

# On the distribution of links in the interbank network: evidence from the e-MID overnight money market

Daniel Fricke · Thomas Lux

Received: 20 August 2013 / Accepted: 17 December 2014 / Published online: 13 February 2015  
© Springer-Verlag Berlin Heidelberg 2015

**Abstract** Previous literature on statistical properties of interbank networks 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 reports, we find no evidence in favor of a power-law characterizing the degree distribution. Rather, the data are best described by negative Binomial distri-

---

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 as well as those of two anonymous reviewers. Thomas Lux also gratefully acknowledges financial support of this research from the European Union Seventh Framework Programme under Grant Agreement No. 619255 while Daniel Fricke gratefully acknowledges FP-7 support under Grant Agreement CRISIS-ICT-2011-288501.

---

D. Fricke (✉)

Institute for New Economic Thinking, Oxford Martin School, Oxford, UK  
e-mail: daniel.fricke@sbs.ox.ac.uk

D. Fricke

CABDyN Complexity Centre, Saïd Business School, University of Oxford, Park End St, Oxford OX1 1HP, UK

D. Fricke

Kiel Institute for the World Economy, Hindenburgufer 66, 24105 Kiel, Germany

T. Lux (✉)

Department of Economics, University of Kiel, Olshausenstr. 40, 24118 Kiel, Germany  
e-mail: lux@economics.uni-kiel.de

T. Lux

Banco de España Chair in Computational Economics, University Jaume I, Campus del Riu Sec, 12071 Castellón, Spain

butions. For quarterly data, Weibull, Gamma, and Exponential distributions tend to provide comparable fits. We find similar 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 previous claims that these interbank data fall into the category of scale-free networks.

**Keywords** Interbank market · Network models · Financial network · Systemic risk · Power-law

**JEL Classification** G21 · G01 · E42

## 1 Introduction and existing literature

Since the onset of the global financial crisis (GFC) in 2007/2008, 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 relatively 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.<sup>1</sup>

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.<sup>2</sup> In this regard, the understanding of the structure and functioning of complex networks has advanced significantly in recent years. In particular, uncovering the working of economic networks is extremely useful, as it sheds light on the structure of interactions among economic agents (cf. Schweitzer et al. 2009). Examples include the world trade network (Fagiolo et al. 2010), job information networks (Ioannides and Loury 2004), and bank-firm networks (De Masi and Gallegati 2012).

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 (Erdős 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

<sup>1</sup> See e.g., Upper and Worms (2004), Nier et al. (2007) and Gai et al. (2011).

<sup>2</sup> See Haldane and May (2011) and Albert et al. (2000).

fraction of nodes with degree  $k$  is proportional to  $k^{-\alpha}$ , where  $\alpha$  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 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. An important feature of such scale-free networks is that they can be described as robust-yet-fragile,<sup>3</sup> 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, many interbank networks have been reported to resemble scale-free networks with power-law exponents lying in a relatively narrow range around 2.3 (cf. [Boss et al. 2004](#); [Soramäki et al. 2007](#); [De Masi et al. 2006](#); [Cont et al. 2013](#)), even though most papers lack a thorough statistical analysis of the issue,<sup>4</sup> 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. It needs to be emphasized, though, that there is no strict link between the finding of a power-law degree distribution and the fragility of a (banking) network: First, mechanisms like preferential attachment could be also at work for broad degree distributions that are not exactly of a power-law type. Second, for different definitions of ‘fragility’ (number of contagious defaults or total loss of bank capital), the complex interaction between network topology, the structure of shocks as well as the balance sheet structure could lead to different rankings of various network classes in terms of their resilience (cf. [Roukny et al. 2013](#)).

The distribution of network degrees is just one example among many phenomena in the natural sciences as well as from the socioeconomic 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 ([Castaldi and Dosi 2009](#); [Fagiolo et al. 2008](#)), or the distribution of large asset returns ([Mandelbrot 1963](#); [Krämer and Runde 1996](#); [Lux 2000](#)). 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

<sup>3</sup> See [Albert et al. \(2000\)](#).

<sup>4</sup> This is not to say these papers do not yield very useful phenomenological information about the pertinent network structures. The point here is rather that there seems to exist a convention in the network literature to produce a power-law graph without a deeper statistical analysis of this issue and also often in spite of a relatively small sample of observations that additionally impedes such a strong inference on the underlying distribution.

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 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. It is actually a method that is 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 extend 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; 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 Fréchet, 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. For example, the outer part approaches either a power-law decay, an exponential decay, or a decay toward 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 functions. Since degree distributions are discrete (degrees being integer numbers), it 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.

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. \(2013\)](#) have shown recently that the networks' properties depend on the aggregation period.<sup>5</sup> We will, therefore, not confine our analysis to daily data (the basic frequency of our dataset), 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 typically outperforms the power-law. Overall, these findings indicate that the power-law is typically a poor description of the data, implying that scale-free networks are not an appropriate model for the structure of the Italian interbank network. Moreover, the networks contain a substantial level of asymmetry, related 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: Sect. 2 gives a short introduction into (interbank) networks, Sect. 3 briefly introduces the Italian e-MID trading system and gives an overview of the dataset we have access to, Sect. 4 describes our findings, and Sect. 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) toward all other banks. The diagonal elements  $d_{ii}$  are zero, since a

<sup>5</sup> 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.

bank will not trade with itself.<sup>6</sup> 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.<sup>7</sup> 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  $a_{ij}^u = \max(a_{ij}, a_{ji})$ .

As usual, some data aggregation is necessary to represent the system as a 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.<sup>8</sup> 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 are 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. \(2013\)](#).<sup>9</sup> 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), and 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.

<sup>6</sup> This is of course only true when taking banks as consolidated entities.

<sup>7</sup> 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 for the presence or absence of a claim, respectively.

<sup>8</sup> The vast majority of trades (roughly 95%) is conducted in Euro.

<sup>9</sup> See also the e-MID website <http://www.e-mid.it/>.

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

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

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  $i$  ( $k_i^{\text{in}}$  and  $k_i^{\text{out}}$ ) as

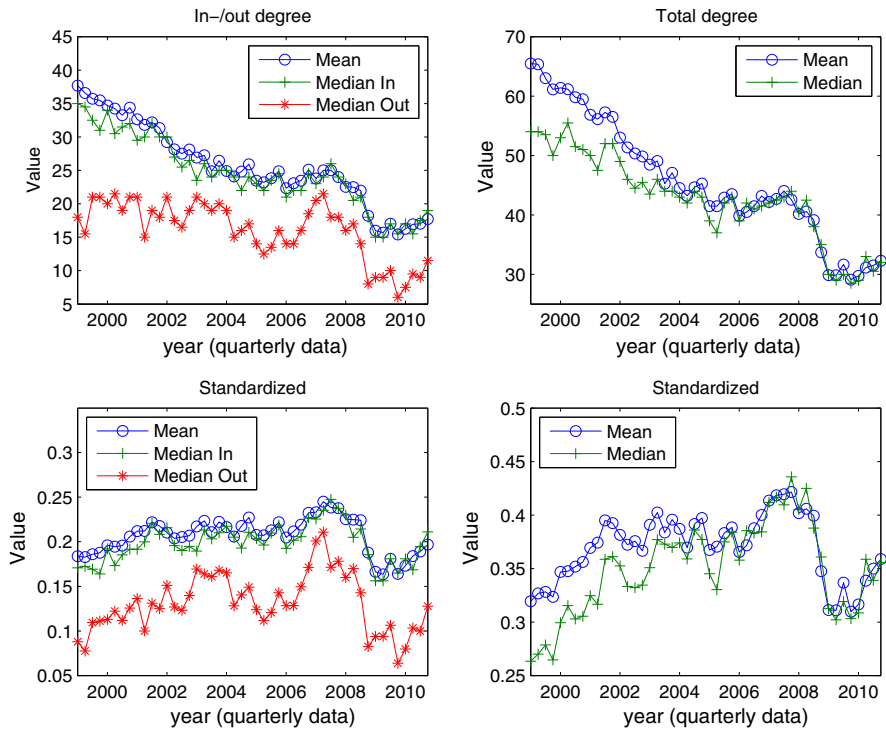
$$\begin{aligned} k_i^{\text{in}} &= \sum_{j \neq i} a_{ji} \\ k_i^{\text{out}} &= \sum_{j \neq i} a_{ij}, \end{aligned} \tag{2}$$

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 daily data, a link  $a_{ij}$  simply equals 1 if there has been a credit extended from bank  $i$  to  $j$  on that specific day. In the rare case of multiple loans granted between identical counterparties in the same direction, the link count also remains equal to 1. However, when counting the number of transactions (see below), multiple loans are summed up to give the statistics of transactions between counterparties during a certain period of time. Moving from daily to longer horizons, links remain dichotomous quantities, i.e.,  $a_{ij} = 1$  means that at least once (but may be multiple times) a loan has been granted by bank  $i$  to bank  $j$ . The number of transactions, however, count the overall number of such loans over a month, quarter, etc.

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 the sum over all loans received or granted by bank  $i$  over a certain time horizon:

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



**Fig. 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)

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

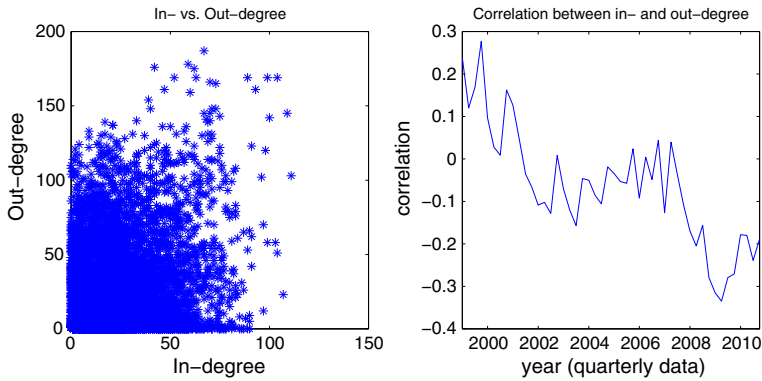
$$n_i^{\text{total}} = n_i^{\text{in}} + n_i^{\text{out}}. \quad (4)$$

#### 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.<sup>10</sup> 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,

<sup>10</sup> Note that the mean in- and out-degree are identical by definition.





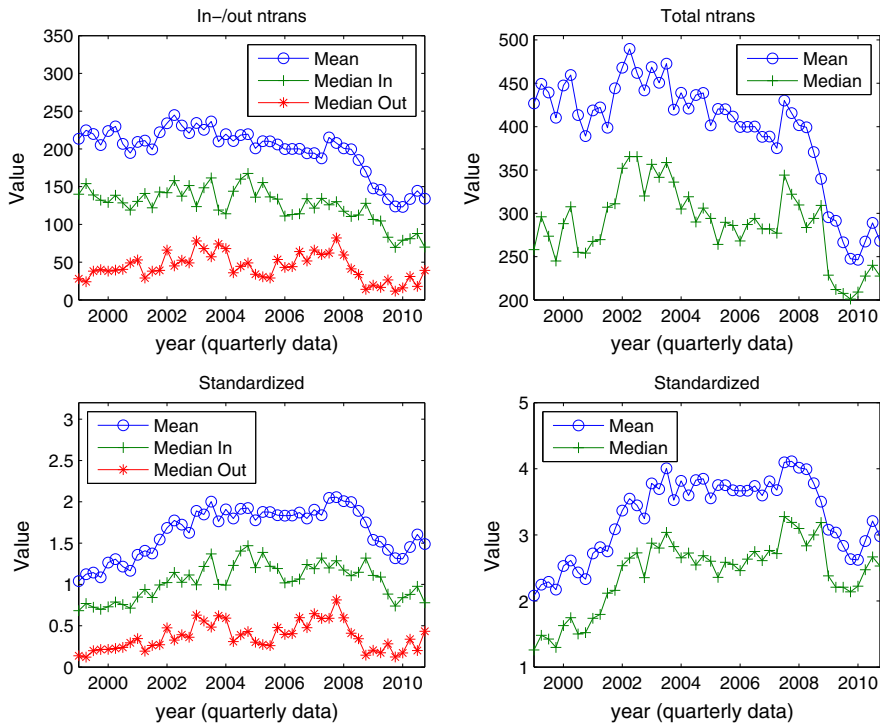
**Fig. 2** Left Scatter plot of in- versus out-degree. Correlation .0899. Right Correlation between individual banks' in- and out-degree over time. Italian banks only

pointing toward 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.<sup>11</sup> 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 Fig. 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, thus, contains a considerable amount of information. The right panel of Fig. 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, Fig. 3 shows the dynamics of the mean and median in-/out-ntrans (left) and the total ntrans (right). The upper left panel shows that the

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



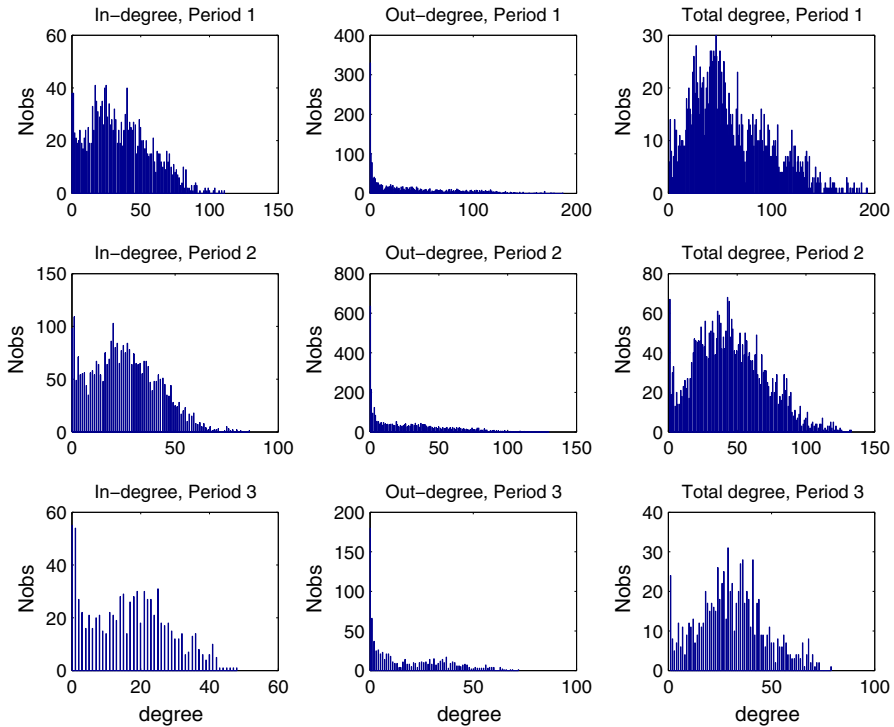
**Fig. 3** Mean and median number of transactions over time. *Left* Directed network. *Right* Undirected network. *Top* absolute levels. *Bottom* standardized values (divided by the number of active banks per quarter)

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 toward 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 dataset into three periods: Period 1 covers quarters 1–10, period 2 covers quarters 11–39, and period 3 covers the remaining quarters 40–48.<sup>12</sup> Assuming that the

<sup>12</sup> Note that the first subsample roughly coincides with the dataset used by De Masi et al. (2006).

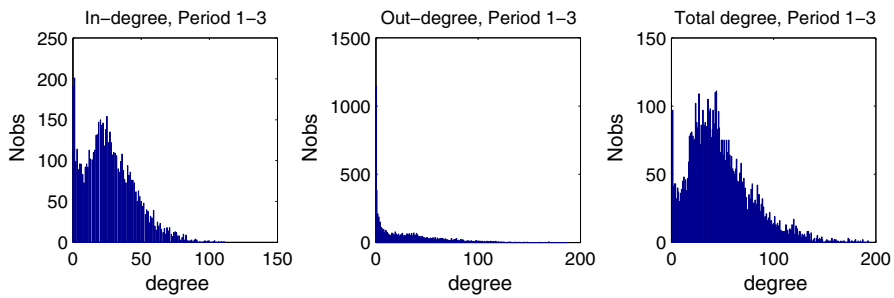


**Fig. 4** Quarterly data, degree. Histograms for in-degree (*left*), out-degree (*center*), and total degree (*right*) for Period 1, 2, and 3

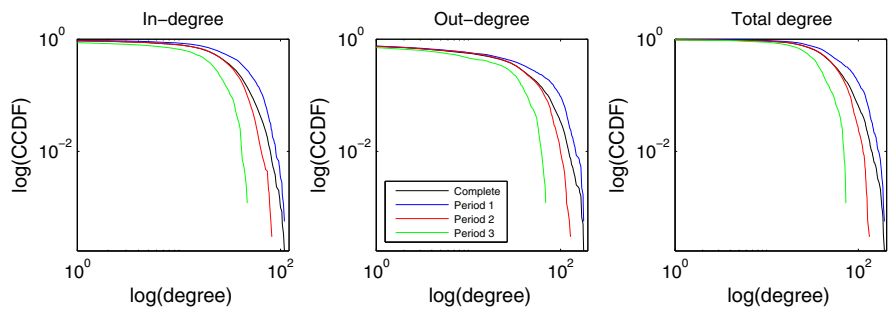
realizations of single days (quarters) are iid draws (or weakly dependent ones) from the same underlying data-generating process, allows us to pool the data of the three subperiods into larger samples for the in-, out-, and total degrees (ntrans) of active banks, respectively. We use both daily and quarterly 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.<sup>13</sup> For the daily (quarterly) data, this amounts to a total of 96,892 (1780), 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

<sup>13</sup> In Appendix 4, we present a similar analysis for the distribution of transaction volumes of individual institutions.



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



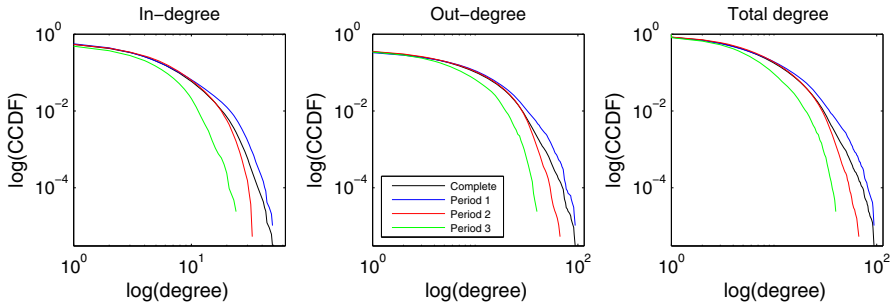
**Fig. 6** Quarterly data, degree. Complementary cumulative distribution functions (ccdf) for degrees: in-degree (*left*), out-degree (*center*), and total degree (*right*) for all time periods on a log–log scale

quarterly data. We see that the histograms look very different when comparing in- and out-degrees for each sample period. 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, and also whether any variable follows a certain specific distribution. In our case, the KS test statistic is calculated as

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

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 Fig. 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



**Fig. 7** Daily data, degree. Complementary cumulative distribution functions (ccdf) for degrees: in-degree (left), out-degree (center), and total degree (right) for all time periods on a log–log scale

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 Figs. 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 Fig. 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 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  $\mu$  (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 nonzero values for the degree and *ntrans* variables in the following.<sup>14</sup> This means, for some distribution functions, we are using truncated variables in general (both for the complete and tail observations) and need to adjust the ML estimators for these distributions accordingly, cf. Appendix 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 ML, where the optimal  $x_m$  is the one corresponding to the lowest KS statistic. Details can be found in Appendix 2.<sup>15</sup> The tail region is then defined by the scale parameter  $x_m$ , and the all candidate tail 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 dataset 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.<sup>16</sup> 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

<sup>14</sup> This is important, since we cannot replicate the large number of zero values that we observe in the empirical data based on these distributions. Ignoring zeros reduces the number of quarterly observations to 1,742, 3,271, and 788 for the in-variables, and 1450, 2733, 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 remains unaffected, since only active banks are in the sample.

<sup>15</sup> 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.

<sup>16</sup> 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.

perform individual Monte Carlo exercises.<sup>17</sup> 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 datasets. 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 remaining ones have already been found to be inferior under the KS criterion. Details on the Monte Carlo design can be found in Appendix 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.

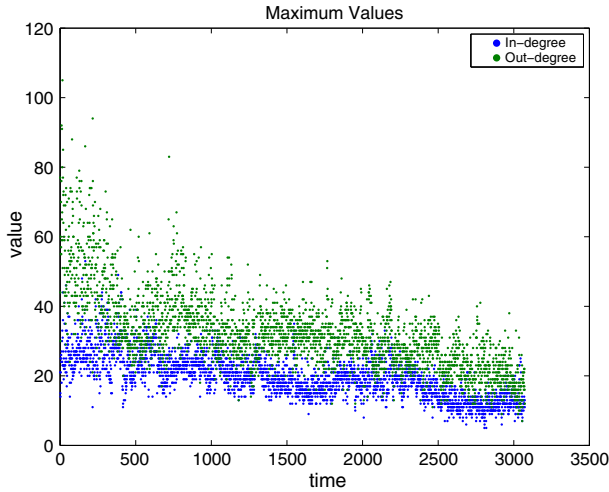
#### 4.2.2 Daily data

We start our analysis with the daily degree data for which earlier studies have reported power-laws (De Masi et al. 2006). 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 datasets, 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; Clauset et al. 2009).

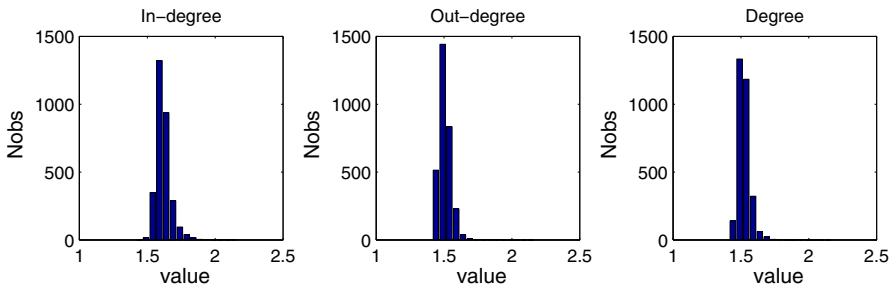
For our data, Fig. 8 shows the maximum in- and out-degrees for the individual days over time. We see that the criterion of Stumpf and Porter (2012) is typically violated. Thus, it should be hard to find evidence in favor of the power-law hypothesis for the complete distributions. Additionally, the number 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, Figs. 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. Fig. 9. In fact, we will see that this stability tends to carry over to the complete distributions of the aggregated data as well. In

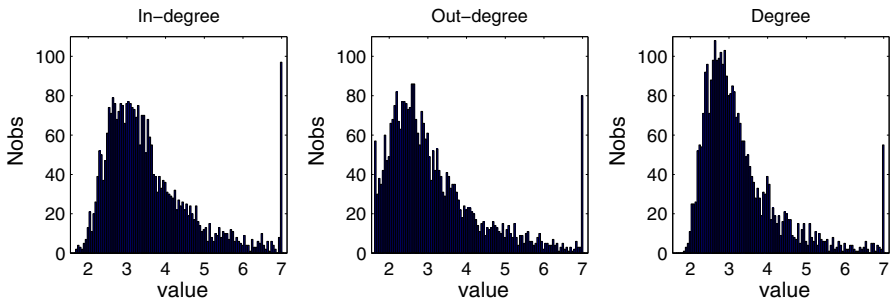
<sup>17</sup> See Clauset et al. (2009) and Stumpf et al. (2005) for similar approaches.



**Fig. 8** Daily data. Maximum in- and out-degrees over time



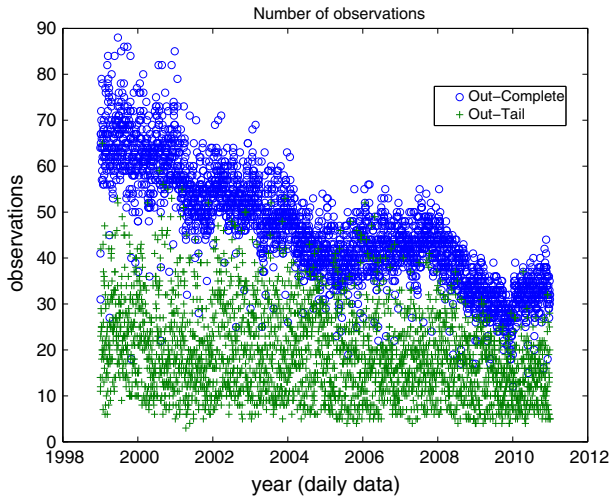
**Fig. 9** Daily data, degree. Histograms for the power-law exponents for the complete distributions, in-, out-, and total degree, respectively



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

contrast, there is a substantial level of heterogeneity for the power-law exponent of the tail observations for the individual days, cf. Fig. 10. Thus, we cannot confirm previously reported findings of ‘interesting’ tail parameters between 2 and 3 for any of the





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

degree variables.<sup>18</sup> While numerous observations lie within this range, for many days, we find substantially larger values, at times as large as 7.<sup>19</sup> Apparently, the daily tail data are too noisy to identify a ‘typical’ tail parameter, cf. Fig. 11.<sup>20</sup> The mismatch between the narrow range of values obtained for the complete dataset 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 on adjacent obser-

<sup>18</sup> De Masi et al. (2006) report power-law exponents between 2 and 3 for total degree, in-degree, and out-degree for daily data of our period 1 of the e-MID overnight record. Soramäki et al. (2007) also report values in this range for daily interbank payments within the U.S. Fedwire system. Boss et al. (2004) fit two power-laws to the more central and the extremal region of the in-degree distribution of monthly Austrian interbank liabilities with the extremal region exhibiting slopes of 1.73 for in-degrees and 2.01 for out-degree. This visual illustration could, however, be as well interpreted as indicating an overall exponential shape. Bech and Atalay (2010) study daily federal fund credits in the U.S. Their study constitutes a rare example of a comparative fit of a number of candidate distributions. While the overall shape of the distribution does not quite show a straight linear shape in a log–log plot, they report that the power-law gave the best fit among the candidates for out-degrees while the negative Binomial did best fit the in-degree distribution.

<sup>19</sup> 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 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.

<sup>20</sup> 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.

**Table 1** Daily data, degree. KS statistic for the candidate distributions (complete data)

Period	In				Out				Total			
	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.	<b>.0063</b>	<b>.0177</b>	<b>.0082</b>	<b>.0224</b>	<b>.0114</b>	<b>.0160</b>	<b>.0138</b>	<b>.0105</b>	<b>.0153</b>	<b>.0208</b>	<b>.0115</b>	<b>.0103</b>
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

Minimum values in bold indicate the best fitting distribution. Asterisks would indicate non-rejection of this distribution at the 5 % significance 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

**Table 2** Power-law parameters and standard errors, degree

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

Values obtained via numerical maximization of the log-likelihood for discrete data  
Standard errors (in parentheses) approximated as  $(\alpha - 1)/\sqrt{T}$ , with  $T$  being the number of observations  
*Top* daily data, *bottom* quarterly data

vations. While it is not straightforward to check this assumption for complete daily ensembles (as opposed 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 (2014) demonstrate that the e-MID network is indeed structurally stable along many dimensions.

We report our estimation results for the pooled daily data in Tables 1, 2, and 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 complete samples and the

**Table 3** Daily data, degree. KS statistic for the candidate distributions (tail)

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

Minimum values in bold indicate the best fitting distribution

Asterisks indicate non-rejection of this distribution at the 5% significance level, where the critical values were obtained from a Monte Carlo exercise as described in the main text

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 datasets. In contrast to the complete distributions, the GOF experiments suggest that the estimated distributions are mostly not rejected for the tail observations.<sup>21</sup> 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  $\alpha \rightarrow \infty$ . 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 the visual inspection of absence of a linear slope over the complete support and very fast decline at the end in Fig. 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

<sup>21</sup> 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.

**Table 4** Quarterly data, degree. KS statistic for the candidate distributions (complete)

Period	In				Out				Total			
	1-3	1	2	3	1-3	1	2	3	1-3	1	2	3
Exponential	.1474	.1544	.1661	.1520	.0797	.0932	<b>.0710</b>	.0925	.1740	.1675	.1887	.2171
Gamma	.0573	.0543	.0723	.0942	<b>.0595</b>	<b>.0514</b>	.0738	.0961	.0414	<b>.0284*</b>	.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	<b>.0380</b>	<b>.0342*</b>	<b>.0456</b>	<b>.0689</b>	.0624	.0609	.0736	<b>.0912</b>	<b>.0246</b>	.0361	<b>.0325</b>	<b>.0495*</b>

Minimum values in bold indicate the best fitting distribution  
 Asterisks indicate non-rejection of this distribution at the 5% significance level, where the critical values were obtained from a Monte Carlo exercise as described in the main text

**Table 5** Quarterly data, degree. KS statistic for the candidate distributions (tail)

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

Minimum values in bold indicate the best fitting distribution  
 Asterisks indicate non-rejection of this distribution at the 5% significance level, where the critical values were obtained from a Monte Carlo exercise as described in the main text

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 an earlier study. It is not clear how these estimates were obtained.

4.2.3 Quarterly data

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 typically rejected at the 5% significance level, 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 in 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% 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 convincing evidence in favor of a power-law distribution.

#### 4.2.4 Robustness and discussion

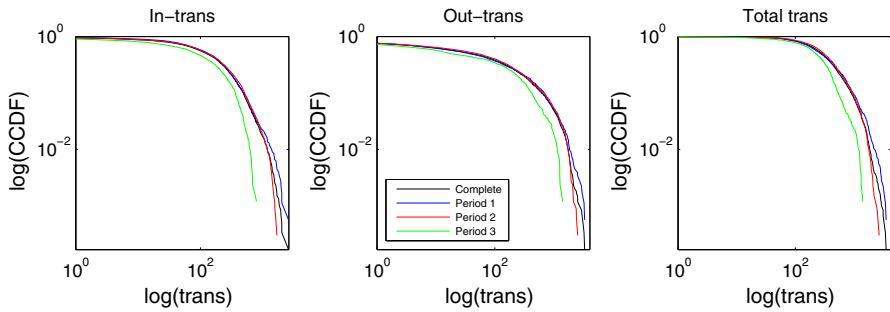
A reason for not finding evidence for power-law distributions may be the fact that we focus on the subnetwork formed by Italian banks only. [Stumpf et al. \(2005\)](#) have shown that (randomly chosen) subnetworks of scale-free networks are in fact not scale-free. Therefore, we also checked the distributions 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). 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. \(2013\)](#) 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.

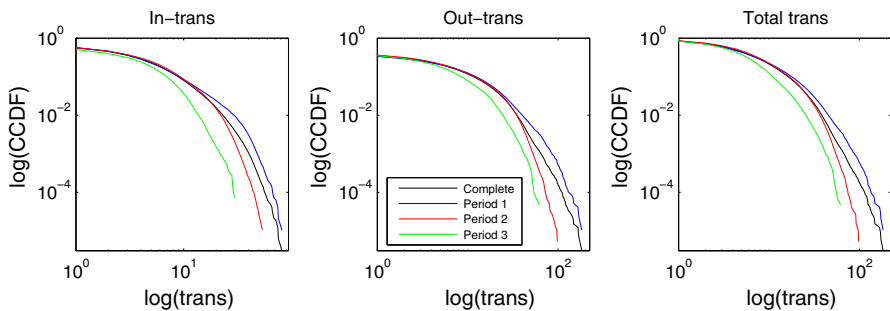
Summing up, our results indicate that the power-law hypothesis needs to be tested more thoroughly for other networks in general and the interbank network in particular, with the power-law being one of many candidate distributions. The findings are in line with other studies casting doubts on certain 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



**Fig. 12** Quarterly data, ntrans. Complementary cumulative distribution functions (ccdf) for the number of transactions: in-trans (*top*), out-trans (*center*), and total trans (*bottom*) for all time periods on a log–log scale



**Fig. 13** Daily data, ntrans. Complementary cumulative distribution functions (ccdf) for the number of transactions in-trans (*top*), out-trans (*center*), and total trans (*bottom*) for all time periods on a log–log scale

‘weak’ dependency.<sup>22</sup> 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 combined with the relatively narrow range of observed values. 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.

<sup>22</sup> 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.

**Table 6** Power-law parameters and standard errors, ntrans

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

Values obtained via numerical maximization of the log-likelihood for discrete data  
 Standard errors (in parentheses) approximated as  $(\alpha - 1)/\sqrt{T}$ , with  $T$  being the number of observations.  
 Quarterly data

**Table 7** Quarterly data, ntrans. KS statistic for the candidate distributions (complete data)

Period	In				Out				Total			
	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	<b>.0153</b>	.0287	.0199	<b>.0388</b>	<b>.0167</b>	<b>.0165</b>	<b>.0185</b>	<b>.0498</b>	.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	<b>.0353</b>	<b>.0477</b>	<b>.0321</b>	.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	<b>.0244</b>	<b>.0186</b>	.0420	.0383	.0389	.0392	.0668	.0391	.0481	.0386	<b>.0437*</b>

Minimum values in bold indicate the best fitting distribution  
 Asterisks indicate non-rejection of this distribution at the 5% significance level, where the critical values were obtained from a Monte Carlo exercise as described in the main text

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

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

**Table 8** Quarterly data, ntrans. KS statistic for the candidate distributions (tail)

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

Minimum values in bold indicate the best fitting distribution

Asterisks indicate non-rejection of this distribution at the 5% significance level, where the critical values were obtained from a Monte Carlo exercise as described in the main text

is quite surprising, given that the scaling parameters now lie 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 for testing the power-law hypothesis.

As another robustness check, we investigated the distribution of transaction volumes (tvol), cf. Appendix 4, 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 suggested 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 data-generating processes are not supported by the empirical shape of the degree distribution that appears very far from the implied power-law structure. Note that these findings are also not in line with a large part of the empirical (interbank) network literature for other datasets, raising doubts on the universality of scale-free behavior of interbank networks. 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. 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; 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.

**Appendix 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.

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

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

or

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

which can be easily evaluated.

### 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  $\theta$  the vector of parameters, can be written as

$$L(\theta|x_1, \dots, x_n) = p(x_1, \dots, x_n|\theta) = \prod_i^n p(x_i|\theta), \quad (12)$$

or in logs

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

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

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

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

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

which can be written as

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

The second part of this equation 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 ML estimator. Therefore, we need to find the  $\theta$  that maximizes Eq. (16). The standard ML estimator would not be efficient.

### Appendix 2: Discrete power-laws and parameter estimation

This presentation is mostly based on [Clauset et al. \(2009\)](#).

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.<sup>23</sup>

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.

Estimation of  $\alpha$  and  $x_m$

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

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

<sup>23</sup> [Clauset et al. \(2007\)](#) show that this is necessary for datasets 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.

where  $n$  is the number of observations.<sup>24</sup> For simplicity, we approximate the standard error of the estimated  $\hat{\alpha}$  (for  $\hat{\alpha} > 1$ ) using the closed-form solution based on continuous data.<sup>25</sup> Neglecting higher-order terms, this can be calculated as

$$\sigma = \frac{\hat{\alpha} - 1}{\sqrt{n}}. \quad (21)$$

However, the equations assume that  $x_m$  is known in order to obtain an accurate estimate of  $\alpha$ .<sup>26</sup> 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  $\alpha$  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.

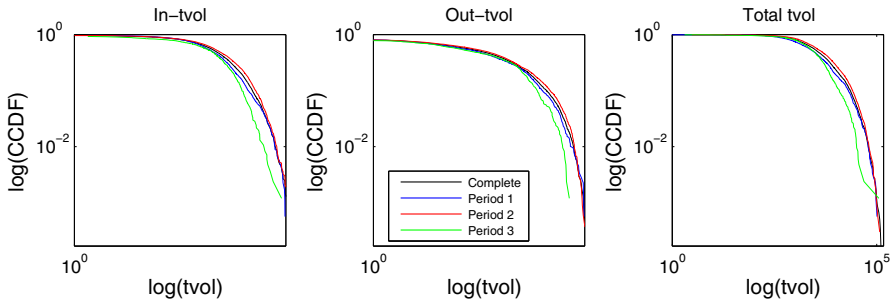
### Appendix 3: Goodness-of-fit test for the estimated distributions

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 datasets from the estimated distribution, compute the distribution of KS statistics, and compare the results to the observed value for the original dataset. If the KS statistic of the empirical dataset is beyond the  $\alpha$  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.

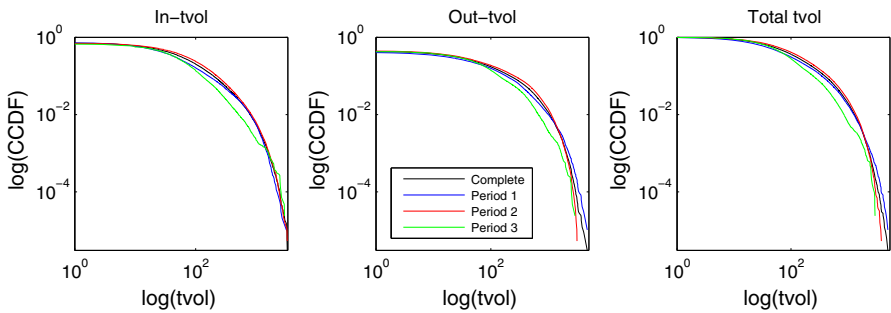
<sup>24</sup> Using a quadratic approximation of the log-likelihood at its maximum, [Clauset et al. \(2009\)](#) also derive an approximate closed-form solution for the estimate of  $\alpha \simeq 1 + n / \left( \sum_{i=1}^n \ln \left[ \frac{x_i}{x_m - .5} \right] \right)$ . This can be seen as an adjusted Hill-estimator, see [Hill \(1975\)](#). While we always report the exact ML estimator, we checked that the approximation is typically not too bad.

<sup>25</sup> [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.

<sup>26</sup> See [Clauset et al. \(2007, 2009\)](#) for an extensive discussion.



**Fig. 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



**Fig. 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

#### Appendix 4: 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 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 .05.<sup>27</sup> We then round the tvol variable toward 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

<sup>27</sup> Note that the maximum daily (quarterly) transaction volumes were 3.75 bn (113.46 bn) Euros for in-tvol, 4.96 bn (111.93 bn) Euros for out-tvol, and 5.32 bn (146.06 bn) 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.

**Table 9** Daily data, tvol. KS statistic for the candidate distributions (complete)

Period	In				Out				Total			
	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	<b>.0274</b>	<b>.0253</b>	<b>.0276</b>	<b>.0340</b>	<b>.0319</b>	<b>.0344</b>	<b>.0310</b>	<b>.0357</b>	<b>.0231</b>	<b>.0248</b>	<b>.0224</b>	<b>.0259</b>
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

Minimum values in bold  
Significance tests not carried out

**Table 10** Power-law parameters and standard errors, tvol. Values obtained via numerical maximization of the log-likelihood for discrete data

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

Standard errors (in parentheses) approximated as  $(\alpha - 1)/\sqrt{T}$ , with  $T$  being the number of observations. Top: daily data, bottom: quarterly data

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.

Daily data

Tables 9, 10, and 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

**Table 11** Daily data, tvol. KS statistic for the candidate distributions (tail)

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

Minimum values in bold  
Significance tests not carried out

**Table 12** Quarterly data, tvol. KS statistic for the candidate distributions (complete data)

Period	In				Out				Total			
	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	<b>.0313</b>	.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	<b>.0332</b>	<b>.0212</b>	<b>.0316</b>	.0810
Neg. bin.	.0491	.0701	.0404	.0314	<b>.0298</b>	<b>.0319</b>	.0417	<b>.0414</b>	.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	<b>.0202</b>	<b>.0378</b>	<b>.0165</b>	.0462	.0308	.0334	<b>.0260</b>	.0585	.0447	.0529	.0430	<b>.0566</b>

Minimum values in bold  
Significance tests not carried out

**Table 13** Quarterly data, tvol. KS statistic for the candidate distributions (tail)

Period	In				Out				Total			
	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	<b>.0271</b>	<b>.0542</b>	<b>.0267</b>	<b>.0521</b>	<b>.0674</b>	<b>.0407</b>	<b>.0475</b>	<b>.0459</b>	<b>.0353</b>	<b>.0483</b>	<b>.0695</b>	<b>.0394</b>
Power-law	.0642	.0588	.0672	.0703	.0785	.0733	.0813	.0790	.0788	.0778	.0812	.0484

Minimum values in bold  
Significance tests not carried out

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.

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 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.

## References

- Albert R, Jeong H, Barabasi A-L (2000) Error and attack tolerance of complex networks. *Nature* 406:378–482
- Alderson DL, Li L (2007) Diversity of graphs with highly variable connectivity. *Phys Rev E* 75:046102
- Anderson CW (1970) Extreme value theory for a class of discrete distributions with applications to some stochastic processes. *J Appl Prob* 7(1):99–113
- Avnir D, Biham O, Lidar D, Malcai O (1998) Is the geometry of nature fractal? *Science* 279(5347):39–40
- Axtell RL (2001) Zipf distribution of U.S. firm sizes. *Science* 293(5536):1818–1820
- Beaupain R, Durré A (2012) Nonlinear liquidity adjustments in the Euro area overnight money market. Working paper series 1500, European Central Bank
- Bech M, Atalay E (2010) The topology of the federal funds market. *Phys A* 389(22):5223–5246
- Boss M, Elsinger H, Summer M, Thurner S (2004) Network topology of the interbank market. *Quant Finance* 4(6):677–684
- Caldarelli G (2007) *Scale-free networks*. Oxford University Press, Oxford
- Castaldi C, Dosi G (2009) The patterns of output growth of firms and countries: scale invariances and scale specificities. *Empir Econ* 37(3):475–495
- Clauset A, Young M, Gleditsch KS (2007) On the frequency of severe terrorist events. *J Confl Resolut* 51(1):58–87
- Clauset A, Rohilla Shalizi C, Newman MEJ (2009) Power-law distributions in empirical data. *SIAM Rev* 51:661–703
- Cont R, Santos EB, Moussa A (2013) Network structure and systemic risk in banking systems. In: Fouque J, Langsam J (eds) *Handbook of systemic risk*. Cambridge University Press, Cambridge
- Danielsson J, de Haan L, Peng L, de Vries C (2001) Using a bootstrap method to choose the sample fraction in tail index estimation. *J Multivar Anal* 76(2):226–248
- De Masi G, Iori G, Caldarelli G (2006) Fitness model for the Italian interbank money market. *Phys Rev E* 74(6):66112
- De Masi G, Gallegati M (2012) Bank-firms topology in Italy. *Empir Econ* 43(2):851–866
- Drees H, Kaufmann E (1998) Selecting the optimal sample fraction in univariate extreme value estimation. *Stoch Processes Appl* 75(2):149–172
- Erdős P, Renyi A (1959) On random graphs. *Publ Math* 6:290–297
- European Central Bank (2007) Euro money market study 2006. Final report, ECB
- Fagiolo G, Napoletano M, Roventini A (2008) Are output growth-rate distributions fat-tailed? Some evidence from OECD countries. *J Appl Econ* 23(5):639–669
- Fagiolo G, Reyes J, Schiavo S (2010) The evolution of the world trade web: a weighted-network analysis. *J Evolut Econ* 20(4):479–514
- Finger K, Fricke D, Lux T (2013) Network analysis of the e-MID overnight money market: the informational value of different aggregation levels for intrinsic dynamic processes. *Comput Manag Sci* 10(2–3):187–211
- Fricke D, Lux T (2014) Core-periphery structure in the overnight money market: evidence from the e-MID trading platform. *Comput Econ*. doi:10.1007/s10614-014-9427-x
- Gabaix X (1999) Zipf's law for cities: an explanation. *Q J Econ* 114(3):739–767
- Gai P, Haldane A, Kapadia S (2011) Complexity, concentration and contagion. *J Monet Econ* 58(5):453–470
- Haldane AG, May RM (2011) Systemic risk in banking ecosystems. *Nature* 469(7330):351–355
- Hill BM (1975) A simple general approach to inference about the tail of a distribution. *Ann Stat* 3(5):1163–1174
- Ioannides YM, Loury LD (2004) Job information networks, neighborhood effects, and inequality. *J Econ Lit* 42(4):1056–1093



- Krämer W, Runde R (1996) Stochastic properties of German stock returns. *Empir Econ* 21(2):281–306
- Leadbetter M (1983) Extremes and local dependence in stationary sequences. *Zeitschrift fuer Wahrscheinlichkeitstheorie und Verwandte Gebiete* 65:291–306
- Lux T (2000) On moment condition failure in German stock returns: an application of recent advances in extreme value statistics. *Empir Econ* 25(4):641–652
- Mandelbrot B (1963) The variation of certain speculative prices. *J Bus* 36:394
- Nier E, Yang J, Yorulmazer T, Alentorn A (2007) Network models and financial stability. *J Econ Dyn Control* 31(6):2033–2060
- Reiss R-D, Thomas M (2007) *Statistical analysis of extreme values: with applications to insurance, finance, hydrology and other fields*, 3rd edn. Birkhäuser Verlag, Switzerland
- Roukny T, Bersini H, Piroette H, Caldarelli G, Battiston S (2013) Default cascades in complex networks: topology and systemic risk. *Sci Rep* 3(2759):1–8
- Schweitzer F, Fagiolo G, Sornette D, Vega-Redondo F, Vespignani A, White DR (2009) Economic networks: the new challenges. *Science* 325(5939):422–425
- Silverberg G, Verspagen B (2007) The size distribution of innovations revisited: an application of extreme value statistics to citation and value measures of patent significance. *J Econom* 139(2):318–339
- Soramäki K, Bech ML, Arnold J, Glass RJ, Beyeler W (2007) The topology of interbank payment flows. *Phys A* 379:317–333
- Stumpf MPH, Ingram PJ (2005) Probability models for degree distributions of protein interaction networks. *Europhys Lett* 71(1):152–158
- Stumpf MPH, Ingram PJ, Nouvel I, Wiuf C (2005) Statistical model selection applied to biological network data. *Proc Comput Syst Biol* 3:65–73
- Stumpf MPH, Wiuf C, May RM (2005) Subnets of scale-free networks are not scale-free: sampling properties of networks. *Proc Natl Acad Sci USA* 102(12):4221–4224
- Stumpf MPH, Porter MA (2012) Critical truths about power laws. *Science* 335(6069):665–666
- Upper C, Worms A (2004) Estimating bilateral exposures in the German interbank market: is there a danger of contagion? Cross-border bank contagion in Europe. *Eur Econ Rev* 48(4):827–849

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.