

Soramäki, Kimmo; Cook, Samantha

Article

SinkRank: An algorithm for identifying systemically important banks in payment systems

Economics: The Open-Access, Open-Assessment E-Journal

Provided in Cooperation with:

Kiel Institute for the World Economy – Leibniz Center for Research on Global Economic Challenges

Suggested Citation: Soramäki, Kimmo; Cook, Samantha (2013) : SinkRank: An algorithm for identifying systemically important banks in payment systems, Economics: The Open-Access, Open-Assessment E-Journal, ISSN 1864-6042, Kiel Institute for the World Economy (IfW), Kiel, Vol. 7, Iss. 2013-28, pp. 1-27,
<https://doi.org/10.5018/economics-ejournal.ja.2013-28>

This Version is available at:

<https://hdl.handle.net/10419/77858>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<http://creativecommons.org/licenses/by/3.0/>

SinkRank: An Algorithm for Identifying Systemically Important Banks in Payment Systems

Kimmo Soramäki and Samantha Cook

Abstract

The ability to accurately estimate the extent to which the failure of a bank disrupts the financial system is very valuable for regulators of the financial system. One important part of the financial system is the interbank payment system. This paper develops a robust measure, SinkRank, that accurately predicts the magnitude of disruption caused by the failure of a bank in a payment system and identifies banks most affected by the failure. SinkRank is based on absorbing Markov chains, which are well-suited to model liquidity dynamics in payment systems. Because actual bank failures are rare and the data is not generally publicly available, the authors test the metric by simulating payment networks and inducing failures in them. They test SinkRank on several types of payment networks, including Barabási-Albert types of scale-free networks modeled on the Fedwire system, and find that the failing bank's SinkRank is highly correlated with the resulting disruption in the system overall; moreover, the SinkRank algorithm can identify which individual banks would be most disrupted by a given failure.

Published in Special Issue [Coping with Systemic Risk](#)

JEL C63 E58 G01 G28

Keywords Payment systems; systemic importance; graph theory; Markov distance; absorbing system

Authors

Kimmo Soramäki, ✉ Founder and CEO, Financial Network Analytics,
kimmo@soramaki.net

Samantha Cook, Chief Scientist, Financial Network Analytics

Citation Kimmo Soramäki and Samantha Cook (2013). SinkRank: An Algorithm for Identifying Systemically Important Banks in Payment Systems. *Economics: The Open-Access, Open-Assessment E-Journal*, Vol. 7, 2013-28. <http://dx.doi.org/10.5018/economics-ejournal.ja.2013-28>

1 Introduction

The ability to accurately estimate the extent to which the failure of a bank disrupts the financial system is very valuable for financial regulators. This paper develops a robust measure based on absorbing Markov chains, SinkRank, that accurately predicts the magnitude of disruption caused by the failure of a bank in an interbank payment system and identifies the banks most affected by a failure.

Interbank payment systems provide the backbone for all financial transactions. Virtually all economic activity is facilitated by transfers of claims by financial institutions. In turn, these claim transfers generate payments between banks whenever they are not settled across the books of a single bank. These payments are settled in interbank payment systems. In 2010, the annual value of interbank payments made e.g. in the Pan-European system TARGET2 was \$839 trillion. In the corresponding Fedwire system in the United States, the amount was \$608 trillion - over 40 times its annual GDP (BIS, 2010). Due to the sheer size of the transfers, and their pivotal role in the functioning of financial markets and the implementation of monetary policy, payment systems are central for policymakers and regulators.

Systemic risk in payment systems has been studied since Humphrey (1986) who found significant risk in the U.S. Fedwire system in the mid 1980s. Subsequent studies by Angelini et al. (1996), Bech and Soramäki (2002) and Galos and Soramäki (2005) found the risks to be limited. Since then, however, most payment systems have switched from net settlement to real-time gross settlement (RTGS; Bech et al., 2008), transforming credit risk into liquidity risk as gross settlement eliminates the former at the cost of the latter. Various works have since used simulations to study risks and liquidity needs in RTGS systems, either by creating entirely simulated systems or by introducing changes in data from real payment systems. A growing body of work (Schulz, 2011; Grat-Osinka and Pawliszyn, 2007; Arjani, 2006) uses simulation to study the relationship between liquidity requirements and delays in payment systems. Simulations of failures in payment systems generally focus on system-wide risks and liquidity effects (Glaser and Haene, 2009; McAndrews and Wasilyew, 2005; Ledrut, 2007; Ball and Engert, 2007; Docherty and Wang, 2010). Schmitz and Pühr (2009) studied network structure in payment systems with induced shocks, but found that network

properties were of limited use for stability analysis. Here we use network methods to develop a metric that not only identifies systemically important banks but can also predict the banks most affected by a failure, and validate the metric using simulated payment systems.

The paper is organized as follows. In the next section we discuss existing measures of centrality in network theory and introduce the new centrality metric SinkRank. Section 3 describes the model that is used to simulate bank failures for testing SinkRank and Section 4 presents simulation results that evaluate the accuracy of SinkRank for forecasting the impact of failures and the banks most affected. Section 5 concludes.

Technical details and computer code for reproducing all calculations presented in this paper are given in the Annexes. Interactive versions of the charts are available at www.fna.fi/sinkrank.

2 Centrality in Network Theory

In the past decade, significant progress towards understanding the structure and functioning of complex networks has been made within the fields of statistical mechanics and social network analysis.

A multitude of centrality measures has been developed, each with an explicit or implicit network process in mind. Borgatti (2005) identifies several stylized processes. According to his typology, a process can progress in the network through geodesic paths, paths, trails or walks. Processes that travel via geodesic (shortest) paths are, for example, problems of the type “traveling salesman,” i.e. they always take the shortest route between two nodes. Processes that travel via paths need not necessarily use the shortest route, but do not visit any node more than once. These can be, for example, viral infections (a person becomes immune once infected) or the routing of internet traffic.

Processes that travel along trails do not visit any given link more than once. Such a process is for example the spread of gossip where a person may forward it to several other people or hear the same news from several different people, but a person typically does not hear the same news more than once from the same person. Processes that are characterized as walks are not restricted in their behavior.

An example of such are the money flows studied here, where everyone can pay everyone multiple times.

Further on, Borgatti characterizes processes in the dimensions of parallel duplication, serial duplication and transfers. In parallel duplication the process spreads at the same time from a node to all its neighbors. In serial duplication it duplicates one link at a time. An example of the former is an e-mail broadcast and of the latter viral infection (assuming no one infects multiple people at exactly the same time). Instead, in transfer the process moves something in the network. When it is moved, it leaves the originating node and is now possessed by the receiving node. This is the case with payments.

The most commonly used centrality measures are Degree, Closeness and Betweenness proposed by Freeman (1978) and different variations of Eigenvector centrality which was pioneered by Katz (1953) and Bonacich (1972, 1987).

Degree centrality (or simply Degree) counts the number of neighbors of each node. It is a local measure that only takes the immediate neighborhood of the node into account. It can count neighbors with incoming links, outgoing links or either, and can be weighted by link properties; for example, the weighted out-degree is referred to as out-strength.

The insight underlying *Closeness centrality* is that nodes with shorter geodesic paths to other nodes are more central. It is generally calculated as the average length of geodesic paths from a node to each other node in the network. *Betweenness centrality* defines as central those nodes through which a high share of geodesic paths pass.

What is known today as *Eigenvector centrality* encapsulates the idea that the centrality of a node depends directly on the centrality of the nodes that link to it (or that it links to). Eigenvector centrality measures assume parallel duplication along walks. A famous commercialization of Eigenvector centrality is Google's PageRank algorithm (Page et al., 1999), which adds a random jump probability for 'dangling' nodes and thus allows the measure to be calculated for all types of networks. PageRank and Eigenvector centrality can be thought of as the proportion of time spent visiting each node in an infinite random walk through the network. For calculating Eigenvector centrality, the network must be strongly connected (i.e. the underlying transition matrix must be nonsingular).

In payment networks banks (nodes) transfer payments related to customer requests or their own trading along directed links of the network. When a payment is made the money is no longer available to the sender, and the receiver of the funds can use the funds to make a payment to any other bank in the system. Using the terminology of Borgatti (2005), the *transfer* process takes place along *walks* in the network as any bank can both make payments to and receive payments from any other bank multiple times (assuming the paying bank has sufficient funds or credit).

Payment networks are accompanied with liquidity and risk-management constraints and exhibit feedback loops. Banks may not have enough liquidity to settle a payment or may decide to postpone a payment due to liquidity and risk management concerns. These decisions again depend on the state of the system at that time, and also influence the state of the system. Traditional measures of centrality that have been developed with other types of processes in mind (e.g. processes transmitted along geodesic paths or trails or processes based on duplications instead of transfer) may not be able to accurately identify central nodes in payment systems.

Liquidity constraints may make banks unable to make payments and may alter the unconstrained process significantly. When the constraints are hard, the system may become very unpredictable and be governed by a process of congestion and cascades. When liquidity is scarce, the settlement process loses correlation with the process of payments that would need to be settled. These dynamics are described in Beyeler et al. (2007). Recently-developed centrality measures created for the financial domain include Battiston et al.'s DebtRank (2012) and Craig and von Peter's core-periphery model (2010).

The new centrality measure proposed here, SinkRank, is based on absorbing Markov chains, which are well-suited to model transfers along walks. A Markov system is a system that can be in one of several states, and can pass from one state to another at each time step according to fixed probabilities. If a Markov system is in state i , there is a fixed probability, p_{ij} , of its going into state j at the next time step; p_{ij} is called a transition probability.

An absorbing random walk is a random walk that starts from a node and eventually terminates at an absorbing node. In terms of centrality our interest is the expected number of steps that are taken before termination when the walk

starts from another randomly-chosen node. Absorbing nodes that require a smaller expected number of steps are considered more central than absorbing nodes that require a large number of steps.

Any network can be represented with an adjacency matrix and such a matrix can be turned into a transition matrix. The transition matrix for $M = [s_{ij}]_{n \times n}$ is defined by dividing each element by the row sum, $P = \left[\frac{s_{ij}}{\sum_j s_{ij}} \right]_{n \times n}$, where the transition probabilities for a random walk are defined by the link weights s_{ij} . Here the links represent payments made between banks and the link weights are the payment values.

An absorbing state is a state from which there is a zero probability of exiting. An absorbing Markov system is a Markov system that contains at least one absorbing state, and is such that it is possible to get from each non-absorbing state to each other non-absorbing state and to some absorbing state in one or more time-steps (i.e. the network is strongly connected except for the absorbing states). An absorbing Markov system reflects the process taking place when a bank fails in a payment system: Any payments sent to the failing bank remain in the failing bank's account and don't exit until recovery.

When analyzing an absorbing system, we first number the states so that the absorbing states come last in the matrix. The transition matrix P of an absorbing system is:

$$P = \begin{bmatrix} S & T \\ 0 & I \end{bmatrix}$$

where I is an $m \times m$ identity matrix (m = the number of absorbing states), S is a square $(n - m) \times (n - m)$ matrix (n = total number of states, so $n - m$ = the number of non-absorbing states), 0 is a zero matrix and T is an $(n - m) \times m$ matrix. Here we consider only single failures, i.e. $m = 1$, but, as detailed above, the measure can be easily extended to analyze multiple simultaneous failures as well.

The matrix S gives the transition probabilities for movement among the non-absorbing states. To obtain information about the time to absorption in an absorbing Markov system, we first calculate the fundamental matrix Q .

$$Q = (I - S)^{-1}$$

The i, j th entry of Q defines the number of times, starting in state i , a process is expected to visit state j before absorption. The total number of steps expected before absorption equals the total number of visits a process is expected to make to all the non-absorbing states. This is the sum of all the entries in the i th row of Q . We call this the ‘*Sink Distance*’ of the node.

In calculating SinkRank, we calculate the Sink Distance of each non-absorbing node and take an average. The measure is analogous to distances along paths except that the process is based on the number of steps in walks defined by the transition matrix and ending at the absorbing node. Finally, we invert the average so that larger values of the metric correspond to more central nodes. Thus, our measure of SinkRank is defined by:

$$\text{SinkRank} = \frac{n - m}{\sum_i \sum_j q_{ij}}.$$

Note that Sink Distance can only be calculated when a path exists between the absorbing node and the non-absorbing node being considered. Thus, SinkRank can only be calculated for strongly connected components, and is most useful as a centrality metric for networks that are strongly connected (as payment systems generally are). For networks that are not strongly connected, we can first add a small constant to the zero elements of the transition matrix, equivalent to the random jump probability used in the PageRank algorithm. If we denote the random jump probability as $1 - \alpha$ and the transition matrix as $P = [p_{ij}]_{n \times n}$, adding the random jump probability is equivalent to replacing each element p_{ij} of P with $\alpha p_{ij} + \frac{1-\alpha}{n}$. In other words, each zero element in the transition matrix is replaced with $\frac{1-\alpha}{n}$ and each non-zero element is multiplied by α and added to $\frac{1-\alpha}{n}$. SinkRank can then be calculated for the transition matrix with random jumps included.

SinkRank is an intuitively meaningful metric in a payment system as it can measure how close a failing bank is to the other banks in the system via payment flows. We expect failures to be more disruptive when they occur in banks that are more central, i.e. banks that have higher SinkRank. The SinkRank of a node denotes the inverse average value of payments that need to be made for a unit of liquidity anywhere in the network to reach the node, and takes possible values in the range $(0, 1]$. The maximum possible SinkRank of 1 is obtained, for example, in

the case of the center node of a star network: Payments made by the spokes of the star reach the center in one step.

3 Simulation model of payment system

Because bank failures are rare and the data is not generally publicly available, we test the SinkRank metric by simulating payment networks and inducing failures in them. The simulation model incorporates both liquidity constraints and a queuing mechanism for payments that cannot be settled due to the liquidity constraint. For the payment simulations we use the FNA payment simulator¹ which has previously been used *inter alia* in Berge and Christophersen (2012) and McLafferty and Denbee (2012). Given a payment network, that is, a set of banks and the times and values of payments made between them, the FNA payment simulator records the balance of each bank throughout the payment period, as well as information on payment delays, average balances, total number of payments made and received, and can also simulate bank failures by restricting the outgoing payments made by a specified bank. The payment network itself was generated based on the Barabási-Albert (BA) model (Barabási and Albert, 1999) for scale-free networks. BA networks have a few very highly connected nodes and many more less-connected nodes, a structure that well represents payment systems with a few very central banks and many more peripheral ones. Properties of the BA network used in the simulations are summarized in Table 1² and more details on generating the network can be found in Annex 1.

The payment data used in the simulation is randomly generated as detailed in the Annexes. The generating process is able to produce payment flows with close resemblance to the Fedwire payment network in the United States.

Visualizations of the BA network are shown in Figures 1 and 2; each node (circle) represents a bank and the arcs between them represent payments. Figure 1 shows the entire network. Node sizes are scaled by Out-strength (that is, larger nodes represent banks that make more total payments), and arc width is scaled by the number of payments (that is, thicker arcs represent more payments made). The

¹ See www.fna.fi/solutions/payment-simulator

² Fedwire network as described in Soramäki et al. (2007). Model network is one realization.

Property	Value in simulated BA network
Number of nodes	100
Number of links	1220
Connectivity	0.123
Reciprocity	0.317
Degree (k)	24.4
Max (k -in)	57
Max (k -out)	56
Number of payments	5000
Value paid ('1000)	604

Table 1: Properties of BA network topology

network is characterized by a few large well-connected banks with high centrality and many more small banks, as is typical in scale-free networks.

Figure 2 shows the maximally-connected subgraph of the BA network; that is, the largest subgraph that contains a link between each pair of nodes. The maximally connected subgraph has 17 nodes, which represent the core of the network. The maximally connected subgraph is a subgraph of the strong components of a network: The nodes in the maximally connected subgraph are strongly connected, and meet the additional constraint that each node is directly connected to each other node.

In the simulation model, each bank starts the day with a given opening balance. Payments are tested for settlement as they are fetched from the file of generated payments. If the value of the payment is larger than the available balance of the sending bank, the payment is put in the sending bank's queue of pending payments. If the value is smaller than the available balance, the payment is settled and the account of the sending bank is debited and the account of the receiving bank is credited.

The bank whose account was credited may now be able to settle some of its previously queued payments, if any such payments exist. Queued payments are released on a 'First-in-First-Out' (FIFO) basis. If a payment from the queue can be

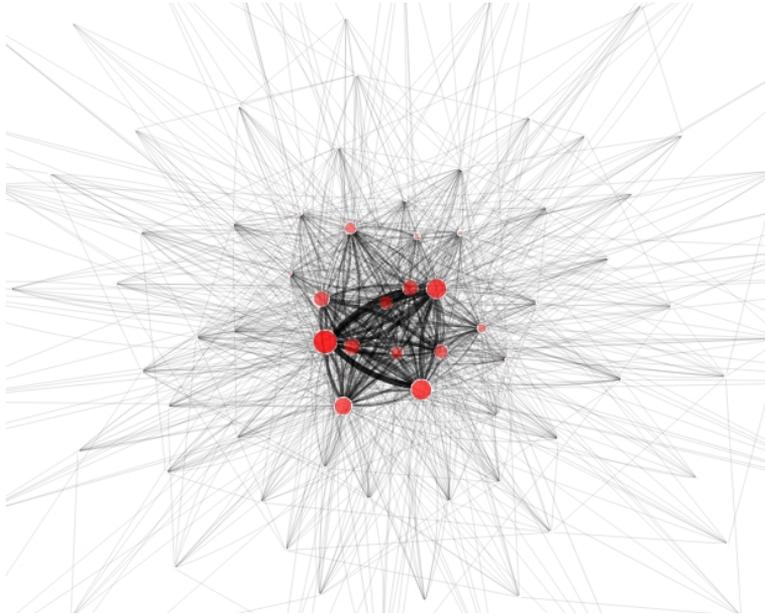


Figure 1: BA network visualization, entire network

settled, the recipient of the newly-released payment may now be able to release its first queued payment - i.e. a single payment can cause the release of many queued payments in a cascade. At the aggregate level this creates a process where the system may become congested, manifesting as an increase in queued payments, and occasionally the queued payments are settled in cascades when payments that can be settled due to incoming funds from previously settled payments are released to others. The behavior of such a system is described in detail in Beyeler et al. (2007).

In the failure simulations we set each bank in turn to be unable to send any payments during the day; that is, we set each bank in turn to be an absorbing state. The failing bank continues, however, to receive payments and will therefore trap some of the system's total liquidity in its account. As a consequence other banks will run short of liquidity and queues will build, first causing existing liquidity buffers to be used more and eventually causing payments to be delayed. The

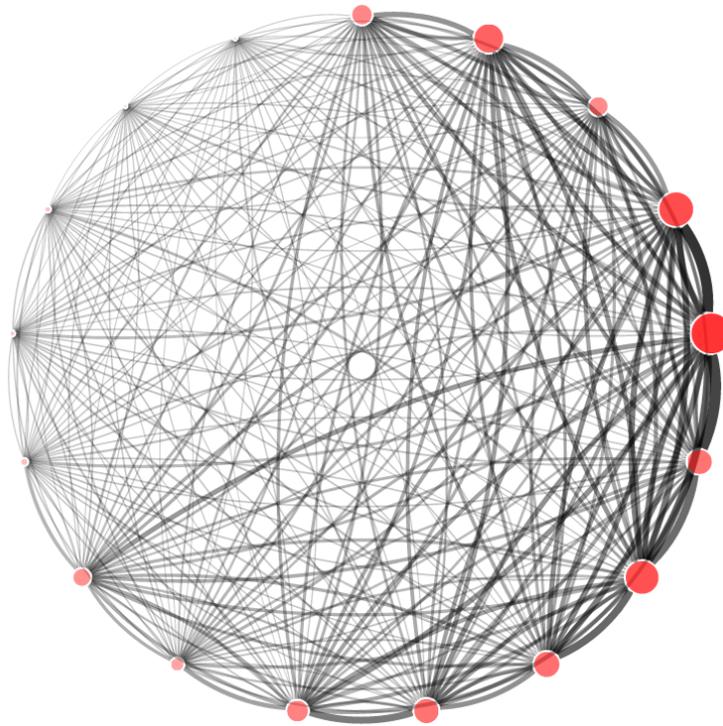


Figure 2: BA network visualization, maximally connected subgraph

assumption that a failing bank continues to receive payments is accurate at least in the short run, as there is certainly some delay between the time when a failed bank stops making payments and the time when other banks in the system learn of the failure and potentially change their behavior. Moreover, banks in a payment system have obligations to make their payments on time and may be sanctioned or face other consequences if they fail to do so. Even in the face of a known bank failure, contractual requirements and industry throughput guidelines should encourage banks to continue making payments to all banks in the system, including any distressed or failed banks. The FNA code to replicate these simulations can be found in Annex II.

We calculate duration of delays in the system aggregated over all banks (‘Congestion’) and the average reduction in available funds of the other banks due to the failing bank, (‘Liquidity Dislocation’). We use their duration-weighted sum as a measure of the extent of the disruption caused by the failing bank; we refer to this measure as ‘Impact.’

The magnitude of the Impact of a bank failure is dependent on the level of liquidity in the system. If other banks have enough liquidity to offset the funds that did not arrive from the failing bank, no delays will occur. In the trivial case of unlimited liquidity, no Congestion would ever occur and each bank’s Liquidity Dislocation would be equal to the amount of payments not received from the failing bank.

In the simulations we set the initial balance of each bank at the minimum level that allows all banks to process all payments immediately when no bank failure is present. Thus, when a failure occurs, Congestion will be caused by lack of sufficient liquidity in at least some banks, which in turn will cause Liquidity Dislocation and/or Congestion at other banks.

Figure 3 summarizes the two disruption measures considered and the relationship between them. Each point represents a single failed bank and shows the Liquidity Dislocation and Congestion calculated for all other banks in the network. All bank failures cause at least some Liquidity Dislocation, whereas Congestion only occurs in about half (62 / 100) of the bank failures; if all banks affected by a failure have enough liquidity to make their payments, no delay will occur. The relationship between Liquidity Dislocation and Congestion is convex as theoretically shown in Galbiati and Soramäki (2011). As more liquidity is dislocated, more delays occur that dislocate more liquidity.

4 SinkRank and Failure Distance

For each bank in the network we calculate its SinkRank (as described in Section 2), Out-Strength (that is, the sum of all its outgoing payments - a measure of the size of the bank), and PageRank. These centrality measures are related to the Impact (that is, the duration-weighted sum of Congestion and Liquidity Dislocation) experienced in the simulation. Figure 4 shows that SinkRank, Out-strength, and

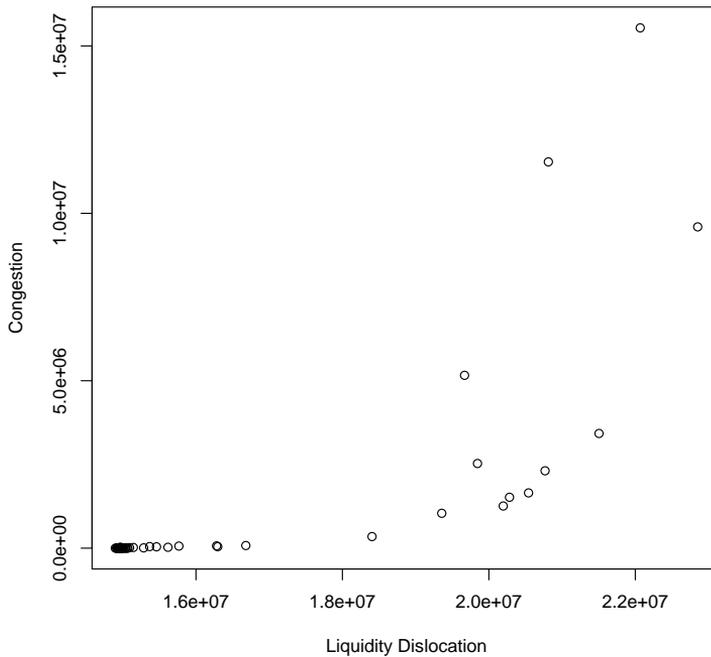


Figure 3: Network-level disruption measures

PageRank are all very strongly related ($r > 0.99$) to Impact in BA networks. Note that different transformations of Impact are shown for each centrality metric, chosen so as to make the relationships between Impact and centrality as linear as possible. SinkRank is most linearly related to Impact on the inverse scale, so its correlation has the opposite sign.

Figure 5 shows the relationship between centrality and Impact in other types of payment network. We consider random and complete networks of the same size (number of links) as the BA network, with link weights assigned randomly and payments generated as in Annex I. The average link weights were set such that the total value of payments in the system is approximately the same across networks.

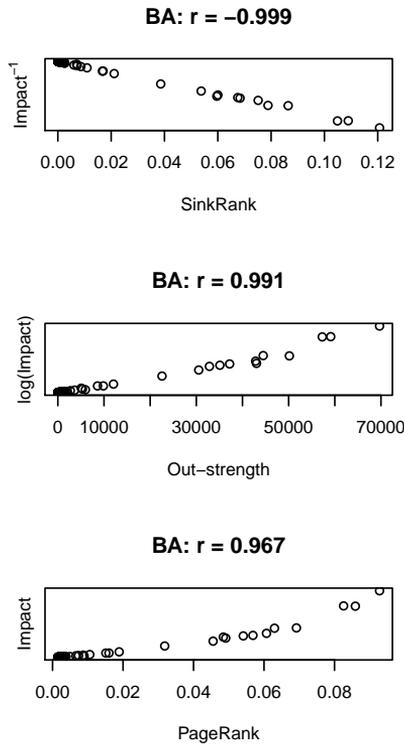


Figure 4: Relationship between Impact and centrality measures in BA networks

SinkRank and PageRank are both strongly related to Impact in all three networks, whereas the relationship between Out-strength and Impact appears to hold only in the BA network which has strong correlations with strength and degree of nodes.

The results in Figures 4 and 5 are for aggregate network properties: A bank's centrality is strongly related to the *overall* impact seen in the system if that bank fails. We can further utilize the SinkRank technology to identify which individual banks are most susceptible to disruption in the case of a bank failure. We define the *Failure Distance* as the Sink Distance *from* a failing bank *to* any other bank. To

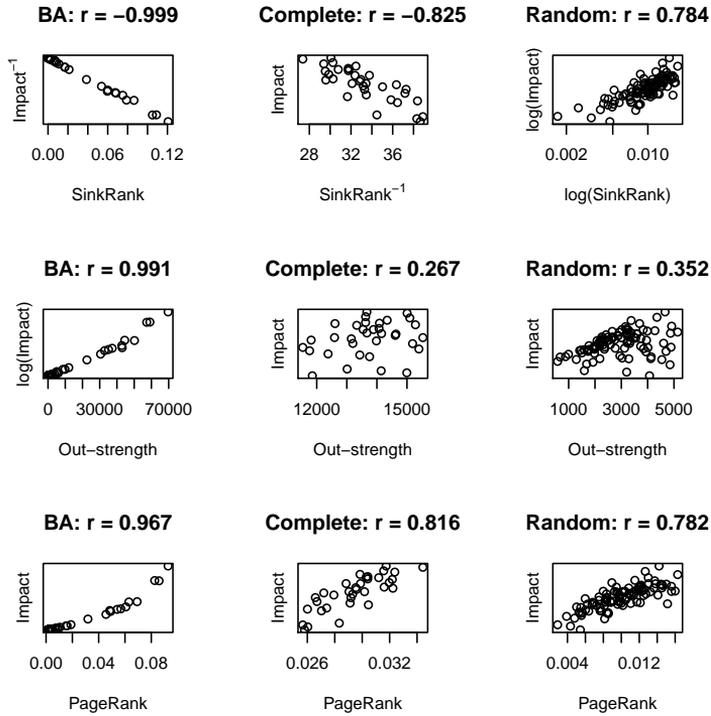


Figure 5: Relationship between Impact and centrality measures in BA, complete, and random networks

calculate Failure Distance, we arrange the system's transition matrix so that the bank whose Failure Distance is being calculated comes last:

$$P' = \begin{bmatrix} S' & T' \\ 0 & 1 \end{bmatrix}.$$

The failing bank is now part of the submatrix S' . From S' we calculate the corresponding fundamental matrix Q' , and the Failure Distance is the sum of the entries of Q' in the row that corresponds to the failing bank. In other words, if bank i fails,

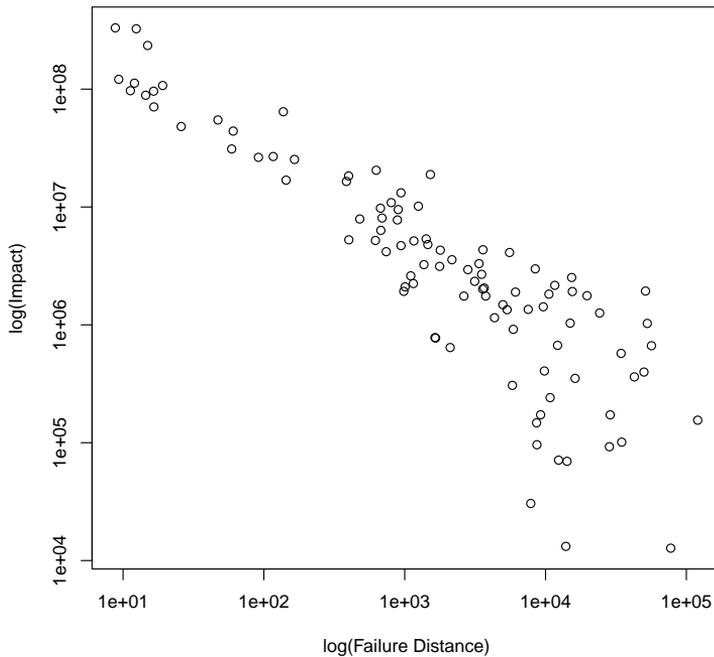


Figure 6: Relationship between Failure Distance and Impact when the most central bank fails

we calculate the Failure Distance of any other bank j by treating bank j as the sink and calculating the Sink Distance of bank i to bank j .

Banks with small Failure Distances are close to the failing bank and downstream from it in the payment chain, and so should be most disrupted by the failure. Figure 6 shows the Failure Distances and Impact for the BA network when the bank with the highest SinkRank (that is, the most central bank) fails. Banks with smaller Failure Distances indeed exhibit larger disruptions, and the relationship is quite strong ($r < -0.85$).

5 Conclusions

This paper developed the new metric SinkRank based on absorbing Markov chains and evaluated its accuracy by comparing it with results from simulated failure scenarios in payment systems modeled after the Fedwire system. This initial analysis has shown that it is possible to accurately rank banks on the basis of metrics calculated from network topology to estimate the potential disruption their failure would cause in the payment system: SinkRank was shown to be predictive of network-level disruption in the case of a bank failure. In addition, the related metric Failure Distance was shown to be predictive of the impact on individual banks in BA networks.

Several possibilities exist for extending the work. First, a more robust analysis with regression models to investigate the explanatory power of different metrics or combined metrics could be carried out. A longer time series of different realizations of the networks and failure simulations would also make the results more robust. More simulations on alternative network topologies with longer path lengths and different correlations among network topology and link values could provide better information on the relative merits of the different metrics across network topologies. These networks could be artificial (lattice, random, etc.) or constructed from real payment data. It may also be possible to further improve the SinkRank metric by taking into account the liquidity distribution at the time of failure.

It is impossible to confirm any metric as the “best” predictor of system disruption, because the appropriate historical data are not available for testing: We can only study simulated systems, and simulations are never 100% representative of reality. However, we can construct payment systems that closely mimic reality and have found that the SinkRank metric performs well in these and other systems. Moreover, SinkRank is theoretically sound, as its calculation mimics the flow of payments in a payment system. The simulation results combined with its strong theoretical underpinnings suggest that SinkRank truly is a more useful centrality measure for banks in payment systems than those others considered here.

Acknowledgements: The authors thank Tor Berge, Kristian Dupont, and Miklos Kalozsi for their valuable contributions to this paper.

Annex I: Algorithm for generating payment data

Payment networks exhibit complex properties. We take as a starting point the Fedwire network which consists of almost 8000 banks processing over 411.000 payments on an average day. The network is described in detail in Soramäki et al. (2007). Due to the highly confidential nature of the data, it is rarely available for research outside central banks and therefore artificial data needs to be used.

There are three main aspects in describing the payment network: the structure of the links, link weight distributions, and individual payment distributions; in other words, who pays whom, how often, and how much. Both link weights and payment values have also correlations with each other. Generating a mechanism that produces all desired aspects of the data is thus challenging.

The main structural characteristic of the network is a power law degree distribution. This means that a few very large banks connect to a large number of very small banks. The in- and out-degrees correlate strongly, i.e. banks that receive payments from many different banks also send payments to many different banks, and vice versa. The largest degree in the Fedwire network on an average day is 1922 for incoming links and 2097 for outgoing links. The network also has a very low connectivity. Only 0.3% of all possible links are present on an average day. In addition, the links have a very high reciprocity of 0.22. This means that 22% of relationships between two nodes are bidirectional - if a link exists from A to B, then a link also exists from B to A. Reciprocity in a random network is on average equal to its connectivity, i.e. over 70 times smaller in this case.

The link weights (number of payments) also follow a power law distribution and have a very high positive correlation with the degree of the node. This means that large banks have both more links and that these links transmit more payments than links of smaller banks. The number of payments in reciprocal links also has a high correlation, denoting strong bi-directional business relationships between banks.

The payment values have a lognormal distribution, and again their value depends on the size of the banks, measured as the number of counterparties or the total value sent (i.e. out-or in-strength).

We develop a simple payment generation process extending the BA model by Barabási and Albert (1999) for generating random scale-free networks. The BA

model is based on two processes, growth and preferential attachment. Growth in the model means that initially the network only has a few nodes and nodes are gradually added to the system. Preferential attachment means that the more connected a node is, the more likely it is to receive new links. Newly added nodes are therefore more likely to connect to nodes with many existing links.

The model developed here aims at reproducing the main statistical properties described for the Fedwire network above. The model applies growth and preferential attachment as the main drivers of the generation process, but instead of adding links, it adds payments. A link is formed when the first payment is drawn from a bank to another. Additional payments between banks with existing payments add to the weight of the link.

We start with an initial number of nodes n_0 . We then draw new payments one by one, m payments for each new node, until we have the desired number of nodes, n . We use a vector $H[h_i]_{i=1,\dots,n}$ to track the amount of preferential attachment strength that has been allocated to each bank (node) i and a matrix $M = [s_{ij}]_{n \times n}$ to track the number of payments created to and from each bank. The matrix M can also be interpreted as a weighted adjacency matrix of the payment network, where the weight is the number of payments.

The other main difference from the BA model is the addition of a parameter α that denotes the “strength of preferential attachment”, i.e. how much is added to h when a payment is sent or received. In the BA model, $\alpha = 1$; because the number of payments here is vastly higher than the number of links to draw, the addition to the preferential attachment must be smaller so as not to skew the degree distribution too much. The pseudo-code for the algorithm is given below.

```

FOR  $i = 1, \dots, n_0$  (add initial banks/nodes)
    SET  $h_i = 1$ 
FOR  $k = n_0 + 1, \dots, n$  (banks)
    FOR  $l = 1, \dots, m$  (average number of payments per bank)
        SELECT random sender  $i^*$  such that bank  $i$  has the probability
         $\frac{h_i}{\sum h_i}$  of being selected as a sender
        SET  $h_i^* = h_i^* + \alpha$  (update preferential attachment strength)
    
```

```

WHILE  $i^* \neq j^*$  (exclude loops, payments to oneself)
    SELECT random receiver  $j^*$  such that bank  $j$  has the probability of  $\frac{h_j}{\sum h_j}$  of being selected as recipient of the payment
    SET  $h_j^* = h_j + \alpha$  (update preferential attachment strength)
    SET  $s_{ij} = s_{ij} + 1$  (create payment/link)

SET  $h_{m_0+k} = 1$  (create new bank/node)
    
```

Table 2 summarizes the comparison.³ The model seems to be able to reproduce the main characteristics of the Fedwire topology very well with $n_0 = 10$ and $\alpha = 0.1$. The parameter n_0 determines the number of core banks, and α the slope of the power law co-efficient in the degree distribution. Both the real and the generated network are sparse, with power law degree distributions and high clustering and reciprocity. In and out degrees are highly correlated and the degrees of the largest bank are very similar.

	Fedwire	Model
nodes	5066	5066
links	75397	70710
connectivity	0.003	0.003
reciprocity	0.215	0.213
degree (k)	14.9	14.0
max (k -in)	2097	2210
max (k -out)	1922	2215
payments ('1000)	411	411

Table 2: Properties of BA network topology

The next step after creating the interaction topology and the number of payments each bank sends to each other is to add time of submission and value to each payment. The time for each payment is drawn from a uniform distribution between 08:00 and 17:00 (opening hours of the simulated payment system) and the value is

³ Fedwire network as described in Soramäki et al. (2007). Model network is one realization.

drawn from a normal distribution with a mean of 1 and standard deviation of 0.2. As the individual payment values were log-normally distributed in the real data, we then exponentiate the drawn values. In addition, larger banks interchange larger payments with each other than do smaller banks. We achieve this by scaling the payment values by $Min(k_{sender}, k_{receiver})$.

In the robustness analysis we also consider random networks and complete networks whose link weights (number of payments) are assigned randomly such that the total value of payments is approximately the same across the different network types. Payment values are generated as detailed above. FNA commands for generating and summarizing the networks and carrying out payment simulations are given below.

Annex II: FNA Commands for Reproducing Results

Generating Networks

```
# Generate Barabasi-Albert (BA) network with link weights showing
# number of payments from each node to the other
ba -nv 100 -m 50 -v0 10 -alpha 0.1 -preserve false -seed 123

# Generate Random network and assign each link with a weight
# drawn from a uniform distribution between 1 and 7
random -nv 100 -na 1200 -preserve false -seed 123
calcap -e [?random:uniform:1,7:123?] -saveas number

# Generate Complete network and assign each link with a weight
# drawn from a uniform distribution between 1 and 8
complete -nv 34 -directed -preserve false
calcap -e [?random:uniform:1,8:123?] -saveas number

#
# Creating payment files (one for each of the above networks)
#
# Create one day (8h - 17h) of payments
# Log payments have mean 1 and sd 0.2
createpayments -number number -open 08:00:00 -close 17:00:00
```

```
-mean 1 -stdev 0.2 -saveas network.csv -seed 123
```

Calculating Network Metrics

```
# Calculate SinkRank of each node, weighted by  
# value of payments  
sinkrank -ap value
```

```
# Calculate weighted out-degree (Out-strength)  
degree -p value -direction out -saveas value
```

```
# Calculate PageRank of each node, weighted by  
# value of payments  
pagerank -p value
```

```
# Calculate number of nodes and links in each network  
order -saveas numnodes  
size -saveas numlinks
```

```
# Calculate connectivity and reciprocity  
connectivity  
reciprocity
```

```
# Calculate average reciprocity of each node  
avgvp -p reciprocity -saveas reciprocity
```

```
# Calculate degree and average degree in network  
degree -direction undirected -saveas degree  
avgvp -p degree -saveas degree
```

```
# Calculate in-degree and average in-degree in network  
degree -direction in -saveas indegree  
maxvp -p indegree -saveas maxindegree
```

```
# Calculate out-degree and average out-degree in network  
degree -direction out -saveas outdegree  
maxvp -p outdegree -saveas maxoutdegree
```

```
# Calculate number of payments in each network
sumap -scope network -p number -saveas numpayments
```

```
# Calculate value of payments in each network
sumap -scope network -p value -saveas value
```

Simulating Payments

```
# Calculate starting values so that balances are never negative
rtgs -paymentsfile network.csv[skiplines=1]
    -openingtime 080000 -closingtime 170000
    -outrecords out_records -dateformat yyyy-MM-dd -dateproperty date
calcvp -e -1*min_rtgs_balance -saveas starting_balance
calcvp -e min_rtgs_balance -saveas overdraft_limit
```

```
# Simulate Payment system without failure
rtgs -paymentsfile network.csv[skiplines=1]
    -fund starting_balance
    -openingtime 080000 -closingtime 170000
    -outrecords out_records -outbanks out_banks_BA_success
    -dateformat yyyy-MM-dd -dateproperty date
```

```
# Simulate Payments system where Bank ID 00001 Fails
# (Repeated for each bank in the network)
rtgs -paymentsfile network.csv[skiplines=1]
    -overdraft overdraft_limit
    -openingtime 080000 -closingtime 170000
    -outrecords out_records -outbanks out_banks_BA00001
    -dateformat yyyy-MM-dd -dateproperty date -strickenbank 00001 -capacity 0
```

References

Angelini, P., Maresca, G., and Russo, D. (1996). Systemic risk in the netting system. *Journal of Banking & Finance*, 20(5): 853–868. URL <http://ideas.repec.org/a/eee/jbfina/v20y1996i5p853-868.html>.

- Arjani, N. (2006). Examining the trade-off between settlement delay and intraday liquidity in Canada's LVTS: A simulation approach. Working Papers 06-20, Bank of Canada. URL <http://ideas.repec.org/p/bca/bocawp/06-20.html>.
- Ball, D., and Engert, W. (2007). Unanticipated defaults and losses in Canada's large-value payment system, revisited. Discussion Papers 07-5, Bank of Canada. URL <http://ideas.repec.org/p/bca/bocadp/07-5.html>.
- Barabási, A.-L., and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439): 509–512. URL <http://www.sciencemag.org/content/286/5439/509.abstract>.
- Battiston, S., Puliga, M., Kaushik, R., Tasca, P., and Caldarelli, G. (2012). DebtRank: Too central to fail? Financial networks, the FED and systemic risk. *Scientific Reports (Nature)*, 2. URL <http://www.nature.com/srep/2012/120802/srep00541/full/srep00541.html>.
- Bech, M., Preisig, C., and Soramäki, K. (2008). Global trends in large-value payments. *Federal Reserve Bank of New York Economic Policy Review*, 14(2): 59–81. URL <http://ideas.repec.org/a/fip/fednep/y2008isepp59-81nv.14no.2.html>.
- Bech, M., and Soramäki, K. (2002). Liquidity, gridlocks, and bank failures in large value payment systems. In *E-money and Payment Systems Review*, pages 111–126. Central Banking Publications.
- Berge, T. O., and Christophersen, C. (2012). Operational problems in banks - Effects on the settlement of payments in Norges Bank. *Norges Bank Economic Bulletin (Penger og Kreditt)*, 83: 36–47.
- Beyeler, W., Glass, R., Bech, M., and Soramäki, K. (2007). Congestion and cascades in payment systems. *Physica A: Statistical Mechanics and its Applications*, 384: 693–718.
- BIS (2010). Statistics on payment, clearing and settlement systems in the CPSS countries - Figures for 2010 - preliminary release. Discussion paper, Bank for International Settlements, Committee for Payment and Settlement Systems, Basel. URL <http://www.bis.org/publ/cpss98.htm>.

- Bonacich, P. (1972). Factoring and weighting approaches to status scores and clique identification. *Journal of Mathematical Sociology*, 2(1): 113–120.
- Bonacich, P. (1987). Power and centrality: A family of measures. *American Journal of Sociology*, 92(5): 1170–1182. URL <http://www.jstor.org/stable/2780000>.
- Borgatti, S. P. (2005). Centrality and network flow. *Social Networks*, 27: 55–71.
- Craig, B., and von Peter, G. (2010). Interbank Tiering and Money Center Banks. Discussion paper 322, BIS Monetary and Economic Department.
- Docherty, P., and Wang, G. (2010). Using Synthetic Data to Evaluate the Impact of RTGS on Systemic Risk in the Australian Payments System. *Journal of Financial Stability*, 6(2): 103–117. URL <http://ideas.repec.org/a/eee/finsta/v6y2010i2p103-117.html>.
- Freeman, L. (1978). Centrality in social networks: I. Conceptual clarification. *Social Networks*, 1: 215–239.
- Galbiati, M., and Soramäki, K. (2011). An agent-based model of payment systems. *Journal of Economic Dynamics and Control*, 35(6): 859–875. URL <http://ideas.repec.org/a/eee/dyncon/v35y2011i6p859-875.html>.
- Galos, P., and Soramäki, K. (2005). Systemic risk in alternative payment system designs. Working Paper Series 508, European Central Bank. URL <http://ideas.repec.org/p/ecb/ecbwps/20050508.html>.
- Glaser, M., and Haene, P. (2009). Liquidity effects of a participant-level operational disruption in the Swiss Interbank Clearing System. In H. Leinonen (Ed.), *Simulation analyses and stress testing of payment networks*, Bank of Finland Scientific Monograph E:42, pages 59–82.
- Grat-Osinka, A., and Pawliszyn, M. (2007). Liquidity levels and settlement delays in the Sorbnet system - Simulation-based approach with application of the BOF-PSS2 Payment System Simulator. *Financial Markets and Institutions*, pages 53–66.

- Humphrey, D. (1986). Payments finality and risk of settlement failure. In A. Saunders, and L. White (Eds.), *Technology and the Regulation of Financial Markets*, pages 97–112. Lexington, MA: Heath.
- Katz, L. (1953). A new index derived from sociometric data analysis. *Psychometrika*, 18: 39–43. URL <http://ideas.repec.org/a/spr/psycho/v18y1953i1p39-43.html>.
- Ledrut, E. (2007). How can banks control their exposure to a failing participant? In H. Leinonen (Ed.), *Simulation studies of liquidity needs, risks and efficiency in payment networks*, Bank of Finland Scientific Monograph E:39, pages 227–252.
- McAndrews, J., and Wasilyew, G. (2005). Simulations of failure in a payment system. In H. Leinonen (Ed.), *Liquidity, risks and speed in payment and settlement systems - a simulation approach*, Bank of Finland Scientific Monograph E:31, pages 229–248.
- McLafferty, J., and Denbee, E. (2012). Liquidity Saving in CHAPS: A Simulation Study. In B. Alexandrova-Kabadjova, S. Martinez-Jaramillo, A. Garcia-Almanza, and E. Tsang (Eds.), *Simulation in Computational Finance and Economics: Tools and Emerging Applications*, pages 120–142. Hershey, USA: IGI Global.
- Page, L., Brin, S., Motwani, R., and Winograd, T. (1999). The PageRank Citation Ranking: Bringing Order to the Web. Technical Report 1999-66, Stanford InfoLab. URL <http://ilpubs.stanford.edu:8090/422/>.
- Schmitz, S. W., and Pühr, C. (2009). Structure and Stability in Payment Networks - A Panel Data Analysis of ARTIS Simulations. In H. Leinonen (Ed.), *Simulation Studies of the Liquidity Needs, Risks and Efficiency in Payment Networks*, Bank of Finland Scientific Monograph E:42, pages 145–188.
- Schulz, C. (2011). Liquidity requirements and payment delays Participant type dependent preferences. Working Paper Series 1291, European Central Bank. URL <http://ideas.repec.org/p/ecb/ecbwps/20111291.html>.

Soramäki, K., Bech, M., Arnold, J., Glass, R., and Beyeler, W. (2007). The topology of interbank payment flows. *Physica A*, 379: 317–333. URL <http://www.sciencedirect.com/science/article/pii/S0378437106013124>.

Please note:

You are most sincerely encouraged to participate in the open assessment of this article. You can do so by either recommending the article or by posting your comments.

Please go to:

<http://dx.doi.org/10.5018/economics-ejournal.ja.2013-28>

The Editor