Y. (2018). Intermediation in the interbank lending market. *Manuscript, Stanford Graduate School of Business.*

De Paula, A. (2017). Econometrics of network models. In *Advances in Economics and Econometrics: Theory and Applications, Eleventh World Congress*, pages 268–323. Cambridge University Press Cambridge.

Denbee, E., Julliard, C., Li, Y., and Yuan, K. (2017). Network risk and key players: A structural analysis of interbank liquidity.

Duffie, D. (2017). Financial regulatory reform after the crisis: An assessment. *Management Science*, 64(10):4835–4857.

Eisfeldt, A. L., Herskovic, B., Rajan, S., and Siriwardane, E. (2018). OTC intermediaries.

Elliott, M., Golub, B., and Jackson, M. O. (2014). Financial networks and contagion. *American Economic Review*, 104(10):3115–53.

Elliott, M., Hazell, J., and Georg, C.-P. (2018). Systemic risk-shifting in financial networks.

European Council (2018). Banking: Council agreement on measures to reduce risk. Technical report, European Council.

Farboodi, M. (2017). Intermediation and voluntary exposure to counterparty risk.

Gale, D. and Yorulmazer, T. (2013). Liquidity hoarding. *Theoretical Economics*, 8(2):291–324.

Glasserman, P. and Young, H. P. (2015). How likely is contagion in financial networks? *Journal of Banking & Finance*, 50:383–399.

Gofman, M. (2017). Efficiency and stability of a financial architecture with too-interconnected-to-fail institutions. *Journal of Financial Economics*, 124(1):113–146.

Greenwood, R., Stein, J. C., Hanson, S. G., and Sunderam, A. (2017). Strengthening and streamlining bank capital regulation. *Brookings Papers on Economic Activity*, 2017(2):479–565.

Hull, J. et al. (2009). *Options, futures and other derivatives/John C. Hull.* Upper Saddle River, NJ: Prentice Hall,.

Iyer, R. and Peydro, J.-L. (2011). Interbank contagion at work: Evidence from a natural experiment. *The Review of Financial Studies*, 24(4):1337–1377.

Kashyap, A. K., Stein, J. C., and Hanson, S. (2010). An analysis of the impact of 'substantially heightened' capital requirements on large financial institutions. *Booth School of Business, University of Chicago, mimeo*, 2.

Kotidis, A. and Van Horen, N. (2018). Repo market functioning: The role of capital regulation.

Modigliani, F. and Miller, M. H. (1958). The cost of capital, corporation finance and the theory of investment. *The American*, 1:3.

Robles-Garcia, C. (2018). Competition and incentives in mortgage markets: The role of brokers.

Welfens, P. J. (2011). The transatlantic banking crisis: lessons, EU reforms and G20 issues. In *Financial Market Integration and Growth*, pages 49–126. Springer.

Yellen, J. (2013). Interconnectedness and systemic risk: Lessons from the financial crisis and policy implications. *Board of Governors of the Federal Reserve System, Washington, DC.*

A First stage regression results

Table 5: First stage: Default risk

	p_{it}
X^1_{it}	-0.82***
	(-2.61)
FE	i
Other X	Y
R^2	0.82
No. obs	378

Note: Figures in parentheses are t-statistics. ***, **, * indicate difference from 0 at 1%, 5% and 10% significance, respectively. X^1_{it} is a revenue-weighted average of stock market indices and the other fundamentals include the Morgan Stanley World Index, VIX and the first two principal components of World Bank macroeconomic data, as we describe in the text.

Table 6: First-stage: Network formation results

	Estimate	t statistic
X_{it}	-0.57***	-3.73
X_{jt}	0.22	1.51
X_{kt}	0.35***	11.90
X_{it}^2	0.01	0.18
X_{jt}^2	0.28*	1.88
X_{kt}^2	-0.39***	-10.05
X_{jt}/X_{it}	0.01	1.42
X_{jt}/X_{kt}	-0.46	-0.55
X_{it}/X_{kt}	-1.44*	-1.69
$\lambda_{it}X_{it}$	13.86***	7.31
$\lambda_{it}X_{jt}$	-2.94	-1.57
$\lambda_{it}X_{kt}$	-10.69***	-13.64
$\lambda_{it}X_{it}^2$	-0.27	-0.35
$\lambda_{it}X_{jt}^2$	-9.24***	-5.38
$\lambda_{it}X_{kt}^2$	11.70***	8.84
$\lambda_{it}X_{jt}/X_{it}$	-0.192**	-2.05
$\lambda_{it}X_{jt}/X_{kt}$	10.32	0.93
$\lambda_{it}X_{it}/X_{kt}$	-21.27*	-1.89
FE	ij	
R^2	0.70	
No. obs	6,426	

B Mathematical appendix

B.1 EQC

In this appendix, we derive the equilibrium quantity condition, EQC. The first order supply condition is:

$$r_{ijt} = -\frac{\partial r_{ijt}}{\partial C_{ijt}} C_{ijt} + mc^e_{ijt} + \frac{\partial p_{it}}{\partial C_{ijt}} \sum_k \frac{\partial mc^e_{ikt}}{\partial p_{it}} C_{ikt} - \frac{\partial \Pi^D_{it}}{\partial p_{it}} \frac{\partial p_{it}}{\partial C_{ijt}} + r_{i0t} + e^S_{ijt}$$

It follows immediately from DRP that $\frac{\partial p_{it}}{\partial C_{ijt}} = \tau_t \Gamma_{ij} p_{jt}$, from our assumed cost function that $\frac{\partial mc^e_{kjt}}{\partial p_{it}} = \phi_1 \lambda_{kjt}$ and from our demand model that $\frac{\partial r_{ijt}}{\partial C_{ijt}} = -B$ and $\frac{\partial \Pi^D_{it}}{\partial p_{it}} = -\sum_k \frac{\partial r_{kit}}{\partial p_{it}} C_{kit}$:

$$r_{ijt} = BC_{ijt} + \phi_1 \lambda_{ijt} p_{it} + \phi_1 \tau_t \Gamma_{ij} p_{jt} \sum_m \lambda_{imt} C_{imt} + \frac{\partial p_{it}}{\partial C_{ijt}} \sum_m \frac{\partial r_{mit}}{\partial p_{it}} C_{mit} + r_{i0t} + e^S_{ijt}$$

For ease of exposition we then repeat the same equation for supply from bank k to bank i:

$$r_{kit} = BC_{kit} + \phi_1 \lambda_{kit} p_{kt} + \phi_1 \tau_t \Gamma_{ki} p_{it} \sum_m \lambda_{kmt} C_{kmt} + \frac{\partial p_{kt}}{\partial C_{kit}} \sum_m \frac{\partial r_{mkt}}{\partial p_{kt}} C_{mkt} + r_{k0t} + e^S_{kit}$$

When bank i considers how much to supply to bank j, it takes into account the impact of the resulting increase in p_{it} on its profits from being supplied exposures: this is the penultimate term in the equations above. That is, it takes into account the effect of its supply on r_{kit}. We assume that bank i takes the interest rates of transactions involving other parties as given, such that:

$$\frac{\partial r_{kit}}{\partial p_{it}} = \phi_1 \tau_t \Gamma_{ki} \sum_m \lambda_{kmt} C_{kmt}$$

Substitute this and the equation for demand into supply, and we obtain the EQC:

$$\begin{aligned} 0 \quad = \quad & \delta_{jt} + \zeta_{ij} + e^D_{ijt} - 2BC_{ijt} - \sum_{k \neq i}^{N} \theta_{ik} C_{kjt} + e^S_{ijt} \\ & -\lambda_{ijt} \phi_1 p_{it} - \phi_1 \tau_t \Gamma_{ij} p_{jt} \sum_{k \neq i}^{N} C_{ikt} \lambda_{ikt} - r_{i0t} \\ & -\phi_1 \tau_t^2 \Gamma_{ij} p_{jt} \sum_k C_{kit} \Gamma_{ki} \sum_m C_{kmt} \lambda_{kmt} \end{aligned}$$

B.2 Equilibrium links are non-linear in fundamentals

Consider a simplified version of the model in which banks do not consider the impact of their supply decisions on Π^D; that is, they consider the impact on their funding costs when supplying on the interbank market, but not on their funding costs when demanding from the interbank market. This means that the EQC is linear in C. Furthermore, for simplicity of exposition (and without loss of generality regarding the form of equilibrium C) suppose $\zeta = e^D = e^S = r_0 = 0$, $2B = \phi_1 = \lambda = 1$, $\theta_{ij} = \theta$, $\Gamma_{ij} = \Gamma$ for all banks and parameters are such that all equilibrium exposures are strictly positive. The EQC is then as follows:

$$0 = \delta_{jt} - C_{ijt} - \theta \sum_{k \neq i}^{N} C_{kjt} - p_{it} - \Gamma p_{jt} \sum_{k \neq i}^{N} C_{ikt}$$

In this case an analytical expression for equilibrium exposures exists, where $\boldsymbol{C}$ is a $N(N-1) \times 1$ vector of endogenous exposures, $\boldsymbol{p}$ is a $N \times 1$ vector of default probabilities, $\boldsymbol{X}$ is a $N \times 1$ vector of fundamentals, $\boldsymbol{M_i}$, $\boldsymbol{M_j}$, $\boldsymbol{M_{\sum i}}$ and $\boldsymbol{M_{\sum j}}$ are matrices that select and sum the appropriate elements in $\boldsymbol{C}$ and $\boldsymbol{p}$ and . and $\circ$ signify matrix multiplication and the Hadamard product, respectively:

$$\boldsymbol{C} = \Big[\boldsymbol{I} + \theta \boldsymbol{M_{\sum j}} + (\boldsymbol{M_j}.\boldsymbol{p}) \circ \boldsymbol{M_{\sum i}}\Big]^{-1} \Big[\boldsymbol{M_j}.\boldsymbol{\delta} - \boldsymbol{M_i}.\boldsymbol{p}\Big]$$

Given that p is a linear function of X, as set out in the DRP, it follows that equilibrium C is a non-linear function of X.

C Additional post-estimation tests

C.1 Default risk and cost of equity

In this sub-section, we show test our parameterisation of a bank's cost of equity as a function of its default risk is reasonable. We run a linear regression of a bank's cost of equity, taken from Bloomberg and based on a simple CAPM model, on its default risk.

$$c^e_{it} = \phi p_{it} + FE_i + FE_t + e^e_{it}$$

As we set out below, we find that the relationship between the two is positive and significant, as expected. Riskier banks face a higher cost of capital, even when controlling for time fixed effects.

Table 7: Cost of equity and default risk

	c^e_{it}
p_{it}	1.31***
	(2.94)
FE	i,t
R^2	0.69
No. obs	346

Note: Figures in parentheses are t-statistics. ***, **, * indicate difference from 0 at 1%, 5% and 10% significance, respectively.

C.2 Testing heterogeneous contagion intensity

We set out above three motivations for heterogeneous contagion intensity Γ_{ij}: (1) correlations in fundamentals (risk sharing, in other words), (2) variations in product and (3) other pairwise variations, including common jurisdiction. We estimate general Γ_{ij} without imposing any of these motivations in estimation, meaning we can test them post-estimation. In particular, risk sharing implies a relationship between $\mathbf{X}\boldsymbol{\beta}$ and Γ_{ij}, which we test in the following way.

As bank-specific fundamentals we use equity indices weighted by the geographic revenues of each bank, as we describe above. This implies that banks that get their revenues from the same geographic areas will have positively correlated fundamentals, and banks that have differing geographic revenue profiles will have less correlated fundamentals. For each pair of banks we calculate the empirical correlation coefficient as $\widehat{\rho}_{ijt} = Corr(\mathbf{X_{it}}\widehat{\boldsymbol{\beta}}, \mathbf{X_{jt}}\widehat{\boldsymbol{\beta}})$.

We then divide our bank pairs into two groups, "more correlated" and "less correlated", by defining the dummy variable $1_{\rho_{ij}} = 1$ if $\widehat{\rho}_{ij} > median(\widehat{\rho}_{ij})$ and $1_{\rho_{ij}} = 0$ otherwise. We divide bank pairs similarly regarding Γ_{ij}, into "safe links" and "risky links", by defining the dummy variable $1_{\Gamma_{ij}} = 1$ if $\widehat{\Gamma}_{ij} > median(\widehat{\Gamma}_{ij})$ and $1_{\Gamma_{ij}} = 0$ otherwise. Risk sharing implies that safe links should be less correlated, and risky links should be more correlated. Risk sharing is, however, difficult to separately identify from other motivations for heterogeneous contagion intensity. In particular, less correlated links are more likely to go across jurisdictions than more correlated links, where going across jurisdictions may make links less safe. We test this by identifying the home jurisdiction of each of the $N = 18$ banks in our sample and classifying each as being in the UK, North America, Europe or Asia. We then define the dummy variable $1_G = 1$ if they share the same home jurisdiction, and 0 otherwise. We do not attempt to test the effect of product variations, as there are many product characteristics and we do not have a clear ranking of their relative riskiness.

We run the following linear regression:

$$1_{\Gamma_{ij}} = \alpha_0 + \alpha_1 1_{\rho_{ij}} + \alpha_2 1_G + \alpha_3 1_G 1_{\rho_{ij}} + e_t^\alpha$$

Table 8: Drivers of heterogeneous contagion intensity

	[**1**]
$1_{\rho_{ij}}$	-0.280***
	(-4.44)
1_G	-0.204
	(-1.61)
$1_G 1_{\rho_{ij}}$	0.600***
	(3.96)
R^2	0.09
No. obs	153

Note: Figures in parentheses are t-statistics. ***, **, * indicate difference from 0 at 1%, 5% and 10% significance, respectively.

The coefficient on the interaction term is positive and significant: where banks are in the same jurisdiction, then more correlated links are less safe. We interpret this as evidence in support of a risk sharing motivation for heterogeneous contagion intensity. The coefficient on 1_G is the right sign (indicating that links within the same jurisdiction are safer), but insignificant. The coefficient on $1_{\rho_{ij}}$ is negative and significant: this suggests that when links go across jurisdictions, less correlated links are actually less safe. This could still be because of confounding jurisdictional effects: within the set of links that cross jurisdictions, more distant links will be riskier but also less correlated.