



Quantitative Finance

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/rquf20>

Filling in the blanks: network structure and interbank contagion

Kartik Anand^a, Ben Craig^{bc} & Goetz von Peter^d

^a Financial Stability Department, Bank of Canada, Ottawa, ON, K1A 0G9Canada.

^b Deutsche Bundesbank, Wilhelm-Epstein-Straße 14, 60431Frankfurt am Main, Germany.

^c Federal Reserve Bank of Cleveland, Cleveland, OH, 44101-1387USA.

^d Bank for International Settlements, Centralbahnplatz 2, Basel, Switzerland.

Published online: 28 Nov 2014.

To cite this article: Kartik Anand, Ben Craig & Goetz von Peter (2014): Filling in the blanks: network structure and interbank contagion, Quantitative Finance, DOI: [10.1080/14697688.2014.968195](https://doi.org/10.1080/14697688.2014.968195)

To link to this article: <http://dx.doi.org/10.1080/14697688.2014.968195>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

Filling in the blanks: network structure and interbank contagion

KARTIK ANAND[†], BEN CRAIG^{‡§} and GOETZ VON PETER^{*¶}

[†]Financial Stability Department, Bank of Canada, Ottawa, ON K1A 0G9, Canada

[‡]Deutsche Bundesbank, Wilhelm-Epstein-Straße 14, 60431 Frankfurt am Main, Germany

[§]Federal Reserve Bank of Cleveland, Cleveland, OH 44101-1387, USA

[¶]Bank for International Settlements, Centralbahnplatz 2, Basel, Switzerland

(Received 15 September 2013; accepted 19 June 2014)

The network pattern of financial linkages is important in many areas of banking and finance. Yet, bilateral linkages are often unobserved, and *maximum entropy* serves as the leading method for estimating counterparty exposures. This paper proposes an efficient alternative that combines information-theoretic arguments with economic incentives to produce more realistic interbank networks that preserve important characteristics of the original interbank market. The method loads the most probable links with the largest exposures consistent with the total lending and borrowing of each bank, yielding networks with *minimum density*. When used in a stress-testing context, the minimum-density solution overestimates contagion, whereas maximum entropy underestimates it. Using the two benchmarks side-by-side defines a useful range that bounds the cost of contagion in the true interbank network when counterparty exposures are unknown.

Keywords: Interbank markets; Networks; Entropy; Intermediation; Systemic risk

JEL Classification: G21, L14, D85, C63

1. Introduction

Interbank contagion is a fundamental channel in many of the stress tests gauging systemic risk. Yet, in practice, the interbank network at the core of these simulations often remains unobserved: because interbank loans are generally arranged over the counter, the bilateral exposures are often known only by the immediate counterparties of each trade. In some jurisdictions, bilateral positions can be obtained from regulatory filings or credit registers. More often, however, central banks and regulators do not observe the network because banks do not report their bilateral exposures. In those cases, the leading method is for researchers to fill in the blanks as evenly as possible, using the available information on each bank's total interbank lending (Upper 2011, Elsinger *et al.* 2013). This approach, known as *maximum entropy*, effectively assumes that banks diversify their exposures by spreading their lending and borrowing across *all* other active banks.

The maximum-entropy (ME) approach for filling in the blanks, however, can be misleading when the result is employed in network analysis or financial stress tests. The practical shortcoming is that applying ME tends to create *complete* networks which obscure the true structure of linkages in the

original network. Key concepts in network analysis, such as degree (number of connections), become meaningless as a result of using ME. Furthermore, when such estimated networks are used for purposes of stress testing, ME introduces a bias (Mistrulli 2011, Upper 2011, Markose *et al.* 2012). The related conceptual shortcoming is that applying ME is optimal from an information-theoretic perspective only if *nothing else* is known about the network. But we do know that interbank networks are typically sparse, because interbank activity is based on relationships (Cocco *et al.* 2009) and smaller banks use a limited set of money center banks as intermediaries (Craig and von Peter 2014). Indeed, most banks would find it prohibitively costly in terms of information processing and risk management to lend to *every* active bank in the system.

This paper proposes an alternative benchmark, one that minimizes the number of links necessary for distributing a given volume of loans on the interbank market. Our *minimum-density* (MD) method, in contrast to ME, is based on the economic rationale that interbank linkages are costly to add and maintain. Our procedure determines a pattern of linkages for allocating interbank positions that is efficient in the sense of minimizing these costs. Intuitively, our approach identifies the most probable links and loads them with the largest possible exposures consistent with the total lending and borrowing observed for

*Corresponding author. Email: Goetz.von.Peter@bis.org

each bank that could be obtained from the bank balance sheet. This link prediction method uses only this balance-sheet data, combined with elements of information theory and economic rationale, to determine a sparse network without using any information from unobserved bilateral positions of the original network.

Even as the stress-testing literature recognizes that the use of ME induces a bias, few viable alternatives have been proposed. In their study of the credit default swap market, [Markose et al. \(2012\)](#) distribute exposures over a subset of possible linkages, where—arbitrarily—the degree of each bank is proportional to its market share. This method leads to a bilateral network that is more concentrated than the ME solution. Technically, our paper is more closely related to [Mastromatteo et al. \(2012\)](#), who provide an efficient algorithm to sample from a distribution of potential networks that satisfy structural constraints.[†] Our paper, in contrast, constructs an optimal multinomial logit sampling distribution that draws on ideas from network analysis and information theory. We provide a heuristic algorithm to reconstruct interbank networks that incorporate priors over the set of links between banks that are more likely to arise.[‡]

In section three, we confront these two estimated benchmarks with the German interbank network where we have detailed data on the linkages between 1800 banks. We find that our MD solution preserves some of the true network's structural features better than maximum entropy. The match can be improved by considering 'low-density' solutions that load bilateral links less aggressively than the MD solution does. This makes our method a reasonable alternative benchmark for estimating missing counterparty exposures, one that does not wipe out the sparse structure of linkages that is so central to network analysis.

The final section shows that systemic risk clearly depends on the pattern of interlinkages. We contrast the results from a standard stress test on the German banking system with those obtained using the two alternative benchmarks instead. We find that the ME approach underestimates contagion, whereas our MD method overestimates it, often to a lesser extent. Using the two benchmarks in this context thus helps identify a range of possible stress-test outcomes and also facilitates robust systemic risk analysis through repeated application. This makes the case for using both benchmarks when gauging how financial linkages affect systemic risk.

While these findings are relevant for central banks and regulators, our approach may be of independent interest. In finance and various other disciplines, situations arise where a network of interest is not fully observed, or has yet to be designed. Networks in transportation, financial markets or international trade are much sparser—and for good reason—than ME would have us believe. Our MD method may thus provide a meaningful alternative in various areas, guided by the simple economic rationale of minimizing the cost of additional links.

[†]Assuming a multinomial logit, or Gibbs-Boltzmann, probability distribution, they employ a belief-propagation algorithm to compute 'marginal' probabilities for the likelihood that a link exists between two banks.

[‡][Wang \(2013\)](#) also proposes methods for estimating interbank networks using daily balance-sheet data for Singapore banks.

2. The minimum density method

2.1. The case for an alternative to maximum entropy

Consider a system of N banks engaged in interbank lending and borrowing. The matrix $\mathbb{X} \in [0, \infty)^{N \times N}$ represents gross interbank positions, where the typical element X_{ij} represents the amount bank i lends to bank j . Such networks are directed and valued. For each bank i , the row sum of \mathbb{X} shows total interbank assets, and the column sum tallies bank i 's interbank liabilities vis-à-vis all other banks,

$$\begin{cases} \text{Interbank assets:} & A_i = \sum_{j=1}^N X_{ij} \\ \text{Interbank liabilities:} & L_i = \sum_{j=1}^N X_{ji} \end{cases} \quad (1)$$

Matrix \mathbb{X} is the network of bilateral exposures that is needed for additional analysis. In many situations, however, the bilateral linkages are unknown. National authorities in most jurisdictions do not observe the full interbank network, because banks do not report bilateral positions or only disclose their largest exposures. However, the authorities generally do have balance-sheet information at their disposal, including the total interbank assets A_i and liabilities L_i of each bank.

Suppose, for ease of exposition, that no bilateral positions are observed.[§] In this simple case, the authorities only know the 'marginals': how much each bank, i , lends and borrows in aggregate on the interbank market, A_i and L_i , but not to and from whom. Since information on counterparty exposures (X_{ij}) is essential for further analysis, it is necessary to resort to a method for filling in the interbank matrix.

The standard approach in the literature is to estimate a matrix by ME ([Sheldon and Maurer 1998](#), [Upper 2011](#), [Elsinger et al. 2013](#)). Intuitively, this approach spreads exposures as evenly as possible, consistent with the marginals, and thereby fills all cells between active banks. Formally, the method solves for the bilateral exposures $\{E_{ij}\}$ that minimize the relative entropy function $\sum_{i,j} E_{ij} \ln(E_{ij}/Q_{ij})$ subject to constraints (1), and relative to prior information, $\{Q_{ij}\}$ on bilateral exposures, if available. Since entropy is a measure of probabilistic uncertainty, this approach is optimal when selecting a solution in the sense of using least information ([MacKay 2003](#)). Entropy optimization is widely used across disciplines ([Fang et al. 1997](#)), and is straightforward to implement using a standard iterative algorithm (RAS), which can be generalized to handle additional constraints ([Blien et al. 1998](#), [Elsinger et al. 2013](#)).

In the interbank context, the ME approach delivers an unrealistic network structure, and one that tends to understate contagion in systemic stress tests. In the simple case where only the marginals are known, the ME solution takes the form of a gravity equation where the estimate E_{ij} is proportional to the product of marginals $A_i L_j$. To the extent that these marginals are positive (that is, all banks both borrow and lend to at least some bank in the network), ME produces a complete network where each bank lends to all other banks. Such a network structure is rather atypical for interbank markets, and when used in stress testing it tends to introduce a bias that underestimates the true extent of contagion ([Mistrulli 2011](#), [Markose et al. 2012](#)). There are theoretical reasons why a

[§]This is without loss of generality, since the method below can accommodate more information with exclusion restrictions on the prior distribution Q below.

complete market structure tends to be more robust to contagion (Allen and Gale 2000). Simulations on more realistic networks show that greater diversification through interlinkages indeed lowers the probability of a crisis, but may also raise its severity when a crisis occurs (Nier *et al.* 2007).

Our approach for deriving an opposite benchmark starts from the premise that establishing and maintaining network linkages is costly. Banks do not spread their borrowing and lending across the entire system, since the costs in terms of information processing, risk management and creditworthiness checks would be prohibitive for all but the largest banks. In reality, interbank activity occurs through relationships (Cocco *et al.* 2009), and interbank networks are *sparse* as a result, often with less than 1% of potential bilateral linkages in active use (see Bech and Atalay 2010 for the United States, and Craig and von Peter 2014 for Germany). These relationships are also *disassortative*: less-connected banks are more likely to trade with well-connected banks than with other less-connected banks (Iori *et al.* 2008, Bech and Atalay 2010). This reflects the economic rationale that smaller banks, rather than transacting with each other, typically use a small set of money center banks as intermediaries (Craig and von Peter 2014). Similar observations hold for dealer networks in financial markets (Li and Schürhoff 2012) or in international trade (Helpman *et al.* 2008, Ahn *et al.* 2011). For these reasons, this paper proposes MD as an alternative benchmark for estimating counterparty exposures.

2.2. An illustration

A simple example helps to illustrate the arguments made so far. Imagine a simple interbank market consisting of seven banks. In figure 1, this market is represented by the top-left matrix; out of seven banks, two only lend (D,E), one only borrows (F) and the remaining four banks intermediate.† The authorities only observe individual bank balance sheets, but no bilateral transactions—they only see what each bank lends or borrows in total on the interbank market (the marginals, bottom left-hand matrix). If they estimate counterparty exposures by ME, they obtain the matrix at the top right. That solution differs substantially from the original \mathbb{X} ; while it preserves the marginals of each bank, it fills all the linkages between active banks with positive exposures. The more banks are active, the less ME will preserve of the original network structure. Our MD solution, on the other hand, minimizes the number of linkages while also respecting all marginals and net positions, and leads to a sparse and concentrated interbank structure.

The original network in this example has a density of 33%, with 14 out of the 42 potential bilateral links being used. The ME solution shows far higher density (62%), whereas our solution reaches the lowest density (21%) attainable with those marginals. While ME spreads interbank activity as evenly as possible, MD does the opposite by concentrating exposures on the smallest possible set of links. This illustrates that the two benchmarks depart from the true network in opposite directions in how they trade off the number vs. the size of interbank linkages.

†Banks A–C are core banks that play a central role in intermediating between periphery banks (Craig and von Peter 2014).

2.3. The formal problem

As an efficient alternative to ME, our approach minimizes the total number of linkages necessary for allocating interbank positions, consistent with total lending and borrowing observed for each bank.‡ Let c represent the fixed cost of establishing a link. Then the MD approach can be formulated as a constrained optimization problem for the matrix \mathbb{Z} ,

$$\begin{aligned} \min_{\mathbb{Z}} \quad & c \sum_{i=1}^N \sum_{j=1}^N \mathbf{1}_{[Z_{ij}>0]} \quad \text{s.t.} \\ & \sum_{j=1}^N Z_{ij} = A_i \quad \forall i = 1, 2, \dots, N \\ & \sum_{i=1}^N Z_{ij} = L_j \quad \forall j = 1, 2, \dots, N \\ & Z_{ij} \geq 0 \quad \forall i, j, \end{aligned} \quad (2)$$

where the integer function $\mathbf{1}$ equals one only if bank i lends to bank j . The economic nature of this problem shares similarities with network design problems in transportation science and communication networks. Minimizing the cost of transporting a given volume of goods from origins to destinations appears analogous to moving money between banks. In our case, the capacities of transportation hubs (banks) are constrained by two marginals, A_i and L_i , and the fixed cost of building new roads (credit relationships) must be considered as well. Such network design problems have been studied for decades and are known to be non-deterministic polynomial-time hard except in very special cases (Campbell and O’Kelly 2012 provide a survey).

We propose a link prediction method that combines elements from information theory and economic rationale. In problem (2), the *value* for any configuration of linkages \mathbb{X} can be expressed as follows. In our objective function, which is used to assign links, we first soften the constraints by assigning penalties for *deviations* from the marginals,

$$\begin{aligned} AD_i &\equiv \left(A_i - \sum_j Z_{ij} \right), \\ LD_i &\equiv \left(L_i - \sum_j Z_{ji} \right), \end{aligned}$$

where LD_i measures bank i ’s current deficit; i.e. how much its bilateral borrowing falls short of the total amount it needs to raise, L_i , which is also the amount to be matched by the solution being constructed, \mathbb{Z} . Defining these deviations helps to make the optimization problem smooth, so that the only non-smooth part lies in the cost of links. When they are introduced into the objective function, the problem maximizes

$$V(\mathbb{Z}) = -c \sum_{i=1}^N \sum_{j=1}^N \mathbf{1}_{[Z_{ij}>0]} - \sum_{i=1}^N \left[\alpha_i AD_i^2 + \delta_i LD_i^2 \right]. \quad (3)$$

Sparse \mathbb{Z} networks that minimize the deviations from the marginals are more efficient and attain higher values in the objective function $V(\mathbb{Z})$.

In addition to being sparse, interbank networks are typically disassortative: small banks seek to match their lending and

‡Related to this, minimally connected networks arise as the *efficient* solution in economic models where agents trade off the costs and rewards of forming links (Goyal 2007, Jackson 2008).

Actual Data								
True Network								
	A	B	C	D	E	F	G	A_i
A	0	3	1	0	0	1	2	7
B	2	0	2	0	0	0	1	5
C	1	1	0	0	0	1	0	3
D	1	0	0	0	0	0	0	1
E	0	0	2	0	0	0	1	3
F	0	0	0	0	0	0	0	0
G	0	1	0	0	0	0	0	1
L_i	4	5	5	0	0	2	4	20

Estimated Networks								
Maximum Entropy Solution								
	A	B	C	D	E	F	G	A_i
A	0	2.53	2.18	0	0	0.74	1.55	7
B	1.72	0	1.6	0	0	0.54	1.14	5
C	0.98	1.06	0	0	0	0.31	0.65	3
D	0.25	0.27	0.23	0	0	0.08	0.17	1
E	0.75	0.81	0.7	0	0	0.24	0.5	3
F	0	0	0	0	0	0	0	0
G	0.3	0.32	0.28	0	0	0.09	0	1
L_i	4	5	5	0	0	2	4	20

Observable Interbank Market								
	A	B	C	D	E	F	G	A_i
A								7
B								5
C								3
D								1
E								3
F								0
G								1
L_i	4	5	5	0	0	2	4	20

Minimum Density Solution								
	A	B	C	D	E	F	G	A_i
A	0	3	0	0	0	0	4	7
B	3	0	2	0	0	0	0	5
C	0	2	0	0	0	1	0	3
D	0	0	0	0	0	1	0	1
E	0	0	3	0	0	0	0	3
F	0	0	0	0	0	0	0	0
G	1	0	0	0	0	0	0	1
L_i	4	5	5	0	0	2	4	20

Figure 1. Illustrative example comparing the ME and MD solutions for a hypothetical interbank market.

borrowing needs through relationships with larger banks that are well placed to satisfy those needs. Such behavior contributes to sparsity, since most small banks can satisfy their needs with a single large counterparty. The relevant measure of size here is a bank's current surplus AD_i and deficit LD_i to be met in the interbank market. We codify this information through the set of probabilities $Q \equiv \{Q_{ij}\}$ for relationships between i and j . The probability that i lends to j increases if either i is a large lender to a small borrower j , or i is a small lender to a large borrower j . This is achieved by the following criterion:

$$Q_{ij} \propto \max \left\{ \frac{AD_i}{LD_j}, \frac{LD_j}{AD_i} \right\}. \quad (4)$$

Thus, the criterion gives a higher probability to a link being chosen if either the asset deficit of bank i is high relative to the liability deficit of bank j or the liability deficit of bank j is high relative to the asset deficit of bank i . While the prior Q codifies the probabilities for picking links, it is agnostic as to the values one should allocate to selected links.[†]

According to the objective function, equation (3), networks with a lower density have higher values. At the same time, the priors Q specify the characteristic interbank feature that small banks typically have links with large banks. These two ingredients are at the crux of our link prediction procedure. In order to generate networks that are both sparse and disassortative, we propose a simple trade-off. Let $P(\mathbb{Z})$ be the probability distribution over all possible network configurations. We derive this distribution by maximizing the sum of

two terms (as in Weisbuch *et al.* 2000). The first is the expected value of networks, $\sum_{\mathbb{Z}} P(\mathbb{Z}) V(\mathbb{Z})$ —based on our formulation in equation (3), networks that have few links and a high value should be more likely. At the same time, to ensure that the most likely network solutions are disassortative, the probability distribution P should be *close* to our prior Q . Formally, this is achieved by maximizing the relative entropy $R(P \parallel Q) = \sum_{\mathbb{Z}} P(\mathbb{Z}) \log(P(\mathbb{Z})/Q(\mathbb{Z}))$ between the distribution P and prior. Thus, the distribution, P , over possible network configuration is derived as the solution to

$$\max_P \sum_{\mathbb{Z}} P(\mathbb{Z}) V(\mathbb{Z}) + \theta R(P \parallel Q).$$

The scaling parameter, θ , emphasizes the weight placed on finding solutions with similar characteristics to the prior matrix Q . The solution to this problem can be obtained from the first-order conditions as

$$P(\mathbb{Z}) \propto Q(\mathbb{Z}) e^{\theta V(\mathbb{Z})}, \quad (5)$$

which, when appropriately normalized over all possible network configurations, yields the multinomial logit choice function.[‡] What this expression states is analogous to other settings of probabilistic choice: a candidate \mathbb{Z} has a higher likelihood of being chosen than the prior Q specifies if the departure from Q raises the value of the objective (3) which defines the maximization problem.

The solution would be straightforward if one could enumerate all possible network configurations \mathbb{Z} , and rank them according to $P(\mathbb{Z})$. But the computational complexity of such problems rises exponentially with N due to the number of possible subsets (2^N), even before allocating monetary values

[†]There are many ways in which such a prior can be formulated, with this choice being most parsimonious in using only information on the two banks' marginals. The prior can be refined as more information about the interbank market becomes available. For instance, the probability of the link can be made conditional on the type of banks involved, their respective balance-sheet variables and their proximity to each other.

[‡]This objective function may also be derived by drawing on the insights of Hansen and Sargent (2001) where there is uncertainty about the prior Q and one seeks a robust choice function P . See also Mattsson and Weibull (2002) and Strzalecki (2011).

to each link. Since exhaustive search is impossible—certainly for our application with 1800 banks—we develop a heuristic procedure for allocating links.

2.4. Implementation

The pseudo-code is provided in the appendix, and the Matlab code is available from the authors upon request. The MD heuristic proceeds as follows:

- At each iteration, a link (i, j) is selected with probability Q_{ij} , where Q_{ij} is defined in (4).
- The exposure Z_{ij} is loaded with the maximum value that this pair of banks can transact given their current asset and liability positions, i.e. $Z_{ij} = \min\{AD_i, LD_j\}$.
- If adding this link increases the value function, $V(\mathbb{Z} + Z_{ij}) > V(\mathbb{Z})$, the allocation is retained.
- If, however, the addition of Z_{ij} diminishes the value function:
 - we retain the link as long as the network including Z_{ij} is more likely than without the link, i.e. with probability $P(\mathbb{Z} + Z_{ij})/P(\mathbb{Z}) \simeq \exp\{\theta[V(\mathbb{Z} + Z_{ij}) - V(\mathbb{Z})]\}$.
 - The link is otherwise rejected.
- Finally, once positions have been updated, we proceed to the next iteration until the total interbank market volume has been allocated.[†]

The MD solution \mathbb{Z} that this procedure yields and the ME solution are at opposite extremes of the spectrum, both potentially far from the true \mathbb{X} . One advantage of our method is that we can depart parametrically from the MD benchmark and obtain less aggressive solutions that distribute interbank exposures over more links. This is achieved by scaling the loadings for selected links, $\lambda \times \min\{AD_i, LD_j\}$, using the parameter $\lambda < 1$. This forces the procedure to select more links until the overall interbank volume is reached. Such a ‘low-density solution,’ which we label \mathbb{Y} , features lower concentration and more density than the strict MD benchmark \mathbb{Z} , and as such may be closer to the original network \mathbb{X} .

3. Comparing estimated benchmarks with the true network

3.1. The true interbank network

This section assesses the ME and MD solutions against a large real-world network, the German interbank market. Our reference network \mathbb{X} contains the ‘true’ bilateral interbank positions observed for the German banking system. The Deutsche Bundesbank compiles a set of comprehensive banking statistics on large loans and concentrated exposures (Gross- und Millionenkreditstatistik) comprising all positions between financial institutions in the amount of at least 1.5 million euros or 10% of their liable capital. To obtain a consistent and self-contained network, we consolidate banks by ownership at the

bank holding company level (Konzern) and exclude cross-border linkages. We use data from the second quarter of 2003, a choice that avoids the crisis period and issues with backward data revisions.

The resulting interbank network \mathbb{X} is a square matrix with 3.16 million cells containing the observed bilateral interbank exposures among 1779 active banks, amounting to 855 billion euros in total value. The network is sparse, with a density of 0.59% (18 624 active links). It is best described as a core-periphery structure in which most banks do not lend to each other directly but through core banks acting as intermediaries (Craig and von Peter 2014 elaborate).

3.2. The estimated benchmarks

The bilateral counterparty exposures can also be estimated by the two benchmark methods described above, using only information on every bank’s total interbank lending and borrowing (the marginals). The ME solution (\mathbb{E}) yields an almost complete network with 92.8% of potential links being used. This is a density 158 times that of the original network. Since 92% of banks are active both as lenders and borrowers in the interbank market (and thus have positive marginals), ME fails to place any zeros between these pairs. The original structure of linkages is essentially lost as a result, in the way that figure 1 had illustrated.

Our MD approach, at the other extreme, yields a solution (\mathbb{Z}) with density of 0.11% (3450 links). This alternative benchmark thus allocates 100% of interbank volume using just a sixth of the links in the original network. The MD solution also respects all the marginals in equation (2), so that each bank lends and borrows as much in total as in the true interbank network that \mathbb{Z} is trying to match (i.e. $\sum_j Z_{ij} = A_i$ and $\sum_i Z_{ij} = L_i$, so $AD_i = LD_i = 0 \forall i$).

The reason for this efficiency is that the MD algorithm identifies the most probable links and loads them with the largest possible exposures consistent with banks’ total lending and borrowing in the interbank market (using $\lambda = 1$). As a result, the largest links consistently account for more value than was the case in the original network. Figure 2 shows how aggressively \mathbb{Z} concentrates exposures on the largest links in the network, so that the total interbank volume is already fully allocated by the time we reach 0.185, or 18.5% of the number of links in the original network \mathbb{X} . It is apparent that the ME algorithm takes far longer to allocate the total interbank volume, since it places a positive value on virtually all pairs of banks. This illustrates how ME and MD differ in trading off the number vs. the size of interbank linkages.

This obviously affects the most fundamental statistic in network analysis, the *degree* of a node (number of connections). Degree distributions are widely used as a diagnostic tool to characterize different types of networks, e.g. for distinguishing scale-free networks from Erdős–Rényi random graphs. The degree of individual nodes, and the degree distribution of the network as a whole, typically become meaningless after using ME. Whereas banks in the true network have only 11 counterparties on each side of their balance sheet on average, the ME solution \mathbb{E} has nearly all banks lending to each other. This makes its degree distribution basically degenerate, a vertical

[†]We also include a small probability of link deletion to allow the algorithm to cover the entire space of possible network configurations. This ensures that the state space is ergodic (see Proposition 1 in Anand et al. 2012).

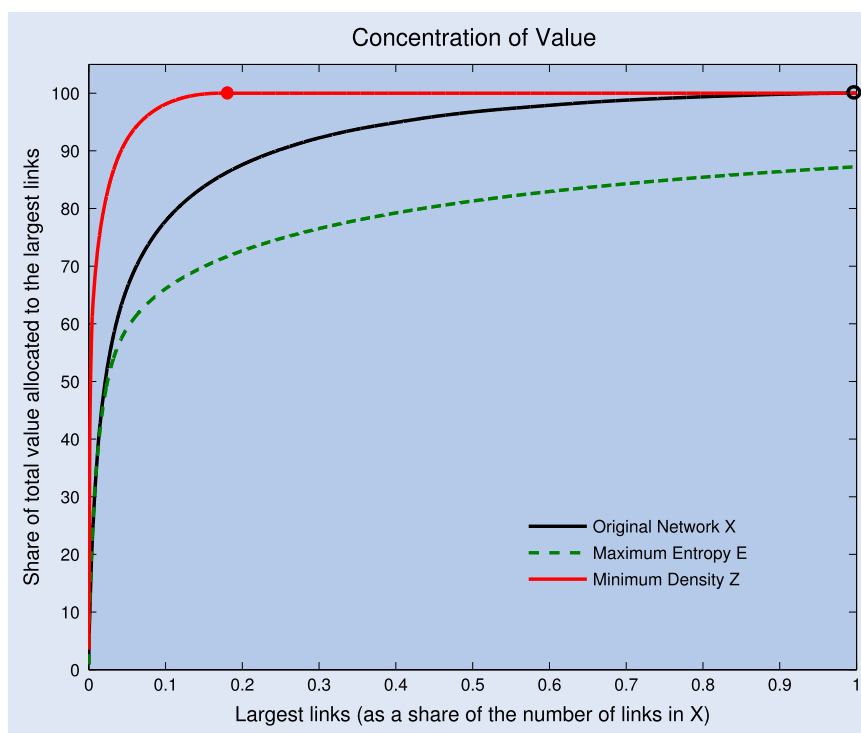


Figure 2. The figure shows the concentration of value on the largest links for the different networks. The x -axis ranks bilateral linkages (in descending order of size) and expresses the first n links as a share of the total number of links in the original network \mathbb{X} (18 624). The y -axis shows the cumulative share of value allocated to the largest n links, relative to the total interbank volume. The dots indicate at which point 100% of volume has been reached. For \mathbb{X} this is at unity, for \mathbb{Z} this occurs at 0.185, whereas \mathbb{E} needs 158 times the number of links in \mathbb{X} before reaching 100% of interbank volume.

line in figure 3. The MD solution \mathbb{Z} errs on the low side, but also assigns a high degree to the most important banks in the market. As a result, its degree distribution broadly retains the shape of the true interbank network \mathbb{X} , up to a factor reflecting its lower density. The mismatch is lessened when considering a less-aggressive solution to the MD problem, a ‘low-density’ solution with density similar to that of \mathbb{X} .

3.3. Comparing network features

A more formal comparison is provided in table 1, where the ME and MD benchmarks are evaluated against the original interbank market on several key network statistics. The end result is that MD preserves key features of the original network better than ME, especially when a low-density solution is selected. The first three lines of the table restate what was already apparent from figures 2 and 3, namely that ME distributes exposures so widely that the estimate \mathbb{E} has orders of magnitude more links than the original network \mathbb{X} . In so doing, ME also distorts a number of other characteristic features.

The fact that interbank relations often involve smaller less-connected banks trading with larger well-connected banks gives rise to negative *assortativity*.[‡] In the original interbank network, assortativity is clearly negative (−0.53). Similar results were found for the federal funds market (Bech and Atalay 2010) and for the Italian interbank market (Iori et al. 2008). Disassortativity is nearly as strong for \mathbb{Z} , hence the MD solu-

tion preserves this important characteristic of interbank markets.[‡]

The focus on a few select counterparties comes out most clearly in the *dependence* measures listed next in table 1.§ Under MD, the largest linkage on average accounts for close to 100% of a bank’s total interbank borrowing. This reflects the efficiency with which the MD algorithm identifies counterparties that can single-handedly satisfy the funding needs of smaller banks. In the original network, high dependence on a single counterparty is also quite frequent when borrowing, though less so when lending. Since the ME solution spreads links indiscriminately across counterparties, it preserves very little of this feature—only to the extent that banks with large L_i borrow more from banks with large A_i .

One network feature that appears to be lost under MD is *local clustering*, the propensity of nodes to form cliques. The local clustering coefficient averages the probabilities that two neighbors of a node are themselves connected (Jackson 2008).¶ In matching big lenders with small borrowers and vice versa, MD tends to generate star-like networks where the extent of clustering is low. Indeed, it is part of the efficiency of the MD solution that transitive relationships are replaced by a single

[‡]The assortativity coefficient is the Pearson correlation coefficient of degree between pairs of linked nodes.

[‡]This is not the case by construction. The assortativity built into the link prediction model related banks with large or small surplus or deficits, which need not coincide with high or low degree.

[§]Our dependence measure calculates what share of each bank’s total borrowing (or lending, respectively) is transacted with its single largest counterparty, and averages this ratio across all active banks.

[¶]The measure has been computed with MIT’s Matlab Tools for Network Analysis available at: <http://strategic.mit.edu/downloads.php?page=matlabnetworks>.

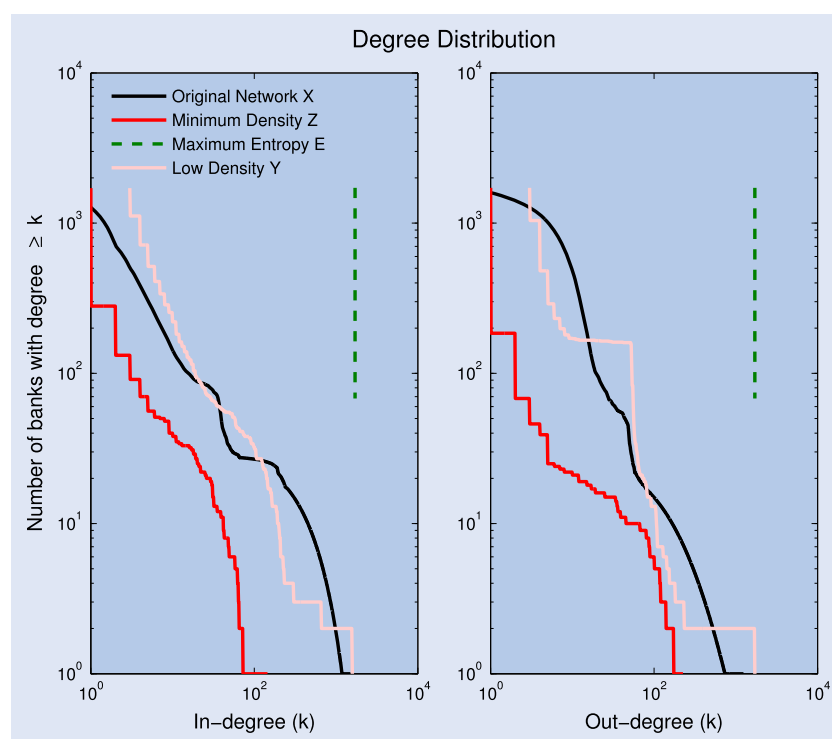


Figure 3. The figure displays the degree distribution in its cumulative form, showing the number of banks with a degree greater than the number shown on the x -axis, on a double log scale. A straight line would indicate a Pareto cumulative indicative of a power law distribution. The degree distribution of the original network \mathbb{X} has been smoothed to preserve the confidentiality of individual bank data, and shows averages at the end points, instead.

Table 1. Comparing basic network features of benchmark estimates with those of the original German interbank network.

Network Characteristic	\mathbb{E} Max entropy	\mathbb{X} True network	\mathbb{Z} Min density	\mathbb{Y} Low density
Density, in %	92.8	0.59	0.11	0.61
Degree (average)	1649	10.5	1.94	10.9
Degree (median)	1710	6	1	4
Assortativity	-0.03	-0.53	-0.40	-0.32
Dependence when borrowing, %	12.2	84.7	97.3	93.4
Dependence when lending, %	7.2	45.1	97.4	87.2
Clustering local average, %	99.8	33.4	0.03	7.62
Core size, % banks	92.6	2.5	1.1	2.1
Error score, % links	14.6	9.2	41.2	35.7

link where possible, obviating the need for smaller banks to form additional local relationships for their lending or funding needs. At the other extreme, clustering under ME trivially equals 100% among active banks. Therefore, both ME and MD fail to preserve local clustering.

The final rows of table 1 consider the extent to which the different networks exhibit *tiering*.[†] The German interbank network is highly tiered, with only 9.2% of network links inconsistent with a perfectly tiered structure (error score). Indeed, 98% of interbank volume has one of the 44 core banks (2.5% of banks in the system) on either side of the transaction. The MD

network retains this structure, even if its core is less than half of the original size because exposures are concentrated on fewer links. Since core banks continue to be linked to each other, however, the interbank network remains a single market in the sense of a giant connected component. In the ME network, on the other hand, the core-periphery structure largely disappears: the core trivially includes all active banks (more than 90% of the banks in the system), leaving only inactive banks in the periphery. The hierarchical structure of interbank networks is largely erased when applying ME.

We briefly elaborate the point that our method also allows one to derive solutions that are less aggressive than the MD benchmark \mathbb{Z} . These low-density solutions, \mathbb{Y} are obtained by lowering $\lambda < 1$, the extent to which selected links are loaded in $\lambda \times \min\{AD_i, LD_j\}$. Figure 4 illustrates one way of doing so, in order to show how selected network properties can be brought closer to those of the original network \mathbb{X} . In particular, it is possible to raise density monotonically (thick solid line) simply by varying the parameter λ , or the time at which it is set equal to one in the algorithm. The network \mathbb{Y} with density similar to \mathbb{X} also resembles \mathbb{X} in terms of other network features (as in table 1). This can be useful provided the researcher estimating counterparty exposures has information about the density of the interbank market \mathbb{X} , or the average number of links that banks maintain. However, there are many ways in which departures from the MD benchmark can be specified and programmed, leading to different results—so we do not pursue this refinement of MD further in what follows.

We conclude that the MD approach preserves some of the original network's structural features better than the ME benchmark does. This makes the MD approach a promising

[†]An interbank market is tiered when few banks intermediate between other banks that do not transact directly with each other; how close an interbank market is to such a core-periphery structure can be measured using block modelling methods (Craig and von Peter 2014).

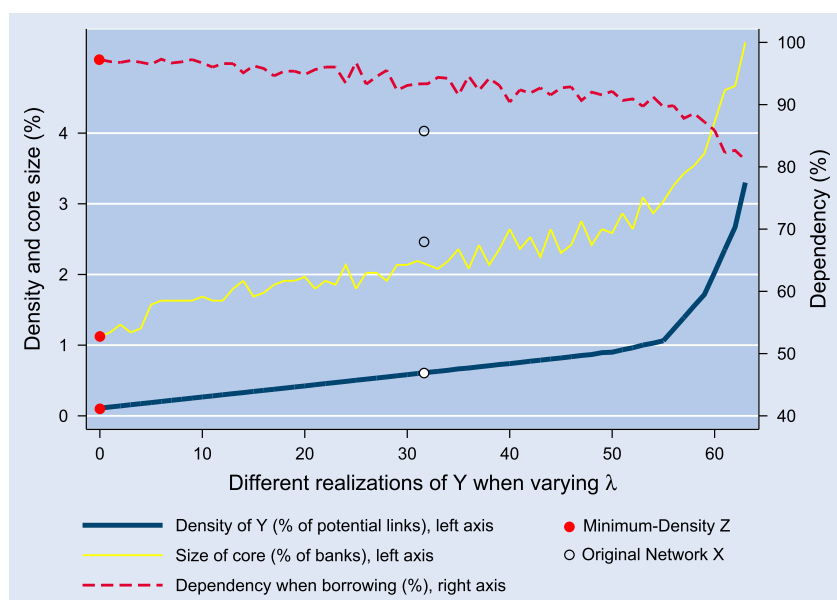


Figure 4. This figure shows three network features for 65 different low-density solutions \mathbb{Y} . The implementation here sets $\lambda = 0.5$ for the first k links being filled by the algorithm and $\lambda = 1$ thereafter, with k raised from 0 to 100 000 in 65 (unequally spaced) steps. The first realization (at $k = 0$) is the MD network \mathbb{Z} with the network features shown as red dots (as in table 1). The black circles indicate the values for the original network \mathbb{X} , plotted at the point where a comparable low-density network \mathbb{Y} reaches a density similar to \mathbb{X} (at $k = 16000$).

alternative to ME for filling in missing counterparty exposures, especially if the estimated matrix is used in network analysis where the pattern of linkages is of central interest. The next section explores an important application where both the pattern and size of linkages matter.

4. Performance in system stress tests

In this section, we contrast the results from a systemic stress test on the German banking system with those obtained using the alternative estimated networks. We opt for a standard simulation methodology, since the focus here is on variation in the simulation *inputs*, not on breaking new ground in stress methodology. Three building blocks in any stress-testing exercise are (i) the trigger event, (ii) the contagion mechanism and (iii) data on bank balance sheets and counterparty exposures, where we consider the estimated networks using maximum entropy and MD, alongside the ‘true’ observed interbank network.

4.1. Stress-test ingredients

Regarding the trigger event, the majority of studies focus on the unanticipated failure of individual banks (Mistrulli 2011, Upper 2011), and we follow this approach here. There are many ways to design a stress test in which banks hold common exposure to a variety of factors, such as stock market declines, credit risk, etc. Since all of these approaches cause losses to all banks simultaneously, we take a shortcut and directly hit their capital ratios. Our experiment combines a single failure with an across-the-board decline of regulatory capital as a percentage of risk-weighted assets by 2 percentage points. Banks go into default when their regulatory capital dips below a common regulatory requirement of 6%.

We run four sets of simulations, one for each interbank network separately, using ME (\mathbb{E}) and MD (\mathbb{Z}), as well as a low density (\mathbb{Y}) solution at $k=16000$, and compare the results to those from the true interbank market (\mathbb{X}). Each set of simulations proceeds as follows: in each run, we let a single bank i fail exogenously to trigger the contagion process, and solve for the clearing vector to obtain (a) the number and identity of banks that default as a result of contagion (excluding the initial failure) and (b) the total assets and interbank liabilities of the defaulting banks. From these 1779 runs (one for each bank), we report the average number of contagious defaults and total bank assets affected by contagion (as in Mistrulli 2011).

We use two methods for simulating how losses cascade through the banking system. The first method is the sequential default algorithm traditionally used in the stress-testing literature (including Furfine 1999, Mistrulli 2011). It assigns a deadweight loss proportional to loss-given-default (LGD) after each default to each of the surviving holders of an asset, where their capital is reduced by the exposure times the LGD. While easy to implement, this method ignores the fact that subsequent defaults induce additional losses to banks that had already defaulted in earlier rounds of the process. Hence we also employ the clearing vector methodology of Eisenberg and Noe (2001), which has gained traction in recent years (as discussed in Upper 2011, Elsinger et al. 2013). The Eisenberg–Noe methodology is internally consistent: it determines the fixed point at which contagion comes to a halt, and in solving the simultaneity problem also delivers a unique solution. However, in determining the system’s LGD endogenously, the Eisenberg–Noe methodology assumes that all assets can be liquidated at book value to meet the liabilities of defaulting banks—the more contagion, the less tenable this assumption becomes. To capture possible distress selling and bankruptcy costs, we allow for an additional deadweight loss of β percent assessed on the liabilities of defaulting banks, and step up its

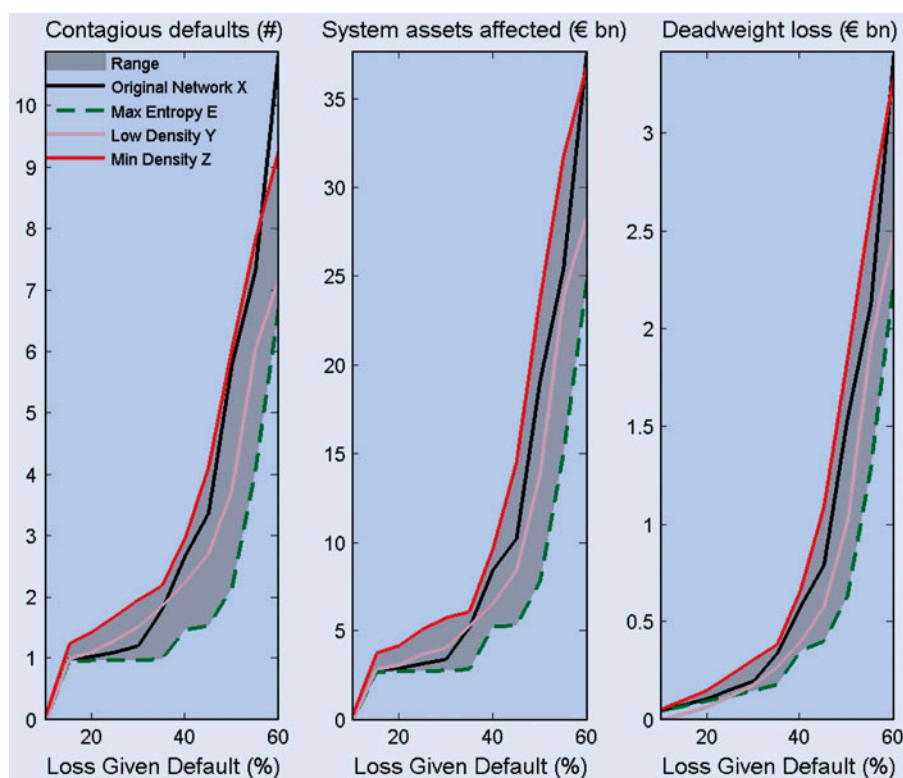


Figure 5. Results of the first stress test using the *sequential default algorithm*. The three lines compare the different average results obtained by using the ‘true’ interbank network (black) and the ME (green dashed), MD (red) and low-density (light red) solutions. The stress tests are run separately for the different levels of bankruptcy costs shown on the x -axis. The left panel plots the number of banks that default as a result of contagion (excluding the initial failure). The center panel expresses the extent of contagion by aggregating the total assets of the banks that end up in default. The right panel displays the deadweight loss arising from distress selling and bankruptcy costs assessed on the liabilities of defaulting banks according to a LGD. In all cases, individual bank failures are accompanied by a decline in the regulatory capital ratio by 2 percentage points at all banks, and banks default when their regulatory capital ratio falls below 6%. The losses are passed on to the other banks in the interbank market through a straight proportional LGD.

value from 0 to 35% in separate simulations (where $\beta = 0$ leads to the original Eisenberg–Noe fixed point). This generalizes the results and allows one to compute the overall liquidation and bankruptcy cost to the system as a whole. This is a true deadweight loss, net of all the redistributions that occur through the contagion process.

4.2. Results from the sequential default algorithm

Figure 5 summarizes the results of the standard stress test with single bank failures and an exogenous LGD. We first describe the reference results for the German interbank network (black bold lines), before turning to those for the estimated benchmarks. As LGD increases (horizontal axes), so do the number of defaults, the value of affected system assets, and the deadweight loss. The steepness after LGD of 30% is in line with earlier findings that the severity of contagion jumps once LGD exceeds 40% (Upper and Worm 2004). The simulations suggest that about one bank defaults on average as a result of the failure of an arbitrary single bank, or 2–3 banks once LGDs exceed a few percent (left panel). This low average is largely due to the fact that the results reported are averaged over 1779 banks. In contrast with the average, the worst case produces more than 1200 contagious defaults (70% of banks in the system). It is a well-known feature of complex networks with highly connected hubs that they are robust to random

failure but vulnerable to targeted attack (Albert *et al.* 2000). For the vast majority of individual bank failures, contagion does not result, in that the initial bank failure did not cause any other bank to default. If one were to look at the number of banks failing *conditional* on contagion occurring, the numbers on the left-hand axis increase by a factor of 100 (for low LGDs) down to a factor of 50 (for high LGDs): for high LGDs, more than 500 banks default on average.[†]

When the same stress test is run on the counterparty exposures estimated by ME, the results *underestimate* the extent of contagion (figure 5, green dashed lines). The entropy solution spreads exposures so widely that more small banks fail at low LGDs, if contagion occurs, but they will incur smaller losses on the system. The entropy solution seems to optimize a network in a direction that makes contagion happen more frequently (and with small shocks to the system), but mostly involves smaller banks that rarely bring down the entire system.[‡]

The MD solution \mathbb{Z} (as well as most variants of \mathbb{Y}) almost always *overestimates* the true extent of contagion (red and light red lines). The bias is small for the number of banks,

[†]At the highest LGDs, 29 cases produce more than 3 contagious defaults, including 19 with more than 100 banks, and 14 of these cause more than 1200 defaults.

[‡]At the highest LGDs, the maximum number of banks failing is less than 1160, which happens in only nine cases. Indeed, in the 35 cases with contagion, only the top nine cases involve contagion affecting more than three banks.

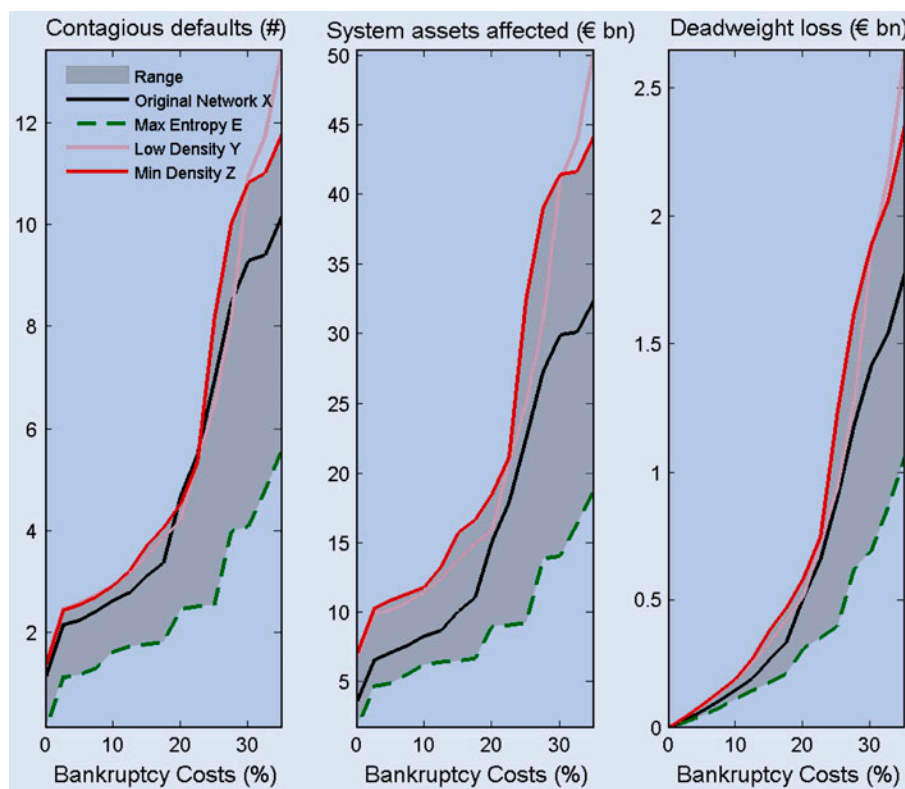


Figure 6. Results of the second stress-test using the *Eisenberg–Noe clearing vector methodology*, with an additional bankruptcy shown on the x -axis. The left and center panels are analogous to figure 5. The right panel shows the deadweight loss associated with the bankruptcy costs passed on to all of the banks of the system (including defaulting banks).

but somewhat greater when measuring contagion by ‘system assets affected’, i.e. the total assets of defaulting banks (figure 5, center panel). The same ranking holds for the system-wide deadweight loss (right panel). *Conditional* on contagion, more banks fail in network \mathbb{Z} than in the true interbank network \mathbb{X} , for most of the LGD range.[†] This illustrates a fairly general property of the MD solution: it produces more contagion than both the original network and the entropy estimate. However, for the highest LGDs representing extreme stress, the more diversified entropy estimate \mathbb{E} lies far below the original \mathbb{X} result, whereas our \mathbb{Z} estimate tracks it quite closely while slightly overstating contagion. The fact that our \mathbb{Z} has fewer links actually reduces some of the banks’ exposure to contagion.[‡] Moderate contagion is more likely to happen, but there is sometimes a buffer from complete system collapse, although this buffer is small.

As a general rule, our MD solution \mathbb{Z} provides an upper bound on system-wide stress, whereas the entropy estimate \mathbb{E} yields a lower bound, at least for the standard contagion exercises we consider here. These two bounds are illustrated in the figures by the grey shaded area denoted ‘Range’. Only for extreme scenarios with LGDs approaching 60% does the original network move outside the range defined by these two

benchmarks. The \mathbb{Z} (and the \mathbb{Y} variants) that we tested provided an effective upper bound for system-wide stress.

4.3. Results from the clearing vector methodology

Figure 6 shows that the Eisenberg–Noe clearing vector methodology produces similar stress-test results, with the same qualitative character of the results from the sequential default algorithm. In particular, ME underestimates systemic stress substantially, while MD overestimates the true extent of contagion: this again gives rise to a useful range that contains the estimate of systemic risk in the true interbank market. The exposures ME places between the major banks are not as large as the true exposures in the interbank market or in our MD estimate (recall figure 2). For this reason, our \mathbb{Z} generally overestimates contagion, providing an upper bound for virtually all stress-test results.[§] The intermediate results for our low-density solution \mathbb{Y} sometimes mitigate this bias—the results are close to the estimate of systemic risk in the true interbank network.^{||}

[†]The gentler stress tests with 10% LGD show contagion with 72 banks under \mathbb{Z} , compared to 6 for \mathbb{X} , and only 4 with the widely diversified entropy estimate \mathbb{E} .

[‡]All cases cause less than 1000 defaults in the most extreme scenarios, whereas 13 cause over 900 defaults, and 99 cause over 3 defaults.

[§]There is a small region where the number of banks affected under \mathbb{Z} is slightly smaller than the number in the original network.

^{||}However, as is true of several \mathbb{Y} matrices we computed, the results were not uniform. While \mathbb{Y} often comes close to the results for \mathbb{X} for some range of bankruptcy costs, it can also diverge in the highest range, as shown in figure 6.

4.4. Interpretation

Our findings suggest that the extent of contagion in the true German interbank market lies between the estimates obtained for the MD solution (lower bound) and the ME benchmark (upper bound) for both the sequential default and clearing vector methodologies. In Mistrulli's direct comparison, the ME solution also underestimates systemic risk at low loss-given-default while overstating it at the higher end. We confirm his first result, and find that it is robust to a change in methodology from the sequential default algorithm to a (generalized) clearing vector method. There are several reasons in our context why there is more contagion under MD than under ME. By minimizing density, the MD network focuses the concentration of exposures onto fewer links; this also means that the greater loss transmitted by a given link is more likely to exceed the capital of the lending bank and thereby cause its default.[†] To some extent, this *concentration effect* is balanced by the fact that the scope of contagion is somewhat limited by the *sparsity effect*: a lower number of linkages in \mathbb{Z} also reduces the number of conduits for the propagation of losses. This countervailing effect does not offset higher concentration owing to the *negative assortativity* in both the MD and the true network. The failure of small banks is largely inconsequential, because they do not cause systemic stress in any of the networks. The stress following a large bank failure, on the other hand, affects many small banks beyond their ability to withstand the losses, and this outcome dominates the averaged results in figures 5 and 6. In our tests, greater diversification (as in the entropy estimate \mathbb{E}) does not raise the severity of a crisis if one occurs—unlike what is predicted by some simulations (Nier *et al.* 2007) or theoretical results on ‘tipping points’ (Gai *et al.* 2011).

Overall, our stress-test findings suggest that the MD approach, while delivering an economically meaningful alternative to ME, also leads to a reasonable estimate of *overall systemic risk* in stress tests. Using both benchmarks helps identify a useful range of possible stress-test results when the true counterparty exposures are unknown (shaded areas in figures 5 and 6). The alternatives can also be compared in terms of the rankings they predict on the systemic importance of individual banks. Consider the 100 largest banks ranked by the damage their failure causes (as estimated from the true interbank network) in terms of contagious failures, total affected assets and other measures of contagion. The MD and ME results generally lead to similar rankings, with Kendall and Spearman rank correlations close to 80% relative to the ‘true’ ranking. This suggests that both benchmarks deliver fairly reliable rankings of *systemically important banks*.

5. Conclusion

The pattern and size of linkages are of central importance in many areas of network analysis and finance. Yet, bilateral

counterparty exposures are often unknown, and ME serves as the leading method for filling in the blanks. This paper proposes an economically meaningful alternative that combines known network features with information-theoretic arguments to produce more realistic interbank networks. Our MD solution preserves some characteristic features of the original interbank market. Moreover, in the context of systemic stress testing, the use of a MD network as input yields an upper bound on the cost of systemic stress where the ME benchmark produces a lower bound. The resulting range contains the true cost of systemic stress of the German banking system, and such a range can be useful in applications where the true counterparty exposures are unknown. Indeed, the MD approach is more suitable for robustness analysis: whereas the ME network is unique, our link prediction method is stochastic and thus capable of generating many low-density networks for repeated application in stress tests. This is a useful feature in a context where networks are not fully observed yet the structure and size of linkages is of great consequence for financial stability.

Acknowledgements

This paper was prepared for the conference ‘Interlinkages and Systemic Risk’ in July 2013 in Ancona. We would like to thank Jean-Cyprien Heam, Iman van Lelyveld, Fabrizio Lillo, Rosario N. Mantegna, Sheri Markose, Kostas Tsatsaronis and Christian Upper, as well as our discussant Fabio Caccioli and two anonymous referees for helpful comments. The comments of seminar participants at the Bank of Canada, the Bank for International Settlements and Deutsche Bundesbank are also gratefully acknowledged. The views expressed in this paper do not necessarily reflect those of the institutions the authors are affiliated with.

References

- Ahn, J., Khandelwal, A. and Wei, S.J., The role of intermediaries in facilitating trade. *J. Int. Econ.*, 2011, **84**(1), 73–85.
- Albert, R., Jeong, H. and Barabasi, A.L., Error and attack tolerance of complex networks. *Nature*, 2000, **406**, 378–382.
- Allen, F. and Gale, D., Financial contagion. *J. Polit. Econ.*, 2000, **108**(1), 1–33.
- Anand, K., Gai, P. and Marsili, M., Rollover risk, network structure and systemic financial crises. *J. Econ. Dyn. Control*, 2012, **36**(8), 1088–1100.
- Bech, M. and Atalay, E., The topology of the federal funds market. *Physica A*, 2010, **389**(22), 5223–5246.
- Blien, U. and Graef, F., Entropy optimization methods for the estimation of tables. In *Classification, Data Analysis, and Data Highways*, edited by I. Balderjahn, R. Mathar and M. Schader, pp. 3–15, 1998 (Springer: Berlin).
- Campbell, J. and O’Kelly, M., Twenty-five years of hub location research. *Transp. Res.*, 2012, **46**(2), 153–169.
- Cocco, J., Gomes, F. and Martins, N., Lending relationships in the interbank market. *J. Financ. Intermed.*, 2009, **18**(1), 24–48.
- Craig, B. and von Peter, G., Interbank tiering and money center banks. *J. Financ. Intermed.*, 2014, **23**(3), 322–347.
- Eisenberg, L. and Noe, T., Systemic risk in financial systems. *Manage. Sci.*, 2001, **47**(2), 236–249.
- Elsinger, H., Lehar, A. and Summer, M., Network models and systemic risk assessment. In *Handbook on Systemic Risk*, edited by J.P. Fouque and J. Langsam, pp. 287–305, 2013 (Cambridge University Press: New York).

[†]The MD solution has the single-minded goal of minimizing the system-wide cost of linkages, but ignores the value of diversifying exposures. Banks, in reality, would individually take this value of diversification into account in order to limit their exposure to systemic risk. In a more general optimization problem, one might find that banks that economize on links also choose to limit their exposure to systemic banks.

- Fang, S.C., Rajasekera, J. and Tsao, J., *Entropy Optimization and Mathematical Programming*, 1997 (Kluwer Academic: Norwell, MA).
- Furfine, C.H., The microstructure of the federal funds market. *Financ. Markets Inst. Instrum.*, 1999, **8**(5), 24–44.
- Gai, P., Haldane, A. and Kapadia, S., Complexity, concentration and contagion. *J. Monetary Econ.*, 2011, **58**(5), 453–470.
- Goyal, S., *Connections – An Introduction to the Economics of Networks*, 2007 (Princeton University Press: Princeton, NJ).
- Hansen, L. and Sargent, T., Robust control and model uncertainty. *Am. Econ. Rev.*, 2001, **91**(2), 60–66.
- Helpman, E., Melitz, M. and Rubinstein, Y., Estimating trade flows: Trading partners and trading volumes. *Q. J. Econ.*, 2008, **123**(2), 441–487.
- Iori, G., de Masi, G., Precup, O., Gabbi, G. and Caldarelli, G., A network analysis of the Italian overnight money market. *J. Econ. Dyn. Control*, 2008, **32**(1), 259–278.
- Jackson, M., *Social and Economic Networks*, 2008 (Princeton University Press: Princeton, NJ).
- Li, D. and Schürhoff, N., Dealer networks. Mimeo, HEC Lausanne, 2012.
- MacKay, D., *Information Theory, Inference, and Learning Algorithms*, 2003 (Cambridge University Press: Cambridge).
- Markose, S., Ginsante, S. and Shaghghi, A.R., ‘Too Interconnected to Fail’ financial network of US CDS market: Topological Fragility and Systemic Risk. *J. Econ. Behav. Organiz.*, 2012, **83**(3), 627–646.
- Mastromatteo, I., Zarinelli, E. and Marsili, M., Reconstruction of financial networks for robust estimation of systemic risk. *J. Stat. Mech.: Theory Exp.*, 2012, P03011.
- Mattsson, K.G. and Weibull, J., Probabilistic choice and procedurally bounded rationality. *Games Econ. Behav.*, 2002, **41**(1), 61–78.
- Mistrulli, P., Assessing financial contagion in the interbank market: Maximum entropy versus observed interbank lending patterns. *J. Bank. Finance*, 2011, **35**(5), 1114–1127.
- Nier, E., Yang, J., Yorulmazer, T. and Alentorn, A., Network models and financial stability. *J. Econ. Dyn. Control*, 2007, **31**(6), 2033–2060.
- Sheldon, G. and Maurer, M., Interbank lending and systemic risk: An empirical analysis for Switzerland. *Swiss J. Econ. Stat.*, 1998, **134**(4), 685–704.
- Strzalecki, T., Axiomatic foundations of multiplier preferences. *Econometrica*, 2011, **79**(1), 47–73.
- Upper, C., Simulation methods to assess the danger of contagion in interbank markets. *J. Financ. Stab.*, 2011, **7**(3), 111–125.
- Upper, C. and Worm, A., Estimating bilateral exposures in the German interbank market: Is there a danger of contagion? *Eur. Econ. Rev.*, 2004, **48**, 827–849.
- Wang, B., Complex network-based intervention and mitigation approaches during financial crisis. PhD Thesis, Nanyang Technological University, 2013.
- Weisbuch, G., Kirman, A. and Herreiner, D., Market organisation and trading relationships. *Econ. J.*, 2000, **110**(463), 411–436.

Appendix 1. MD Pseudo-code

The primitives of our procedure are two Markov processes—one for adding new links and value to the interbank network (with probability $1 - \epsilon$), and the second to delete links, and correspondingly value, from the network (with probability ϵ). The intuition for this design is based on the notion of ergodicity, in that from any network structure, \mathbb{X} , through a finite series of link additions and deletions, we can obtain any other network \mathbb{X}' . As demonstrated in Proposition 1 of Anand et al. (2012), this ensures that the distribution $P(\mathbb{X})$ over network configurations is stationary, and hence our Metropolis algorithm is well defined. As a consequence of being stochastic, our algorithm can generate multiple realizations of MD networks, which allows for robustness in conducting stress-test exercises. Algorithm 1 provides the pseudo-code for our procedure.

begin

(S) : $\mu^{(0)} = \{(i, j)\}_{i,j=1}^N$ and $v^{(0)} = \{\emptyset\}$.

(P) : $Q_{ij}^{(0)} = \max \left\{ \frac{AD_i^{(0)}}{LD_j^{(0)}}, \frac{LD_j^{(0)}}{AD_i^{(0)}} \right\}$, $\forall (i, j) \in \mu^{(0)}$.

(C) : $\tau = 1$.

while

$$V(\mathbb{Z}) < (1 - \epsilon) \times \sum_{i=1}^N AD_i^{(0)}$$

do

(D) : $\rho \in [0, 1]$ at random.

if $\rho < \epsilon$ **then**

Remove link.

(L) :

$(i, j) \in v^{(\tau-1)}$ with probability $1 / |v^{(\tau-1)}|$.

(M) : $AD_i^{(\tau)} = AD_i^{(\tau-1)} + Z_{ij}$ and $LD_j^{(\tau)} = LD_j^{(\tau-1)} + Z_{ij}$.

(A) : $Z_{ij} = 0$.

(S) : $\mu^{(\tau)} = \mu^{(\tau-1)} \cup (i, j)$ and $v^{(\tau)} = v^{(\tau-1)} \setminus (i, j)$.

end

else

Add link.

(P) : $Q_{ij}^{(\tau-1)} =$

$\max \left\{ \frac{AD_i^{(\tau-1)}}{LD_j^{(\tau-1)}}, \frac{LD_j^{(\tau-1)}}{AD_i^{(\tau-1)}} \right\}$, $\forall (i, j) \in \mu^{(\tau-1)}$.

(L) : (i, j) with probability $Q_{ij}^{(\tau-1)}$.

(A) : $Z_{ij} \leftarrow \lambda \times \min \left\{ AD_i^{(\tau-1)}, LD_j^{(\tau-1)} \right\}$.

(A) : $\mathbb{Z}' = \mathbb{Z} + Z_{ij}$.

(D) : $\psi \in [0, 1]$ at random.

if $V(\mathbb{Z}') > V(\mathbb{Z}) \parallel \psi < \exp(\theta [V(\mathbb{Z}') - V(\mathbb{Z})])$ **then**

(A) : $\mathbb{Z} = \mathbb{Z}'$.

(M) :

$AD_i^{(\tau)} = AD_i^{(\tau-1)} - Z_{ij}$ and $LD_j^{(\tau)} =$

$LD_j^{(\tau-1)} - Z_{ij}$. (S) : $\mu^{(\tau)} =$

$\mu^{(\tau-1)} \setminus (i, j)$ and $v^{(\tau)} = v^{(\tau-1)} \cup (i, j)$.

end

end

(P) :

$Q_{ij}^{(\tau)} = \max \left\{ \frac{AD_i^{(\tau)}}{LD_j^{(\tau)}}, \frac{LD_j^{(\tau)}}{AD_i^{(\tau)}} \right\}$, $\forall (i, j) \in \mu^{(\tau)}$.

(C) : $\tau \leftarrow \tau + 1$.

end

end

Algorithm 1: Minimum-Density algorithm to allocate interbank networks. The labels for the different lines are, (S) : update sets, (P) : update priors, (C) : update counter, (D) : draw random number, (L) : pick link, (A) : assign value and (M) : update marginals. The variable $\epsilon \ll 1$ is the allocation error tolerance.