Kiel

Working Papers

Kiel Institute
for the World Economy

On the Distribution of Links in the Interbank Network: Evidence from the e-Mid Overnight Money Market

by Daniel Fricke and Thomas Lux

No. 1819 | 01/2013

Kiel Working Paper No.1819 | 01/2013

Title: On the Distribution of Links in the Interbank Network: Evidence from the e-Mid Overnight Money Market

Author: Daniel Fricke and Thomas Lux

Abstract: Previous literature on statistical properties of interbank loans has reported various power-laws, particularly for the degree distribution (i.e. the distribution of credit links between institutions). In this paper, we revisit data for the Italian interbank network based on overnight loans recorded on the e-MID trading platform during the period 1999-2010 using both daily and quarterly aggregates. In contrast to previous authors, we find no evidence in favor of scale-free networks. Rather, the data are best described by negative Binomial distributions. For quarterly data, Weibull, Gamma, and Exponential distributions tend to provide comparable fits. We find comparable results when investigating the distribution of the number of transactions, even though in this case the tails of the quarterly variables are much fatter. The absence of power-law behavior casts doubts on the claim that interbank data fall into the category of scale-free networks.

Keywords: interbank market, network models
JEL classification: G21, G01, E42

Prof. Dr. Thomas Lux
Kiel Institute for the World Economy
Telephone: +49 431 8814 278
E-mail: Thomas.lux@ifw-kiel.de
24100 Kiel, Germany

Daniel Fricke
Kiel Institute for the World Economy
Telephone: +49 431 8814
Email: Daniel.Fricke@ifw-kiel.de
24100 Kiel, Germany

Christian-Albrechts-University Kiel
Department of Economics
Chair of Monetary Economics and International Finance
24098 Kiel, Germany

Bank of Spain chair of Computational Finance
Department of Economics, University Jaume I
Castellón, Spain

This paper is part of a research initiative launched by the Leibniz Community. We are grateful for helpful comments by Aaron Clauset and Michael Stumpf

On the Distribution of Links in the Interbank Network: Evidence from the e-MID Overnight Money Market.[†]

Daniel Fricke[‡§] Thomas Lux[‡§¶]

This version: January 2013

Abstract

Previous literature on statistical properties of interbank loans has reported various power-laws, particularly for the degree distribution (i.e. the distribution of credit links between institutions). In this paper, we revisit data for the Italian interbank network based on overnight loans recorded on the e-MID trading platform during the period 1999-2010 using both daily and quarterly aggregates. In contrast to previous authors, we find no evidence in favor of scale-free networks. Rather, the data are best described by negative Binomial distributions. For quarterly data, Weibull, Gamma, and Exponential distributions tend to provide comparable fits. We find comparable results when investigating the distribution of the number of transactions, even though in this case the tails of the quarterly variables are much fatter. The absence of power-law behavior casts doubts on the claim that interbank data fall into the category of scale-free networks.

JEL-Code: G21, G01, E42
Keywords: interbank market, network models

[†]The article is part of a research initiative launched by the Leibniz Community. We are grateful for helpful comments by Aaron Clauset and Michael Stumpf.

[‡]Department of Economics, University of Kiel, Olshausenstr. 40, 24118 Kiel
[§]Kiel Institute for the World Economy, Hindenburgufer 66, 24105 Kiel.
[¶]Banco de España Chair in Computational Economics, University Jaume I , Campus del Riu Sec, 12071 Castellon.

1 Introduction and Existing Literature

Since the onset of the global financial crisis (GFC) in 2007/08, the analysis of network structures formed by interbank liabilities has received increasing attention. Considering an ensemble of financial institutions, individual banks are connected to each other through some of their activities (usually credit flows) and the bilateral exposures can be mapped into a credit network. Such a perspective is useful in order to study the knock-on effects on other banks due to disruptions of the system caused by the failure of individual nodes (e.g. insolvency of one bank). A new strand of literature has started to construct financial networks based on empirical data available at supervisory authorities or hypothetical network structures to investigate the contagious effects of failures of single banks.[1] A basic finding of network theory is that the topology of a network is important for its stability, with the interbank network obviously being no exception.[2] In this regard, the understanding of the structure and functioning of complex networks has advanced significantly in recent years.

In this paper, we focus on one of the most prominent network characteristics, namely the degree distribution, where the degree is the number of (incoming/outgoing) connections per node. Even though the degree distribution does not provide sufficient information for all facets of the structure of the network (Alderson and Li, 2007), it is often considered as one of the defining characteristics of different network types. For example, networks with random link formation (Erdös and Renyi, 1959, or ER random networks) display Poisson degree distributions, i.e. most nodes have degrees within a relatively narrow range. In contrast, many real-world networks have been reported to display fat-tailed degree distributions: most nodes have a very small degree, but the tail contains nodes with substantially larger degrees (cf. Clauset *et al.*, 2009). This feature is shared by the important class known as scale-free (SF) networks, in which the fraction of nodes with degree k is proportional to $k^{-\alpha}$, where α is the so-called scaling parameter. The term scale-free indicates that there is no typical scale of the degrees, i.e. the mean may not be representative. These networks received considerable attention in the literature due to a number of interesting properties (cf. Caldarelli, 2007). One important feature of scale-free networks is that they can be described as robust-yet-fragile,[3] indicating that random disturbances are easily absorbed (robust) whereas targeted attacks on the most central nodes may lead to a breakdown of the entire network (fragile). Quite interestingly,

[1] See e.g. Upper and Worms (2004), Nier *et al.* (2007), and Gai *et al.* (2011).

[2] See Haldane and May (2011) and Albert *et al.* (2000).

[3] See Albert *et al.* (2000).

many interbank networks have been reported to resemble scale-free networks (cf. Boss *et al.*, 2004, Soramäki *et al.*, 2006, De Masi *et al.*, 2006, and Iori *et al.*, 2008). If the network of credit relationships had such a structure, this would carry important policy implications. For instance, such a network might experience long stable periods, during which disruptions are confined to peripherical banks and can be absorbed easily within the entire system. However, such periods could be a misleading indicator of the overall stability of the system as problems affecting the most central nodes could suddenly cause a breakdown of the entire network, cf. Haldane (2009).

The distribution of network degrees is just one example among many phenomena in the natural sciences as well as from the socio-economic sphere that have been claimed to follow a scaling law (power-law or Pareto-law). Other well-known examples include: Zipf's law for the city size distribution (Gabaix, 1999), the distribution of firm sizes (Axtell, 2001), the size distribution of innovations (Silverberg and Verspagen, 2007), the distribution of output growth-rates (Fagiolo *et al.*, 2008) or the distribution of large asset returns (Mandelbrot, 1963, Lau *et al.*, 1990, and Jansen and de Vries, 1991). While these examples appear to be supported by empirical evidence and meanwhile count as stylized facts, a variety of other findings of power-laws seem more questionable. It appears from a number of recent reviews of power-law methodology and power-law findings (cf. Avnir *et al.*, 1998, Stumpf and Porter, 2012) that there had been an over-emphasis on scaling laws and often too optimistic interpretation of statistical findings in the literature of the natural sciences. For instance, in a meta-study of power-laws reported in publications in the main physics outlet *Physical Review* between 1990 and 1996, Avnir *et al.* (1998) found that most claims of power-laws (aka scaling or fractal behavior) had a very modest statistical footing. As they say '... the scaling range of experimentally declared fractality is extremely limited, centered around 1.3 orders of magnitude.' In terms of statistics jargon this means that the more typical declaration of a power-law in these publications is based on a partially linear slope in a relatively small intermediate range of the empirical cumulative distribution of some observable.

The power-law exponent (like the ones reported for the degree distribution) is typically obtained by a linear regression in a log-log plot of the cumulative distribution. Obviously, this approach suffers from a number of shortcomings: (i) even if the hypothetical data-generating process is a Pareto distribution, this log-log fit would not be an efficient way to extract the parameter of the underlying distribution.[4] It is actually a method that is

[4]See Goldstein *et al.* (2004). Gabaix and Ibragimov (2011) improve the regression method by shifting the rank observations.

definitely inferior to maximum likelihood (which is easy to implement), and results are hard to interpret as, due to the dependency of observations in the log-log plot of the *cumulative* distribution, the statistical properties of this estimator are not straightforward, (ii) the implicit censoring of the data that is exerted by selecting a *scaling* range makes it easy to deceive oneself. Many distributions might actually have some intermediate range in their 'shoulders' where their cdf looks appropriately linear. But their remaining support (small and large realizations) might display a completely different behavior. Since power-laws in the natural sciences are thought to be interesting if they extent over several orders of magnitude, it is unclear what the interpretation of such an intermediate power-law approximation would be.

Statistical extreme value theory (EVT) provides yet another perspective on power-law behavior. The basic result of this branch of statistics is a complete characterization of the limiting distributions of extremes (maximum or minimum) of time series of iid observations (where results for the iid case have been generalized for dependent processes under relatively mild conditions, cf. Leadbetter, 1983, and Reiss and Thomas, 2007, for details). According to EVT, the appropriately scaled minimum or maximum of a series of observations converges in distribution to one of only three functional forms: the Fréchét, Gumbel or inverted Weibull distribution. Since extremes are by definition very rare, it is often even more relevant, that the tail of a distribution converges in distribution in a similar way to one of three adjoint functional forms. Namely, the outer part approaches either a power-law decay, an exponential decay or a decay towards a fixed endpoint for the three types of extremal behavior, respectively. Power-law behavior is, therefore, a very general form of limiting behavior for the large realizations of a stochastic process. EVT has originally been developed for continuous distribution function. Since degree distributions are discrete (degrees being integer numbers), its is worthwhile to note that corresponding limit laws for discrete variables are available as well, cf. Anderson (1970). In our context this might imply that very large realizations of the degree distribution could still decay like a power-law even if the bulk of the distribution does not appear to follow such a distribution (and the implications for the fragility of the system might be similar as for 'true' scale-free networks). It is important to emphasize that both the limiting behavior of extremes and tails are stable under aggregation. Hence, data at different levels of (time-) aggregation should obey the same extreme value and tail behavior.

One reason for the 'popularity' of power-laws in the natural sciences is that they are often the signatures of relatively simple and robust generating mechanisms that might apply to a variety of phenomena. In the case of networks, a power-law distribution of degrees is the imprint of so-called scale-free networks. Reported power-laws for interbank networks have been within a relatively narrow range around 2.3 both for the in- and out-degree distributions (see e.g. Boss *et al.*, 2004, Soramäki *et al.*, 2006, and De Masi *et al.*, 2006), even though most papers lack a thorough statistical analysis of the issue, with Bech and Atalay (2010) being a notable exception. If these findings were robust, the known generating mechanisms for scale-free networks would be strong candidates as mechanisms for the formation of interbank links. Furthermore, the well-known reactions of scale-free networks to disturbances would be of immediate concern for macro-prudential regulation. Thus, taking into account the relevance of such topological features, and the documented over-emphasis on power-law behavior, a more rigorous statistical analysis of the distributional properties of interbank network data should be worthwhile. Similar approaches have revealed that numerous previous claims of power-law behavior were not supported by the data (Stumpf and Porter, 2012).

In this paper, we consider interbank networks based on the Italian e-MID (electronic market for interbank deposits) data for overnight loans during the period 1999-2010. Our main focus is to fit a set of different candidate distributions to the degrees for different time horizons. Using daily data over the period 1999-2002, De Masi *et al.* (2006) reported power-laws for the distribution of in- and out-degrees, with tail parameters 2.7 and 2.15, respectively. Finger *et al.* (2012) have shown recently that the networks' properties depend on the aggregation period.[5] We will, therefore, not confine our analysis to daily data (the basic frequency of our data set), but also look at the distribution of in- and out-degrees for networks constructed on the base of aggregated data over longer horizons. Quite surprisingly in view of the previous literature, we find hardly any support in favor of previously reported power-laws: at the daily level the degrees are usually fit best by negative Binomial distributions, while the power-law may provide the best fit for the tail data. However, we typically find very large power-law exponents (with values as large as 7), i.e. levels where the power-law is virtually indistinguishable from exponential decay. At the quarterly level, Weibull, Gamma, and Exponential distributions tend to provide comparable fits for the complete degree distribution, while the tails again tend to display exponential decay. We find comparable results when investigating the distribution of the number of transactions, even though in this case the tails of the quarterly variables are somewhat fatter. However, the Log-normal distribution

[5]Since we cannot easily observe the state of a hypothesized network of interbank links at a given point in time, some data aggregation is necessary. Usually, for time-aggregated data a link is assumed to exist between two banks, if there has been a trade at any time during the aggregation period.

typically outperforms the power-law. Overall these findings indicate that the power-law is typically a poor description of the data, implying that preferential attachment and other generating mechanisms for scale-free networks are unsuitable explanatory mechanisms for the structure of the Italian interbank network. Moreover, the networks contain a substantial level of asymmetry, due to the low correlation between in- and out-degrees. Additionally, we find that the two variables do not follow identical distributions in general.

The remainder of this paper is structured as follows: Section 2 gives a short introduction into (interbank) networks, section 3 briefly introduces the Italian e-MID trading system and gives an overview of the data set we have access to. Section 4 describes our findings and section 5 concludes and discusses the relevance of these findings for future research.

2 Networks

A network consists of a set of N nodes that are connected by M edges (links). Taking each bank as a node and the interbank positions between them as links, the interbank network can be represented as a square matrix of dimension $N \times N$ (data matrix, denoted $\mathbf{D}$). An element d_{ij} of this matrix represents a gross interbank claim, the total value of credit extended by bank i to bank j within a certain period. The size of d_{ij} can thus be seen as a measure of link intensity. Row (column) i shows bank i's interbank claims (liabilities) towards all other banks. The diagonal elements d_{ii} are zero, since a bank will not trade with itself.[6] Off-diagonal elements are positive in the presence of a link and zero otherwise.

Interbank data usually give rise to directed, sparse and valued networks.[7] However, much of the extant network research ignores the last aspect by focusing on binary adjacency matrices only. An adjacency matrix $\mathbf{A}$ contains elements a_{ij} equal to 1, if there is a directed link from bank i to j and 0 otherwise. Since the network is directed, both $\mathbf{A}$ and $\mathbf{D}$ are asymmetric in general. In this paper, we also take into account valued information by using both the raw data matrix as well as a matrix containing the number of trades between banks, denoted as $\mathbf{T}$. In some cases it is also useful to work with the undirected version of the adjacency matrices, $\mathbf{A}^u$, where $a^u_{ij} = \max(a_{ij}, a_{ji})$.

As usual, some data aggregation is necessary to represent the system as a

[6] This is of course only true when taking banks as consolidated entities.

[7] Directed means that $d_{i,j} \neq d_{j,i}$ in general. Sparse means that at any point in time the number of links is only a small fraction of the $N(N-1)$ possible links. Valued means that interbank claims are reported in monetary values as opposed to 1 or 0 in the presence or absence of a claim, respectively.

network. In the following, we define interbank networks by aggregating over daily as well as quarterly lending activity.

3 The Italian Interbank Market (e-MID)

The Italian electronic market for interbank deposits (e-MID) is a screen-based platform for trading of unsecured money-market deposits in Euros, US-Dollars, Pound Sterling, and Zloty operating in Milan through e-MID SpA.[8] The market is fully centralized and very liquid; in 2006 e-MID accounted for 17% of total turnover in the unsecured money market in the Euro area, see European Central Bank (2007). Average daily trading volumes were 24.2 bn Euro in 2006, 22.4 bn Euro in 2007 and only 14 bn Euro in 2008. We should mention that researchers from the European Central Bank have repeatedly stated that the e-MID data is representative for the interbank overnight activity, cf. Beaupain and Durré (2012).

Detailed descriptions of the market and the corresponding network properties can be found in Finger *et al.* (2012).[9] In this paper we used all registered trades in Euro in the period from January 1999 to December 2010. For each trade we know the banks' ID numbers (not the names), their relative position (aggressor and quoter), the maturity and the transaction type (buy or sell). The majority of trades is conducted overnight and due to the global financial crisis (GFC) markets for longer maturities essentially dried up. We will focus on all overnight trades conducted on the platform, leaving a total number of 1,317,679 trades. If not stated otherwise, the reported results are based on trades conducted between Italian banks only, reducing the total number of trades to 1,215,759.

4 Results

In this section we present empirical results on the dynamics and distribution of the number of links (degrees) and the number of transactions (ntrans) of individual institutions. The degree of a node gives the total number of links that a bank has with all other banks and can thus be seen as a measure for the importance of individual nodes. Undirected networks imply symmetric adjacency matrices. In this case bank i's total degree k_i is simply the number

[8] The vast majority of trades (roughly 95%) is conducted in Euro.

[9] See also the e-MID website http://www.e-mid.it/.

of relationships bank i has with other banks, i.e.

$$k_i^{total} = \sum_{j \neq i} a_{ij}^u. \tag{1}$$

For directed networks, we differentiate between incoming links (bank i borrows money from other banks) and outgoing links (i lends money to other banks), and define the in- and out-degree of i (k_i^{in} and k_i^{out}) as

$$\begin{aligned} k_i^{in} &= \sum_{j \neq i} a_{ji} \\ k_i^{out} &= \sum_{j \neq i} a_{ij}, \end{aligned} \tag{2}$$

respectively. Note that our networks contain only banks with at least one (directed) link. In this way, the total degree of a sample bank is always at least equal to one, while it may be the case that either the in- or out-degree equals zero for a particular bank. Since we ignore zero values in the distribution fitting approach, this affects the number of observations for the different variables.

For the number of transactions, we use similar definitions based on the $\mathbf{T}$ matrix, with each element $t_{i,j}$ giving the number of trades with credit extended from bank i to bank j. To be precise, we calculate the number of in-/out-transactions as

$$\begin{aligned} n_i^{in} &= \sum_{j \neq i} t_{ji} \\ n_i^{out} &= \sum_{j \neq i} t_{ij}. \end{aligned} \tag{3}$$

Additionally, we analyze the total number of transactions, for simplicity defined as the sum of in- and out-transactions

$$n_i^{total} = n_i^{in} + n_i^{out}. \tag{4}$$

4.1 Dynamics of the Degrees and Number of Transactions

Before investigating the distribution of the variables under study, we provide a brief overview of their dynamics over time, restricting ourselves to quarterly data here. Figure 1 shows the in-/out-degrees (left) from the directed networks and the total degrees from the undirected networks (right). The upper left panel shows the mean and median in-degree and out-degrees over time.[10] Clearly, the mean values are decreasing over time, and so does the median in-degree which is mostly very close to the mean value. For both series we find a significant structural break after quarter 10. In contrast, the median out-degree fluctuated around an average value of roughly 17 over most of the sample period, but with a significant structural break after quarter 39 due to the GFC. These values are considerably smaller than the values for the in-degree, pointing towards a substantial level of skewness in the out-degree distribution. Thus, the distributions of in- and out-degrees are likely to be not identical. The lower left panel shows the relative mean and median degree over time, i.e. the values in the upper panel standardized by the number of nodes active in each quarter. We see that the negative trend in the upper panel is mostly driven by the negative trend in the number of active banks. Thus, the standardization appears to make the in-degrees of different quarters comparable. This is less so for the median out-degree, which is far more volatile over the sample period.[11] For the sake of completeness, the corresponding values for the degrees from the undirected networks are shown on the right-hand side. Both for the absolute and relative values the mean and median values are very similar, except for the beginning of the sample period. This is driven by the high level of asymmetry in the out-degree distribution for the first half of the sample, which appears to decrease later on.

What does the evidence on the differences between the in- and out-degree distributions imply? Given that many studies on interbank markets work with undirected networks, these studies entail the implicit assumption of a high correlation between in- and out-degrees of individual banks. The left panel of Figure 2 shows a scatter-plot of in-degree against out-degree for Italian banks, showing a small correlation of .0899 for all observations. For single quarters, we find that the correlation between these measures may be very small, at times even negative. Thus, banks with a high in-degree do not necessarily have a high out-degree and vice versa. The directed version of the network contains a considerable amount of information. The right panel of Figure 2 indeed shows a relatively monotonic decline of the correlation over time. This implies that banks have become more ‘specialized’, i.e. in any quarter they appear to enter the market predominantly as lenders or borrowers.

For the number of transactions, Figure 3 shows the dynamics of the mean and median in-/out-ntrans (left) and the total ntrans (right). The upper left

[10] Note that the mean in- and out-degree are identical by definition.

[11] Interestingly, after standardizing the degrees, we find structural breaks in all three time series close to quarter 39, i.e. around the GFC.

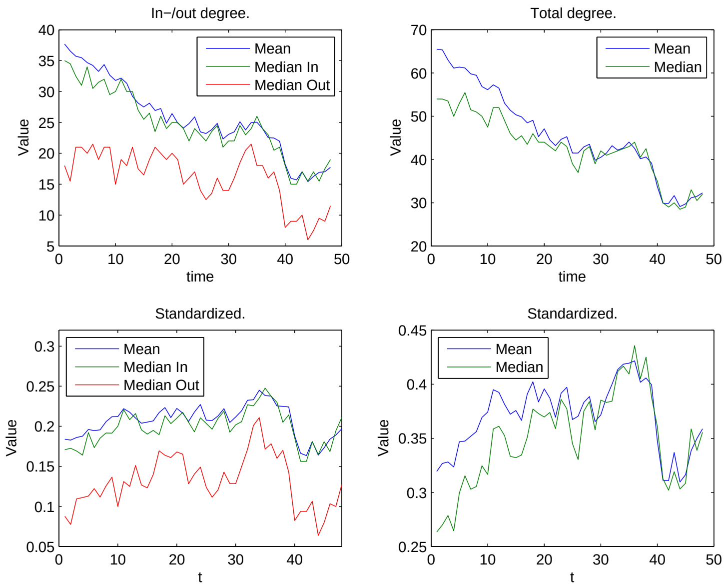

Figure 1: Mean and median degree over time. Left: in- and out-degree. Right: total degree. Top: absolute levels. Bottom: standardized values (divided by the number of active banks per quarter).

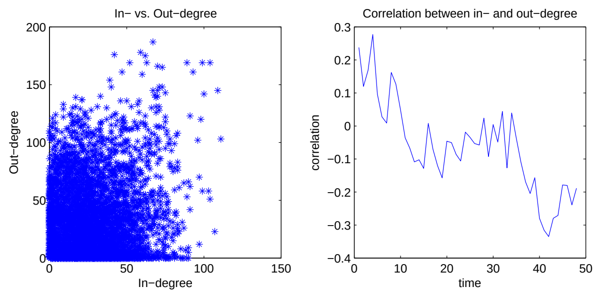

Figure 2: Left: Scatter plot of in- vs. out-degree. Correlation: .0899. Right: Correlation between individual banks' in- and out-degree over time. Italian banks.

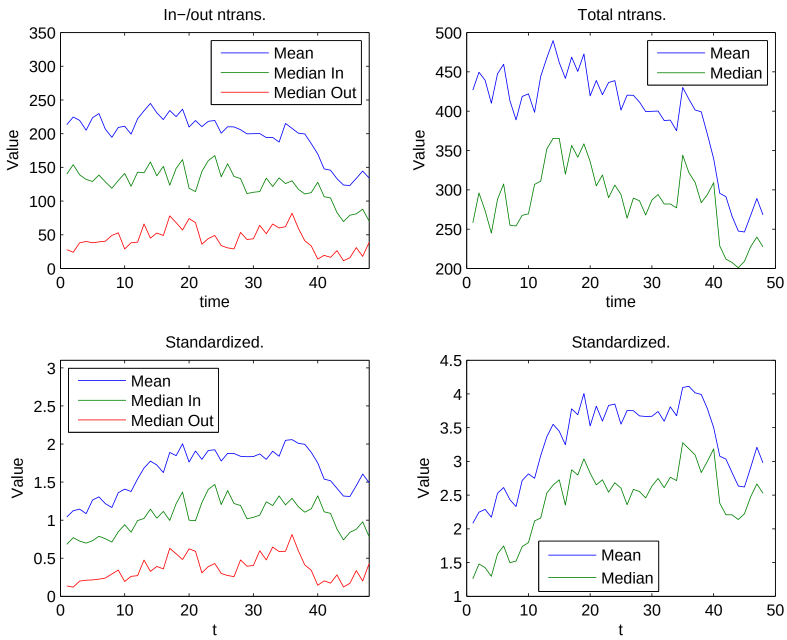

Figure 3: Mean and median number of transactions over time. Left: Directed Network. Right: Undirected Network. Top: absolute levels. Bottom: standardized values (divided by the number of active banks per quarter).

panel shows that the average number of transactions per bank is close to 200 during most quarters, but significantly decreases during and after the GFC. For both variables, the median values are substantially smaller than the mean, which hints towards a high level of skewness. Again, substantial differences in the median values indicate that the in- and out-variables are unlikely to follow identical distributions. The bottom left panel shows the standardized mean and median values. Quite interestingly, the somewhat negative trend of the variables vanishes, except for the GFC period. The same observation applies to the total number of transactions on the right panels. The results concerning the correlation between in- and out-transactions are comparable to those for the degrees (not reported).

4.2 The Degree Distributions

Due to the change in the size of the Italian interbank network, and the detection of two candidates for significant structural breaks during our sample period, we split the data set into three periods: Period 1 covers quarters 1-10, period 2 covers quarters 11-39, and period 3 covers the remaining quarters 40-48.[12] Assuming that the realizations of single days (quarters) are iid draws (or weakly dependent ones) from the same underlying data generating process, allows us to pool the data of the three subperiods into larger samples for the in-, out-, and total degrees (ntrans) of active banks, respectively. We use both daily and quarterly aggregates, i.e. construct variables that count the number of unique counterparties (degree) and total number of transactions (ntrans) for each bank within each day and quarter, respectively.[13] For the daily (quarterly) data this amounts to a total of 96,892 (1,780), 188,582 (3,369), and 41,775 (843) pooled observations for the three periods, respectively. For the sake of completeness, we also show the results when pooling all observations for the three time periods (1-3) for each degree measure. We should stress that pooling observations from several periods is crucially necessary in order to obtain reliable parameter estimates, in particular for daily data. We will elaborate on this issue in more detail in the next section.

As a first step, we compare the in- and out-degree distributions and check whether they could be realizations from the same underlying distribution. Figure 4 shows the histograms of the in-, out-, and total degrees for the different time-periods using quarterly data. We see that the histograms look very different when comparing in- and out-degrees for each sample period.

[12] Note that the first subsample roughly coincides with the data set used by De Masi *et al.* (2006).

[13] In Appendix B we present a similar analysis for the distribution of transaction volumes of individual institutions.

We should note that a substantial fraction of observations equals zero, both for in- and out-degrees. While the in-degree histograms appear to have a certain hump-shape, the out-degrees look more like a slowly decaying function with monotonic decline of probability from left to right. Furthermore, the L-shaped form of the out-degree distributions appears to be more stable over time, even though the scale on the x-axis changes substantially. Individual Kolmogorov-Smirnov (KS) tests provide further evidence against the equality of in- and out-degree distributions for all sample periods. The KS test allows to check whether two variables follow the same probability distribution, but also whether one variable follows a certain specific distribution. In our case, the KS test statistic is calculated as

$$\mathrm{KS}_n = \sup_x |F_{1,n}(x) - F_{2,n}(x)|, \tag{5}$$

where $\sup_x$ denotes the supremum of all possible values, while $F_{1,n}(\cdot)$ and $F_{2,n}(\cdot)$ are the empirical distribution functions of the sample of in-degrees and out-degrees, respectively. At all sensible significance levels, we have to reject the null hypothesis of the equality of both distributions. Similar observations can be made when pooling all observations across the three subperiods, see Figure 5.

Figure 6 shows the complementary cumulative distribution functions (ccdf) for the quarterly degree measures for all sample periods on a log-log scale, the typical way to represent data when suspecting power-law decay. Note that for a power-law, these ccdfs would be straight lines, which upon inspection seems unlikely to provide a good approximation to any of our subsamples, even for the tail regions. Again the distributions of in- and out-degrees look quite different in general, even though the shapes of the tail regions appear to be more homogeneous than what one might have expected after inspection of the raw data in Figures 4 and 5. Similar arguments hold for the distribution of total degrees, which has a somewhat similar shape as the in-degree distribution. For this reason, we will mostly restrict ourselves to comment on the results for the in- and out-degrees, respectively. We also show the ccdfs for the daily observations in Figure 7. Again, it is hard to detect a linear decay for most samples, at least not over several orders of magnitude.

4.2.1 Distribution Fitting Approach

Our basic approach is to fit a number of candidate distributions in order to investigate which distribution describes the data 'best' in a statistical sense. We should note that, similar to the approach in Stumpf and Ingram (2005), we use both discrete and continuous candidate distributions, implying that

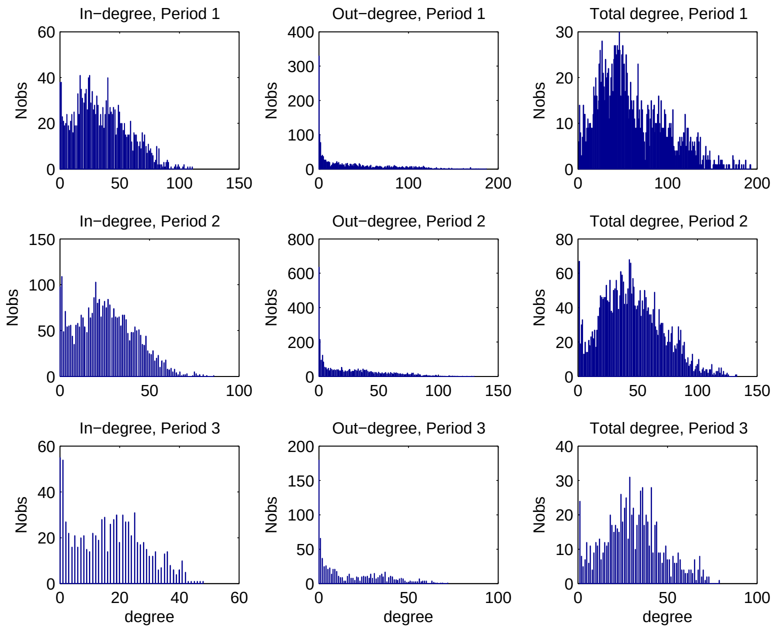

Figure 4: Quarterly data, degree. Histograms for in-degree (left), out-degree (center), and total degree (right) for Period 1, 2, and 3.

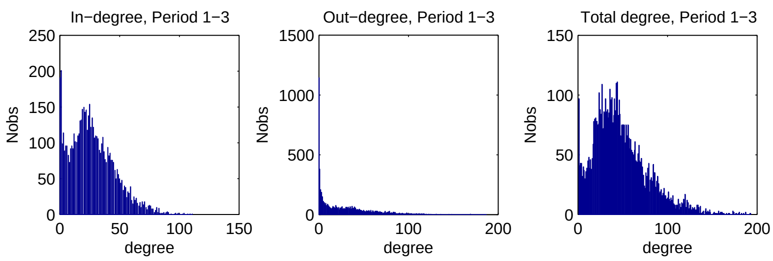

Figure 5: Quarterly data, degree. Histograms for in-degree (left), out-degree (center), and total degree (right) using all observations.

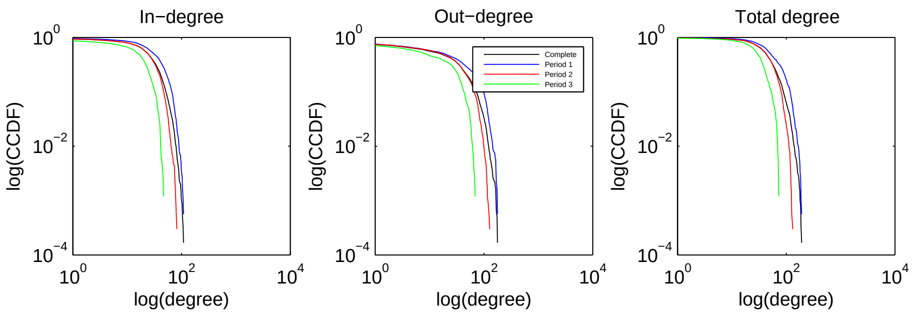

Figure 6: Quarterly data, degree. Complementary cumulative distribution functions (ccdf) in-degree (top), out-degree (center), and total degree (bottom) for all time periods on a log-log scale.

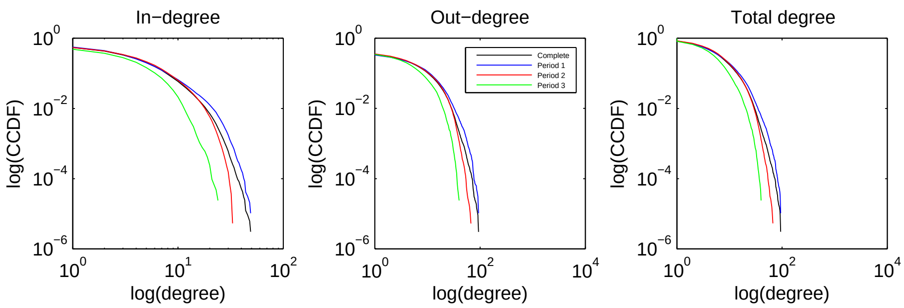

Figure 7: Daily data, degree. Complementary cumulative distribution functions (ccdf) in-degree (top), out-degree (center), and total degree (bottom) for all time periods on a log-log scale.

for the latter we treat the degrees as continuous variables. The candidate distributions, always fitted using maximum likelihood (ML), are:

- the Exponential distribution, with parameter $\lambda > 0$ (rate),
- the Gamma distribution, with parameters $k > 0$ (shape) and $\theta > 0$ (scale),
- the Geometric distribution, with probability parameter $p \in [0, 1]$,
- the Log-normal distribution, with parameters μ (scale) and $\sigma > 0$ (shape),
- the negative Binomial distribution, with parameters $r > 0$ (number of failures) and $p \in [0, 1]$ (success probability),
- the Poisson distribution, with parameter $\lambda > 0$,
- the discrete power-law or Pareto distribution, with parameters $x_m > 0$ (scale) and $\alpha > 0$ (shape),
- the Weibull or stretched exponential distribution, with parameters $\lambda > 0$ (scale) and $k > 0$ (shape).

We should note that a large part of the literature focuses on fitting the power-law only, in particular when the ccdfs have an apparently linear shape. Given that this is not the case here, we test a number of alternative distributions to find the distributions that fit the data best. Nevertheless, even though the power-law might not be a good description of the complete distribution, it could still provide a good fit of the (upper) tail region. Therefore, we conducted two sets of estimations of the above distributions for each sample: first, we fitted the complete distribution using all entries of our samples. Here we should stress, that several of the distributions have strictly positive support, while the others also allow for the occurrence of zero links. For the sake of consistency we will therefore only use non-zero values for the degree and ntrans variables in the following.[14] This means, for some distribution functions, we are using truncated variables in general (both for the complete

[14] This is important, since we cannot replicate the large number of zero values based on these distributions that we observe in the empirical data. Ignoring zeros reduces the number of quarterly observations to 1,742, 3,271, and 788 for the in-variables, and 1,450, 2,733, and 663 for the out-variables, respectively. For the daily data this leaves 70,584, 133,280, and 28,093 for the in-variables, and 39,619, 83,723, and 17,961 for the out-variables, respectively. The number of observations for the total degree and ntrans variables remain unaffected, since only active banks are in the sample.

and tail observations) and need to adjust the ML estimators for these distributions accordingly, cf. Appendix A.1. In a second step, we explicitly fitted three of the eight candidate distributions, namely the Exponential, the Log-normal, and the power-law, to a certain upper tail region for each period and variable (the other candidates would obviously make little sense as tail distributions). There are different possibilities to identify the 'optimal' tail region. Here we employ the approach of Clauset *et al.* (2009), which has been demonstrated to yield reliable estimates of both power-law parameters for certain distributions converging to Paretian tail behavior. The basic idea of this approach is to find the optimal tail parameter for all possible cutoff points using maximum likelihood, where the optimal x_m is the one corresponding to the lowest KS statistic. Details can be found in Appendix A.2.[15] The tail region is then defined by the scale parameter x_m, and the other distributions are fitted to all observations where $x \geq x_m$. Note that this approach gives an obvious advantage to the fit of the power-law in the 'tail' region. Quite surprisingly, however, in many cases the power-law is not the best description of the data tailored in this way as we will see below.

In these goodness-of-fits (GOF) experiments, we first estimate the parameters for each candidate distribution, both for the complete data set and the upper tail region, respectively, using ML. Using these parameters, we calculate the KS test-statistic for each candidate distribution and take the one with the lowest value as the 'best' fit of the respective data.[16] As a last step, we evaluate the GOF of this candidate distribution based on the KS test statistic. Given that the critical values of the KS distribution are only valid for known distributions (i.e. without estimating parameters), we have to perform individual Monte-Carlo exercises.[17] In these exercises, we randomly sample many degree sequences from the best fitting distributions with their estimated parameter values and then calculate the KS test statistic of these synthetic data sets. The reported p-values count the relative fraction of observations larger than the observed ones, such that low p-values (say 5%) indicate that the pertinent distribution can be rejected. We should stress that we carry out this analysis only for the best fitting distribution, since the

[15]There exist a number of alternative approaches in statistical extreme value theory for determining the optimal tail size. The approaches by Danielsson *et al.* (2001) and Drees and Kaufmann (1998) yielded results very similar to those reported in the text. We also checked certain fixed thresholds for identifying the tail region. The results remain qualitatively the same as long as the chosen upper quantile is reasonably large.

[16]In principle, we could also use likelihood-based criteria, e.g. AIC or BIC. However, Clauset *et al.* (2009) provide some evidence that the KS statistic is preferable as it is more robust to statistical fluctuations.

[17]See Clauset *et al.* (2009) and Stumpf *et al.* (2005) for similar approaches.

remaining ones have already been found to be inferior under the KS criterion. Details on the Monte-Carlo design can be found in Appendix A.3.

In the following we will use this approach to investigate the distribution of degrees and number of transactions for both daily and quarterly aggregates. Already at this point we should stress that the GOF tests mostly indicate that the distributions have to be rejected at traditional levels of significance for the complete samples, while the fits to the tail tend to perform better. This finding is, however, strongly driven by the significantly smaller number of observations for the tail data, which yields relatively large and more volatile KS statistics compared to the complete distributions.

4.2.2 Daily Data

We start our analysis with the daily degree data for which earlier studies have reported power-laws (De Masi *et al.*, 2006, and Iori *et al.*, 2008). Before turning to the results, we need to stress several complicating issues arising from network data in general, and our data in particular. For example, Stumpf and Porter (2012) note that '[a]s a rule of thumb, a candidate power-law should exhibit an approximately linear relationship on a log-log plot over at least two orders of magnitude in both the x and y axes. This criterion rules out many data sets, including just about all biological networks'. In this sense, finite and possibly very small network sizes make it hard to provide evidence for scale-free networks (Avnir *et al.*, 1998, and Clauset *et al.*, 2009).

For our data, Figure 8 shows the maximum in- and out-degrees for the individual days over time. We see that the criterion of Stumpf and Porter (2012) is typically violated. Thus, it should be hard to find evidence in favor of the power-law hypothesis for the complete distributions. Additionally, the number of observations in the 'tail' of the data for a single day becomes very small leading to large fluctuations of estimates across days and large error bands of single estimates. These issues highlight the importance of applying rigorous statistical methods to identify the best fitting distributions, i.e. simply identifying a linear slope of the ccdf on a log-log scale might easily be misleading. Similar remarks also apply for the daily ntrans variables (see below), while quarterly data are typically slightly less problematic.

To highlight our previous comments, Figures 9 and 10 show the distribution of the estimated daily power-law parameters for the complete and tail observations, respectively, for all sample days. For the complete daily samples, the results are very stable over time and across types of degrees, cf. Figure 9. In fact, we will see that this stability tends to carry over to the complete distributions of the aggregated data as well. In contrast, there is a substantial level of heterogeneity for the power-law exponent of the tail

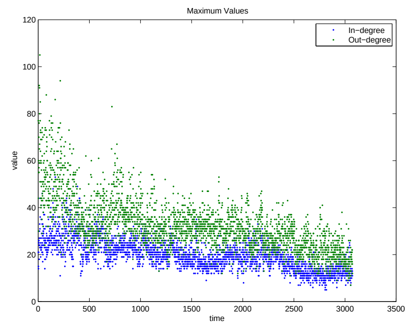

Figure 8: Daily data. Maximum in- and out-degrees over time.

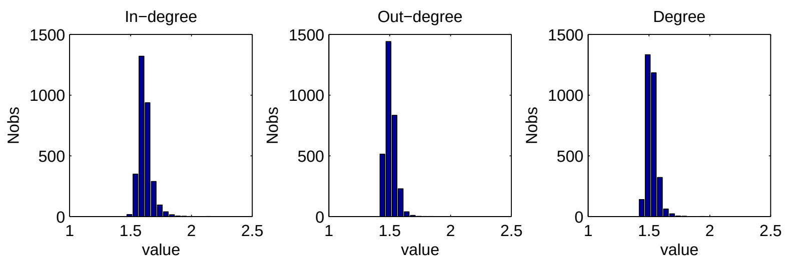

Figure 9: Daily data, degree. Histograms for the power-law exponents for the complete distributions, in-, out- and total degree, respectively.

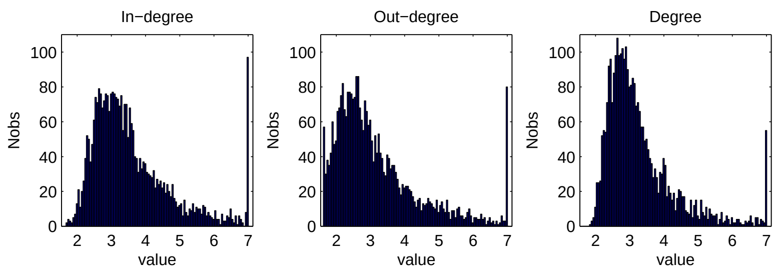

Figure 10: Daily data, degree. Histograms for the power-law exponents for the tail observations, in-, out- and total degree, respectively.

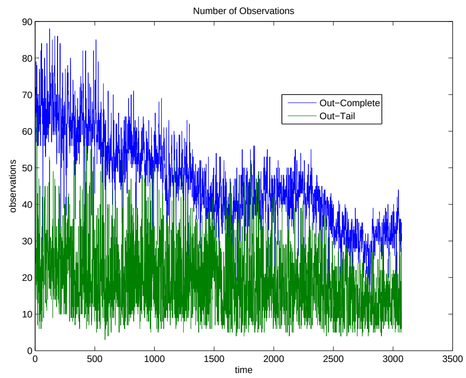

Figure 11: Daily data, degree. Total number of observations (complete) and number of tail observations for out-degree.

observations for the individual days, cf. Figure 10. Thus, we cannot confirm previously reported findings of 'typical' tail parameters between 2 and 3 for any of the degree variables.[18] While numerous observations lie within this range, for many days we find substantially larger values, at times as large as 7.[19] Apparently, the daily tail data are too noisy to identify a 'typical' tail parameter, cf. Figure 11.[20] The mismatch between the narrow range of values obtained for the complete data set of single days and the broad range of estimates for their tail might also indicate that the former are mainly determined by the more central part of the distribution.

Since data for single days are too scarce to allow reliable parameter estimation, pooling observations over longer horizons might be advisable to obtain better estimates. This, of course, requires the assumption of daily data being drawn independently from the same underlying distribution, or only with weak dependence of adjacent observations. While it is not straightforward to check this assumption for complete daily ensembles (as opposed

[18]The results are very similar when focusing on the individual period 1-3 as defined before.

[19]We have set 7 as the upper bound of the power-law parameter in our numerical ML implementation. For larger values the evaluation of the zeta function appearing in the discrete Pareto law, cf. Appendix A.2, is not accurate enough to obtain reliable estimates. The fact that the estimated values hit the upper bound quite frequently indicates that the estimated values may become even larger when increasing the upper bound.

[20]We also generated synthetic power-law distributed random draws and estimated their scaling parameters based on the algorithm for the selection of the tail region detailed above (not reported). For the small sample sizes of the typical daily data, the tail parameter of these synthetic data is highly volatile as well, even though the very large values observed for the actual data are very rare. As usual, however, increasing the number of observations (say more than 500), typically yields estimates very close to the true parameters.

to a time-series of univariate daily data), we have made some attempt at checking for statistical breaks for averages of degree statistics and have cut our complete sample into subsamples accordingly. Note also that any analysis of a network structure would be more or less futile, if we could not assume some stationarity of the structural characteristics of the network. Fricke and Lux (2012) demonstrate that the e-MID network is indeed structurally stable along many dimensions.

	In				**Out**				**Total**			
Period	1-3	1	2	3	1-3	1	2	3	1-3	1	2	3
Exponential	.0465	.0627	.0448	.0498	.0789	.0911	.0777	.0447	.0488	.0764	.0374	.0406
Gamma	.0627	.0670	.0661	.0870	.0515	.0511	.0536	.0559	.0562	.0637	.0512	.0592
Geometric	.0132	.0250	.0129	.0299	.0608	.0759	.0595	.0224	.0214	.0510	.0127	.0289
Log-normal	.0814	.0748	.0816	.1001	.0725	.0722	.0736	.0746	.0631	.0605	.0641	.0701
Neg. Bin.	**.0063**	**.0177**	**.0082**	**.0224**	**.0114**	**.0160**	**.0138**	**.0105**	**.0153**	**.0208**	**.0115**	**.0103**
Poisson	.2313	.2409	.2347	.1715	.3500	.3774	.3476	.2678	.2973	.3318	.2892	.2087
Power-law	.2099	.2151	.2107	.1985	.2077	.2024	.2079	.2140	.2366	.2219	.2427	.2373
Weibull	.0591	.0630	.0646	.0872	.0547	.0552	.0574	.0555	.0522	.0575	.0481	.0581

Table 1: Daily data, degree. KS statistic for the candidate distributions (complete). Minimum values in bold indicate the best fitting distribution. Asterisks would indicate non-rejection of this distribution at the 5% confidence level, where the critical values were obtained from a Monte-Carlo exercise as described in the main text. There is, however, no such case in this table.

	In				**Out**				**Total**			
Period	1-3	1	2	3	1-3	1	2	3	1-3	1	2	3
Daily												
Complete	1.61	1.61	1.60	1.67	1.50	1.48	1.50	1.54	1.52	1.51	1.51	1.57
	(.001)	(.003)	(.002)	(.005)	(.002)	(.003)	(.002)	(.005)	(.001)	(.002)	(.001)	(.003)
Tail	7.00	7.00	7.00	7.00	5.93	4.43	7.00	5.53	6.03	4.70	7.00	7.00
	(.175)	(.300)	(.160)	(.260)	(.110)	(.078)	(.170)	(.161)	(.100)	(.071)	(.146)	(.393)
Quarterly												
Complete	1.28	1.26	1.28	1.33	1.29	1.28	1.29	1.33	1.24	1.23	1.24	1.27
	(.004)	(.006)	(.005)	(.001)	(.004)	(.008)	(.006)	(.001)	(.003)	(.006)	(.004)	(.009)
Tail	5.13	7.00	7.00	4.63	7.00	6.90	7.00	4.82	5.20	6.90	7.00	5.01
	(.134)	(.460)	(.325)	(.233)	(.482)	(.532)	(.412)	(.306)	(.145)	(.421)	(.330)	(.261)

Table 2: Power-law parameters and standard deviations, degree. Values obtained via numerical maximization of the log-likelihood for discrete data. Standard deviations (in parentheses) approximated as $(\alpha - 1)/\sqrt{T}$, with T being the number of observations. Top: daily data, bottom: quarterly data.

We report our estimation results for the pooled daily data in Tables 1-3. Our main finding is that the negative Binomial distribution provides the best fits (in bold) for all daily degree measures and for all samples (i.e. the

	In				Out				Total			
Period	1-3	1	2	3	1-3	1	2	3	1-3	1	2	3
Exponential	**.0357***	**.0354***	**.0355***	.0954	.0642	.0580	**.0353***	**.0388***	.0685	.0637	**.0300***	.0457
Log-normal	.0664	.0639	.0762	.1036	.0479	.0479	.0541	.0771	.0484	.0455	.0521	.0927
Power-law	.0372	.0376	.0400	**.0203***	**.0129***	**.0305***	.0392	.0455	**.0114***	**.0192***	.0352	**.0382***

Table 3: Daily data, degree. KS statistic for the candidate distributions (tail). Minimum values in bold indicate the best fitting distribution. Asterisks indicate non-rejection of this distribution at the 5% confidence level, where the critical values were obtained from a Monte-Carlo exercise as described in the main text.

complete samples and the three subsamples identified via tests for structural breaks), cf. Table 1. The results from the GOF experiments indicate, however, that the best fitting candidate distributions have to be rejected. Therefore, even the winner among the candidate distributions appears to be an unlikely description of the data. We should also stress that the fit of the power-law is usually rather poor, competing with the Poisson distribution for the worst description of the data. Similar to the findings for the individual days, the estimated tail parameters are between 1.5 and 1.6, cf. Table 2 (top, complete). Figure 7 together with the relatively poor KS statistics for estimated power-laws suggests that estimates in the scaling range 1-2 are obtained as very inaccurate straight lines fitted to a strongly curved distributional shape. Moving to the tail observations, we find that exponential and power-law distributions tend to provide the best fit for all variables, cf. Table 3. Thus, it appears that the power-law is a better description of the tail observations - a usual finding for many data sets. In contrast to the complete distributions, the GOF experiments suggest that the estimated distributions are mostly not rejected for the tail observations.[21] Upon closer inspection, however, we see the KS statistics of the exponential and the power-law are typically close to each other, in particular when the tail exponents are very large, cf. Table 2 (top, complete). Even though the power-law appears to provide the best fit for some of the tail data, the very large parameter values (larger than 4, often close to 7) are in a range where the power-law becomes almost undistinguishable from exponential decay. Often such high values would be obtained spuriously from distributions with an exponential decline as semi-parametric estimators of the tail index would not be able to 'identify' the limit of $\alpha \to \infty$. The huge difference in estimated power-law parameters for the complete sample compared to the tail also indicates that the empirical distribution shows pronounced curvature (actually confirming

[21] This result is driven by the higher noise level in the tail data due to a smaller number of observations compared to the complete distributions.

the visual inspection of absence of a linear slope over the complete support and very fast decline at the end in Figure 7). On the other hand, it is also interesting to remark that the estimated coefficients are relatively uniform for both the complete sample and the tail, respectively, across periods and for all the measures of degree. This speaks of relatively uniform shapes of the distributions, at least in view of this simple statistic. Summing up, the power-law distribution appears to be a poor description of the data, both for the complete distribution and the tail observations (where it more or less coincides with an exponential for the high estimates of the tail index). We also need to stress that the identified power-law exponents, both for individual days and pooled observations, are far off from those reported in earlier studies. It is not clear how these estimates were obtained.

4.2.3 Quarterly Data

	In				**Out**				**Total**			
Period	1-3	1	2	3	1-3	1	2	3	1-3	1	2	3
Exponential	.1474	.1544	.1661	.1520	.0797	.0932	**.0710**	.0925	.1740	.1675	.1887	.2171
Gamma	.0573	.0543	.0723	.0942	**.0595**	**.0514**	.0738	.0961	.0414	**.0284***	.0673	.0943
Geometric	.1533	.1586	.1723	.1619	.0778	.0918	.0728	.1009	.1771	.1696	.1920	.2223
Log-normal	.1141	.0972	.1274	.1377	.1063	.0984	.1164	.1226	.0972	.0760	.1185	.1453
Neg. Bin.	.0601	.0580	.0729	.1025	.0708	.0615	.0836	.1081	.0395	.0318	.0627	.0881
Poisson	.3117	.3462	.2753	.2561	.4367	.4707	.4116	.4183	.3601	.4115	.3183	.2489
Power-law	.3828	.4023	.3849	.3522	.2727	.2728	.2842	.2608	.4376	.4546	.4387	.4291
Weibull	**.0380**	**.0342***	**.0456**	**.0689**	.0624	.0609	.0736	**.0912**	**.0246**	.0361	**.0325**	**.0495***

Table 4: Quarterly data, degree. KS statistic for the candidate distributions (complete). Minimum values in bold indicate the best fitting distribution. Asterisks indicate non-rejection of this distribution at the 5% confidence level, where the critical values were obtained from a Monte-Carlo exercise as described in the main text.

	In				**Out**				**Total**			
Period	1-3	1	2	3	1-3	1	2	3	1-3	1	2	3
Exponential	**.0248***	.0466	**.0325***	**.0628***	**.0395***	.0394	**.0352***	**.0459***	**.0331***	**.0315***	**.0437***	**.0530***
Log-normal	.0379	**.0441***	.0431	.0766	.0526	.0533	.0663	.0756	.0451	.0494	.0769	.0794
Power-law	.0651	.0748	.0471	.0949	.0502	**.0384***	.0559	.0918	.0515	.0405	.0602	.0778

Table 5: Quarterly data, degree. KS statistic for the candidate distributions (tail). Minimum values in bold indicate the best fitting distribution. Asterisks indicate non-rejection of this distribution at the 5% confidence level, where the critical values were obtained from a Monte-Carlo exercise as described in the main text.

The results for the quarterly data are shown in Tables 4 and 5. Weibull distributions typically provide the best fits for the in- and total degrees, while Exponential and Gamma distributions yield comparable fits as the Weibull for the out-degrees. Similar to the complete distributions for the daily data, the optimal fits are insignificant, except for three cases. Again, the best fits appear to be unlikely descriptions of the observed data. The power-law exponents for the complete sample are again quite small, typically between 1.25 and 1.3, cf. Table 2 (bottom, complete). Turning to the tail observations, we find that Exponential distributions provide the best fits in all but two cases (in-degree in period 1, out-degree period 1). Similar to the daily observations, for the best fits of the tail data, the pertinent distributions cannot be rejected as the 'true' data-generating process at the 95 percent significance level. The poor fit of the power-law again comes along with relatively large tail exponents, cf. Table 2 (bottom, tail). In summary, similar to the daily data, we do not find evidence in favor of scale-free networks.

4.2.4 Robustness and Discussion

A reason for not finding evidence for power-law distributions may be the fact that we focus on the subnetwork formed by Italian banks only. Stumpf *et al.* (2005) have shown that (randomly chosen) subnetworks of scale-free networks are in fact not scale-free. Therefore we also checked the distributions including foreign banks as well, similar to the existing papers using the e-MID data. We found that the results (including the tail parameters) remain qualitatively unaffected (not reported). In terms of Stumpf *et al.* (2005), these findings indicate again that the networks including all banks are unlikely to be scale-free, and that our previous findings for Italian banks alone are not biased due to random sampling from a larger scale-free network (indeed, it seems very unlikely that the Italian banks should constitute a random set from the overall sample of all banks). Then it comes as no surprise that the subnetworks formed by Italian banks only are also not scale-free. In fact, it is remarkable that there appears to be no significant qualitative effect of incorporating foreign banks or not.

In Finger *et al.* (2012) it has been shown that the quarterly e-MID networks are more complete representations of the underlying 'latent' network structure, whereas daily networks might be seen as random activations of parts of the more complex, hidden structure. Under this perspective, the lack of coincidence of the fitted distributions for different levels of time aggregation might not be too surprising.

Summing up, our results indicate that the power-law hypothesis needs to be tested more thoroughly for other networks in general and the interbank network in particular,[22] with the power-law being one of many candidate distributions. The findings are in line with other studies casting doubts on certain claims of power-law and scaling behavior in a broad range of empirical studies (cf. Avnir *et al.*, 1998).

4.3 The Distribution of the Number of Transactions

Note that the quarterly degree of a given bank is not the simple sum of its daily degrees, since a link that has been activated many times over a quarter, is counted only as one link on the quarterly level of network activity. If we consider the number of links in the daily data as (possibly power-law distributed) random variables, the number of transactions over a longer time horizon is, in fact, what we obtain from simple aggregation of the daily degrees observed for any bank i. Assuming that the degrees of all banks are drawn from the same distribution, we obtain in this way a sample of sums of random variables following the same underlying distribution. Note that we would expect a power-law at the daily level to survive in the aggregation process for an iid random process of link formation as well as for various extensions allowing for 'weak' dependency.[23] The extremal behavior of the distribution of degrees should, therefore, be preserved in the distribution of the number of transactions over longer horizons. We turn to the analysis of this quantity in this subsection.

Note also that the finite size of the network might pose a problem due to the effective imposition of an upper limit on the observable degrees. It might, therefore, be the case that a scale-free distribution is just hard to verify because of the small number of observations. In contrast, the aggregated ntrans variables have the advantage that they have no obvious upper bound, so testing the power-law hypothesis might be more sensible in this case.

Figures 12 and 13 show the ccdfs of the quarterly and daily ntrans variables. Again, linear decay over several orders of magnitude is hard to detect visually. However, at least for the quarterly data we see that the variables under study span several orders of magnitude, making the data more useful candidates for our distribution fitting approach.

For daily data, the range of the observed variables remains rather limited, even though the maximum value is roughly twice the one for the degrees (not reported). Since for daily realizations of the number of transactions

[22] See Stumpf and Porter (2012).

[23] The stability under aggregation of power-laws characterizing the tails of iid random variables is one of the basic tenets of the statistical theory of extremes, cf. Reiss and Thomas (2007). In this sense, summing up daily power-law networks should preserve the tail index for different frequencies.

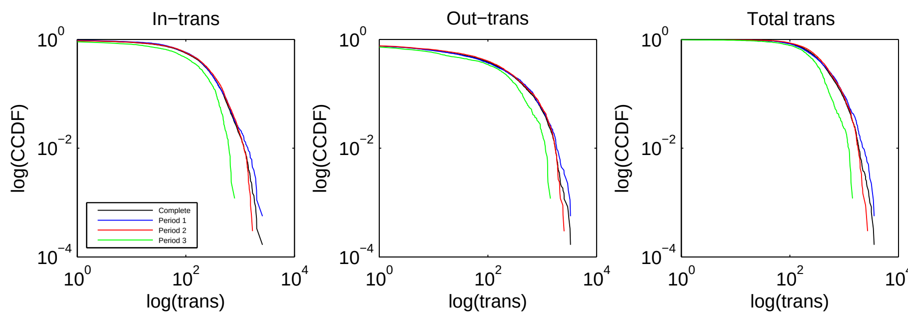

Figure 12: Quarterly data, ntrans. Complementary cumulative distribution functions (ccdf) in-trans (top), out-trans (center), and total trans (bottom) for all time periods on a log-log scale.

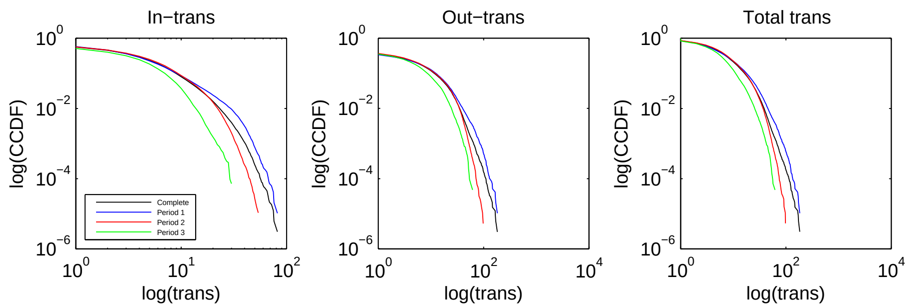

Figure 13: Daily data, ntrans. Complementary cumulative distribution functions (ccdf) in-trans (top), out-trans (center), and total trans (bottom) for all time periods on a log-log scale.

	In				Out				Total			
Period	1-3	1	2	3	1-3	1	2	3	1-3	1	2	3
Quarterly												
Complete	1.19	1.19	1.19	1.21	1.21	1.22	1.21	1.23	1.16	1.16	1.16	1.17
	(.003)	(.005)	(.003)	(.008)	(.003)	(.006)	(.004)	(.009)	(.002)	(.004)	(.003)	(.006)
Tail	3.48	2.78	3.54	3.06	3.16	2.99	2.21	2.95	2.77	3.03	2.76	3.68
	(.091)	(.082)	(.116)	(.136)	(.094)	(.143)	(.044)	(.180)	(.040)	(.106)	(.050)	(.200)

Table 6: Power-law parameters and standard deviations, ntrans. Values obtained via numerical maximization of the log-likelihood for discrete data. Standard deviations (in parentheses) approximated as $(\alpha - 1)/\sqrt{T}$, with T being the number of observations. Quarterly data.

we find virtually identical results to those of the daily degrees, we abstain from presenting these here, and immediately turn to quarterly aggregated observations.

	In				Out				Total			
Period	1-3	1	2	3	1-3	1	2	3	1-3	1	2	3
Exponential	.0452	.0516	.0467	.0748	.2334	.2598	.2131	.2716	.0612	.0524	.0761	.0955
Gamma	**.0153**	.0287	.0199	**.0388**	**.0167**	**.0165**	**.0185**	**.0498**	.0356	.0485	.0325	.0549
Geometric	.0450	.0511	.0465	.0744	.2331	.2595	.2129	.2713	.0614	.0525	.0763	.0960
Log-normal	.0801	.0786	.0843	.0953	.0712	.0610	.0725	.1013	.0701	.0508	.0738	.1207
Neg. Bin.	.0205	.0280	.0249	.0419	.0205	.0188	.0223	.0586	**.0353**	**.0477**	**.0321**	.0543
Poisson	.5744	.5859	.5775	.5527	.6706	.6860	.6710	.6221	.5896	.6164	.5813	.5216
Power-law	.3610	.3865	.3658	.3384	.2381	.2397	.2607	.2209	.4497	.4488	.4648	.4236
Weibull	.0185	**.0244**	**.0186**	.0420	.0383	.0389	.0392	.0668	.0391	.0481	.0386	**.0437***

Table 7: Quarterly data, ntrans. KS statistic for the candidate distributions (complete). Minimum values in bold indicate the best fitting distribution. Asterisks indicate non-rejection of this distribution at the 5% confidence level, where the critical values were obtained from a Monte-Carlo exercise as described in the main text.

We show the results in Tables 7 and 8, finding that negative Binomial, Gamma and Weibull distributions appear among the best fits, depending on the concept (in-, out-, or total transactions) and the period considered. However, their KS statistics are typically at a comparable level. The results from the GOF experiments show that the best fitting distributions are nevertheless rejected as data-generating processes (exception: total ntrans in period 3). Again, the fit of the power-law is very poor in general, with tail parameters around 1.20, cf. Table 6, and KS statistics that consistently come in second to last (with the Poisson distribution performing worst). Moving to the quarterly tail data, we find that in most cases the Log-normal provides the best fit (exceptions: out-degree for the complete sample and total degree in period 3). This is quite surprising, given that the scaling parameters now lie

	In				Out				Total			
Period	1-3	1	2	3	1-3	1	2	3	1-3	1	2	3
Exponential	.0684	.0874	.0709	.0687	**.0393***	.0805	.0532	.0751	.0528	.0616	.0492	.0948
Log-normal	**.0258***	**.0202***	**.0441***	**.0560***	.0408	**.0594***	**.0495**	**.0700***	**.0274***	**.0336***	**.0362***	.0579
Power-law	.0476	.0570	.0520	.1219	.0988	.0741	.1024	.0718	.0676	.0604	.0683	**.0426***

Table 8: Quarterly data, ntrans. KS statistic for the candidate distributions (tail). Minimum values in bold indicate the best fitting distribution. Asterisks indicate non-rejection of this distribution at the 5% confidence level, where the critical values were obtained from a Monte-Carlo exercise as described in the main text.

in the 'typical' range for power-laws, here between 2.21 and 3.68. Therefore, even though the power-law estimates appear more sensible, the power-law distribution is inferior by some margin in fitting the tail data (with a cut-off determined by the best-fitting Pareto law) to the Log-normal, and sometimes also to the Exponential. As with the previous cases, the results from the GOF experiments indicate that the best fitting tail distributions usually cannot be rejected via KS tests with Monte-Carlo distributions. While it is well-known that it is hard to distinguish Log-normal from power-law tails, these findings raise doubts on the universality of power-law tails and highlight the need for thorough statistical approaches of testing the power-law hypothesis.

As another robustness check, we investigated the distribution of transaction volumes (tvol), cf. Appendix B, again differentiating between in-tvol, out-tvol and their sum (total tvol), respectively. While the tails of the tvol variables are typically much fatter compared to the degree and ntrans variables, the power-law remains a poor description both for daily and quarterly data.

5 Conclusions

In this paper, we have revisited the distributional properties of interbank loans for the Italian interbank network during the years 1999-2010. Using both the degrees and the number of transactions, we fitted a set of different candidate distributions to these data for daily and quarterly aggregates, respectively. Given that the daily networks have previously been claimed to be scale-free (De Masi *et al.*, 2006), it comes as a surprise that we find no evidence in favor of the power-law hypothesis: at the daily level the degrees are usually fit best by negative Binomial distributions, while the tails tend to decay exponentially, i.e. the fitted power-laws display very large tail parameters. At the quarterly level, Weibull, Gamma, and Exponential distributions tend to provide comparable fits for the complete degree distribution, while the tails again tend to display exponential decay. For the number of transactions we find comparable results, even though the tails of the quarterly data appear to be fatter. However, in this case the Log-normal distribution usually outperforms the power-law. Moreover, we found that the networks are characterized by a substantial level of asymmetry, as exemplified by the low correlation between in- and out-degrees. We also find that the two variables do not follow identical distributions in general.

Overall these findings indicate that the power-law is typically a poor description of the data. This implies that preferential attachment and related mechanisms (see e.g. De Masi *et al.*, 2006), are unlikely explanations for the formation of the Italian interbank network. Note that these findings are also not in line with a large part of the empirical (interbank) network literature for other data sets, putting doubts on the universality of scale-free behavior of interbank networks. Our results also indicate that the power-law hypothesis needs to be tested more thoroughly for other networks in general and the interbank network in particular. The findings are related to other studies casting doubts on certain claims of power-law and scaling behavior in a broad range of empirical studies (cf. Avnir *et al.*, 1998, and Stumpf and Porter, 2012), and it seems possible that claims of scale-free behavior of interbank lending activity may not survive under closer statistical scrutiny.

Bibliography

Albert, R., H. Jeong, and A.-L. Barabasi (2000): "Error and Attack Tolerance of Complex Networks," *Nature*, (406), 378–482.

Alderson, D. L., and L. Li (2007): "Diversity of Graphs with Highly Variable Connectivity," *Physical Review E*, 75, 046102.

Anderson, C. W. (1970): "Extreme Value Theory for a Class of Discrete Distributions with Applications to Some Stochastic Processes," *Journal of Applied Probability*, 7(1), 99–113.

Avnir, D., O. Biham, D. Lidar, and O. Malcai (1998): "Is the Geometry of Nature Fractal?," *Science*, 279(5347), 39–40.

Axtell, R. L. (2001): "Zipf Distribution of U.S. Firm Sizes," *Science*, 293(5536), 1818–1820.

Beaupain, R., and A. Durré (2012): "Nonlinear Liquidity Adjustments in the Euro Area Overnight Money Market," Working Paper Series 1500, European Central Bank.

Bech, M., and E. Atalay (2010): "The Topology of the Federal Funds Market," *Physica A*, 389(22), 5223–5246.

Boss, M., H. Elsinger, M. Summer, and S. Thurner (2004): "Network Topology of the Interbank Market," *Quantitative Finance*, 4(6), 677–684.

Caldarelli, G. (2007): *Scale-Free Networks*. Oxford University Press.

Clauset, A., C. Rohilla Shalizi, and M. E. J. Newman (2009): "Power-law Distributions in Empirical Data," *SIAM Review*, 51, 661–703.

Clauset, A., M. Young, and K. S. Gleditsch (2007): "On the Frequency of Severe Terrorist Events," *Journal of Conflict Resolution*, 51(1), 58–87.

Danielsson, J., L. de Haan, L. Peng, and C. de Vries (2001): "Using a Bootstrap Method to Choose the Sample Fraction in Tail Index Estimation," *Journal of Multivariate Analysis*, 76(2), 226 – 248.

De Masi, G., G. Iori, and G. Caldarelli (2006): "Fitness Model for the Italian Interbank Money Market," *Phys. Rev. E*, 74(6), 66112.

Drees, H., and E. Kaufmann (1998): "Selecting the Optimal Sample Fraction in Univariate Extreme Value Estimation," *Stochastic Processes and their Applications*, 75(2), 149 – 172.

Erdös, P., and A. Renyi (1959): "On Random Graphs," *Publicationes Mathematicae*, 6, 290–297.

European Central Bank (2007): "Euro Money Market Study 2006," Discussion paper, ECB.

Fagiolo, G., M. Napoletano, and A. Roventini (2008): "Are Output Growth-Rate Distributions Fat-Tailed? Some Evidence from OECD Countries," *Journal of Applied Econometrics*, 23(5), 639–669.

Finger, K., D. Fricke, and T. Lux (2012): "Network Analysis of the e-MID Overnight Money Market: The Informational Value of Different Aggregation Levels for Intrinsic Dynamic Processes," Kiel Working Paper 1782, Kiel Institute for the World Economy.

Fricke, D., and T. Lux (2012): "Core-Periphery Structure in the Overnight Money Market: Evidence from the e-MID Trading Platform," Kiel working paper 1759, Kiel Institute for the World Economy.

Gabaix, X. (1999): "Zipf's Law for Cities: An Explanation," *The Quarterly Journal of Economics*, 114(3), 739–767.

Gabaix, X., and R. Ibragimov (2011): "Rank -1/2: A Simple Way to Improve the OLS Estimation of Tail Exponents," *Journal of Business & Economic Statistics*, 29(1), 24–39.

Gai, P., A. Haldane, and S. Kapadia (2011): "Complexity, Concentration and Contagion," *Journal of Monetary Economics*, 58(5), 453–470.

Goldstein, M. L., S. A. Morris, and G. G. Yen (2004): "Problems with Fitting to the Power-law Distribution," *The European Physical Journal B - Condensed Matter and Complex Systems*, 41, 255–258, 10.1140/epjb/e2004-00316-5.

Haldane, A. (2009): "Rethinking the Financial Network," Speech delivered in April 2009 at the Financial Student Association, Amsterdam. http://www.bankofengland.co.uk/publications/speeches/2009/speech386.pdf.

Haldane, A. G., and R. M. May (2011): "Systemic Risk in Banking Ecosystems," *Nature*, 469(7330), 351–355.

Hill, B. M. (1975): "A Simple General Approach to Inference About the Tail of a Distribution," *The Annals of Statistics*, 3(5), 1163–1174.

Iori, G., G. De Masi, O. V. Precup, G. Gabbi, and G. Caldarelli (2008): "A Network Analysis of the Italian Overnight Money Market," *Journal of Economic Dynamics and Control*, 32(1), 259–278.

Jansen, D. W., and C. G. de Vries (1991): "On the Frequency of Large Stock Returns: Putting Booms and Busts into Perspective," *The Review of Economics and Statistics*, 73(1), 18–24.

Lau, A. H.-L., H.-S. Lau, and J. R. Wingender (1990): "The Distribution of Stock Returns: New Evidence against the Stable Model," *Journal of Business & Economic Statistics*, 8(2), 217–23.

Leadbetter, M. (1983): "Extremes and Local Dependence in Stationary Sequences," *Zeitschrift fuer Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 65, 291–306.

Mandelbrot, B. (1963): "The Variation of Certain Speculative Prices," *The Journal of Business*, 36, 394.

Nier, E., J. Yang, T. Yorulmazer, and A. Alentorn (2007): "Network Models and Financial Stability," *Journal of Economic Dynamics and Control*, 31(6), 2033–2060.

Reiss, R.-D., and M. Thomas (2007): *Statistical Analysis of Extreme Values: with Applications to Insurance, Finance, Hydrology and Other Fields (3rd Edition)*. Birkhäuser Verlag.

Silverberg, G., and B. Verspagen (2007): "The Size Distribution of Innovations Revisited: An Application of Extreme Value Statistics to Citation and Value Measures of Patent Significance," *Journal of Econometrics*, 139(2), 318–339.

Soramaki, K., M. L. Bech, J. Arnold, R. J. Glass, and W. Beyeler (2006): "The Topology of Interbank Payment Flows," Staff Reports 243, Federal Reserve Bank of New York.

Stumpf, M. P. H., and P. J. Ingram (2005): "Probability Models for Degree Distributions of Protein Interaction Networks," *Europhysics Letters*, 71(1), 152–158.

Stumpf, M. P. H., P. J. Ingram, I. Nouvel, and C. Wiuf (2005): "Statistical Model Selection Applied to Biological Network Data," *Proceedings in Computational Systems Biology*, (3), 65–73.

Stumpf, M. P. H., and M. A. Porter (2012): "Critical Truths About Power Laws," *Science*, 335(6069), 665–666.

Stumpf, M. P. H., C. Wiuf, and R. M. May (2005): "Subnets of Scale-free Networks are not Scale-free: Sampling Properties of Networks," *Proceedings of the National Academy of Sciences of the United States of America*, 102(12), 4221–4224.

Upper, C., and A. Worms (2004): "Estimating Bilateral Exposures in the German Interbank Market: Is there a Danger of Contagion? Cross-Border Bank Contagion in Europe," *European Economic Review*, 48(4), 827–849.

A Technical Appendix

A.1 Truncated Distributions and Maximum Likelihood

The distribution fitting approach described in the main text involves fitting a set of candidate distributions with possibly differing support. For example, some distributions have support at zero, while others do not. Similarly, when focusing on the tail observations we have to get rid of the probability mass below the cutoff point in order to accurately calculate the statistics. Therefore, we describe the use of truncated distributions and ML fitting in this Appendix in more detail.

A.1.1 Normalization

When working with truncated variables, we need to make sure to use the correct pdfs and cdfs, since the ML estimation and the evaluation of the fit (KS statistic) depend on them. In order to illustrate this issue, let variable x have the pdf $p(x)$ with support $[0, \infty]$. As usual, the cdf is defined as

$$P(a) = P(X \leq a) = \int_0^a p(x)dx. \tag{6}$$

Now, suppose the data are (left-)truncated at some value x_m, i.e. the variable $\tilde{x}$ follows the same distribution as x, but the pdf has limited support $[x_m, \infty]$ with minimum value $x_m > 0$. For our purposes, it is therefore useful to define the quantity

$$P_<(a) = P(X < a) = 1 - \int_a^\infty p(x)dx, \tag{7}$$

or more compactly

$$P_<(a) = P(a) - p(a). \tag{8}$$

We can properly construct the pdf of $\tilde{x}$, say $\tilde{p}$, as

$$\tilde{p}(x) = \begin{cases} \frac{p(x)}{1-P_<(x_m)}, & \text{if } x \geq x_m \\ 0, & \text{else.} \end{cases} \tag{9}$$

where the denominator distributes the probability mass of $p(x)$ among the support of $\tilde{x}$.

For the calculation of the KS statistics, we also need the adjusted cdf. For the supported values of $\tilde{x}$ it takes the form

$$\tilde{P}(x) = \int_{x_m}^x \frac{p(x)}{1 - P_<(x_m)}dx = \frac{1}{1 - P_<(x_m)} \int_{x_m}^x p(x)dx, \tag{10}$$

or

$$\tilde{P}(x) = \frac{P(x) - P_<(x_m)}{1 - P_<(x_m)}, \tag{11}$$

which can be easily evaluated.

A.1.2 Maximum Likelihood for Truncated Variables

Using the previous definitions, we can show that the ML estimator for left-truncated variables does not coincide with the standard estimator. The standard ML estimator, i.e. using a sample of n observations of x and denoting by θ the vector of parameters, can be written as

$$L(\theta|x_1, \cdots, x_n) = p(x_1, \cdots, x_n|\theta) = \prod_i^n p(x_i|\theta), \tag{12}$$

or in logs

$$\ln(L) = \sum_i^n \ln[p(x_i|\theta)]. \tag{13}$$

Using the definitions from above, we can show that the ML estimator for left-truncated variables differs from the one in Eq. (13). Using Eq. (9), we can write the likelihood as

$$L = \prod_i^{\tilde{n}} \tilde{p}(x_i|\theta) = \prod_i^{\tilde{n}} \frac{p(x|\theta)}{1 - P_<(x_m|\theta)}, \tag{14}$$

where x ignores those observations smaller than x_m and the total number of observations is $\tilde{n}$ instead of n. Taking logarithms we obtain

$$\ln(L) = \sum_i^{\tilde{n}} \ln\left[\frac{p(x_i|\theta)}{1 - P_<(x_m|\theta)}\right] = \sum_i^{\tilde{n}} \ln[p(x_i|\theta)] - \sum_i^{\tilde{n}} \ln[1 - P_<(x_m|\theta)], \tag{15}$$

which can be written as

$$\ln(L) = -\tilde{n}\ln[1 - P_<(x_m|\theta)] + \sum_i^{\tilde{n}} \ln[p(x_i|\theta)]. \tag{16}$$

The second part of this Eq. looks familiar, as it corresponds to Eq. (13) for the $\tilde{n}$ observations with values $\geq x_m$. However, the normalization term on the left does not vanish (as it depends on the parameter vector) and affects the location of the maximum likelihood estimator. Therefore, we need to find the θ that maximizes Eq. (16). The standard ML estimator would not be efficient.

A.2 Discrete Power-laws and Parameter Estimation

This presentation is mostly based on Clauset *et al.* (2009).

A.2.1 Discrete Power-laws

A power-law distributed variable x obeys the pdf

$$p(x) \propto x^{-\alpha}, \tag{17}$$

where $\alpha > 0$ is the tail exponent with 'typical' interesting values in the range between 1 and 3. In many cases, however, the power-law only applies for some (upper) tail region, defined by the minimum value x_m. While it is common to approximate discrete power-laws by the (simpler) continuous version, for our (integer-valued) data, we employ the more accurate discrete version in the paper.[24]

In the discrete case, the cdf of the power-law can be written as

$$P(x) = \frac{\zeta(\alpha, x)}{\zeta(\alpha, x_m)}, \tag{18}$$

where

$$\zeta(\alpha, x_m) = \sum_{n=0}^{\infty}(n + x_m)^{-\alpha} \tag{19}$$

is the generalized or Hurwitz zeta function.

A.2.2 Estimation of α and x_m

For a given lower bound x_m, the ML estimator of α can be found by direct numerical maximization of the log-likelihood function

$$\mathcal{L}(\alpha) = -n\ln[\zeta(\alpha, x_m)] - \alpha\sum_{i=1}^{n}\ln[x_i], \tag{20}$$

where n is the number of observations.[25] For simplicity, we approximate the standard error of the estimated $\hat{\alpha}$ (for $\hat{\alpha} > 1$) using the closed-form

[24] Clauset *et al.* (2007) show that this is necessary for data sets from the social sciences, where the maximum value is usually only a few orders of magnitude larger than the minimum, i.e. the tail is heavy but rather short. In such cases the estimated exponents can be biased severely when using the continuous approximation.

[25] Using a quadratic approximation of the log-likelihood at its maximum, Clauset *et al.* (2009) also derive an approximate closed-form solution for the estimate of $\alpha \simeq 1 + n/\left(\sum_{i=1}^{n}\ln\left[\frac{x_i}{x_m - 0.5}\right]\right)$. This can be seen as an adjusted Hill-estimator, see Hill (1975). While we always report the exact ML estimator, we checked that the approximation is typically not too bad.

solution based on continuous data.[26] Neglecting higher-order terms, this can be calculated as

$$\sigma = \frac{\hat{\alpha} - 1}{\sqrt{n}}. \tag{21}$$

However, the equations assume that x_m is known in order to obtain an accurate estimate of α.[27] When the data span only a few orders of magnitude, as usual in many social or complex systems, an underpopulated tail would come along with little statistical power. Therefore, we employ the numerical method proposed by Clauset *et al.* (2007) for selecting the x_m that yields the best power-law model for the data. To be precise, for each x_m over some reasonable range, we first estimate the scaling parameter using Eq. (20) and calculate the corresponding KS statistic between the fitted data and the theoretical distribution with the estimated parameters. The reported x_m and α are those that minimize the KS statistic, i.e. minimize the distance between the observed and fitted probability distribution. According to Clauset *et al.* (2007; 2009), minimizing the KS statistic is generally superior to other distance measures, e.g. likelihood-based measures such as AIC or BIC.

A.3 Goodness-of-Fit Test for the Estimated Distributions

Since the distribution of the KS statistics is unknown for the comparison between an empirical subsample and a hypothetical distribution with estimated parameters, we carry out a Monte Carlo approach. We sample synthetic data sets from the estimated distribution, compute the distribution of KS statistics and compare the results to the observed value for the original data set. If the KS statistic of the empirical data set is beyond the α percent quantile of the Monte Carlo distribution of KS values, we reject the pertinent distribution at the $1-\alpha$ level of significance. In our results, we indicate significant fits at the 5% confidence level using asterisks. We should stress that we carry out this (very time-consuming) GOF experiment only for the distribution with the minimum KS statistic for each sample and variable, respectively. This can be justified by the fact that, even though other candidate distributions may not be rejected as well, they are clearly inferior to the optimal distribution in terms of the KS statistic.

[26] Clauset *et al.* (2009) also derive an (approximate) estimator for the standard error based on discrete data, which is, however, much harder to evaluate as it involves derivatives of the generalized zeta function.

[27] See Clauset *et al.* (2007; 2009) for an extensive discussion.

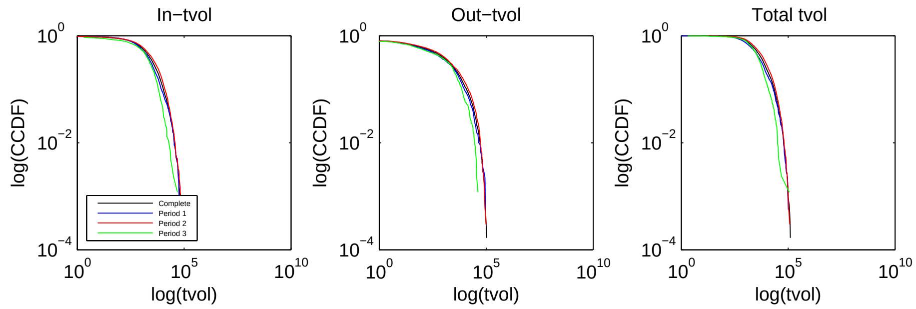

Figure 14: Quarterly data, tvol. Complementary cumulative distribution functions (ccdf) in-tvol (top), out-tvol (center), and total tvol (bottom) for all time periods on a log-log scale.

B Distributional Properties of Transaction Volumes

Here we report the results using another important measure for interbank networks, namely the transaction volumes (tvol). We use the same distribution fitting approach as before, differentiating between in-tvol, out-tvol and their sum (total tvol), respectively. Figures 14 and 15 show the ccdfs on a for the quarterly and daily variables on a log-log scale. We should stress that the minimum trade size on the e-MID market is 50,000 Euros. In order to run our estimation procedure in a reasonable amount of time, we rescale the tvol variables by a factor of 10^{-6} such that a transaction size of 50,000 is represented by a value of 0.05.[28] We then round the tvol variable towards the nearest integer (otherwise the discrete candidate distributions could not be accurately evaluated), again ignoring zero values. In this way, we restrict our samples to relatively large transaction volumes with at least 500,000 Euros, represented by positive integer values. Note that, besides the upward bias of the data and the fact that the data now span several orders of magnitude, it is again hard to visually detect linear decay over several orders of magnitude in the ccdfs. We should also stress that we did not perform the GOF exercise for the tvol variables, since it is too time-consuming in this case.

[28] Note that the maximum daily (quarterly) transaction volumes were 3.75bn (113.46bn) Euros for in-tvol, 4.96bn (111.93bn) Euros for out-tvol and 5.32bn (146.06bn) Euros for total tvol, respectively. For such huge numbers, the estimation procedure, in the numerical optimization for the power-law parameters, tends to take a very long computation time. Therefore, the results in this section should be treated with care, since the rescaling might affect our statistical analysis.

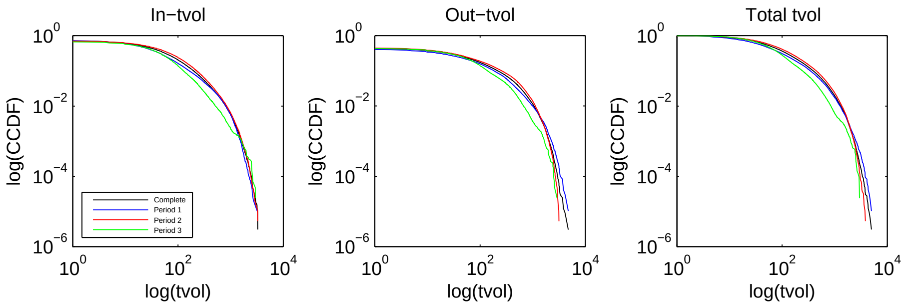

Figure 15: Daily data, tvol. Complementary cumulative distribution functions (ccdf) in-tvol (top), out-tvol (center), and total tvol (bottom) for all time periods on a log-log scale.

B.1 Daily Data

	In				**Out**				**Total**			
Period	1-3	1	2	3	1-3	1	2	3	1-3	1	2	3
Exponential	.1485	.1868	.1252	.1141	.1926	.1981	.1986	.1294	.1737	.2118	.1578	.1242
Gamma	.0765	.0972	.0675	.0821	.0685	.0566	.0690	.0834	.0833	.0876	.0766	.0942
Geometric	.1471	.1851	.1241	.1120	.1919	.1975	.1980	.1281	.1729	.2108	.1570	.1226
Log-normal	**.0274**	**.0253**	**.0276**	**.0340**	**.0319**	**.0344**	**.0310**	**.0357**	**.0231**	**.0248**	**.0224**	**.0259**
Neg. Bin.	.0686	.0876	.0604	.0753	.0599	.0489	.0598	.0766	.0751	.0779	.0685	.0880
Poisson	.6552	.6878	.6423	.6253	.6780	.6785	.6728	.6618	.6803	.6971	.6715	.6527
Power-law	.3587	.3324	.3643	.3836	.3416	.3320	.3490	.3774	.3678	.3406	.3693	.3897
Weibull	.0499	.0648	.0442	.0602	.0427	.0355	.0482	.0590	.0536	.0561	.0496	.0698

Table 9: Daily data, tvol. KS statistic for the candidate distributions (complete). Minimum values in bold. Significance tests not carried out.

Tables 9-11 show the results for the daily data. The complete distributions are now usually fitted best by Log-normal distributions, whereas the fit of the power-law is very poor in general. The power-law parameters are again very small, with typical values around 1.22, cf. Table 10 (top, complete). For the tail observations, the best fit again is always provided by Log-normal distributions, cf. Table 11. Interestingly, the tail exponents of the daily data are within the typical range of meaningful power-laws, cf. Table 10 (top, tail), but the power-law is still not the best description of the data. In the end, for the transaction volumes we find no evidence in favor of power-laws.

B.2 Quarterly Data

Tables 12 and 13 show the results for the quarterly data. The complete in-, out-, and total degree distributions are now fit best by Weibull, negative

Period	In 1-3	1	2	3	Out 1-3	1	2	3	Total 1-3	1	2	3
Daily												
Complete	1.23	1.24	1.22	1.24	1.21	1.22	1.21	1.22	1.22	1.23	1.21	1.22
	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.002)	(.001)	(.001)	(.001)	(.001)
Tail	2.66	2.15	2.72	2.86	3.23	2.63	4.88	3.39	3.33	3.27	3.35	3.38
	(.010)	(.009)	(.013)	(.042)	(.022)	(.023)	(.239)	(.079)	(.019)	(.036)	(.023)	(.069)
Quarterly												
Complete	1.13	1.15	1.13	1.16	1.14	1.15	1.14	1.15	1.12	1.12	1.14	1.14
	(.002)	(.004)	(.002)	(.006)	(.002)	(.004)	(.003)	(.006)	(.002)	(.003)	(.002)	(.005)
Tail	2.53	2.31	2.55	2.81	3.37	1.97	3.43	2.11	2.02	2.02	3.36	2.46
	(.050)	(.074)	(.060)	(.155)	(.346)	(.050)	(.284)	(.077)	(.020)	(.040)	(.315)	(.091)

Table 10: Power-law parameters and standard deviations, tvol. Values obtained via numerical maximization of the log-likelihood for discrete data. Standard deviations (in parentheses) approximated as $(\alpha - 1)/\sqrt{T}$, with T being the number of observations. Top: daily data, bottom: quarterly data.

Period	In 1-3	1	2	3	Out 1-3	1	2	3	Total 1-3	1	2	3
Exponential	.0752	.1339	.0720	.1270	.0306	.0851	.0442	.0887	.0315	.0547	.0201	.0782
Log-normal	**.0194**	**.0235**	**.0175**	**.0169**	**.0141**	**.0170**	**.0425**	**.0245**	**.0133**	**.0162**	**.0170**	**.0279**
Power-law	.0516	.0536	.0526	.0226	.0587	.0466	.0627	.0301	.0557	.0455	.0645	.0351

Table 11: Daily data, tvol. KS statistic for the candidate distributions (tail). Minimum values in bold. Significance tests not carried out.

Binomial, and Log-normal distributions, respectively. In many cases, these distributions yield comparable KS statistics, but the clear advantage of the Log-normal distribution for the daily data does not carry over to the quarterly level in all cases. Similar to the daily estimates, the power-law parameters are within the usual range of empirical power-laws. As before, however, the tails are best described by Log-normal distributions. Therefore, while the tails of the tvol variables are somewhat fatter compared to the degree and ntrans variables, the power-law remains a poor description of the data.

	In				Out				Total			
Period	1-3	1	2	3	1-3	1	2	3	1-3	1	2	3
Exponential	.1741	.1940	.1722	.1258	.3419	.3656	.3383	.3139	.1520	.1808	.1390	.0887
Gamma	.0511	.0713	.0429	**.0313**	.0342	.0370	.0465	.0449	.0722	.0881	.0631	.0651
Geometric	.1741	.1940	.1722	.1258	.3418	.3655	.3383	.3139	.1519	.1808	.1390	.0886
Log-normal	.0619	.0486	.0626	.1075	.0610	.0635	.0566	.0969	**.0332**	**.0212**	**.0316**	.0810
Neg. Bin.	.0491	.0701	.0404	.0314	**.0298**	**.0319**	.0417	**.0414**	.0712	.0867	.0626	.0649
Poisson	.7128	.7366	.6989	.6848	.7579	.7574	.7516	.7264	.7139	.7277	.6969	.6907
Power-law	.4030	.3888	.4106	.3765	.2730	.2752	.2910	.2930	.4733	.4614	.4545	.4588
Weibull	**.0202**	**.0378**	**.0165**	.0462	.0308	.0334	**.0260**	.0585	.0447	.0529	.0430	**.0566**

Table 12: Quarterly data, tvol. KS statistic for the candidate distributions (complete). Minimum values in bold. Significance tests not carried out.

	In				Out				Total			
Period	1-3	1	2	3	1-3	1	2	3	1-3	1	2	3
Exponential	.0855	.1380	.0759	.1476	.0961	.1529	.0567	.1563	.1071	.1434	.0721	.1508
Log-normal	**.0271**	**.0542**	**.0267**	**.0521**	**.0674**	**.0407**	**.0475**	**.0459**	**.0353**	**.0483**	**.0695**	**.0394**
Power-law	.0642	.0588	.0672	.0703	.0785	.0733	.0813	.0790	.0788	.0778	.0812	.0484

Table 13: Quarterly data, tvol. KS statistic for the candidate distributions (tail). Minimum values in bold. Significance tests not carried out.