Structure and Stability in Payment Networks – A Panel Data Analysis of ARTIS Simulations

Stefan W. Schmitz[1] and Claus Puhr[2]

The purpose of this study is the investigation of the importance and impact of network structure, both, at the network level across days and at the node level across days and scenarios (stricken ARTIS participants) for the stability of payment systems in the face of operational shocks. The analysis is based on a large number of simulations of the Austrian large-value payment system ARTIS that quantify the contagion impact of operational shocks at participants' sites. The analysis uncovers that only few payment system participants are systemically important and that contagion displays substantial variation across time and across scenarios. A subsequent panel data investigation tries to explain the variation across time and network participants by structural differences of the payment network across time and the position of the stricken account within the network. It uncovers that (i) standard variables such as liquidity and liquidity loss can explain a substantial fraction of variation, both, across time and across scenarios, that (ii) the structure of the network itself adds very little and (iii) the position of the stricken account within the network indeed contributes somewhat to explaining the variations of contagion. Relative explanatory power is higher when the analysis focuses on contagion measured by the number of banks with unsettled payments or the value of unsettled payments than in the case of the measure based on the number of unsettled payments. In light of the fact that those structural indicators add only little – in terms of explanatory power – to the more traditional measures of the role of an individual participants in the payment system (value and volume of payments) we conclude that at this stage network indicators seem to be of limited use for stability analysis.

JEL: E50, G10.

1 Introduction

Recent work on the stability of banking systems suggests a systematic relationship between network structure, system stability and contagion.[3] Similarly, a recent study conjectures that network structure might be relevant for the stability of payment systems.[4] In previous research we uncovered a large variation of the contagion impact of an individual bank's failure to process payments across banks, across days, and across scenarios.[5] Here we investigate whether the position of the stricken bank within the network helps to explain contagion across scenarios and whether daily variations in network structure contribute to understanding the variation of contagion across days.

Studies concerning network stability in the real world[6] focused on the often observed fact that a few nodes have a large number of links, while most nodes have only few. The

[1] Corresponding author, email: stefan.schmitz@oenb.at.
[2] The authors thank DI Alfred Muigg and Mag. Wolfgang Draxler for providing data, Sylvia Kaufmann, Harry Leinonen, Johannes Lindner, Rainer Puhr and participants of the Bank of Finland Payment System Simulator Seminar 2008 in Helsinki as well as two anonymous referees for comments; and Valentina Metz for very able research assistance. All remaining errors are our own. The views expressed in this paper are those of the authors and do not necessarily reflect those of the OeNB or the Eurosystem.
[3] Inter alia Boss et al. 2004.
[4] Soramäki et al. 2007.
[5] Schmitz, Puhr 2007.
[6] E.g. the Internet, but also large value payment systems such as FedWire and BOJ-NET, or the Austrian interbank market.

reason a lot of attention has been placed on these so called scale-free networks[7] is their robustness to random node removal (the common way to assess instability). A targeted attack, however, in which the most highly connected nodes are removed, leads to quick disintegration. In financial stability analysis, this framework and focus might be relevant for interbank credit (where establishing a credit relation – a link – is costly). The physical network structure of ARTIS, however, is not scale-free but complete[8]. Connectivity is hence not the relevant conceptualisation of stability.

The stability problem in ARTIS is not that bank A cannot make a payment to bank B because the two are not linked anymore. The problem is that bank A might not have the liquidity to make the payment, because it didn't receive payments from, say, bank C (or any other bank) in the first place. As connectivity relates to the flow of liquidity in the system and the liquidity flows through central nodes are higher than that through peripheral nodes, it plays an indirect role for the analysis of stability. Therefore, our measures of the contagion impact of shocks focus on the impact of the shock on the flow of liquidity (i.e. unsettled payments) rather than on the disintegration of the network.

To quantify the contagion effect following the failure of an individual bank we conduct about 30 000 simulations based on actual ARTIS transaction data from 16 November 2005 to 16 November 2007, following the methodology presented in our previous work on ARTIS.[9] In addition to this quantification of (contagiously) unsettled payments in case of an operational incident we calculate a large number of network indicators on the network (44) as well as the node-level (71) for each scenario (stricken ARTIS participants) and for each day in the sample.

In the main body of this paper we investigate whether the variation of network indicators can explain the variation of contagion. We start out with a *univariate analysis* at the network-level, regarding variation across days and at the node-level, regarding variation across days and scenarios (stricken ARTIS participants). In a second, *multivariate* step we conduct an exlporatory panel data analysis which includes network indicators at both levels and we show how well they explain the variation of our three measures of contagion across scenarios and days (the *number of banks with unsettled payments*, the *volume of unsettled payments*, and the *value of unsettled payments*).

The remainder of the paper is structured along the following lines: In section 2 we present data on the network structure of ARTIS. Section 3 introduces the simulations. Based on the results we discuss the scale of contagion in ARTIS and try to provide a means to determine systemically important banks. Section 4 covers the univariate analysis and provides a first glance at the relation of network and node-level indicators to the contagion effects in the simulations. In section 5 we cover the multivariate analysis and present the results of our panel data analysis. Section 6 wraps up our findings.

[7] Scale-free networks are a special case of the aforementioned networks with few important and many minor nodes (in terms of links), where their degree distribution follows a power law $P(k) \sim k^{-\gamma}$.

[8] Participants do not have to submit payments to each other via hubs; they can do so via direct links. The only exception would be a failure of the entire payment system infrastructure, but this question is beyond the scope of this paper.

[9] Schmitz, Puhr (2007) and Schmitz et al. (2008).

2 Measures of Network Structure

The definition of the network under investigation is not trivial in empirical network analysis. We focus on the Giant Strongly Connected Component (GSCC) of ARTIS.[10] The GSCC is the largest component of the network in which all nodes connect to each other via directed paths[11]. We have chosen this definition of the network for two reasons: first, ARTIS contains a comparatively large number of accounts which are not related to financial stability (i.e. offset accounts of OeNB's cash distribution subsidiary) and which are not active on most of the days in the sample. Second, we want to ensure the comparability of our data with that reported for FedWire in Soramäki et al. (2006) which refers to the GSCC.[12]

A related question is the selection of the appropriate indicator of network structure as the number of available indicators is large. At the network-level we calculate 44 network indicators[13]. Similarly, the number of available indicators at the node-level comes to 71. We composed our set of indicators to include, on the one hand, those used in comparable studies and, on the other hand, those suggested by the underlying theory for selecting appropriate indicators for specific typologies of payment flows:

Boss et al. (2004) relate contagion in the interbank market to *betweenness centrality*[14] at the node-level, because this measure has a higher explanatory value than the alternative network indicators in their data set. They uncover a dented linear relationship. Banks with *betweenness centrality* $0 \leq C_B(h) \leq 2$ do not cause any contagious defaults. For $C_B(h) > 2$ they find a linear relationship with a slope of about 0.8.

Borgatti (2005) studies the selection of the appropriate centrality measure for various typologies of flow processes. He classifies flows along two dimensions: the characteristics of the route through the network and the characteristics of the transfer mode. The first dimension considers the constraint on the sequences in which links and nodes are (repeatedly) passed. Liquidity can be transferred to any other node in the network (including the submitter of the first payment). Hence, it is unconstrained (referred to as *walk*). The second dimension refers to the way in which the flowing good is passed on along the route from one node to another. In the case of liquidity the initial holder has to part with it (referred to as *transfer*).

What does that imply for the flow of liquidity in ARTIS? In a physically complete network banks do not have to make payments to other banks via third parties. They transfer directly to the ultimate receiver. However, the flow of liquidity does not stop there. Where it ultimately ends up, is beyond the control (and interest) of the initial submitter of a payment. Given that *betweenness centrality* is based on the share of all

---

[10] For comparable data on the network of all active accounts see Schmitz, Puhr (2007). For a description of the Austrian banking system see OeNB and FMA (2004) The Austrian Financial Markets, Vienna, pp. 50-55.

[11] A directed path is a path that connects to nodes without passing any node or link more than once.

[12] For a comparison as well as a more detailed account of ARTIS network indicators refer to Schmitz et al. (2008)

[13] This includes the directed and/or value/volume weighted and/or average/maximum values for select indicators. Kyriakopulos et al. (forthcoming) find a strong dependence of network characteristics on aggregation time. The large number of network indicators and the critical role of aggregation time pose the problem of data-mining in network topology studies.

[14] For the definitions, formulas, and graphical illustrations of the network indicators see Appendix 2.

shortest paths through a node, it is not a good measure of centrality in the study of liquidity flows. *Degree centrality* is more suitable for this purpose.

Besides considering the most meaningful indicators we want to ensure a high degree of comparability of our results with other papers that use different network indicators. Moreover, we want to investigate whether network indicators in general add value to the more traditional measure used in comparable simulation studies (i.e. the size of the individual node in terms of value and volume of transactions). Therefore we focus on the measures *value* and *volume* as well as on the network indicators *average path length, degree*, *connectivity*, *clustering*, *betweenness centrality* and *dissimilarity index* as provided in Table 1 (for the network-level averages across participants).

Table 1: ARTIS Network Indicators (Network-level)

	Mean	Median	Min.	Max.	Std.Dev.
Payments					
Volume	15 380	15 436	9 786	25 000	2 019
Value (EUR bn)	48.5	46.9	22.6	84.9	10.6
Average (EUR mn)	3.20	3.00	1.90	5.90	0.70
Size					
Nodes	133.2	132	112	159	9.3
Links	1 376	1 376	1 222	1 602	69
Distance Measure					
Avg. Path Length	2.4	2.4	2.2	2.6	0.08
Connectivity					
Average Degree	15.6	15.5	14.2	17.8	0.6
Connectivity (%)	7.9	7.9	5.9	9.9	0.8
Clustering (%)	58.3	58.3	51	63.7	2.3
Others					
Betweenness Cent. (%)	0.8	0.8	0.6	0.9	0.1
Dissimilarity Index	0.47	0.47	0.39	0.60	0.03

Source: Own calculations based on daily averages of the ARTIS GSCC from 16 November 2005 to 16 November 2007 (excluding Austrian holidays).

For our observation period from 16 November 2005 to 16 November 2007 the *average volume* of transactions per day is 15 380 in the GSCC of ARTIS. The *average value* of transactions per day comes to EUR 48.5 billion. The average transaction size amounts to 3.2 million EUR. The size of the network is defined by the number of nodes. On average there are 133.2 nodes in the GSCC during the sample period of which 63 are in the GSCC on all days. The active nodes are linked by an average of 1 376.1 directed links.[15]

An indicator of the distance between nodes is the shortest path length. We calculate the *average shortest path length* for each originating node by averaging across terminating nodes and then averaging across originating nodes to derive the average path length of the entire network. Across days this value equals 2.4.

How well nodes are connected in the network is captured by the *average degree* of the network. It is calculated by summing across all (undirected) links originating from each

[15] The average number of nodes in ARTIS active each day was 209.8 and the number of directed links was 1 637.5.

node and than averaging across nodes.[16] Averaged also across days, it amounts to 15.6 in the ARTIS system. However, the most active nodes have a much larger number of links originating and terminating at them.[17] The *connectivity* of the network is captured by the number of actual links relative to the number of possible links. Connectivity averages 7.9 percent. The *clustering coefficient* provides a measure of the average connectivity of the neighbours of all nodes in the GSCC. On average about 58 percent of the neighbours of each node are also linked.

Betweenness centrality measures how many shortest paths pass through the average node. The value of 8 percent is quite low and stems from the centrality of a few nodes with high betweenness centrality and a large number of nodes with low values. The *dissimilarity index* captures the relative viewpoints of the network from two neighbouring nodes. If the network looks very similar from both nodes, the dissimilarity index is small. In the GSCC it amounts to 0.47 which implies that on average the perspectives of the GSCC differ substantially from any two neighbouring nodes. A lot of nodes link that otherwise do not share many network characteristics. In sum, we interpret the data on network indicators as corroborating previous evidence that a few large nodes dominate the payment system and that many of the smaller nodes connect to the largest nodes at the centre of the network.

3 Measures of Network Stability

As argued in the introduction, connectivity is not an adequate criterion to capture the systemic impact of an operational problem at one of the nodes in a large value payment system. Alternatively, we suggest defining a threshold based on the average contagion effect following the failure of an individual payment system participant. As operational failures, let alone such with systemic impact, are few and far between we resort to simulations. These provide us with what we call contagious defaults, which can be measured in three ways.

First, the *number of participants* (banks or transfer accounts) *with unsettled payments* at end of day measures how many participants (banks or transfer accounts) faced liquidity problems due to the operational incident at another participant. Second, the *number of unsettled payments* at end of day is the total volume of all payments that could not be settled by the participants that were not subject to an operational incident. Third, the *value of unsettled payments* at the end of day is the total value of all payments that could not be settled by the participants that did not experience an operational problem.

We conduct 31 311 simulations based on 63 different scenarios for 497 transaction days from 16 November 2005 to 16 November 2007 (excluding Austrian holidays) which yield some 620 million simulated transactions.[18] These simulations are calculated with a

[16] The out-degree refers to the number of links originating at the node while the in-degree is based on to the number of links terminating at the node. Across the network the average out- and in-degree are equal to m/n.

[17] For the analysis of the degree distribution see Schmitz et al. (2008) where the hypothesis of a Power Law distribution is rejected for the monthly network and Kyriapopulos et al. (forthcoming) who find degree distributions might seems to have a Power Law distribution in daily networks (in ARTIS), but that this property vanishes in longer aggregation times.

[18] For more details on simulations, their motivation, and their design see Schmitz, Puhr (2007). The operation of ARTIS was discontinued after 16 November 2007, due to the introduction of TARGET2.

self-implemented Matlab-based software tool (inspired by Bank of Finland Payment System Simulator), which was tailored to ARTIS particularities.

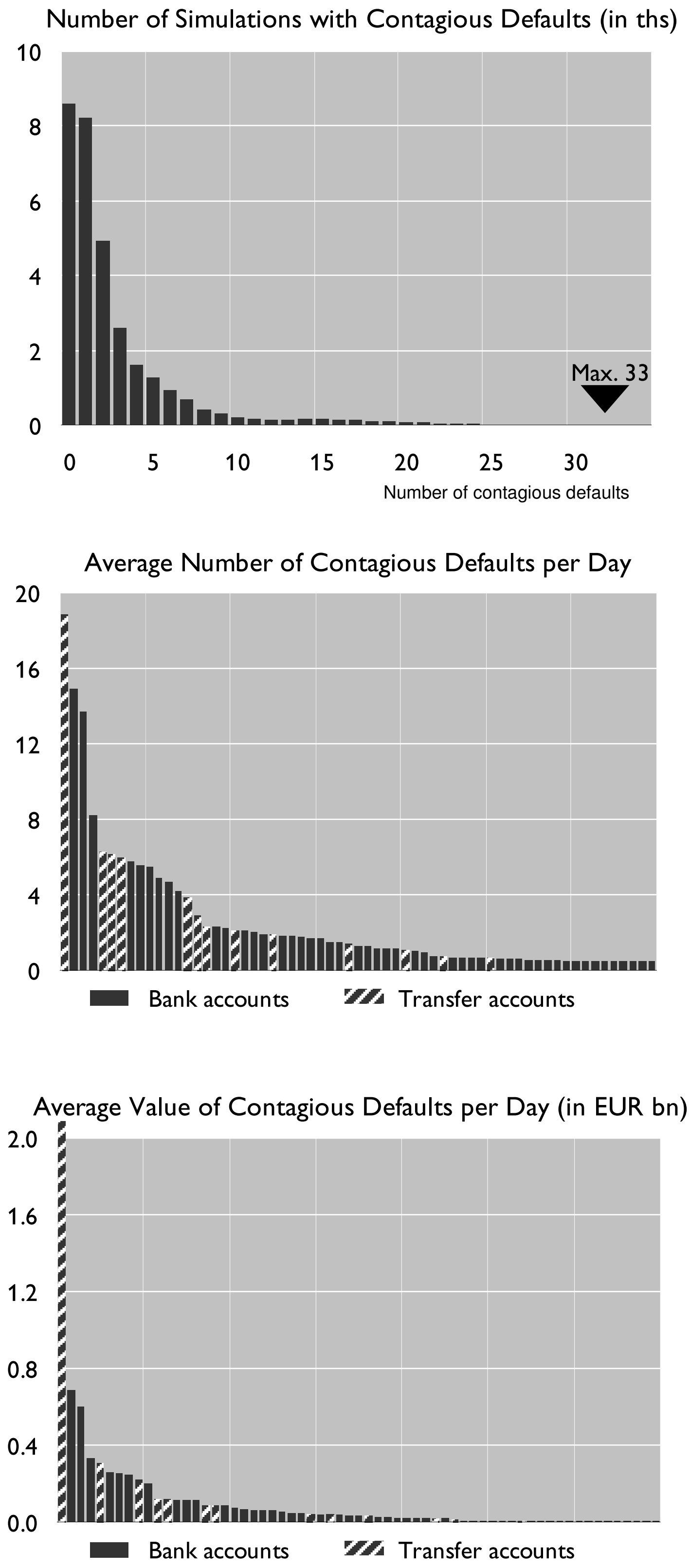
For this paper we run simulations for all 50 banks that are in the GSCC on all Austrian working days throughout the sample period and all 13 Transfer accounts[19] that form part of the system on all days in the sample period. We assume an operational incident that hits one participant (banks or transfer accounts) in each simulation. The operational incident is mapped into the simulation as the incapacitation of the participant to process outgoing payments, i.e. the inability to submit transactions for the whole day.[20] This assumption is extreme but plausible. As shown in Schmitz, Puhr (2007) shorter outages of participants may lead to payment delays but not to unsettled payments.

In Chart 1, the upper panel shows that about 27.5 percent of all simulations (8 604) do not lead to any contagion at all. A further 26.3 percent (8 230) yield one contagious default and 33.1 percent (10 375) two to five. 13.1 percent (4 102) lead to more than five contagious defaults with a maximum across the 31 311 simulations of 33.

[19] Transfer accounts are ARTIS accounts held by other ESCB central banks at OeNB. All national TARGET components are directly linked by transfer accounts. All transactions to and from the respective country and Austria are routed via these transfer accounts.

[20] It is assumed that the resulting illiquidity of the participant is not interpreted as potential insolvency by other participants of the payment system and the financial system at large. In addition, ARTIS provides business continuity arrangements for participants. We tested their impact in Schmitz and Puhr (2007), but disregard them in this paper, as they are of little relevance for the interaction between network topology and contagion.

Chart 1: Contagious Defaults in ARTIS



Source: Own calculations based on daily simulations of the 50 banks and 13 Transfer accounts that formed part of the ARTIS GSCC from 16 November 2005 to 16 November 2007 (excluding Austrian holidays).

The other two panels in Chart 1 show the average contagious defaults per simulation (the former in terms of number of *participants* (banks or transfer accounts) with unsettled payments and the latter in terms of average value of unsettled payments, both due to contagious defaults per simulation). As argued above, we suggest using this information to derive a set of systemically relevant ARTIS participants. If we set the threshold[21] for example in terms of the value of contagious defaults, e.g. only participants that cause at least an average value of EUR 48.5 million of unsettled payments (or 0.1 percent of average value of transactions settled across days), we see the number of systemically relevant participants shrink to 24 (17 banks, seven Transfer accounts). That equals about seven percent of the average of 230 banks in ARTIS (during the sample period) and to about two percent of the average of 850 banks in Austria. These results suggest that the supervision of operational risk in banks' payment processing capacity could focus on a relatively small set of systemically relevant banks in Austria and on their business continuity arrangements.

4 Univariate Analysis of Structure and Stability

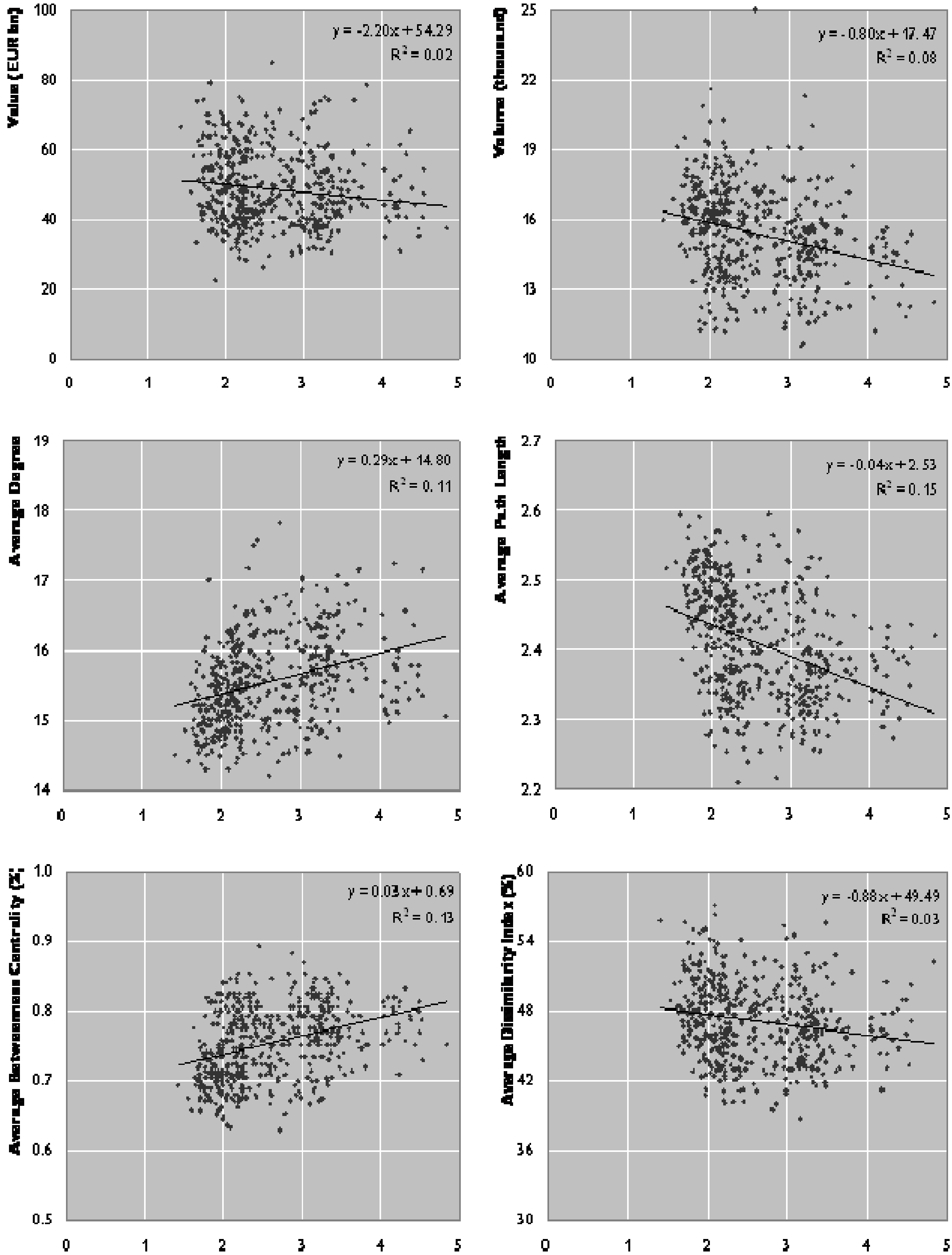
Following the argument in Section 2 (choice of structural measures) and in Section 3 (choice of stability measures), we provide a selection of univariate results that in turn provide the intuition for our multivariate analysis in Section 5. We look at whether the variation of network indicators at the network-level across days (4.1) and at the node-level across stricken participants (4.2) explain the variation of contagion across days and across stricken participants.

4.1 Network-level

In the top two panels of Chart 2 we depict the scatter plot for the value (left hand panel) and the volume of all payments (right hand panel) submitted to ARTIS on the y-axis and the number of contagious defaults in terms of the *number of participants with unsettled payments* (daily averages across scenarios) per day on the x-axis. The variation of *value* explains 2 percent and the variation of *volume* accounts for 8 percent of the variation of the contagion impact per day.[22]

[21] To some extent that threshold is arbitrary and depends on the risk aversion of the supervisory authority.
[22] Neither *volume* nor *value* is significant at the common confidence levels.

Chart 2: ARTIS Network Indicators (Network-level) vs. Contagion (Daily Average of Number of Participants with Unsettled Payments across Scenarios)



Source: Own calculations based on daily ARTIS GSCC network indicator averages and on daily simulations of the 50 banks and 13 Transfer accounts that were part of the ARTIS GSCC from 16 November 2005 to 16 November 2007 (excluding Austrian holidays).

The explanatory value of the variables *value* and *volume* is low. Do network indicators perform any better? In the four other panels of Chart 2 we look at the following indicators (in and out, unweighted, undirected): *degree*, *average path length*, *betweenness centrality* and *dissimilarity index*. The *average path length* (15 percent) and *betweenness centrality* (13 percent) have the highest explanatory values. The daily variation in *degree* accounts for 10 percent of the variation in contagion and that of the *dissimilarity index* for only 3 percent. Although the explanatory power of three of the network indicators is higher than that of *value* and *volume*, the levels are still low and tests for significance fail at all common confidence levels.

The highest explanatory power of any of the remaining 39 indicators is 15.4 percent (*average number-weighted clustering coefficient*), while a number of indicators have no explanatory power at all. The univariate analysis suggests that daily variations in network indicators at the network-level across days are of limited use in the stability analysis of ARTIS. However, that does not preclude that either (i) network indicators at the node-level or (ii) structural differences across networks might influence their (relative) resilience.

Regarding the latter (structural differences) we lack data for an in depth analysis, but will study it in further research. Given the fact that other large value payment systems which display considerable differences in size share notable structural commonalities with ARTIS[23], some doubt is justified whether network indicators at the network-level could explain contagious effects in other large value payment systems. That leaves the question whether the different positions of the nodes (that experience the operational incident) in the network account for this variation?

4.2 Node-level

In the top two panels of Chart 3 we plot the *value* and *volume* of payments of the stricken node in each simulation against its contagion effect in terms of the *number of participants with unsettled payments*. The variations of *value* and *volume* across simulations explain 73 percent and 68 percent of the variation of the contagion impact across simulations.[24] The slopes have the expected signs: more active nodes cause more contagion.

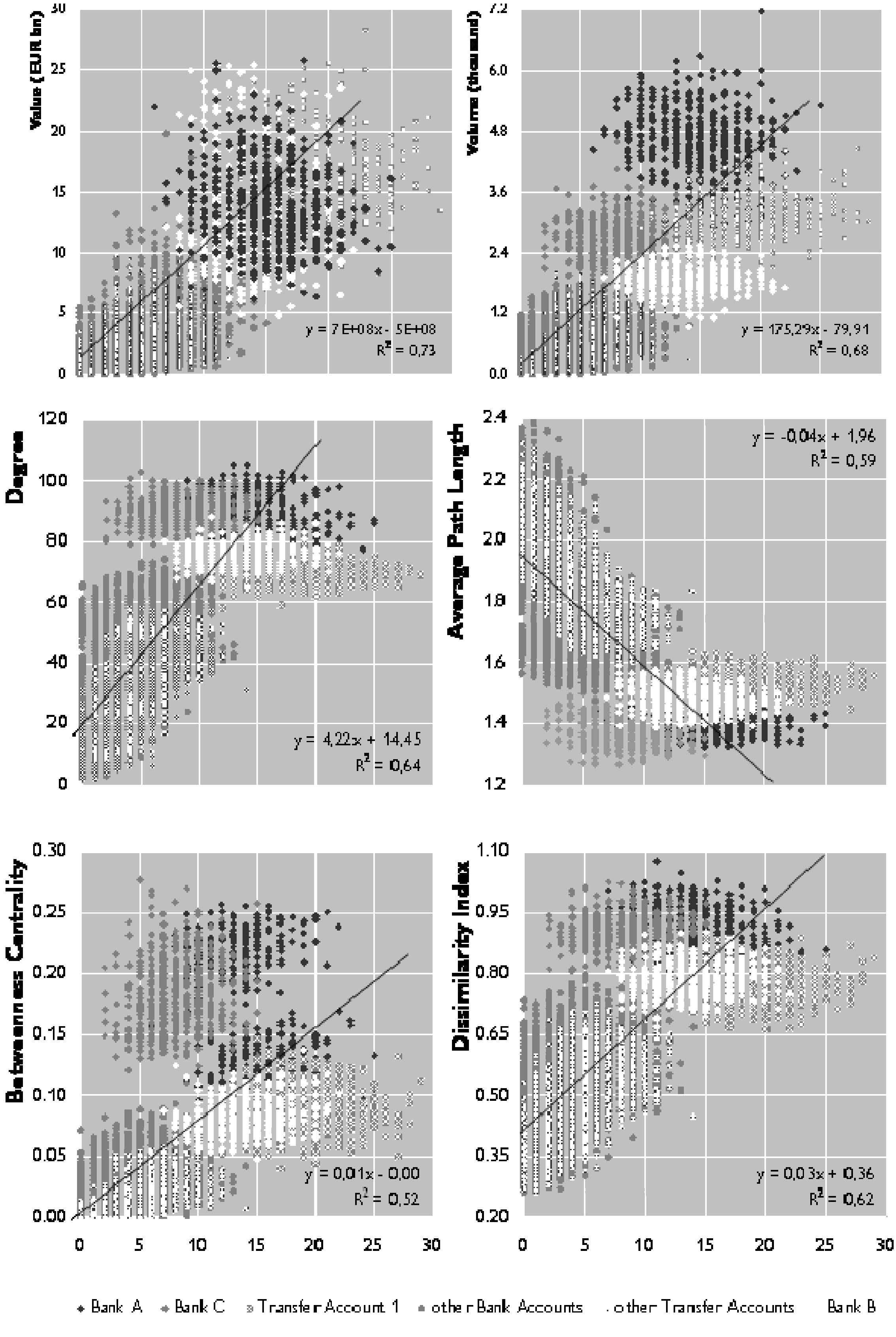
Given the large number of data points (31 311) and the variation of the stability measure across ARTIS participants[25], we differentiate in Chart 3 between shocks to banks and Transfer accounts. In addition we highlight the three most active banks and the most active Transfer account. The differentiation reveals a pronounced grouping in both panels. In the right hand panel it also points to structural differences in contagion impact not accounted for by variations in *volume*. The most active Transfer account and one of the three banks tend to group below the regression line (i.e. they causes more contagion than estimated by their volumes of transactions) and the other two banks group above the regression line (i.e. they cause less contagion than estimated by their volumes of transactions).

[23] As shown e.g. in the comparison of FedWire and ARTIS in Schmitz, Puhr (2007).

[24] Both *volume* and *value* are highly significant at all common confidence levels.

[25] As shown in Section 3, see for instance Chart 1.

Chart 3: ARTIS Network Indicators (Network-level) vs. Contagion (Daily Average of Number of Participants with Unsettled Payments across Scenarios)



Source: Own calculations based on daily node-level ARTIS GSCC network indicators and on daily simulations of the 50 banks and 13 Transfer accounts that were part of the ARTIS GSCC every day from 16 November 2005 to 16 November 2007 (excluding Austrian holidays).

In the face of low explanatory values of the variables *value* and *volume* at the network-level we asked whether network indicators at the network-level perform any better and found out they actually did, albeit at sill at modest levels. Given the already high explanatory power of *value* and *volume* at the node-level, we look at whether our four previously selected network indicators at the node-level[26] can again add to that.

We find that the explanatory values of all four network indicators are actually quite high;[27] the most simple measure *degree* yields an R^2 of 64 percent, variations in *average path length* across simulations account for 59 percent of the variation of the number of contagious defaults across simulations. The more complex measures *betweenness centrality* and *dissimilarity index* yield R^2s of 52 and 62 percent, respectively. The signs of all slopes are in line with expectations: simulations, in which more active and more central nodes are shocked, feature a higher contagion impact. Moreover, all indicators are highly significant at common confidence levels and also in the order of magnitude of the reported interaction between *betweenness centrality* and contagious defaults for the Austrian interbank market.[28]

The remaining 65 network indicators yield explanatory values between nil (*number-weighted average path length based on payments received*) and 77 percent (*relative volume of payments received*).[29] The results demonstrate that network indicators at the node-level seem to explain large parts of the variation in contagion across stricken participants in a univariate setting. However, they seem to add little to the high explanatory values of the traditional measures of activity (*value* and *volume*). Furthermore, the large set of available indicators and the huge differences in their explanatory values pose the problem of data mining.

The aforementioned differentiation according to the stricken ARTIS participant (bank or Transfer account) confirms the pronounced structural differences in contagion impact not accounted for by variations in *volume*. In all four network indicator based panels of Chart 3 simulations based on the most active Transfer account cluster at the right hand side of the regression line, while those based on the two aforementioned banks to its left.[30] This finding points at structural differences in contagion impact which are not accounted for by measures of activity or network indicators and which warrant further research.

We also investigate the interaction between network structure and network stability for other measures of contagion, as an example we present the *value of unsettled payments*. Just as above we start with an analysis of the explanatory value of node size (*value* and *volume* of payments originating at the node). Both values are lower than the respective previous results presented in Chart 3. Moreover, only *value* is significant, explaining 54 percent of

[26] Previously selected network indicators include: *Degree*, *average path length*, *betweenness centrality* and *dissimilarity index*.

[27] We present the simple linear regression results in order to provide an indication of relative performance and to motivate our approach in the panel data analysis rather than suggesting that OLS is appropriate per se.

[28] See Boss et al. 2004.

[29] Due to the large number of observations and the ensuing degrees of freedom, any indicator with an R^2 of 0.51 or higher is significant at all common confidence levels, whereas for indicators with an R^2 of 0.50 or below the null hypothesis cannot be rejected.

[30] The graphs might also be read as suggesting non-linearity; but we prefer the interpretation of structural differences which we can then exploit in the panel data analysis.

variation in contagion. Albeit an explanatory power of 39 percent, *volume* is not significant at common confidence levels.

How well do the network indicators at the node-level fare in comparison? The R^2s of the four previously presented network indicators range between 24 and 29 percent and are therefore considerably lower than (i) the respective values for the measures of node size above, but (ii) also their respective values when explaining contagion as measured by the *number of participants with unsettled payments*.[31]

We conclude that if we measure contagion by the *value of unsettled payments*, network indicators are clearly dominated by the traditional measures of size. Comparing the results for the two measures of contagion, *number of participants with unsettled payments* versus *value of unsettled payments*, reveals that contagion under the latter measure is much harder to explain even by the superior traditional variables.

Table 2: Correlations between ARTIS Network Indicators (Node-level)

	Volume	Value	Avg. PL	Degree	Conn.	Clust.	Btw. C.	Dissim.
Volume	100.0%	89.0%	84.0%	83.0%	-77.0%	-57.0%	89.0%	85.0%
Value		100.0%	76.0%	75.0%	-70.0%	-52.0%	77.0%	78.0%
Avg. PL			100.0%	99.0%	-96.0%	-72.0%	85.0%	95.0%
Degree				100.0%	-97.0%	-72.0%	85.0%	93.0%
Conn.					100.0%	62.0%	-79.0%	-85.0%
Clust.						100.0%	-56.0%	-78.0%
Btw. C.							100.0%	87.0%
Dissim.								100.0%

Source: Own calculations based on daily averages of the ARTIS GSCC from 16 November 2005 to 16 November 2007 (excluding Austrian holidays). Network indicators include: Average Path Length (Avg. PL), Connectivity (Conn.), Clustering Index (Clust.), Betweenness Centrality (Btw. C.), Dissimilarity Index (Dissim.).

In order to corroborate our findings from the univariate analysis, that network indicators at the node-level do not add much value to stability analysis, we present the correlations between the traditional measures of activity (*value* and *volume*) and selected network indicators in Table 2. The data reveals that particularly indicators of centrality are highly correlated with *value* and *volume*. Nevertheless, the question remains open whether these indicators add some explanation in a multivariate setting.

5 Multivariate Analysis of Structure and Stability

In this section we study the robustness of our findings in the univariate setting in section 4 in an exploratory multivariate study. We focus on combining one of the traditional measures of node size (value) with network indicators at the node and at the network-level as well as additional control variables (e.g. beginning of day liquidity at individual nodes, dummy variable for Transfer accounts) in a panel data setting to answer the following four questions:

[31] Individual explanatory values are as follows: *degree* 28 percent, *average path length* 25 percent, *betweenness centrality* 24 percent and *dissimilarity index* 29 percent, none of which are significant at common confidence levels.

1. What explains the variations in contagion across days within scenarios?
2. What explains the variations in contagion across scenarios on each day?
3. Are network indicators at the network and/or at the node-level significant in this context?
4. What is the explanatory contribution of the network indicators at the network and/or at the node-level in the context of questions 1 and 2?

In Section 5.1 we introduce our measures of contagion as dependent variables. We try to explain their variation by three groups of independent variables discussed in Section 5.2: first, the independent variables at the network-level, which are constant across panels but vary across time; second, the independent variables at the node-level that vary across time and across scenarios; third, we add a dummy variable for Transfer accounts to corroborate the findings of some of the hitherto unexplained structural particularities uncovered in the scatter plots of Section 4. In Section 5.3 we introduce our model as well as its assumptions and the estimation method. In Section 5.4 we present the results.

5.1 Dependent Variables

As dependent variables we focus on our three measures of contagion (they exclude the impact on the stricken bank). Three different measures of contagion provide the unique opportunity to check the robustness of the models and the parameter estimates. Table 3 shows that the means and the standard deviations of the dependent variables differ substantially over time and across scenarios. First, the *number of participants with unsettled payments* amounts to an overall daily average of 2.6 (*overall standard deviation* 3.8) with a minimum of 0 and a maximum of 33. The variation between scenarios (*between standard deviation* 3.5) is much higher than the variation across time within scenarios (*within standard deviation* 1.4). Second, the average daily *volume of unsettled payments* is 7.6 per day (*overall standard deviation* 21.7). The lowest value is 0 the highest value 1 172. In this case the standard deviation between scenarios (14) is slightly lower than the one within scenarios (16.7). Third, the daily *value of unsettled payments* averages EUR 112 million with a range from 0 to 10.7 billion. The overall standard deviation is EUR 335 million and the between standard deviation is much higher (EUR 284 million) than the within one (EUR 181 million).

Table 3: Dependent Variables (Measures of Contagion)

Variable		Mean	Std.Dev.	Min.	Max.	Obs.
Number of Participants with Unsettled Payments	overall	2.6	3.8	0	33	N=31311
	between		3.5			n=63
	within		1.4			T=497
Volume of Unsettled Payments	overall	7.6	21.7	0	1172	N=31311
	between		14			n=63
	within		16.7			T=497
Value of Unsettled Payments (in EUR billion)	overall	0.11	0.34	0	10.73	N=31311
	between		0.28			n=63
	within		0.18			T=497

Source: Own calculations based on ARTIS GSCC data from 16 November 2005 to 16 November 2007 (excluding Austrian holidays). Standard Deviation (Std.Dev.), Observations (Obs.).

5.2 Independent Variables

We group the independent variables in two groups, namely those at the network- and those at the node-level.

First, the independent variables at the network-level are constant across panels but vary across time ([*Z*] in the model below): They include aggregate liquidity (*Liquidity*[32]), a traditional measure of network size (aggregate value of all transactions (*Value (Network*)), and a range of network indicators at the network-level.

Table 4 displays the independent variables at the network level – *Liquidity*, *Value(Network)*, and the network indicators at the network-level (including mean and standard deviation for all variables).We define aggregate liquidity (*Liquidity*) in the system (mean: EUR 18.3 billion, standard deviation: EUR 3.2 billion, Table 4) as the sum of beginning of day balances (EUR 7.5 billion; EUR 0.8 billion) and unencumbered collateral[33] (EUR 10.8 billion; EUR 2.9 billion) across participants in the system.[34]

Turning to the network indicators at the network level, it is apparent that the relative standard deviations of the network indicators across days – often only small fractions of the respective means – are much lower than that of the measures of contagion (exceeding the respective means).[35]

[32] It corresponds to our traditional measure of size value in the univariate analysis. We chose to rename here to facilitate economic interpretation and intuition.

[33] We simply aggregate across beginning of day balances and unencumbered collateral because the latter can be liquidised via interest free Daylight Overdrafts at OeNB within minutes (for details see Schmitz, Puhr 2007).

[34] Focusing on real historical data might restrict the generalisation of the results to other payment systems in which participants might follow different liquidity policies. E.g. in systems, that experience frequent operational outages, it might be rationale for participants to hold sufficient liquidity to settle all outgoing payments. As a consequence there would be no contagion at all.

[35] While standard deviation is not the ideal parameters to describe the distributions of the network indicator, they are helpful in pointing out the differences between within and between scenario variation in our data set.

Table 4: Independent Variables at the Network Level: Liquidity, Netvalue, and Network Indicators at Network-Level

Variable		Mean	Std.Dev.	Min.	Max.	Obs.
Liquidity (in EUR billion)	overall	18.27	3.23	11.68	32.37	N=31311
	between		0.00			n=63
	within		3.23			T=497
Value (Network) (in EUR billion)	overall	48.90	10.62	22.88	85.60	N=31311
	between		0.00			n=63
	within		10.62			T=497
Average Degree (Network)	overall	12.3	0.4	11.3	14.4	N=31311
	between		0.0			n=63
	within		0.4			T=497
Average Connectivity (Network)	overall	0.040	0.003	0.030	0.050	N=31311
	between		0.000			n=63
	within		0.003			T=497
Average Path Length (Network)	overall	2.50	0.06	2.40	2.70	N=31311
	between		0.00			n=63
	within		0.06			T=497
Average Cluster Index (Network)	overall	0.40	0.03	0.40	0.50	N=31311
	between		0.00			n=63
	within		0.03			T=497
Average Betweeness Centrality (Network)	overall	0.0050	0.0003	0.0040	0.0060	N=31311
	between		0.0000			n=63
	within		0.0003			T=497
Average Dissimilarity Index (Network)	overall	1.3	0.9	0.6	5.2	N=31311
	between		0.0			n=63
	within		0.9			T=497

Source: Own calculations based on ARTIS data for the independent variables from 16 November 2005 to 16 November 2007 (excluding Austrian holidays). Standard Deviation (Std.Dev.), Observations (Obs.). N.B. Data differs from that in Table 1 (network indicators for all accounts in the GSCC).

Second, the independent variables at the node-level (*Liquidity Loss* and the network indicators at the node level; Table 5 plus a dummy variable for Transfer accounts) vary, both, across time and across scenarios ([*X*] in the model below). *Liquidity Loss* is measured as the value of payments that were due by the stricken bank but were not processed in the simulations due to operational problems at the stricken bank.[36] The dummy variable *D* for Transfer accounts took the values 0 or 1 and entered the models as *D*×*Liquidity Loss* to measure the deviation of the impact of the *Liquidity Loss* variable in the case of Transfer accounts from the average across all participants.

[36] We have also used alternative proxies for the impact of the operational problem of the stricken bank on liquidity in the system, such as liquidity drain and liquidity sink, which yield similar results both in terms of sign and significance.

Table 5: Independent Variables at the Node Level: Liquidity Loss and Network Indicators at Node-Level

Variable		Mean	Std.Dev.	Min.	Max.	Obs.
Liquidity Loss	overall	0.76	1.69	0.00	22.05	N=31311
(in EUR billion)	between					n=63
	within					T=497
Average	overall	25.4	19.9	2	105	N=31311
Degree (Node)	between		19.8			n=63
	within		3			T=497
Average	overall	0.200	0.15	0.015	0.800	N=31311
Connectivity (Node)	between		0.150			n=63
	within		0.02			T=497
Average	overall	1.90	0.18	1.20	2.30	N=31311
Path Length (Node)	between		0.18			n=63
	within		0.02			T=497
Average	overall	0.50	0.2	0.13	1.00	N=31311
Cluster Index (Node)	between		0.19			n=63
	within		0.08			T=497
Average Betweeness	overall	0.1360	0.0349	0.0000	0.2760	N=31311
Centrality (Node)	between		0.0341			n=63
	within		0.0086			T=497
Average Dissimilarity	overall	0.44	0.13	0.26	1.07	N=31311
Index (Node)	between		0.1			n=63
	within		0.03			T=497

Source: Own calculations based on ARTIS data for the 63 accounts defining the scenarios from 16 November 2005 to 16 November 2007 (excluding Austrian holidays). Standard Deviation (Std.Dev.), Observations (Obs.).

Table 5 also shows that the standard deviations of the network indicators at the node-level are much higher across scenarios than within scenarios across time. Similarly, their means and standard deviations are much higher than that of the network indicators at the network-level; e.g. the mean of the average degree at the network-level across days is 12.4 (standard deviation 0.4) while the mean of the degree at the node-level across scenarios and across days is 25.4 (standard deviation 19.9).

5.3 Models, Specifications, and Estimation

We estimate the following static fixed effects model:

$$\begin{bmatrix} y_1 \\ y_2 \\ . \\ y_{63} \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \\ . \\ X_{63} \end{bmatrix} \beta_I + [Z]\beta_{II} + \begin{bmatrix} v_1 \\ v_2 \\ . \\ v_{63} \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ \varepsilon_{63} \end{bmatrix}$$

The dependent variables y_1 to y_{63} are 497×1 vectors containing daily values for the dependent variable. The vector $[y_1 \ \ldots \ y_{63}]$ therefore has 31311 elements. In theory, the matrices X_1 to X_{63} would be 497×9 dimensional as they contain daily observations of the independent variables *Liquidity Loss*, *D×Liquidity Loss,* and the six network indicators at the node-level plus the constant term. However, in practice the six network indicators are highly correlated. Similarly, the vector *Z* would contain *Liquidity* and the network indicators at the network-level and would have 497×7 dimensions. But the network indicators at the node-level are also highly correlated (see Table 2) which would lead to multicollinearity.

The vectors of parameters β_I and β_{II} are to be estimated. The 63×1 dimensional vector [ν_1 ... ν_{63}] contains the scenario specific unobservable time-invariant regressors and the 497×1 dimensional vector [ε_1 ... ε_{63}] consists of the standard error term for each observation (31311 elements). Our panel is balanced as we conduct simulations for all scenarios in all periods and the number of simulations is equal for all days in the sample period.

In order to avoid multicolinearity we estimated a basic model for each of the measures of contagion consisting of the independent variables *Liquidity*, *Liquidity Loss* and *D×Liquidity* (specification 1 in Tables 6, 7, and 8). The choice of variables for this basic model rests on previous results and economic intuition which show that (*i.*) the aggregate liquidity of the system (*Liquidity*) reduces contagion[37], that (*ii.*) *Liquidity Loss* is significantly positively correlated with contagion in the univariate analysis, and that (*iii.*) Transfer accounts display interesting particularities in the univariate analysis which warrant further attention. To this model structure we add the variable *Value (Network)* (specification 2) or a particular network indicator at the network and at the node-level (specifications 3 to 8).[38]

The following equation is specification 5 of model 1 and explains the variations of the *number of participants with unsettled payments* across days and scenarios by a constant and the standard ingredients *Liquidity*, *Liquidity Loss* and *D×Liquidity Loss* plus the average path length at the node-level and the network-level.

$$Numberofparticipants_{it} = \beta_1 + \beta_2 Liquidity_t + \beta_3 LiquidityLoss_{it} + \beta_4 D \times LiquidityLoss_{it} + \\ + \beta_5 AvgPathLength(Node)_{it} + \beta_6 AvgPathLength(Network)_t + \nu_i + \varepsilon_{it}$$

Similarly specification 5 of model 2 would explain the variations of the *number of unsettled payments* across days and scenarios with the same independent variables:

$$Numberofunsettledpayments_{it} = \beta_1 + \beta_2 Liquidity_t + \beta_3 LiquidityLoss_{it} + \beta_4 D \times LiquidityLoss_{it} + \\ + \beta_5 AvgPathLength(Node)_{it} + \beta_6 AvgPathLength(Network)_t + \nu_i + \varepsilon_{it}$$

Similarly specification 5 of model 3 would contain the variations of the *value of unsettled payments* across days as dependent variable and the same independent variables.

$$Valueofunsettledpayments_{it} = \beta_1 + \beta_2 Liquidity_t + \beta_3 LiquidityLoss_{it} + \beta_4 D \times LiquidityLoss_{it} + \\ + \beta_5 AvgPathLength(Node)_{it} + \beta_6 AvgPathLength(Network)_t + \nu_i + \varepsilon_{it}$$

There are some basic assumptions regarding static fixed effects models. One states that the error term ε is uncorrelated with past, present, and future values of the independent

[37] See the papers in Leinonen (2005).

[38] Why does specification 2 only contain the network level variable and not the node level variable? Remember that the corresponding node level variable is already contained in the specification as Liquidity Loss.

variables. This *assumption of strict exogeneity* ensures that the agents whose behaviour is modelled are not influenced by past realisations of the error term ε. Moreover, *cross-panel and cross-time conditional homoskedasticity* states that the conditional variance of the error term ε – given the time-invariant unobservable scenario specific effect ν – is constant across scenarios and across time. Furthermore *serial independence* presupposes that the error terms are serially independent within panels and *cross-panel independence* that they are independent across scenarios.

Are these assumptions fulfilled in our dataset? The first one is by virtue of the simulation design, as we do not model behavioural reactions of banks to operational shocks. The values of the explanatory variables are historic observations of a world without operational shocks and hence contagion. Consequently, there are no observable, unexplained variations in contagion – past realisation of the error terms – which could influence banks' behaviour: e.g. banks cannot adjust their liquidity reserves or payment behaviour to account for error terms which are the results of counterfactual simulations. The other three assumptions are not fulfilled[39]. Hence, an ordinary least squares (OLS) estimate of a standard fixed-effects model would yield inconsistent and biased standard errors. We employ an alternative estimator that takes into account conditional heteroskedasticity, serial correlation, and cross-sectional dependence of the error terms ε.

We apply a panel-corrected standard error estimator with panel specific autocorrelations where the parameters are estimated by Prais-Winsten regression. The parameter estimates are conditional on the estimators of the autocorrelation parameters in each panel. The estimator uses a feasible generalised least squares (FGLS) estimate of the variance-covariance matrix which is asymptotically efficient under the assumed covariance structure of the disturbance terms (heteroskedasticity and contemporaneous correlation across panels).

With 63 panels this yields 63 variance estimates and 1 953 covariance estimates. Together with the 63 autocorrelation estimates and (up to) 6 parameters a total of 2 085 parameter estimates are required. The estimation procedure yields consistent standard errors but at the expense of a large loss in degrees of freedom (2 079). However, with 31 311 observations the degrees of freedom are still large.

Beck and Katz (1995) argue that full FGLS estimates were overly optimistic and the Prais-Winsten estimator superior. Although they derived their results for data with some 10 to 20 panels and 10 to 40 time periods we also preferred the Prais-Winsten estimator for our larger data set. Especially, since the large number of observations in each panel (497) supports the asymptotic behaviour of the panel-specific autocorrelations.

5.4 Results of Model 1 (Number of Participants with Unsettled Payments)

We present the results of the panel data estimates in three tables (Table 6, 7, and 8), one for each measure of contagion. Table 6 summarises the results for the 8 specifications of

[39] See Annex – Test Results fort the static fixed effects model.

Model 1. The dependent variable is the *number of participants with unsettled payments*, i.e. the number of banks (or transfer accounts; excluding the stricken bank/transfer accounts) with unsettled payments at the end of the day.

Table 6: Results Model 1 (Number of Participants with Unsettled Payments)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	3.28	4.02	3.38	2.32	14.73	6.97	4.63	-0.93
	*13.71****	*16.5****	*3.8****	*5.14****	*10.1****	*14.84****	*7.25****	*-3.85****
Liquidity	-0.10	-0.10	-0.10	-0.08	-0.09	-0.10	-0.10	-0.12
	*-8.19****	*-8.05****	*-9.16****	*-7.87****	*-7.56****	*-16****	*-7.67****	*-9.65****
Liquidity Loss	1.65	1.66	0.78	0.81	1.01	1.43	1.07	0.94
	*59.21****	*58.92****	*28.64****	*30.8****	*38.61****	*55.18****	*33.78****	*33.17****
D x Liquidity Loss	0.25	0.25	0.61	0.60	0.53	0.26	0.54	0.51
	*8.99****	*8.93****	*23.28****	*23.92****	*20.57****	*9.66****	*19.29****	*19.42****
Degree (Node)			0.09					
			*59.92****					
Connectivity (Node)				12.23				
				*59.88****				
Avg. Path Length (Node)					-8.09			
					*-56.71****			
Cluter Index (Node)						-4.00		
						*-51.65****		
Betw. Cenrality (Node)							33.80	
							*26.37****	
Dissimilarity Index (Node)								11.13
								*40.07****
Value (Network)		-0.02						
		*-7.76****						
Degree (Network)			-0.16					
			*-2.24****					
Connectivity (Network)				-32.51				
				*-3.25****				
Avg. Path Length (Network)					1.51			
					*2.52****			
Cluter Index (Network)						-3.60		
						*-3.42****		
Betw. Cenrality (Network)							-342.26	
							*-3.00****	
Dissimilarity Index (Network)								0.03
								0.82
R^2	**69.23**	**69.35**	**69.96**	**72.41**	**71.58**	**70.88**	**66.54**	**68.69**
Relative Impact of Transfer Account (in %)	15%	15%	78%	74%	52%	18%	50%	54%

Source: Own calculations based on daily network- and node-level ARTIS GSCC indicators and on daily simulations of the 50 banks and 13 Transfer accounts that were part of the ARTIS GSCC from 16 November 2005 to 16 November 2007 (excluding Austrian holidays). Numbers (1) to (8) provide results for the respective model specifications, each including the independent variables for which results are provided. These values are parameter estimates (upper) and corresponding z-values (lower). * denotes significance at the 90 percent confidence level, ** at 95 and *** at 99 percent.

In a first step, we estimate specification 1 – column (1) – as the most parsimonious model with only three explanatory variables: *Liquidity*, *Liquidity Loss,* and the dummy variable (*D×Liquidity Loss*). All three variables are highly significant and have the expected signs: higher values of aggregate liquidity in the system reduce contagion; large values of liquidity loss at the stricken bank increase contagion; and the contagion impact of operational shocks at Transfer accounts is significantly higher than that of the average participant in the system (as suggested by Chart 3 above).

These results are not particularly surprising, but it is reassuring for our approach that they are robust across all specifications. The goodness of fit of specification 1 is high with an R^2 of 69.23.[40] The relative impact of a stricken Transfer account is 15 percent higher than the average across stricken banks and Transfer accounts. The parameter values of the three explanatory variables are very robust across specifications. They remain highly significant in all specifications.

In a second step, we add further explanatory variables which either capture network or node characteristics. Since the various network indicators are highly correlated, we add only one indicator in each specification for, both, the node- and the network-level in Specifications 3 to 8. In Specification 2 we include a traditional measure of network size (*Value* of all transactions settled on a specific day) to contrast the explanatory impact of this traditional measure with more sophisticated network indicators (*degree*, *average path length* etc.). A higher value of transactions in the network is associated with lower contagion. However, the additional variable only has a very small impact on the explanatory power of the model (R^2 increases by 0.12 percentage points).

The network indicators at the node-level are highly significant. Operational shocks at nodes with higher *degree*, higher *connectivity*, higher *betweenness centrality* or higher *dissimilarity indices* cause higher contagion. Similarly, nodes with higher *average path length* and higher *cluster indices* feature lower contagion. In sum, more connected, more central nodes, and nodes with less mutually connected neighbours cause relatively more contagion, even when we compare them with (i) nodes which cause a similar liquidity loss, (ii) nodes that are either also banks / Transfer accounts, and (iii) nodes that experience operational shocks on days with the same level of aggregate liquidity. We conclude that in Model 1 the position of the stricken bank in the network has indeed an impact on contagion.

The network indicators at the network-level are significant as well (except for the average dissimilarity index). Lower *average degree*, *average connectivity*, *average cluster index* and *average betweenness centrality* are associated with a higher contagious impact of an operational problem at a given bank / Transfer account across days. Higher *average path length* implies a higher contagious impact. In sum, the denser the network is on a specific day, the lower the contagious impact on this day in Model 1. In Model 1 the structure of the network on a given day in the sample significantly influences contagious.

[40] The R^2 reported in panel data analysis differs from the OLS R^2. While they are still a useful measure of the model and its specification, they are not equal to the fraction of variation of the dependent variable explained by the estimated equation. They can be interpreted as squared correlations between the estimated and the observed dependent variable.

The additional pairs of explanatory variables do not seem to improve the goodness of fit of the model, though it is somewhat higher than for the simple measure of network activity *Value (Network)*. Adding connectivity at the node-level and at the network-level – specification (4) – increases R^2 from 69.23 to 72.41 percent. In this respect, the multivariate analysis indeed contradicts the univariate analysis in section 4 above.

The relative impact of Transfer accounts varies quite strongly across specifications ranging from roughly 15 percent – specifications (1), (2), and (6) – to almost 80 percent in specification (3). However, the relative impact is positive across all specifications and confirms the observation in Chart 3 above.

5.5 Results of Model 2 (Number of Unsettled Payments)

Table 7 summarises the results for the 8 specifications of Model 2. The dependent variable is the *number of unsettled payments* at the end of the day (excluding those of the stricken bank).

Table 7: Results Model 2 (Number of Unsettled Payments)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	7.32	9.90	-3.06	0.57	35.54	12.55	4.73	-0.36
	7.95*	***10.67****	***-0.93***	***0.34***	***6.38****	***6.93****	***2.13***	***-0.35***
Liquidity	-0.28	-0.25	-0.27	-0.20	-0.22	-0.25	-0.26	-0.29
	-5.65*	***-5.44****	***-6.18****	***-4.57****	***-4.75****	***-5.43****	***-5.29****	***-6.24****
Liquidity Loss	6.61	6.67	4.97	4.94	5.32	6.29	5.60	5.40
	41.91*	***41.84****	***24.31****	***24.17****	***28.22****	***38.61****	***22.59****	***26.62****
D x Liquidity Loss	2.40	2.38	3.10	3.10	2.94	2.45	2.91	2.84
	8.36*	***8.29****	***10.53****	***10.55****	***10.08****	***8.49****	***9.52****	***9.72****
Degree (Node)			0.18					
			13.51*					
Connectivity (Node)				24.59				
				13.85*				
Avg. Path Length (Node)					-17.09			
					-13.92*			
Cluter Index (Node)						-6.25		
						-10.84*		
Betw. Cenrality (Node)							53.96	
							5.65*	
Dissimilarity Index (Node)								19.00
								10.02*
Value (Network)		-0.07						
		-8.15*						
Degree (Network)			0.51					
			1.93					
Connectivity (Network)				30.93				
				0.86				
Avg. Path Length (Network)					1.21			
					0.56			
Cluter Index (Network)						-5.29		
						-1.4*		
Betw. Cenrality (Network)							426.79	
							1.09	
Dissimilarity Index (Network)								0.05
								0.52
R^2	**39.79**	**39.81**	**39.90**	**40.01**	**40.02**	**39.77**	**39.72**	**39.67**
Relative Impact of Transfer Account (in %)	36%	36%	62%	63%	55%	39%	52%	53%

Source: Own calculations based on daily network- and node-level ARTIS GSCC indicators and on daily simulations of the 50 banks and 13 Transfer accounts that were part of the ARTIS GSCC from 16 November 2005 to 16 November 2007 (excluding Austrian holidays). Numbers (1) to (8) provide results for the respective model specifications, each including the independent variables for which results are provided. These values are parameter estimates (upper) and corresponding z-values (lower). * denotes significance at the 90 percent confidence level, ** at 95 and *** at 99 percent.

In specification (1) the variables *Liquidity*, *Liquidity Loss*, and the dummy variable (*D×Liquidity Loss*) are highly significant and carry the expected signs. Again the parameters and the z-values are robust across all specifications. Contagion is lower on days with higher aggregate liquidity. It is higher when the stricken bank planned to transact a higher value of payments on the day of the simulated operational problem. Operational shocks to Transfer accounts have a higher contagion impact than those to bank aacounts. The relative impact is more than twice as high in specification (1) of Model 2 (36 percent) than in the comparable specification of Model 1 (15 percent). The goodness of fit of the specification is lower in Model 2 (39.79 percent) than in Model 1 (69.23 percent). The variation of the number of unsettled payments across days and across scenarios is harder to capture than that of the number of participants affected by contagion.

Similar to Model 1 all network indicators at the node-level are highly significant and have the same signs as in the respective specifications of Model 1. A participant with higher node *degree*, *connectivity*, *betweenness centrality* and *dissimilarity index,* but lower *average path length* and *cluster index* causes a higher contagion effect than participants with similar values of out-going payments on that day. The position of the stricken node within the network has a significant influence on the contagion caused by an operational shock, even after controlling for aggregate liquidity, for liquidity loss due to operational problems at the stricken bank, and for whether it is a bank or a Transfer account.

Turning to the network indicators at the network-level provides the following picture: only *Value (Network)* is significant at the 99 percent confidence level. Again, operational problems cause less contagion on days with more network activity. The *degree* is significant at the 95 percent confidence level, i.e. a more connected network is subject to more contagion. This is contradicts the finding in Model 1 and also the following finding in Model 2: The *clustering index* is significant at 90 percent level which implies that a network with nodes that have more mutually connected neighbours experiences less contagion. The other more sophisticated network indicators are not significant in Model 2.

The explanatory value of the additional explanatory variables in specifications (2) to (8) is very low. It increases from 39.79 percent in specification (1) to at most 40.02 in specification (5). We conclude that network indicators (both at the node and at the network-level) add little to explaining the variations of the number of unsettled payments across days and across scenarios.

The relative impact of an operational shock at a Transfer account varies strongly across specifications with a minimum of 36 and a maximum of 63 percent. The minimum is higher, but the maximum is lower in Model 2 than in Model 1. Nevertheless, the dummy is significant in all specifications.

5.6 *Results of Model 3 (Value of Unsettled Payments)*

Table 8 summarises the results for the 8 specifications of Model 3. The dependent variable is the value of unsettled payments at the end of the day (excluding the payments of the stricken bank).

Table 8: Results Model 3 (Value of Unsettled Payments)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	5.22E-02	7.37E-02	2.47E-02	3.63E-02	2.87E-01	5.94E-02	4.24E-02	5.35E-02
	5.3*	***7.16****	***0.59***	***1.7*****	***4.08****	***2.63****	***1.44****	***4.59****
Liquidity	-2.66E-03	-2.19E-03	-2.70E-03	-2.50E-03	-2.08E-03	-2.68E-03	-2.63E-03	-2.72E-03
	-5.01*	***-4.29****	***-5.08****	***-4.72****	***-3.57****	***-4.77****	***-4.8****	***-4.97****
Liquidity Loss	9.56E-02	9.60E-02	9.22E-02	9.04E-02	9.09E-02	9.72E-02	9.65E-02	9.57E-02
	47.6*	***47.77****	***35.69****	***35.17****	***38.17****	***45.99****	***33.87****	***36.91****
D x Liquidity Loss	1.59E-01	1.59E-01	1.60E-01	1.61E-01	1.61E-01	1.58E-01	1.58E-01	1.59E-01
	34.54*	***34.48****	***34.24****	***34.39****	***34.55****	***34.6****	***33.08****	***34.04****
Degree (Node)			3.33E-04					
			2.34**					
Connectivity (Node)				6.84E-02				
				3.63*				
Avg. Path Length (Node)					-5.76E-02			
					-4.42*			
Cluter Index (Node)						2.78E-02		
						3.66*		
Betw. Cenrality (Node)							-4.65E-02	
							-0.53	
Dissimilarity Index (Node)								-2.52E-03
								-0.12
Value (Network)		-6.24E-04						
		-5.49*						
Degree (Network)			1.76E-03					
			0.52					
Connectivity (Network)				7.28E-02				
				0.16				
Avg. Path Length (Network)					-5.32E-02			
					-1.92**			
Cluter Index (Network)						-5.39E-02		
						-1.11		
Betw. Cenrality (Network)							1.93E+00	
							0.36	
Dissimilarity Index (Network)								5.25E-04
								0.41
R^2	**70.62**	**70.63**	**70.64**	**70.65**	**70.69**	**70.68**	**70.60**	**70.61**
Relative Impact of Transfer Account (in %)	166%	166%	174%	178%	177%	163%	164%	166%

Source: Own calculations based on daily network- and node-level ARTIS GSCC indicators and on daily simulations of the 50 banks and 13 Transfer accounts that were part of the ARTIS GSCC from 16 November 2005 to 16 November 2007 (excluding Austrian holidays). Numbers (1) to (8) provide results for the respective model specifications, each including the independent variables for which results are provided. These values are parameter estimates (upper) and corresponding z-values (lower). * denotes significance at the 90 percent confidence level, ** at 95 and *** at 99 percent.

Specification (1) represents the basic Model 3 with the dependent variable *value of unsettled payments* and the independent variables *Liquidity*, *Liquidity Loss*, and the dummy variable (*D×Liquidity Loss*). All three variables are highly significant and have the expected signs. High aggregate liquidity in the system cushions the contagious effect of an operational shock. It increases with the value of payments that the stricken bank would have transferred under business as usual. Operational problems at Transfer accounts cause significantly more contagion than the average across accounts. The relative impact (166 percent), however, is much larger than in Model 1 (15 percent) and Model 2 (36 percent). Unlike in the other two models the relative impact is quite stable across specifications (1) to (8) ranging from 163 to 174 percent. The goodness of fit of specification (1) is high with an R^2 of 70.62.

Four out of the six network indicators at the node-level are significant, the most sophisticated (*betweenness centrality* and *dissimilarity index*) are not. The stricken banks' / Transfer accounts' *degree*, *connectivity*, and *average path length* have the same signs as in the respective specifications in Models 1 and 2. Again we conclude more central and more connected nodes cause more contagion. However, the cluster index is significant again but changes sign relative to two previous models, i.e. nodes with a higher fraction of mutually connected nodes cause more contagion.

With the exception of the traditional measure *Value (Network)* and *average path length* network indicators at the network-level are not significant. Those two, however, have the same signs as in the respective specifications in Models 1 (and 2). On days with a higher total transactions value and/or more dispersed network structure contagion is lower. However, the more sophisticated network indicators at the node and at the network-level seem to have little bearing on the goodness of fit. R^2 ranges from 70.60 in specification (7) to 70.69 in specification (5). The most parsimonious specification features an R^2 of 70.62.

5.7 Overall Results of the Multivariate Analysis

The explanatory value of all three models is high across all eight respective specifications. It ranges from about 40 percent for Model 2 to roughly 70 percent in Models 1 and 3. In the first case the variation within scenarios is higher than that between scenarios, while in the other two cases the opposite holds true. In conjunction with the higher R^2-values in Models 1 and 3 we conclude that our models are relatively better suited to explain the variation of the contagion impact between scenarios than across time within scenarios.

The three variables aggregate *Liquidity*, *Liquidity Loss*, and the dummy variable for Transfer accounts (*D×Liquidity Loss*) are highly significant across all models and across all specifications. They have the same sign in all cases. We regard the following results as robust:

- The contagion effect is lower on days with a higher aggregate liquidity in the system.[41]

[41] We re-estimated specifications (1) to (8) in Models 1 to 3 in sub-periods of the sample period. In one sub-period, *Liquidity* was not significant in some specifications.

- The contagion effect is lower in scenarios and on days with lower liquidity loss due to operational shocks.
- The system is significantly more vulnerable to operational shocks that hit Transfer accounts. These have special characteristics which are neither fully captured by their *Liquidity Loss* measure nor by the position within the network.

Three network indicators at the node-level are significant and have the same sign in all three Models 1 to 3: *degree* (+), *connectivity* (+), and *average path length* (-). We conclude that operational shocks at more connected and more central nodes cause more contagion even after controlling for variations in liquidity loss (which can also be regarded as indicator of their importance/activity in the payment system) and for whether they are Transfer accounts or banks.

None of the more sophisticated network indicators at the network-level is significant in all three models. Only the traditional measure of network size (*Value (Network)* the aggregate value of payments in the network in a given day) is significant across models and features the same sign. Days with higher transaction activity are associated with lower contagion even after controlling for aggregate liquidity, liquidity loss, and whether the shock hits a Transfer account or a bank.

The additional explanatory value of the network indicators at the node- and at the network-level seems to be very limited. The liquidity situation of the system, the liquidity loss due to the operational shock, and the hitherto unknown special characteristics of Transfer accounts can explain the variation of contagion across days and across scenarios already very well. The position of the stricken bank / Transfer account in the network and the structure of the network on the specific day of the shock add little explanatory value.

In further research we will focus on two main issues: first, the impact of different liquidity strategies of banks; in our approach aggregate liquidity conceals the potential impact of the distribution of liquidity in the system. Furthermore, we could run simulations on different levels of theoretical liquidity at individual banks and test for the impact of liquidity at the individual banks level. Second, we find that network indicators at the network level add little to a stability analysis within a given network. However, combining simulation data from different networks might reveal that network indicators play a more prominent role in stability analysis between networks than within networks.

6 Summary

The analysis of the network indicators of ARTIS shows that the network is compact. This is mostly due to the fact that almost all active nodes are linked to a small number of nodes at the centre of the network (the largest banks and the most active Transfer accounts). This network structure is quite stable across days.

We conducted 31 311 simulations based on 63 different scenarios for 497 transaction days from 16 November 2005 to 16 November 2007 (excluding Austrian holidays). Although the scenarios focus only on the banks and on the Transfer accounts that are part of the GSCC on all days, more than a quarter of all simulations do not lead to contagion (in terms of the number of banks with unsettled payments) at all, and two fifth yield only one or two contagious defaults. We arbitrarily define a conservative threshold for the systemic importance of an account based on the average contagion it causes across days. An account is deemed systemically important, if it causes an average contagious impact of at least EUR 48.5 million of unsettled payments (0.1 percent of average transactions settled across days). We find that only a very small number of participants are systemically important (seven percent of banks in the network and about 50 percent of Transfer accounts). The simulation results suggest that the ARTIS system is remarkably stable with respect to operational incidents at one of its participants. The strong contagion impact of the Transfer accounts is an interesting feature revealed by the simulations and suggests that the removal of Transfer accounts by the single shared platform in TARGET 2 can improve resilience relative to the old TARGET system.

The simulation results reveal that contagion varies strongly across scenarios and across days. In order to explore the determinants of variation we employ a panel data analysis. We test 8 specifications of three models (based on three different measures of contagion). Specification (1) in each model is the most parsimonious one with three independent variables: *Liquidity* (aggregate liquidity in the system on any given day), *Liquidity Loss* (the value of payments due by the stricken account on any given day), and a dummy variable (*D×Liquidity Loss*) for the Transfer accounts in the panel. We find that the contagion effect is lower on days with a higher aggregate liquidity in the system as well as in scenarios and on days with lower liquidity loss due to operational shocks. Operational shocks at Transfer accounts render the network significantly more vulnerable to operational shocks.

Over recent years payment system research has increasingly focused on network analysis. We apply our very rich data set to empirically test the interaction between network structure and network stability for the first time. Specifications (2) to (8) extend the basic model by including network indicators at the node- and at the network level. The results for the network indicators at the node level suggest that operational shocks at more connected and more central nodes cause more contagion. The results for the network indicators demonstrate that operational shocks on days with higher transaction activity cause lower contagion. These results are highly robust across models and across specifications. But more sophisticated network indicators at the network level are insignificant.

Furthermore, we find that the most parsimonious specification (1) features a high goodness of fit in Models 1 and 3 (dependent variables *number of participants with unsettled payments* and *value of unsettled payments*, respectively) and a slightly lower one in Model 2 (dependent variable *value of unsettled payments*). The additional explanatory value of the network indicators at the node and at the network level is very low, though. With respect to the interaction between network structure and network stability we conclude that the position of a stricken node within the network has an impact on network stability in the face of an operational shock, although the explanatory value is small. The results for network indicators at the network seriously question the hypothesis that variations in network structure (within a given payment system) are relevant for network stability.

Bibliography

Albert, R., H. Jeong, A.-L. Barabasi (1999) Diameter of the World Wide Web, Nature 401, 130-131.

Albert, R., H. Jeong, A.-L. Barabasi (2000) Error and attack tolerance of complex networks, Nature 406, 378-381.

Albert, R., A.-L. Barabasi (2002) Statistical mechanics of complex networks, Reviews of Modern Physics 74, 47-97.

Arellano, M. (2003) Panel Data Econometrics, Cambridge University Press, Cambridge.

Baltagi, B. H. (2001) Econometric Analysis of Panel Data, John Wiley & Sons, Chichester.

Beck, N., J. N. Katz. (1995) What to do (and not to do) with time-series cross-section data, American Political Science Review, Vol. 89 No. 4, 634–47.

Borgatti, S. P. (2005) Centrality and network flow, Social Networks 27, 55-71.

Boss, M., H. Elsinger, M. Summer, S. Thurner (2004) An empirical analysis of the network structure of the Austrian interbank market, OeNB Financial Stability Review 7, 77-87.

Breusch, T., A. Pagan (1980) The LM test and its Applications to Model Specification in Econometrics, Review of Economic Studies, Vol. 47 No. 1, 239-53.

DeGroot, M.H. (1985) Probability and Statistics, Second Edition, Addison-Wesley. Reading, Massachusetts.

Frees, E. W. (1995) Assessing cross-sectional correlations in panel data, Journal of Econometrics, Vol. 69 No. 2, 393-414.

Friedman, M. (1937) The use of ranks to avoid the assumption of normality implicit in the analysis of variance, Journal of the American Statistical Association, Vol. 32 No. 200, 675-701.

Inaoka, H, T. Ninomiya, K. Taniguchi, T. Shimizu, H. Takayasu (2004) Fractal Network derived from banking transaction - An analysis of network structures formed by financial institutions, Bank of Japan Working papers No. 04-E-04.

Kyriakopulos, F., S. Thurner, C. Puhr, S. W. Schmitz (2009) Network and eigenvalue analysis of financial transaction networks, European Physical Journal B (forthcoming)

Latzer, M., Schmitz, S. W. (eds.) (2002) Carl Menger and the Evolution of Payment Systems. From Barter to Electronic Money, Edward Elgar, Cheltenham.

Leinonen H. (ed.) (2005) Liquidity, risk and speed in payment and settlement systems – a simulation approach, Bank of Finland Studies E:31, Helsinki.

Newman, M. E. J. (2003) The structure and function of complex networks, (available at http://arxiv.org/abs/cond-mat/0303516).

Newman M. E. J. (2005) Power Laws, Pareto Distributions, and Zipf's Law, Contemporary Physics 46, 323-351.

Oesterreichische Nationalbank and Finanzmarktaufsichtsbehörde (2004) The Austrian Financial Markets, Vienna.

Pesaran, M. H. (2004) General diagnostic tests for cross-sectional dependence in panels, University of Cambridge, Cambridge Working Papers in Economics 0435.

Schmitz, S. W., C. Puhr (2006) Liquidity, Risk Concentration and Network Structure in the Austrian Large Value Payment System. Available at SSRN: http://ssrn.com/abstract=954117.

Schmitz, S. W., C. Puhr (2007) Risk concentration, network structure and the assessment of contagion in the Austrian large value system ARTIS, in: H. Leinonen (ed.), Simulation studies of the liquidity needs, risks and efficiency in payment network, Bank of Finland Scientific Monograph E:39, Helsinki,183-226.

Schmitz, S. W., C. Puhr, M. Boss, G. Krenn, V. Metz (2008), Systemically Important Accounts, Network Topology and Contagion in ARTIS. Available at SSRN: http://ssrn.com/abstract=1137864.

Schmitz, S. W., G. E. Wood (eds.) (2006), Institutional Change in the Payments System and Monetary Policy, Routledge, London.

Soramäki, K., M. L. Bech, J. Arnold, R. J. Glass, W. E. Beyeler (2006) The Topology of Interbank Payment Flows, Federal Reserve Bank of New York Staff Report, New York No. 243.

Soramäki, K., W. E. Beyeler, M. L. Bech, R. J. Glass (2007) New approaches for payment system simulation research, in: H. Leinonen (ed.) Simulation studies of the liquidity needs, risks and efficiency in payment network, Bank of Finland Scientific Monograph E:39, Helsinki, 15-40.

Wooldridge, J. M. (2002) Econometric Analysis of Cross Section and Panel Data, MIT Press, Cambridge.

Zhou, H. (2003) Distance, dissimilarity index, and network community structure, Physical Review E 67, 061901, 1-8.

Annex 1 – Test results

The assumption of conditional homoskedasticity is tested by a likelihood ratio (LR) test which compares the log-likelihoods under the restricted model (homoskedastic errors) and the unrestricted model (heteroskedastic errors). Both models are estimated by iterated generalised least squares (IGLS). The tests statistics clearly reject the assumption of conditional homoskedasticity for all three models in specification (1). (This is to be expected, as some scenarios hardly generate contagion, so that the variance is extremely low.) The resulting test statistics and error probabilities are shown in table 9.

Table 9: Test Results – Likelihood Ratio Test for Conditional Homoskedasticity

	LR Chi²(62)	Prob.
Model 1	18501	0.00
Model 2	103015	0.00
Model 3	74980	0.00

Source: Own calculations based on data and model specifications as presented in Section 5. Likelihood Ratio (LR) and Error Probability (Prob.).

In order to test for the assumption of serial independence we conduct a Wooldridge test for all three models in specification (1). The test is based on residuals of regressions in first differences which are then regressed on their lagged values at t-1. The test is robust to conditional heteroskedasticity. The tests statistics reject the assumption of serial independence for Models 1 and 3 in specification (1). The resulting test statistics and error probabilities are shown in table 10.

Table 10: Test Results – Wooldrige Test

	F(1,62)	Prob.
Model 1	20.3	0.00
Model 2	2.3	0.14
Model 3	13.5	0.00

Source: Own calculations based on data and model specifications as presented in Section 5. Error Probability (Prob.).

Three tests are available to check for cross-panel independence. They were suggested by Friedman (1937), Frees (1995), and Pesaran (2004), respectively. All three test statistics reject the assumption of cross-panel independence for all three models in specification (1) (Table 11).

Table 11: Test Results – Friedman, Fress, and Pesaran Tests

	Friedman	Prob.	Frees	Prob.	Pesaran	Prob.
Model 1	11728.05	0.00	11.12	0.00	363.11	0.00
Model 2	7378.70	0.00	7.06	0.00	147.08	0.00
Model 3	5744.16	0.00	4.81	0.00	120.80	0.00

Source: Own calculations based on data and model specifications as presented in Section 5. Error Probability (Prob.).

The data exhibits high correlation between time-invariant unobservable scenario specific effects ν and the explanatory variables. Consequently, we test for fixed versus random effects. Table 13 presents the results of Breusch-Pagan (1980) likelihood ratio (LR) tests of random effects for all three models in specification (1).

Table 12: Test Results – Breusch-Pagan Likelihood Ratio Test

	LR	Prob.
Model 1	306000	0.00
Model 2	43300	0.00
Model 3	230000	0.00

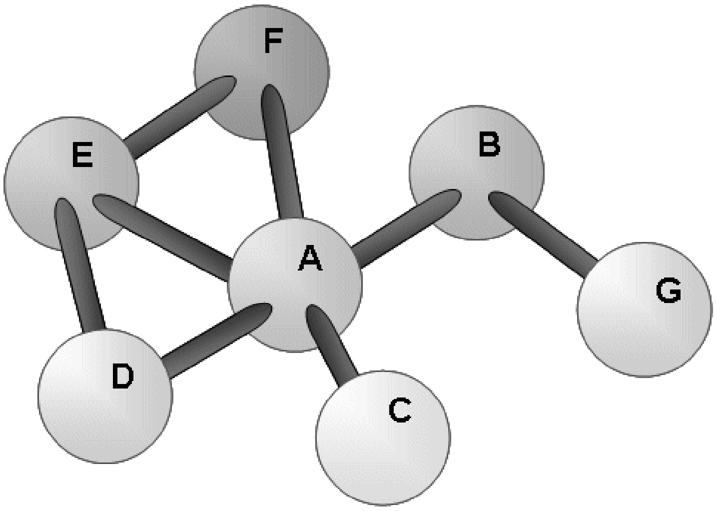
Source: Own calculations based on data and model specifications as presented in Section 5. Likelihood Ratio (LR) and Error Probability (Prob.).

Given the results of all four tests, an ordinary least squares (OLS) estimate of a standard fixed-effects model would yield inconsistent and biased standard errors and we have to apply an estimator that can handle conditional heteroskedasticity, serial correlation, and cross-sectional dependence of the error terms ε.

Annex 2 – Definition of network indicators

This annex summarises the definitions and formulas of the network indicators used in the paper. In addition, we provide a simple illustrative example of a network which allows us to visualise different values of network indicators.

Chart 4: A Simple Network Example



Source: Own calculations.

The network in Chart 4 is an undirected, unweighted network with 7 nodes (*number of nodes* $n=7$) and 8 undirected, unweighted links (*number of links* $m=16$[42]). Table 13 summarizes the relevant network indicators from our paper for this network.

Table 13: A Simple Network Example – Node-Level Network Indicators

	A	B	C	D	E	F	G	Network
(Average)* Degree	5	2	1	2	3	2	1	2.3
(Average)* Connectivity	83.3%	33.3%	16.7%	33.3%	50.0%	33.3%	16.7%	38.1%
(Average)* Clustering Coefficient	20.0%	0.0%	0.0%	100.0%	66.7%	100.0%	0.0%	41.0%
Average Path Length	1.2	1.7	2.0	1.8	1.7	1.8	2.5	1.8
(Average)* Betweenness Centrality	85.7%	35.7%	0.0%	0.0%	7.1%	0.0%	0.0%	18.4%
(Average)* Dissimilarity Index	1.5	1.6	1.5	1.4	1.5	1.4	3.1	1.7

*) While the node level indicator of node h is not an average, the corresponding indicator on the network level is indeed the average across all nodes n.

Source: Own calculations.

The network topology indicators[43]

The ***degree*** k_h of node *h* is measured by the number of links originating (*out-degree*) or terminating (*in-degree*) at node *h*. In our case of an undirected network, *in-* and *out-degree* are actually the same. Take for instance node A, which has links to nodes B, C, D, E, and F, hence node A has a *degree* $k_A=5$ while the respective value for node B is $k_B=2$; i.e. node A is linked to 5 other nodes and B only to two.

[42] There are eight connections in our network, between node i and j, and each of these can be seen as a link form node i to node j as well as from node j to node I, hence *m*, the total *number of links* is 16.

[43] Where possible we follow the notation of Albert, Barabasi 2002, Soramäki et al. 2006, Zhou 2003.

On the network level, the *average degree* k of the network is calculated by summing across all links originating from each node (*out-degree* k_i^{out}) or terminating at each node (*in-degree* k_i^{in}) and than averaging across nodes $k = \frac{1}{n}\sum_i k_i^{out} = \frac{1}{n}\sum_i k_i^{in} = \frac{m}{n}$. In our example the *average degree* of the network is 2.3, based on dividing 16 direct links *m* by our seven nodes *n*.

The ***connectivity*** of node *h* is its *degree* over the *number of nodes n*. In our example, node A has *connectivity*=83% while the respective value for node B is *connectivity*=33%; i.e. *connectivity* puts the absolute value of *degree* in relation to the size of the network (as measured by the *number of nodes n*). On a network level, *average* ***connectivity*** is defined by the number of actual (directed) links *m* over the number of possible (directed) links $n(n-1)$. In our example the *average connectivity* of the network is 38.1%, based on dividing 16 direct links *m* by 42 potential (directed) links between our *n* nodes.

The ***clustering coefficient*** $C_C(h)$ of an individual node *h* with k_h neighbours measures how well the latter are connected among each other. The number of potential links between the k_h neighbours is $k_h(k_h-1)/2$. Let the actual number of nodes between them be E_h so that $C_C = \frac{E_h}{k_h(k_h-1)/2}$.

The *clustering coefficient* of node A $C_C(A)$=20% and that of B $C_C(B)$=0%. I.e. node A has 5 direct neighbours so that the potential number of direct links is 10, but only two direct links exist (D-E and E-F). Therefore, the *clustering coefficient* of A is 20%. B has two neighbours with one possible direct link, but A and G are not linked so that the *clustering coefficient* is 0%.

The *average clustering coefficient* of the network C_C is the average of all individual nodes' *clustering coefficients* $C_C(i)$ and is hence defined as $C_C = \frac{1}{n}\sum_i C_C(i)$. In our example the *average degree* is 41.0%, based on our *n* nodes individual *clustering coefficients*.

An indicator of the *distance* d_{ij} between nodes is the lowest possible number of links that connects each node *i* with each other node *j* in the network. It is referred to as *shortest path length*.

We calculated the ***average path length*** ℓ_h for the originating node *h* by averaging across the *shortest path length* to each terminating node *i*. Therefore ℓ_h is defined as $\ell_h = \frac{1}{n-1}\sum_{h \neq i} d_{hi}$.

In the example the *average path length* of node A is much lower (ℓ_A=1.2) than that of node B (ℓ_B=1.7), i.e. from node A any other node can be reached on the shortest path via an average of 1.2 links, while it takes 1.7 links from node B.

For the entire network, the *average path length* ℓ is defined as the *average path length* across all originating nodes ℓ_i divided by our seven nodes *n*, formally written as $\ell = \frac{1}{n}\sum_i \ell_i$

The ***betweenness centrality*** $C_B(h)$ of node *h* provides a measure of how many *shortest paths* d_{ij} pass through node *h*. Let $s_{ij}(h)$ be the number of *shortest paths* between all pairs of nodes *i* and *j* that pass through the node *h* and let s_{ij} the number of all *shortest paths* between all pairs of nodes *i* and *j* then

$$C_B(h) = \sum_{h \neq i \neq j} \frac{s_{ij(h)}}{s_{ij}}.$$

In our example there are 44 *shortest paths*. The lower boundary is given by 42 possible (directed) links $n(n-1)$. To these 42 we have to add another 2, as the *shortest path* from D-F and vice versa could either pass through A or E. By definition, we then have to exclude those 16 paths that link directly neighbouring nodes, which leaves us with 28 *shortest paths* in the denominator. In our example 24 of those pass through node A and 10 pass through node B. Therefore their betweenness centralities are given by $C_B(A)$=85.7% and $C_B(B)$=35.7%.

For the entire network, the *average betweenness centrality* C_B is defined the average of all individual nodes' *betweenness centralities* $C_B(i)$ and is hence defined as $C_B = \frac{1}{n} C_B(i)$. In our case the *average betweenness centrality* is 18.4%, based on our *n* nodes individual *betweenness centrality*.

Finally, the ***dissimilarity index*** of two neighbours nodes *i* and *j* is defined as

$$\Delta_{ij} = \frac{\sqrt{\left[\sum_{h \neq i,j}^{N} \left[d_{ih} - d_{jh}\right]^2\right]}}{(N-2)},$$

where d_{ih} and d_{jh} are distance measures from nodes *i* and *j* to node *h*. It provides a comparison of the view of the entire network from the perspective of all pairs of neighbouring nodes. For the entire network the *average dissimilarity index* is $\Delta = \frac{1}{n(n-1)/2} \Delta_{ij}$.

Our example serves to illustrate that, although in most other indicators node A is distinctly different from all other nodes in the network (particularly regarding our other "advanced" network indicator, *betweenness centrality*), the view on the rest of the network is quite alike as A is the central node in a cluster. Hence, its *dissimilarity index* of 1.5 aligns it more or less with all of its direct neighbours. Node F, which's remote position might have gone unnoticed so far (at least based on other network indicators), is shown to be vastly different from the other nodes, however. These distinct features of *betweenness centrality* and *dissimilarity index* are also the reason why we introduced them in the first place.