Essays on Computational Economics

Piotr Z. Jelonek

Department of Economics

University of Leicester

A thesis submitted for the degree of

Doctor of Philosophy

September 2013

I dedicate this thesis to my wife Anna.

Abstract

This text consists of two parts. In chapters 2–3 the methods are developed that enable the application of tempered stable distributions to measuring and simulating macroeconomic uncertainties. In contrast to the tools used in finance, these results are applicable to low frequency aggregated data, which typically displays tails of moderate gravity. Thus thay are particularly useful in modelling macroeconomic densities. The new methods may be readily employed in Monte Carlo simulations of possibly skewed, moderately heavy–tailed random variates with arbitrary excess kurtosis. In chapter 4 a computational model of endogenous network formation for the inter–bank overnight lending market is proposed. The structure of this market emerges from interactions of heterogeneous agents who are endowed with assets, liabilities and take into account investment risk. As all the banks are large and their trading affects the prices of risky assets, the costs of price slippage breaks the symmetry of portfolio problem, making inter–bank borrowing and lending more desirable. The model takes into account three channels of contagion – bankruptcy cascades, common component of risky asset returns and erosion of liquidity. The network formation algorithm outputs the ensemble of optimal transactions, the outcome of the corresponding link formation process is pairwise stable. This framework is next employed to investigate the stability of the endogenously generated banking systems.

Acknowledgements

I thank Professor Wojciech Charemza for posing interesting and challenging research problems. I am also grateful to Doctor Dan Ladley for numerous brainstorming sessions we had together. Thanks to him many of my bad ideas did not make it to this thesis. I would also like to thank Doctor Christian Lau for proofreading an early version of chapter 3 and Doctor Hao Tian Zhou for sharing his LaTeXtemplate.

Declaration

I declare that the work presented in this thesis is entirely my own. The chapter titled *Tempered Stable Distribution: Properties, Moments, Estimation, Randomization, Application to Foreign Exchange–rates* has a working paper version uploaded to the Leicester University discussion papers website, available as Department of Economics Working Paper No. 12/14.

Piotr Jelonek

PhD Research Student

Astley Clarke Building LG01

Department of Economics

University of Leicester

LE1 7RH

+44 (0) 1162 231 824

+44 (0) 7858 394 994

Contents

List of Tables

List of Figures

Chapter 1

Introduction

This manuscript consists of two distinct parts. The first part encompasses chapters 2–3. In the initial two chapters the methods are developed that enable the application of tempered stable distributions to measuring and simulating macroeconomic uncertainties. The second part consists of chapter 4. This chapter proposes a computational model of endogenous network formation for the inter–bank overnight lending market.

The first part of the thesis addresses the problem of quantifying uncertainty in macroeconomic data. In order to capture skewness, excess kurtosis and gravity of the tails of macroeconomic densities it employs the recently developed tempered stable distributions. These distributions arise from α–stable densities, proposed by Paul Lévy in 1924. While this latter family displayed excellent theoretic properties, it was hardly tractable due to lack of higher order moments. Furthermore, although α–stable distributions often very well fit to economic data, they frequently generate pseudo–random draws of extraordinary magnitudes, even in small samples. Hence they are not suitable to modelling macroeconomic densities, which are usually close to Gaussian and display the tails of moderate gravity. Tempered stable distributions, however, have (almost) all advantages of the latter and no significant drawbacks. They have all moments finite, may display skewness, arbitrary amount of kurtosis and their multivariate extension may capture interdependence of marginals via correlation matrix. Thus tempered stable densities may be employed to solve a number of practical problems that perplexed applied econometricians since early seventies. Despite of their excellent properties, this work is one of the first applications of tempered stable distributions to economic data.

The second part of the thesis approaches the problem of modelling systemic risk. Since the recent financial crisis the works on systemic stability either contained a theoretical model of a banking sector, or simulated its dynamics. While theoretic models explicitly take into account strategic network formation, they admit only limited degree of heterogeneity (typically just two types of banks). Computational papers, on the other hand, are capable to depict much more realistic picture of the banking system, with individual agents varying with respect to size, risk aversion, portfolio composition or liquidity preference. However, interbank linkages in these models are formed not endogenously, but rather at random or partly random basis. Thus the outcome of any stability analysis depends on the network configurations that may never arise in practice. Hence it can not be regarded as reliable. Endogenous network formation is also vital in quantifying the implications of endogenous bankruptcies in a distressed banking system. The reason is that the systemic impact of such event to large extent results from the characteristics of the counterparts of the insolvent institution. Chapter 4 presents an algorithmic model of endogenous network formation for the overnight inter–bank market. According to our best knowledge this algorithm is not only a new, but also the only known solution to this vital problem.

This thesis contains a number of novel results in a few different areas. Chapter 2 presents an original comparison of different probabilistic approaches to modelling uncertainty of macroeconomic distributions. It also demonstrates that macroeconomic densities are rarely Gaussian. Chapter 3 develops new randomization and estimation techniques for tempered stable densities and investigates some of their properties, which were either unknown, or known only for special cases (all the proofs are relegated to Appendix B). In contrast to the tools used in finance, these results are applicable to low frequency aggregated data, which typically displays tails of moderate gravity. Next the proposed multivariate extension of tempered stable distributions is applied to investigate the probabilities of joint currency crisis on the selected pair of exchange–rates. This application relies on a novel Fourier Transform discretization scheme (Appendix A.2). Chapter 4 introduces the network formation protocol for the inter–bank overnight market. The proposed algorithm employs the optimal solution of a portfolio choice problem, the symmetry of which is broken by price slippage costs (Appendix D).

Both parts of this work may be located in two distant corners of the discipline known as computational economics. Apart from the usage of the methods that are highly computationally intensive, the common thread that binds both parts of this thesis are the concepts of uncertainty and structure. In the case of tempered stable densities the uncertainty is captured by the shape of a distribution, while the structure stems from interdependencies of multivariate data. In the case algorithmic network formation the structure of the interbank market is an emergent phenomenon, driven by the optimal behaviour of individual banks. The uncertainty originates form a fact that the market is a complex system which, depending on heterogeneity of the agents, may assume myriads of possible configurations.

After these preliminary remarks we may proceed with the main body of the thesis.

Chapter 2

Heavy–Tails and Stable Distributions

The second chapter reviews the basic probabilistic, parametric approaches to quantifying uncertainty of macroeconomic distributions.

2.1 Introduction

Macroeconomic distributions are sample distributions of economic data, important form macroeconomic analysis, representing aggregate quantities and quoted at low frequencies. A frequency of the data is considered *low* if the distance in time between the consecutive available observations is at least one month.

Future values of macroeconomic data are to large extent *uncertain*. The sources of this uncertainty are twofold. First, as the time goes by, an unexpected event may take place that alters the dynamics of macroeconomic aggregates. Second, a structural change may occur that affects the way economy operates, and thus results in different levels of the registered quantities. As the values assumed in the future by macroeconomic data are critical to efficiency of the policies, implemented today, it is advisable to quantify this uncertainty *ex ante*. This task poses a serious problem. Uncertainty is unobservable and its domain is the future, while all the available data come from the past. Neither future unexpected events, nor impending structural changes are present in the data. Hence the only feasible approach to assessing macroeconomic uncertainties is to form expectations on the basis of available observations, assuming that the mechanism which generates it is constant over time.

The approach to modelling macroeconomic uncertainty presented in this chapter is *probabilistic* and *parametric*. It is always assumed that the mechanism that generates the investigated data is complex and dependent on a large number of factors, some of which are either unknown or not properly measured. Furthermore, it is constant within the sample and will remain unaltered in the immediate future. Under these two assumptions consecutive observations may be treated as random and coming from the same probability distribution. This distribution both approximates the data generating process and captures the ignorance of the underlying economic mechanism. The two assumptions made above jointly imply that although the main object of our interest is uncertainty, the only thing that actually can be measured is *risk*. In order to make this approach tractable, its is further assumed that the distribution which underlies the data is endowed with a known parametrization.

A probability distribution that attempts to depicts uncertainty, domain of which is the future, has to display one particular property. It needs to occasionally generate large observations with non–negligible probabilities. A distribution which does possess this feature is termed *heavy–tailed*.

While this framework is basic, there are two situations when it is highly relevant. First, when there is no model which approximates the economic mechanism generating the data, yet it is necessary to depict the underlying uncertainty. An example of such data, coming from finance, is daily speculative prices of certain commodities. A two macroeconomic examples are either foreign exchange–rates, quoted at monthly frequency, or forecast errors, obtained from a given econometric model. Second, there are a data whose generation is well approximated with known (possibly non–linear) models. However, while these models perform well on regular days, they often fail to generate as many observations of extreme magnitude, as there are actually recorder. If in this second case it is necessary to evaluate the risk of registering large deviations, then it might be beneficial to replace Gaussian error term distribution in the model with another density, which more faithfully depicts the estimated error values.

The structure of the remaining part of this chapter is the following. The first section argues that financial and macroeconomic distributions are not Gaussian. Section two revisits and formalizes the concept of heavy–tailed distribution. It also reviews the basic literature, regarding the application of heavy–tailed distributions in economics and finance. The third section introduces the concept of α–stable distributions and presents their elementary properties. Alpha stable distributions are the predecessors of the tempered stable distributions, which are treated thoroughly in the next chapter. The final section concludes.

2.2 Macroeconomic distributions are not Gaussian

Neither macroeconomic, nor financial distributions are Gaussian. However, while in finance probabilities of large deviations are usually depicted with heavy–tailed densities, in macroeconomics they are often assumed either to be Gaussian, or to be driven by Gaussian distribution. There seem to be two reasons. First, normal density has remarkable, unique statistical properties that make it very convenient to treat. Next, heavy–tails in macroeconomic distributions are not as extreme or pronounced as in the case of financial data. Despite of this empirical fact the Gaussianity of either financial or macroeconomic densities is not a rule, but a rare exception.

In this section first the evidence is presented that empirical distributions of asset and exchange–rate returns display skewness, excess kurtosis and heavy tails. This introductory part relies on the abundant financial literature on the topic. Next, using the example of low frequency exchange rates, it is demonstrated that macroeconomic data often display exactly the same features. The section is concluded with a discussion of the consequences in relying in economic modelling on Gaussian distribution if the actual distribution of the underlying variable is heavy–tailed. The plotted dynamics low frequency macroeconomic time series was extracted from the Datastream.

A major drawback of Gaussian distribution is that it can not replicate skewness and excess kurtosis of empirical data. Significant positive skewness of daily Dow Jones Industrial stocks' returns was reported by Simkowitz and Beedles (1980). In their study authors relied on asymmetry measure that was applicable to random draws, generated from heavy–tailed distributions. This measure did not rely on 2–nd and higher order moments, which do not exist in case of non–Gaussian α–stable distributions. Positive skewness of stock returns was further confirmed by Rozelle and Fielitz (1980) and Kon (1984), who also found significant excess kurtosis in the daily return rates of common stocks and indexes. Skewness and excess kurtosis of hedge funds returns was reported by Brooks and Kat (2002). Unconditional skewness of market returns from selected stock, typically negative, was demonstrated by Harvey and Siddique (2000). Leptokurtic distributions of stock returns were reported by Blattberg and Gonedes (1974), Hagerman (1978) and Haas (2007). Finally, Bauwens and Laurent (2005) mention heteroskedasticity and leptokurtosis as the two established features of financial time series.

Another important feature of the data that Gaussian distribution is not able to replicate is relative high frequency of the observations of large magnitude. Fama (1965a) found in his extensive study that in every series investigated the extreme tails of sample distributions were heavier than the tails implied by Gaussian hypothesis. Officer (1972) observed heavy tails in American stock market data. Jorion (1988) reported that the distributions of exchange rate he modelled with a mixed jump–diffision process were leptokurtic. Marinelli et al. (2001), who depicted US dollar/Swiss franc exchange rate as a process compound in physical time, reported that the intraday exchange rates were heavy–tailed. Jacquier et al. (2004) report that the evidence for fat–tails is very strong for daily exchange rate and equity indices, but less so for weekly data.

There is also some additional evidence which suggests that Gaussian distribution does not fare well in modelling uncertainty in finance. In order to capture exchange rates dynamics of other European currencies against German mark, Vlaar and Palm (1993) extend random term in MA–GARCH specification by additional jump component. This component contributes to gravity of the tails, when it is omitted the resulting model is misspecified and becomes explosive. Loretan and Phillips (1994) observe that fourth moment for stock market and exchange rate series does not seem to be finite. Sun et al. (2007) claim that if data conforms to fractal scaling, abandoning Gaussian hypothesis is one of the only two routes to explain it. Kim et al. (2008) showed that the distribution of asset returns has heavier tails relative to the normal distribution.

Just as in the case of financial data, light–tails in low frequency macroeconomic aggregates are not a rule, but an exception. Unfortunately, the works on heavy–tails in macroeconomic data are scarce. Therefore instead of being backed by literature, the previous statement will be supported with the data. Monthly returns on foreign exchange–rates constitute a prominent example of macroeconomic distributions. In the period from January 1999 till October 2012 out of 15 investigated currencies only US, Australian and Singapore dollars and Chinese yuan were endowed with skewness and excess kurtosis close to zero. Hence only the growth rates of these four currencies could possibly be Gaussian. The growth rates of the remaining currencies, including Canadian and Hong Kong dollars, Japanese yen, Swiss franc, British pound, Brazilian real, Russian ruble, Swedish krona, Danish and Norwegian krones and Korean won displayed either skewness or excess kurtosis that was markedly different than zero. Within the sample Russian ruble was endowed with the largest skewness of 1.31, the largest sample excess kurtosis amounted to 6.79 and was displayed by Swiss franc. Out of this 11 currencies two – British pound and Russian ruble – will be investigated further. Growth rates of the two currencies against euro are positively skewed and leptokurtic, with skewness and excess kurtosis equal to, respectively 1.05, 2.92 (GBP/EUR) and 1.31, 3.42 (RUB/EUR). When a similar exercise is repeated on the daily foreign exchange returns, recorded from 13–th January

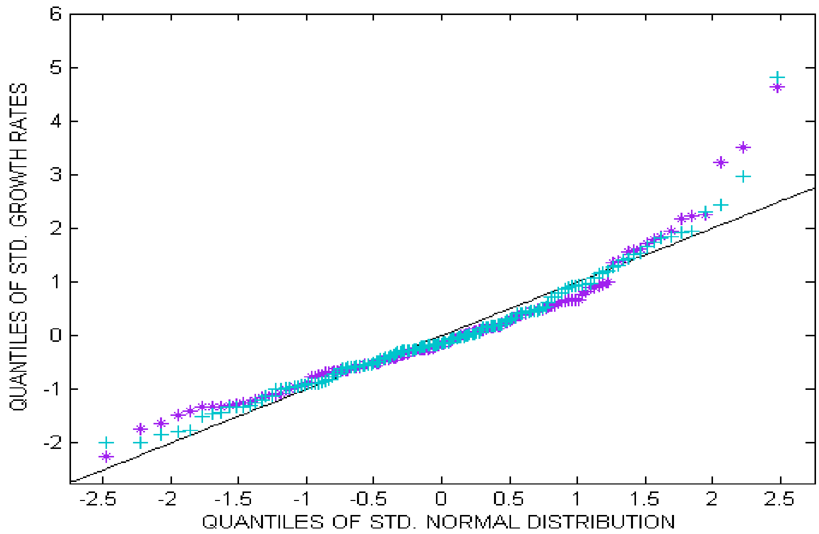
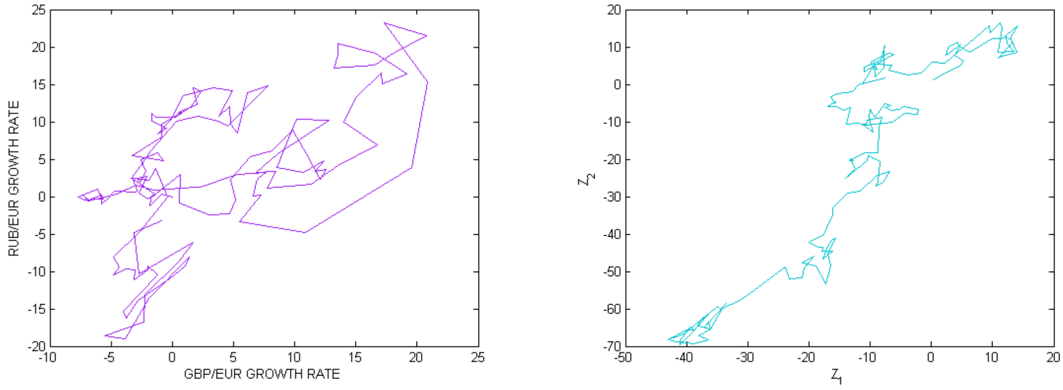


Figure 2.1: QQ plot. On the horizontal axis – (simulated) quantiles of the standard normal distribution. On the vertical axis – sample quantiles of standardized monthly growth rates of RUB/EUR (violet '*') and GBP/EUR (cyan '+') exchange–rates.

2000 till 1–st November 2012, excess kurtosis ranges for different currencies from 2.08 (Canadian dollar) to 68.07 (Swiss franc). At the daily frequency foreign exchange returns against euro of no major currency could be regarded as Gaussian.

Figure 2.1 depicts the quantiles of standardized monthly growth rates of RUB/EUR and GBP/EUR exchange–rates, plotted against the simulated quantiles of standard normal distribution. The data covers the period beginning in January 1999, few months after the Russian Crisis, and running till October 2012 (a total of 152 observations). The number of simulated quantiles is equal to the sample size. The plot reveals heavy tails of both monthly Russian ruble to euro and British pound to euro exchange–rates. Surprisingly, the tails of the latter seem to be heavier. The probability of drawing from standard normal distribution a sample of given size, with the maximum growth rate equal or larger than the actual sample values, amounts to 2.8×10^{-4} for RUB/EUR and 1.2×10^{-4} for British pound to euro exchange–rates. If both samples were obtained from standard Gaussian density, the expected waiting time for the observation equal to or larger than the maximum sample growth rate of Russian ruble to euro exchange–rate would slightly exceed 45,111 years. For British pound to euro exchange–rate this quantity is even more spectacular and amounts to almost 105,931 years.

Figure 2.2 depicts the joint dynamics of the cumulative growth rates of British pound to euro and Russian ruble to euro monthly foreign exchange rates. The chart on the left was obtained from the data quoted from January 1999 till October 2012.



(a) Market data.

(b) Synthetic data.

Figure 2.2: Cumulative GBP/EUR and RUB/EUR forex growth rates plotted for: (a) the actual market data, (b) the data simulated from the fitted bivariate Gaussian distribution.

The figure on the right is plotted for the artificial data, simulated from bivariate Gaussian distribution fitted to the actual data. All the growth rates were centred. In comparison to the real data, the synthetic trajectory is far more erratic. In the actual data monthly growth rates of either small or large magnitude are relatively more common, while growth rates of average magnitude are relatively less frequent than in the synthetic data. Therefore the increments of foreign exchange rates should be depicted by distributions whose tails are heavier than Gaussian. This distribution should simultaneously admit more observations of very small magnitude.

While the literature on large deviations in macroeconomic data is scarce, there are few papers worth mentioning. Balke and Fomby (1994) examined 15 macroeconomic post–war US time series and identified large shocks in each series investigated. The authors claim that as outliers are typically associated with recessions, Gaussian models of business cycle are not appropriate. The widespread evidence in favour of fat–tailed output growth rates is reported by Fagiolo et al. (2008). Finally, Aston (2006) remarks that many macroeconomic time series can be well modelled using heavy tailed densities, especially if the series contain outliers. Hence in case of macroeconomic data normality of error term should not be assumed by default.

The remaining part of this section presents examples of how unrealistic assumption of Gaussian model errors might trigger numerous large scale disasters in the estimation of economic models and the subsequent inference. Although this evidence has been originally collected for normal distribution, it is equally relevant for any other light–tail density. In the following list the feasible issues are ranked according to their relative importance.

Counterfactual assumption of Gaussianity may cause corrupted parameter estimates. If trajectory of investigated time series contains large jumps, then under normal random term many observations are identified as outliers. Vlaar and Palm (1993) report how these spurious outliers resulting from misspecified error term lead to disastrous estimates of model parameters. In their study MA–GARCH models become explosive for three out of six investigated currencies.

Counterfactual assumption of Gaussianity may result in implausible model dynamics. Hartmann et al. (2010) demonstrate that heavy–tailedness of distributions underlying exchange rates is one of sufficient conditions for joint, vehement currency crisis to occur. Joint currency crisis were observed numerous times over the past few decades. However, under normal density of model errors they would be practically impossible.

Counterfactual assumption of Gaussianity may lead to underestimation of risks originating from self–reinforcing or natural phenomena. Balke and Fomby (1994) found large estimated error terms in all the examined time series. As these errors were typically clustered over time and associated with recessions, they concluded that the Gaussian models of business cycle are not appropriate. Pindyck (2011) notes that, as uncertainty about probability distributions of future temperatures is large, heavy–tailed densities are essential in the modelling of climate policy. As mentioned out by Asmussen et al. (2000), due to intrinsic properties of the data heavy–tailed modelling seems to be relevant in areas such as telecommunication and insurance risk.

Counterfactual assumption of Gaussianity may adversely affect model selection. Balke and Fomby (1994) report that in many of the investigated time series much of the evidence of non–linearity is eliminated after controlling for outliers. However, frequent outliers are immediate consequence of assuming distribution of error terms that does not have sufficiently heavy tails.

Counterfactual assumption of Gaussianity may lead to incorrect conclusions drawn from estimated models. Fagiolo et al. (2008) claim that economic models failing to account for fat tails in GDP may deliver invalid implications. This is most likely in the case when the tails are so heavy that higher order moments do not exist.

Counterfactual assumption of Gaussianity may result in incorrect asset pricing. The reason is that heavy–tailed distributions of the underlying assets imply different prices of derivatives than the normal distribution. A seminal paper by Boyarchenko and Levendorskii (2000), who provide a generalized Black–Scholes formula for pricing European options under non–Gaussian price increments, is a notable example.

If the assumption of Gaussianity of error term density is counterfactual, then a variety of viable econometric techniques is greatly reduced. In particular, if the error distribution is a non–normal α–stable density, then the Least Squares procedures become no longer valid. As mentioned by Mandelbrot (1963a) and Fama (1965a), the reason is that the expectation of the minimized sum of squares is infinite. In consequence all the techniques that rely on the concept of variance are no longer applicable. A similar remark is made by Fagiolo et al. (2008) who claim that, due to the widespread evidence in favour of fat–tailed output growth rates, econometric estimation and testing procedures that are heavily sensitive to normality of residuals should be replaced with their robust counterparts. Fama and Roll (1965) indicate that sample standard deviation is unreliable estimator of scale parameter for the data generated from α–stable distribution. Blattberg and Gonedes (1974) noted that if variance of the underlying distribution does not exist, then the methods of spectral analysis can no longer be applied. Another example may be found in Asmussen et al. (2000) who demonstrate that the standard importance sampling technique may no longer be used as a method of rare events generation when the underlying distribution is heavy–tailed.

Finally, if Gaussianity of error term distribution is violated, basic properties of economic models may be inverted. In a brilliant paper Ibragimov (2005) demonstrates that in the extreme heavy–tailed setting probabilistic properties of many economic models may reverse. In particular, a stylized fact that portfolio diversification is always preferable in portfolio Value–at–Risk analysis no longer holds. He also demonstrates that the implications of numerous frameworks, popular in economics and finance, are sensitive to the thickness of the tails of the distributions involved in their assumptions.

The next section is dedicated to a number of feasible routes to modelling large deviations. All these approaches ameliorate the problems, indicated above.

2.3 From heavy–tails to stable distributions

This section first reviews the basic works on modelling large deviations. As the most fundamental of the cited papers concern speculative asset prices, again the bulk of this literature comes from finance. After a brief historical sketch a number of ways to generate distributions with large deviations is investigated. It is next argued that out of the compared distributions the tempered stable densities, whose special case is treated in the next chapter, possess most of the features, desirable in modelling uncertainty.

The following definition will be useful to distinguish between the densities, which may generate observations of large magnitude.

Definition 2.3.1 (Gravity of tails)**.** *The right tail of distribution of random variable X is said to be* ***i)*** fat *if for some $\alpha > 0$ and $x \to +\infty$ it holds that*

$$\mathbb{P}(X > x) \sim x^{-\alpha},$$

ii) heavy *if for all $\lambda > 0$ we have*

$$\lim_{x \to +\infty} e^{\lambda x}\, \mathbb{P}(X \leq x) = +\infty,$$

ii) semi heavy *if it is not heavy and for all $\lambda > 0$ it holds that*

$$\lim_{x \to +\infty} e^{\lambda x^2}\, \mathbb{P}(X \leq x) = +\infty.$$

Part i) of the formalism above is given on p. 9 in Crovella et al. (1998) as a definition of a distribution with heavy right tail. Contrary to their formulation we allow for $\alpha > 2$ to admit the distributions, the tails of which are heavier than the tails of α–stable densities (introduced in the next section). Part ii) of Definition 2.3.1 may be found e.g. in Asmussen (2003). Part iii) was introduced to distinguish the class of distributions whose right tails decay faster than exponentially, but slower than the tails of the Gaussian density.

The definitions of fat, heavy and semi heavy left tails are analogous. The distribution is said to be fat, heavy or semi heavy–tailed if either of its tails is, respectively, fat, heavy or semi heavy. The tails of heavy–tailed distributions decay slower than exponential function. The tails of semi heavy–tailed distributions decay faster than exponential, but slower than Gaussian density. Contrary to heavy–tailed distributions, semi heavy–tailed densities may have all moments finite. After formalizing the basic concepts we may go back to the origins of the literature on large deviations.

The first evidence of significance of heavy tails in economics may be found in the works of Italian sociologist, economist and political scientist Vilfredo Pareto. In the year 1906 he made the observation that in Italy twenty percent of the population owned eighty percent of the land. In 1909 Pareto generalized this remark proposing Pareto distribution, which, as he believed[1], described wealth distribution *"through any human society, in any age, or country"*. The density corresponding to Pareto

[1]See: Mandelbrot and Hudson (2004).

distribution was endowed with a power–law tail, which implied[1] that the tail probabilities were self–scaling. The feature of self–scaling became ever since a landmark of heavy–tailed modelling. The importance of Paretos original contribution has not become apparent for another fifty six years when the seminal works of Benoit Mandelbrot were published.

As noted by Hartmann et al. (2010), it was not until October 1963 when Mandelbrot *"published two remarkable papers [...] which gave financial econometricians enough to work on for subsequent 30 years"*. In *New methods in statistical economics* (Mandelbrot, 1963b) he suggested that in the modelling of asset prices Gaussian distribution of increments should be replaced with a more plausible family of heavy–tailed densities. Since the work of Pareto it was known that the tail probabilities of many economic quantities were self–scaling. Mandelbrot first assumed that this property should hold for the entire distribution. Next he observed that a consistent model of randomness should not depend on the time scale in which the phenomenon of interest is being observed. Finally, he remarked that, for the sake of tractability, the sought class of distributions should be closed with respect to summation of i.i.d. random variables and linear scaling. All three routes led him independently to single conclusion – asset returns were best depicted by an infinite variance sum of α–stable random variates. Simultaneously, in the manuscript titled *The variation of certain speculative prices* (Mandelbrot, 1963a) author presented empirical evidence to support his theory. In this second paper he embarked upon the premise that the increments of logs of speculative prices come from α–stable distribution. Using cotton prices data, Mandelbrot computed the sample second moments of the daily first difference of logs for increasing samples from 1 to $1,300$ observations. Although the sample size was enormous for economic standards, he found that as it increased, the sample standard deviation did not tend to any limiting value. Instead, it varied in an absolutely erratic fashion. Mandelbrot regarded this anomalous behaviour as a signature of infinite variance and thus the vital evidence to support his theory. However, the initial reception of his work was so frigid that he abandoned research of market dynamics for more than a decade. The approach proposed by Mandelbrot was fervently supported by Eugene F. Fama. In order to highlight empirical aspects of Mandelbrots theory of speculative price innovations, Fama (1963) formulated *Stable Paretian Hypothesis*. It assumed that 1) the variances of the empirical distributions behave as if they were infinite, 2) the empirical distributions conform best to the non–Gaussian α–stable densities.

[1] See: Feller (1971), Subchapter VIII.8.

Since the original contribution of Mandelbrot numerous evidence, both data based and theoretic, has been collected to support his findings. Three main empirical arguments were as follows.

First, as already mentioned in the previous section, both financial and economic data is typically endowed with skewness and excess kurtosis. It also contains many more observation of extreme magnitude than normal distribution typically implies. Both remarks point out to the fact that the distribution that generates the data is most likely not Gaussian.

Next, α–stable densities that fit empirical data best are typically endowed with stability index α in–between 1.4 and 1.9. Blattberg and Gonedes (1974) in his stock price analysis obtained $1.65 < \alpha < 1.8$. Marinelli et al. (2001) found that α–stable distributions offer very good fits to exchange rate returns both in quote (intrinsic) time and physical time. Authors observed that exchange rate returns display α–stable increments with index $\alpha \approx 1.4$ at time scales between few minutes and two hours. In longer time scales, such as day or week, α assumed values which ranged from 1.8 to 1.9. While Mandelbrot (1963a) reported $\alpha \approx 1.7$ in his cotton price example, Fama (1965a) found stability indexes clustered around 1.9 for all the investigated stocks. Note that if the data were Gaussian, it would hold that $\alpha \approx 2$, which is the value of stability index that corresponds to normal distribution. As remarked in Fama (1965b), returns on securities in all the cases conformed better to non–Gaussian α–stable distributions with infinite variances.

Finally, if multiple stock return series are analysed simultaneously, they frequently seem to be endowed with the same characteristic exponent. Hagerman (1978) found small differences among the exponents obtained for the investigated AMEX and NYSE securities and attributed these entirely to measurement error. This finding supports portfolio theory, relying on multivariate stable distributions.

The main theoretical argument to support Mandelbrots hypothesis were remarkably good statistical properties of α–stable distributions. In particular, as stated by Akgiray and Booth (1988) 1) only stable laws have domains of attraction and 2) stable distributions belong to their own domains of attraction. The first feature means that α–stable distributions arise as limit densities in a version of Central Limit Theorem where i.i.d. random variables might have infinite variance. In consequence, as observed by Rachev and Mittnik (2000), if the process may be regarded as a sum of many microscopic effects, the only possible limit in the distribution of the sum of i.i.d. random variables is α–stable. The second feature results from the formal definition of stable random variates and implies stability in colloquial sense. In result, if stable distributions emerge as a limit distribution of an additive process, a slight distortion in the fundamentals of this process does not make the outcome fragile. Mandelbrot (1963b) noted that α–stable distributions 3) display self–scaling of tail probabilities, 4) have intrinsic extension to the multivariate case and 5) in multivariate case may be characterized as distributions, for which the density of every linear combination of coordinates is univariate α–stable. Finally, Rachev and Mittnik (2000) remark that 6) α–stable distributions can account for asymmetries and heavy tails.

Although α–stable distributions were *a priori* plausible as models of stock prices, numerous empirical studies provided evidence which tended to contradict Mandelbrots theory. Hsu et al. (1974) found that the standard deviations of American stock data did not display erratic behaviour, predicted by Stable Paretian Hypothesis. Authors observed that longitudinal sums of daily share returns became thinner–tailed for large number of summands, what negated stability of the addends. This finding was confirmed by Blattberg and Gonedes (1974), Hagerman (1978), Fielitz and Rozelle (1983) and Kon (1984) who all noted that fitted stability indexes did not remain constant under temporal aggregation. Praetz (1972) on Australian stock market data demonstrated that t distribution provided better fit than α–stable symmetric density. Blattberg and Gonedes (1974) confirmed this finding on US stock return data. After further investigation they concluded that the region where t distribution performs better than symmetric α–stable density are the tails of sample distribution. Blattberg and Gonedes (1974), Fielitz and Rozelle (1983) and Akgiray and Booth (1988) have all reached the conclusion that second moments for majority of analysed datasets seem to be finite. The claim that the distributions of asset returns have tails thinner the α–stable density was recently supported by Kim et al. (2008). Hsu et al. (1974) noted that when economy may experience dramatic shifts, it is not reasonable to insist on stationarity (homogeneity) of data generating process. Authors demonstrated that if structural breaks of scale parameter are admissible, then rates of return within the periods of homogeneous activity are well depicted by Gaussian distribution with finite variance. Fielitz and Rozelle (1983) observed it is hard to distinguish in sample between a mixture of normal distributions with changing standard deviation and non Gaussian α–stable distribution with varying scale. Kon (1984) concluded that after replacing Gaussian error term with a mixture of normal distributions, the evidence in favour of heavy tails may entirely vanish. Hagerman (1978) rejected a hypothesis that stock returns are endowed with a symmetric α–stable distribution in favour of either: a mixture of Gaussian distributions or t distribution. Both distributions provided better fit to US stock return data. Wang et al. (2009) observed that a model with fixed gravity of tails can not fit the market data well over time. Rachev and

Mittnik (2000) remarked that infinite variance can not be verified in practice. Hartmann et al. (2010) claimed it is sufficient to assume self–scaling holds in the tails of the distributions, thus infinite variance assumption is no longer required.

Since the deficiencies of both Gaussian and stable distributions became apparent, it was clear that to capture uncertainties that underlie economic data, better probabilistic framework is required. Numerous approaches to achieve this aim were attempted over the last thirty years. As the number of proposed probability distributions is vast, the following list is necessarily incomplete. However, the problem of quantifying uncertainty in economic data has been attacked along three main lines.

The first line was to generalize *t distributions*, which are highly tractable, well established in the literature and known to display heavy tails. Applicability of t distribution in modelling heavy tails was reported by Kon (1984) and further confirmed by Aston (2006). To introduce asymmetry to t distribution, Harvey and Siddique (1999) proposed *noncentral t distribution*. However, as their asymmetric density had no third central moment, its skewness did not exist. Another attempt to modify t distribution was made by Hansen (1994) who introduced *skewed t distribution*. His results were further extended as *skewed generalized t distribution* (further abbreviated to sGT) derived by Theodossiou (1998). Different route to deliver a variant of t distribution that could account for sample skewness was taken by Jones and Faddy (2003). The density they proposed will be further termed *skewed factorizable t distribution* (sfT). All four modifications listed above included classic t distribution as a special case. While t density may be generalized to further dimensions as in Glasserman et al. (2002), its most popular extension, known as *multivariate t distribution*, is endowed with symmetric marginal distributions. In general, extension of t distribution to multiple dimensions is not unique. As it is most flexible univariate modification of t density, skewed factorizable t distribution (sfT) will be selected as representative of this class for the sake of comparison with other distributions.

The second approach was to propose a general construction that could introduce skewness into pre–existing family of (possibly heavy–tailed or fat–tailed) symmetric distributions. Few such constructions are known, for the sake of brevity just three will be mentioned here. The first one comes from Azzalini (1985) and was originally applied to normal density. In result *skewed normal* distribution (further abbreviated to sN) was obtained. While normal distribution has light tails, skewed normal density obtained via Azzalini technique may be endowed with arbitrary gravity of tails. As demonstrated in Azzalini and Capitanio (2003), this approach may be extended to multiple dimensions. The second construction was proposed by Fernández and Steel (1998) and was utilized to introduce skewness into t distribution. The output of this procedure will be termed *skewed t distribution* and abbreviated to sT. Note that this density is markedly different than the skewed t distribution, proposed by Hansen (1994). The third method was developed by Bauwens and Laurent (2005) and allows to introduce a vector of skewness parameters to multivariate, spherical symmetric distribution. As in previous case, t distribution serves as example of a practical application.

The third route to generate heavy–tailed densities is to utilize various types of mixtures of normal distributions. Kon (1984) obtained satisfactory results in depicting stock return rates with *ordinary mixture* of normal distributions. In order to model extraordinary movements in stock prices Friedman and Laibson (1989) resorted to *compound (Poisson) mixture* of normal distributions. Finally, Eberlein and Keller (1995) proposed generalized hyperbolic distribution (abbreviated to GH) as a default probability measure in modelling financial data. This semi heavy–tailed distribution is known to be a *variance–mean mixture* of normal distributions. In this case the role of mixing distribution is performed by generalized inverse Gaussian density. he main advantage of mixture techniques is that structure of random variables thus obtained often translates to favourable properties of resulting densities. As GH distribution is the most promising representative of this second class, its properties will be compared further with the main characteristics of other heavy–tailed distributions.

There are two known heavy–tailed distributions that do not follow any of the three routes listed above. These are generalized versions of Pareto and Extreme Value distributions. However, their applicability is severely constrained. What limits the applicability of generalized Pareto distribution is the fact that its domain is censored. In case of Extreme Value distribution the main culprit is the property that generalized Extreme Value distribution has unconstrained domain only if its excess kurtosis is equal to 2.4. Hence none of these two distributions seems to be flexible enough to capture empirical features of economic data.

Despite numerous attempts to replace Gaussian disturbances with more plausible probabilistic framework, each of the distributions listed above lacks some features which makes modelling uncertainty tractable. Table 2.3 the comparison of all the selected densities with infinite variance stable distributions, which were originally proposed by Mandelbrot (1963b) as a default probabilistic model for the increments of speculative prices. Columns in Table 2.3 represent the selected distributions while rows denote certain traits. Sign "+" at the intersection of a row and a column indicates that the distribution, associated with a given column, is known to possess a trait, corresponding to a certain row (and that the prove can be traced in literature). Sign "−" is used in all the remaining cases. The motivation for using the set of properties listed in Table 2.3 is as follows. 1) It is convenient when the underlying distribution displays higher order moments as then many standard tools (software, mathematical approaches) become applicable. 2) Plausible candidate for probabilistic model of uncertainty in economics and finance should display both light and heavy tails, depending on the data. 3) Extension of univariate density to multiple dimensions should be unique. Otherwise the choice of appropriate extension might become an issue. 4) The selected family of distributions should be closed under addition of independent, identically distributed random variates. Otherwise probabilistic features of the corresponding random walks are not consistent over time. 5) The selected family of distributions should be closed under weighted averages of independent, identically distributed random variables. Otherwise it is difficult to utilize in the areas such as asset pricing or portfolio analysis. 6) The long run distributions of random walks endowed with the increments, coming from given family of densities, should be approximately Gaussian. This property is desirable as it corresponds to empirical features of macroeconomic data, such as forecast error distributions. Note that the abilities to display skewness, excess kurtosis or heavy tails are excluded from this comparison. Every density listed in Table 2.3 displays all these features.

As already stated above, the first four distributions being compared are: skewed normal distribution (sN), skewed t distribution (sT), skewed factorizable t distribution (sfT) and generalized hyperbolic distribution (GH). The last two columns correspond to: α–stable distribution (S) and (general) tempered stable distribution (TS). Applying a special case of the latter the economic data is the subject of the next chapter. Daggers † next to the entries, corresponding to stable and (general) tempered stable distribution, denote that these are closed under weighted averages of i.i.d. random variates if the stability indexes of both variables match. Asterisk * in the entry, corresponding to generalized hyperbolic distribution, indicates that the weighted average of i.i.d. random variates is itself generalized hyperbolic if all the variates have normal inverse Gaussian (NIG) distribution (a special case of GH) with identical values of parameter α and the same value of parameter β.

The last column in Table 2.3 refers to (general) tempered stable distribution defined by Rosiński (2007). These distributions are self–decomposable, closed with respect to weighted averages and addition of independent random vectors (if only their stability indexes match). Tempered stable densities may be infinitely divisible and may have higher order moments finite, including exponential moments of some order. Hence, what makes these distributions highly tractable, they might have moment generating functions. They also display skewness and arbitrarily large excess

Properties	Heavy–tailed Distributions sN	sT	sfT	GH	S	TS
Higher order moments may exist	+	+	+	+	−	+
Both light and heavy tails are possible	+	+	+	+	−	+
Natural multivariate extension	+	−	−	+	+	+
Consistent model of additive jumps	−	−	−	+	+	+
Closed under weighted averages	−	−	−	$-^*$	$+^\dagger$	$+^\dagger$
Long–run Gaussianity	+	−	−	+	−	+

Table 2.1: Heavy–tailed distributions and their properties. Sign "+" indicates that the distribution is known to possess a certain trait, "−" is used in all the remaining cases.

kurtosis. Tempered stable distribution may be arbitrarily similar to the underlying stable densities, thus they may (approximately) retain all the favourable properties of the latter family. As demonstrated by Chakrabarty and Meerschaert (2011), tempered stable distributions provide a universal model of accumulated jump. This observation follows from the fact that any random walk with power–law jumps may be approximated with tempered stable densities. Furthermore, random walks with tempered stable increments in long run converge in distribution to Gaussian density.

All the features listed above make tempered stable densities the best known candidate for probabilistic model of macroeconomics uncertainties. The application of tempered stable distributions to macroeconomic data is the main subject of the next chapter. As tempered stable distributions have similar properties to stable distributions, from which they were originally derived, the basic characteristics of the latter family is provided in the next section.

2.4 Stable distributions and their properties

This section is dedicated to stable distributions. It contains the definition of univariate stable random variables along with their basic properties and Fourier transformation of the underlying density. Next the concept of α–stable distribution is introduced. In univariate case it is exactly equivalent to stability of the underlying distribution. Hence the two terms might be used exchangeably. Two alternative definitions of univariate stable variates are given as Fact 2.4.1 in order to formalize the links between stability, Central Limit Theorem and closedness with respect to linear transformations. Finally, the definition of stable random vectors is presented. Many of these properties are parallel to the basic characterization of tempered stable distributions, given in the next chapter.

An extensive overview of the papers on stable distributions and their applications is provided by John P. Nolan on his webpage. In the recent (27–th August 2013) update of this document[1] bibliography alone occupies 130 pages. While the first work on stable distributions was written by Lévy (1924), a number of significant papers were further contributed in nineteen–forties and fifties. However, this topic proved to be notoriously difficult. Rigorous derivation of exact formulas for characteristic functions of what is now known as stable random variates confounded statisticians for another 40 years (for a concise review see Hall (1981)). Numerical aspect of bringing stable distributions to the data posed further serious problems (Nolan, 2013).

Stable distributions constitute a popular family of distributions, renowned for their capacity to reproduce fat tails and accommodate skewness. In univariate case they are endowed with four characteristics: index of stability $\alpha \in (0, 2]$, scale $\delta > 0$, skewness $\beta \in [-1, 1]$ and the location parameter $\mu \in \mathbb{R}$. The main difficulty in handling stable distributions stems from the fact that their probability and cumulated density functions are usually known only in the form of infinite series (see: Feller (1971), Chapter XVII, Section 6, or Palmer et al. (2008) for the case of $\alpha < 1$). Therefore the corresponding theory relies to great extent on characteristic functions and spectral measures. In the remaining part of this paper α–stable distributions will be termed stable and (in univariate case) denoted as $S_\alpha(\beta, \delta, \mu)$. All the definitions and properties presented in this section come from the first two chapters[2] of the book by Samorodnitsky and Taqqu (2000).

There are at least three different, equivalent definitions of stable distributions. The one below allows for the most intuitive interpretation.

Definition 2.4.1 (Stable random variable)**.** *A random variable X has a stable distribution if for any $k \geq 2$ there is $C_k > 0$ and a real number D_k such that*

$$X_1 + X_2 + \ldots + X_k \overset{d.}{=} C_k \cdot X + D_k, \tag{2.1}$$

where $X_1, X_2, \ldots, X_k$ are independent copies of X and $\overset{d.}{=}$ denotes equality in distribution.

[1] Available at: academic2.american.edu/ jpnolan/stable/StableBibliography.pdf

[2] See Definitions: 1.1.4, 1.1.1, 1.1.5, 2.1.3, 2.1.1, Properties: 1.2.1–1.2.3, 1.2.13, 1.2.16 and Theorem 2.1.5 therein.

If $D_k = 0$, random variable X is called *strictly stable*. Stable distribution is Gaussian when $\alpha = 2$ and $\beta = 0$. It can be proved[1] that $C_k = k^{1/\alpha}$, where α amounts to stability index. As α captures the entire information on the scaling of the sums of independent addends, stable distributions are also known as *α–stable*. In univariate case these two terms are equivalent. If $\alpha > 1$ then the shift parameter μ equals to the mean of the distribution. For $\alpha \leq 1$ the mean does not exist. The following statements are equivalent to Definition 2.4.1.

Fact 2.4.1 (Equivalent univariate definitions)**.** *A random variable X is said to be stable if*

i) for any positive numbers A and B there is a positive number C and a real number D such that

$$AX_1 + BX_2 \stackrel{d.}{=} CX + D,$$

where X_1, X_2 are independent copies of X,

ii) it has a domain of attraction, i.e. if there is a sequence of i.i.d. random variables $Y_1, Y_2, \ldots$ and a sequence of real numbers $\{A_k\}$ and positive numbers $\{D_k\}$ such that

$$D_k^{-1}(Y_1 + Y_2 + \ldots + Y_k) + A_k \xrightarrow{d.} X$$

where $\xrightarrow{d.}$ denotes convergence in distribution.

Stable distributions are, by Fact 2.4.1 ii), the family of limit distributions that arise in a version of Central Limit Theorem where the variance of independent addends may be infinite. Due to the properties of Definition 2.4.1 if stable random variates are used to build additive stochastic process, the time scale at which the data is quoted does not affects properties of this process. In particular, the cumulative distribution of its increments remains in the same class regardless of the assumed time scale. This is particularly welcome feature while we model economic processes whose behaviour should not depend on the choice of time units.

Although univariate α–stable densities may be identified via corresponding cumulative distribution and probability density functions, such characterization is of limited practical use. The reason is these are usually known only in the form of series expansions, involving *confluential hypergeometric* function. Hence the most popular tool to treat univariate α–stable variates is Fourier transformation of the underlying distribution. Univariate stable random variables display the following properties.

[1] See: Feller (1971), Theorem VI.1.1, p. 170.

Property 2.4.2 (Univariate characteristic function)**.** *If* $X \sim S_\alpha(\beta, \delta, \mu)$ *with* $0 < \alpha < 2$ *then*

$$\Phi_X(u) = \mathbb{E}e^{iuX} = \begin{cases} \exp\left\{iu\mu - \delta|u|(1 + i\beta\frac{2}{\pi}(\operatorname{sgn} u)\ln|u|)\right\} & \text{if } \alpha = 1, \\ \exp\left\{iu\mu - \delta^\alpha|u|^\alpha(1 - i\beta(\operatorname{sgn} u)\tan\frac{\pi\alpha}{2})\right\} & \text{if } \alpha \neq 1, 2. \end{cases} \tag{2.2}$$

In the remaining part of this chapter whenever random variable X is not explicitly defined, it is assumed that $X \sim S_\alpha(\beta, \delta, \mu)$.

Property 2.4.3 (Absolute moments)**.** *Assume* $\alpha < 2$*. Then*

$$\mathbb{E}|X|^p < +\infty \text{ for any } 0 < p < \alpha,$$

$$\mathbb{E}|X|^p = +\infty \text{ for any } p \geq \alpha.$$

Thus higher order moments of stable random variables do not exist – stable distributions are heavy–tailed. Property 2.4.5. does not apply when $\alpha = 2$ and the corresponding stable random variable is Gaussian.

Property 2.4.4 (Linear transformations)**.** *Let* $Y = aX + b$ *where* $a \neq 0$ *and* b *are real constants. Then*

$$Y \sim S_\alpha(\beta(\operatorname{sgn}\alpha), \delta|a|, a\mu - \frac{2}{\pi}\delta\beta \cdot \ln|a|^a + b) \;\text{ if } \alpha = 1,$$

$$Y \sim S_\alpha(\beta(\operatorname{sgn} a), \delta|a|, a \cdot \mu + b) \;\text{ if } \alpha \neq 1.$$

Property 2.4.5 (Additivity)**.** *Let* X_1 *and* X_2 *be independent random variables with* $X_i \sim S_\alpha(\beta_i, \delta_i, \mu_i)$*,* $i = 1, 2$*. Then* $X_1 + X_2 \sim S_\alpha(\beta, \delta, \mu)$*, with*

$$\beta = \frac{\beta_1\delta_1^\alpha + \beta_2\delta_2^\alpha}{\delta_1^\alpha + \delta_2^\alpha}, \;\; \delta = (\delta_1^\alpha + \delta_2^\alpha)^{1/\alpha}, \;\; \mu = \mu_1 + \mu_2.$$

Hence stable random variables are additive, multiplicative and the stable family is closed under summation.

As demonstrated by Yamazato (1978), all univariate α–stable distributions have unimodal densities. Fast and reliable method to estimate parameters of univariate α–stable distributions was provided by McCulloch (1986).

Note that if X_1, X_2 are independent stable random variables with different stability indexes, then Property 2.4.5 does not apply. Therefore, in general, the expression bX_1+cX_2 does not represent a stable random variable irrespective of the choice of real constants b and c. A notable exception is a situation when X_1 and X_2 have the same distribution. Then the corresponding linear combination is again stable by the virtue of Fact 2.4.1 i). In general, stable distributions are closed under *weighted averages* only if their stability indexes match.

Property 2.4.6 (Mixture representation)**.** *Let $X \sim S_\alpha(\beta, \delta, 0)$ with $\alpha < 2$. Then there exist two i.i.d. random variables Y_1 and Y_2 with common distribution $S_\alpha(1, \delta, 0)$ such that*

$$X \overset{d.}{=} \Big(\frac{1+\beta}{2}\Big)Y_1 - \Big(\frac{1-\beta}{2}\Big)Y_2 + \delta\Big(\frac{1+\beta}{\pi}\ln\frac{1+\beta}{2} - \frac{1-\beta}{\pi}\ln\frac{1-\beta}{\pi}\Big) \text{ if } \alpha = 1,$$

and

$$X \overset{d.}{=} \Big(\frac{1+\beta}{2}\Big)^{1/\alpha}Y_1 - \Big(\frac{1-\beta}{2}\Big)^{1/\alpha}Y_2 \text{ if } \alpha \neq 1.$$

Property 2.4.6 is attributed to Zolotarev, but no direct reference has been traced. It implies that independent random variables $Y \sim S_\alpha(1, \delta, 0)$ may be treated as building blocks of stable distributions.

When $\alpha = 2$, the random stable variable is Gaussian, has all moments finite and its dependence structure is identified by the autocovariance function. However, for $\alpha < 2$ second moments no longer exist and thus this concept is not applicable. Two popular ways[1] to circumvent this obstacle are to introduce the notions of either *covariation* (defined for $0 < \alpha \leq 1$) or *codifference* (for $0 < \alpha \leq 2$). For $\alpha = 2$ both are equivalent to the covariance functions.

Random stable vectors may be defined in the similar way[2] as random variates.

Definition 2.4.2 (Stable random vector)**.** *A random vector $X = (X_1, \ldots, X_n)$ in $\mathbb{R}^n$ is said to be stable if and only if for any $k \geq 2$ there exists a positive constant C_k and a real vector D_k such that*

$$X_1 + X_2 + \ldots + X_k \overset{d.}{=} C_k X + D_k, \tag{2.3}$$

where $X_1, \ldots, X_k$ are independent copies of X and $\overset{d.}{=}$ denotes equality in distribution.

The vector X is said to be α–stable if $C_k = k^{1/\alpha}$. It is called *strictly stable* if formula (2.3) holds with $D_k = 0$. It is called *symmetric stable* if it is stable and satisfies $\mathbb{P}(\{X \in B\}) = \mathbb{P}(\{X \in -B\})$ for any Borel set $B \in \mathcal{B}(\mathbb{R}^n)$.

Contrary to univariate case, multivariate stable vector does not necessarily need to be α–stable. As in the univariate case, multivariate cdfs and pdfs are usually

[1] See: Samorodnitsky and Taqqu (2000), Subchapters 2.7 and 2,10.
[2] See: Samorodnitsky and Taqqu (2000), p. 57–59, Definition 2.1.2, Definition 2.1.4.

known only in the form of series expansions. Definition 2.4.2 may be expressed in the following, equivalent way.

Fact 2.4.7 (Equivalent multivariate definition)**.** *A random vector X is stable if for any positive numbers A and B there is a positive number C and a vector $D \in \mathbb{R}^n$ such that*

$$AX_1 + BX_2 \stackrel{d.}{=} CX + D,$$

where X_1, X_2 are independent copies of X.

Random vector X is called α–stable if the formula above holds with $C = (A^\alpha + B^\alpha)^{1/\alpha}$.

A useful feature of stable random vectors is the fact that they are closed with respect to linear combinations. In three distinct situations the fact that all the linear combinations of given random vector are stable implies that this vector is itself stable.

Property 2.4.8 (Linear combinations)**.** *Let $X = (X_1, \ldots, X_n)$ be a random vector in $\mathbb{R}^n$, define linear combination of the entries of X as $\sum_{k=1}^{n} b_k X_k$ where $b_k \in \mathbb{R}$ for $k \in \{1, \cdots, n\}$.*
"$\Rightarrow$") Let X be stable (respectively, strictly stable, symmetric stable) random vector. Then any linear combination of the entries of X is a stable (respectively, strictly stable, symmetric stable) random variable.
"$\Leftarrow$") Let X be a random vector in $\mathbb{R}^n$. If all linear combinations of the entries of X
i) are α–stable with $\alpha \geq 1$, then X is a stable random vector,
ii) have strictly (respectively, symmetric) stable distributions, then X is a strictly (respectively, symmetric) stable random vector.

Multivariate stable distributions become much less tractable than their univariate counterparts. In particular, their characteristic function no longer admits simple parametrization, similar to equation (2.2). Instead, it is expressed[1] by a multivariate integral with respect to certain *spherical measure* $\sigma(dv)$. The exact form of this measure valid in the case when the marginals are independent may be found in Samorodnitsky and Taqqu (2000) as Example 2.3.5.

Despite the few renowned applications in physics and a number of seminal papers in finance, α–stable distributions did not become popular in Economics. There are two main reasons. First of all, they have no variance and, for $\alpha < 1$, no expected value. Secondly, the magnitude of random values generated from α–stable distributions is unrealistic. In result stable distributions, despite their superior sample fits, are rarely used in (macro)economic simulations.

[1]For more references, see e.g. Sato (2005), p. 83–84.

In order to exploit the ability of α–stable distributions to accommodate skewness or capture inter temporal dependencies, less direct approach has to be taken. Tempered stable distributions introduced in the next chapter are simply α–stable distributions with large jumps being tempered (damped). This procedure makes their all moments finite while simultaneously preserving their desirable properties.

2.5 Conclusions

This chapter makes the following contributions.

First, it demonstrates that 1) heavy–tails in macroeconomic data are not an exception, 2) if not properly taken into account, heavy–tails may cause serious problem in both estimation of and inference from econometric models.

Next, the chapter presents an overview of different ways leading to probability distributions, capable of generating large deviations. It also provides an original comparison of different classes of densities, shedding a new light on their applicability in economics. It is argued that out of the diverse approaches to modelling large deviations, tempered stable distributions (that will be properly defined in the next chapter) are best suited to depict macroeconomic uncertainties.

Third, a concise review of the basic properties of α–stable densities is presented. These properties are useful in the next chapter, where they are compared with characteristics of tempered stable distributions.

Chapter 3 develops new tools that allow for the application of tempered stable densities in modelling macroeconomic distributions.

Chapter 3

Tempered Stable Distribution: Properties, Moments, Estimation, Randomization, Application to Foreign Exchange–rates

This chapter develops the tools that are necessary to depict macroeconomic uncertainties with tempered stable distributions.

3.1 Introduction

The subject of this chapter is the application of tempered stable distributions to quantify macroeconomic risks. Alpha stable distributions, presented in chapter 2, are renown for their ability to accommodate skewness and account for heavy–tails. However, as they possess no higher order moments, they typically generate random numbers of extreme magnitude even in small samples. Hence the idea to dampen probabilities of obtaining large jumps. This operation preserves the desirable properties of α–stable distributions (such as infinite divisibility) and delivers a density with all moments finite. Such semi heavy–tailed distribution, endowed with favourable theoretic properties, finite mean and variance, ability to replicate sample skewness and excess kurtosis, is a natural candidate to depict macroeconomic risks. The density obtained by dampening probabilities of generating large draws from α–stable distribution is further referred to as *tempered stable* (henceforth TS).

There is abundant literature in applied physics and finance on tempered stable processes, tempered stable distributions are the increments of the latter. Hence the task of depicting macroeconomic risks with tempered stable densities appears to be straightforward. This is not the case. The reason lies in differences between characteristics of macroeconomic and financial distributions.

Financial data is typically disaggregated and endowed with high (daily, tick–by–tick) frequencies. As noted by Boyarchenko and Levendorskii (2000), financial distributions in the vicinity of median strongly resemble symmetric α–stable distributions ($\beta \approx 0$). However, as market participants respond differently to gains or losses, the tails of financial distributions are markedly asymmetric. To model this lack of symmetry the probabilities of generating large either positive or negative realizations of financial variables need to be dampened differently. When α–stable distributions are fitted to financial densities, typical estimates of stability index α do not exceed one.

Macroeconomic data, on the other hand, is aggregated and quoted at low frequencies. The corresponding densities are typically moderately skewed in the centre of distribution. As macroeconomic data is driven by fundamentals, the behavioural argument that justifies different ways of dampening probabilities of large positive and negative draws no longer applies. Therefore a more parsimonious setting, where the random draws are being tempered just on the basis of their magnitude and irrespective of their sign, is typically preferred. When α–stable distributions are fitted to macroeconomic densities, the elicited distributions are typically close to Gaussian ($\alpha \approx 2$) and moderately skewed ($|\beta| \neq 1$).

The fact that α–stable distributions fitted to financial and macroeconomic data display different parameter ranges has profound implications. First, it constraints the variety of mathematical methods that can be used to treat the distributions valid in macroeconomics. Any α–stable distribution with $\alpha < 1$ and $\beta = 1$ is supported by a real positive half–line and thus endowed with Laplace transformation. If $\alpha \geq 1$, $\beta = 1$ and the domain of tempered stable random variable is constrained to $[-c, +\infty)$ for some real constant c, then the two–sided Laplace transformation of the underlying probability distribution exists. However, if $\alpha \geq 1$ and $\beta \neq 1$, the ways to treat tempered stable densities lead either through Fourier transformation, or manipulation of spectral measure. Second, different valid parameter ranges limit the applicability of financial literature to macroeconomic modelling. Among financial applications dampened mildly asymmetric α–stable distributions may be obtained only as increments of the processes, considered by Li (2007) and Kim et al. (2010b). However, their results are not directly applicable to the problems, solved in this chapter.

While papers on applied stochastic processes are mostly irrelevant, the literature on tempered stable distributions is scarce. Gravity of tails, moments and variation of certain general families of distribution are investigated by Sztonyk (2010) and Grabchak (2012). Kolossiatis et al. (2011) provide a construction that utilizes tempered stable distributions to generate symplectic measures. The paper by Chakrabarty and Meerschaert (2011) identifies the class of random walks that converge to tempered stable density. Scherer et al. (2010) describe a Fast Fourier Transform (FFT) algorithm, designed for symmetric distributions, which may be used to approximate the density given parameters. The work of Kawai and Masuda (2011) compares a variety of random generators for tempered stable distribution, but only for a case when the stable density being tempered is fully skewed. The framework proposed by Palmer et al. (2008) is even more limited, as only the distributions concentrated on real positive numbers are considered. The tempered stable density investigated by Küchler and Tappe (2011) is the only general enough to befit the purpose, outlined in the previous chapter. At the time when this thesis was written the list above was complete. None of these papers treats the issue of random number generation in non fully–skewed case. There is also no work that would thoroughly investigates parameter estimation for tempered stable distributions.

This chapter develops the missing tools that are needed to depict macroeconomic risks with tempered stable densities. These tools enable the reader to estimate the parameters of and generate pseudo–random numbers from tempered stable distributions, which in their general form are defined by non–probabilistic spectral measures. The assembled toolbox contains a number of theoretic properties of tempered stable distribution that make this density more tractable. It also includes a variety of different parameter estimators and three different randomization procedures. All these results (many of which are new) are provided throughout the sections 3.2–3.5.

The contents of this chapter is as follows. First, the chapter contains a non–standard definition of tempered stable distribution. This definition directly relates tempered stable density to the underlying α–stable distribution, properly handles the case of $\alpha = 1$ and yields centred distribution for nil location parameter. Next, it surveys the main properties of tempered stable densities, some of which were only known in less general setting ($\alpha < 1$, $\alpha \neq 1$, $\beta = 1$). The corresponding proofs may be found in Appendix B. The chapter also presents the complete set formulas for cumulants and moments, previously known only for $\alpha \neq 1$. In a form of Proposition 3.4.1 it proposes a new estimator for parameters of tempered stable distribution, obtained via cumulant matching. It also presents an algorithm for generating random draws from tempered stable density via rejection. The algorithm relies on a novel Propositon 3.5.1, which represents tempered stable random variates as a mixture. This algorithm is valid for all parameter values and may be written in just one line of code if some auxiliary procedures are already available. Out of the three implemented procedures it is also demonstrated to be most accurate in the parameter range, appropriate for macroeconomic distributions. Finally, the pdf charts in this chapter were obtained using a novel Fourier Transform discretization scheme, which generalizes the work of Mittnik et al. (1999). This algorithm, more efficient for asymmetric densities, is provided in Appendix A.

The structure of chapter 3 is the following. Section two formulates a tractable, operational definition of tempered stable distribution that will be used in the remaining part of this text. It also describes its basic properties. The third section provides the formulas for cumulants and moments of tempered stable densities. Section four contains the outline of various parameter estimation techniques for tempered stable distributions. The fifth section presents three methods of random number generation valid in the parameter range, characteristic for macroeconomic data. Section six compares the quality of of pseudo–random draws thus obtained. The final section concludes.

3.2 Definition and properties

A convenient parametric way to define tempered stable distribution is the following.

Definition 3.2.1 (Tempered stable distribution)**.** *Random variable X has tempered stable distribution $TS_\alpha(\beta, \delta, \mu, \theta)$ if its characteristic function takes the form $\Phi_X(u) = e^{\psi_X(u)+i(\mu-\mu_X)u}$ where $\psi_X(u) =$*

$$= \begin{cases} -\frac{1}{2\cos\frac{\pi\alpha}{2}}\delta^\alpha[(1+\beta)(\theta-iu)^\alpha + (1-\beta)(\theta+iu)^\alpha - 2\theta^\alpha] & \alpha \neq 1, \\ \frac{1}{\pi}\delta[(1+\beta)(\theta-iu)\ln(\theta-iu) + (1-\beta)(\theta+iu)\ln(\theta+iu) - 2\theta\ln\theta] & \alpha = 1, \end{cases} \tag{3.1}$$

and the centring term is $\mu_X = \alpha(\cos\frac{\pi\alpha}{2})^{-1}\beta\delta^\alpha\theta^{\alpha-1}$ for $\alpha \neq 1$ and $\mu_X = -\frac{2}{\pi}\beta\delta(\ln\theta+1)$ for $\alpha = 1$. The admissible parameter values are $\alpha \in (0, 2), \beta \in [-1, 1], \delta, \theta > 0, \mu \in \mathbb{R}$.

The formulation above is derived as probably most applicable special case of general definition introduced by Rosiński (2007).

TS distributions evolved from the concept of α–stable distributions, which arise from Central Limit Theorem as limit densities for i.i.d. jumps with heavy power–law tails. Basic properties of α–stable distributions may be found in chapter 2. The idea behind TS density is to alter α–stable distribution so that resulting density had lighter tails. In order to obtain the desired effect, spectral α–stable measure (expressed in polar coordinates) is multiplied by a weighting function which dampens probabilities of generating numbers with large modulus. This approach is known as tempering. The densities obtained this way may retain desirable properties of α–stable distributions, display better fit to the actual data and have higher order moments finite. Chakrabarty and Meerschaert (2011) demonstrate that any random walk with power–law jumps may be approximated with tempered stable density. Hence these distributions provide a universal model of accumulated jump. The density considered here arises when spectral measure of univariate α–stable distribution is weighted with exponent function $e^{-\theta|x|}$. Therefore TS is also known as exponentially tempered stable distribution.

Tempered stable distribution inherits parameters α, β and δ of the α–stable distribution being tempered. The tempering does not affect their qualitative properties: $\alpha \in (0, 2)$ stands for departure from normality (if $\alpha = 2$ the underlying α–stable distribution is Gaussian), $\beta \in [-1, 1]$ governs skewness (if $\beta = 0$ then both α–stable and TS distribution are symmetric), $\delta > 0$ displays scale–like behaviour. Additional parameter $\theta > 0$ measures how far the resulting distribution is from the underlying α–stable density. While $\theta \approx 0$ indicates it is almost exactly α–stable, $\theta \gg 0$ signals a significant departure from the underlying distribution. Parameter $\mu \in \mathbb{R}$ stands for location.

TS distributions constitute densities of Smoothly Truncated Lévy Flights (STLF) stochastic process introduced by Koponen (1995). Boyarchenko and Levendorskii (2000) extended this initial concept, proposing Koponen–Boyarchenko–Levendorskii (KoBoL) process. Finally, Rosiński (2007) defined a general family of tempered stable Lévy processes. The class of infinitely divisible distributions that corresponds to his approach is closed under convolution, self–decomposable, has natural extension to higher dimensions, may display skewness, arbitrary gravity of tails and have all moments finite (CLT may apply).

TS distribution is somehow similar to the well established Carr–Geman–Madan–Yor (CGMY) distribution introduced in Carr et al. (2002). While both distributions are special cases of Classical Tempered Stable (CTS) distribution that represents the increments of KoBoL with single stability index α, CGMY results from asymmetric tempering of symmetric α–stable measure while TS stems from uniform tempering of arbitrary α–stable distribution. This difference translates to diverse tail behaviour.

Contrary to some variants of Definition 3.2.1 used in finance (see e.g.: Boyarchenko and Levendorskii (2000), Kim et al. (2010b)), the formulation above allows to identify the underlying α–stable distribution. While tempered stable distributions utilized in finance are endowed with more parameters, in case of macroeconomic data there is no clear evidence that the reactions of economic agents to upsurges and downturns of indexes are asymmetric. Hence more parsimonious parametrization should be sufficient. If the definition of α–stable random variate in chapter 2 was extended to include Gaussian distribution ($\alpha = 2$), then the tempered stable distribution corresponding to $\alpha = 2$ would again be Gaussian. This is because Lévy measure of Gaussian density is equal to zero and thus it is not affected by tempering. Hence Gaussian distribution would be the only stable density that is simultaneously tempered stable.

Tempered stable distributions constitute a family of possibly skewed, leptokurtic densities with all moments finite that are by construction *infinitely divisible*. They are also endowed with few other useful properties inherited after α–stable distributions. Their main known features are listed in the remaining part of this section. As macroeconomic distributions frequently display moderate skewness and excess kurtosis, TS distribution might be remarkably useful in modelling macroeconomic uncertainty via Monte Carlo simulation. This chapter provides easy to implement randomization method that enables such experiments.

Out of results assembled in this chapter Properties 3.2.1 and 3.2.2 were originally derived in more general forms by, respectively, Rosiński (2007) and Boyarchenko and Levendorskii (2000). The results of Terdik and Woyczyński (2006) are useful in verifying Rosiński existence conditions. Property 3.2.3 follows directly from basic properties of stable distributions and Definition 3.2.1. For $\alpha \neq 1$ general versions of Property 3.2.4 and Corollary 3.3.1 were demonstrated by Kim et al. (2010a) and Terdik and Woyczyński (2006). Property 3.2.5 was formulated for $\alpha < 1$ by Küchler and Tappe (2011). Property 3.2.7 was given in Appendix C of Terdik and Gyires (2009) while Property 3.2.9 was derived by Nakao (2000). Both were only known for $\beta = 1$. Properties 3.2.6 and 3.2.8 could not be traced in the literature, but seem self evident. Definition 3.2.2 is new. All the supplementary proofs are relegated t Appendix B.

In the remaining part of this chapter whenever referring to random variable X it will be assumed that $X \sim TS_\alpha(\beta, \delta, \mu, \theta)$. As for $\mu = 0$ the resulting distribution is centred, the statements collected below take particularly convenient form and are more compact than their α–stable counterparts.

Property 3.2.1 (Existence of moments)**.** *R. v. X has absolute moments about the origin of arbitrary order. Furthermore, for $|u| \leq \theta$ its moment generating function exists and is given by*

$$M_X(u) = \Phi_X(-iu).$$

Property 3.2.2 (Limit distribution)**.** *If* $X \sim TS_\alpha(\beta,\delta,\mu,\theta)$ *then in the limit for* $\alpha \to 0$ *we have*

$$X \sim (V_1 - V_2)/\theta^2$$

where V_1*,* V_2 *are independent copies of* $V \sim Gamma(1/2,\theta)$*.*

Property 3.2.3 (Essential support)**.** *Distribution of random variable* $X \sim TS_\alpha(\beta,\delta,\mu,\theta)$ *is concentrated on* $[-\mu,+\infty)$ *for* $\alpha < 1$ *and* $\beta = 1$*, otherwise it is supported by all real numbers.*

Property 3.2.4 (Cdf)**.** *Cumulative distribution function of random variable* X *is given by*

$$F_X(x) = \frac{1}{\pi} e^{\rho x} Re\Big(\int_0^{+\infty} e^{-ixu} \frac{\Phi_X(u+i\rho)}{\rho - iu}\, du \Big), \quad x \in \mathbb{R}, \tag{3.2}$$

where ρ *is arbitrary real number such that* $\rho \in (0,\theta)$*.*

Property 3.2.5 (Existence of pdf)**.** *Probability density function* $f_X(x)$ *of random variable* X *is well defined, unimodal and smooth, i.e.* $f_X(x) \in C^{+\infty}(\mathbb{R})$*.*

For $\alpha < 1$ and $\beta = 1$ pdf of tempered stable distribution may be elicited by exponential tilting of the corresponding stable pdf, just as in Kawai and Masuda (2011). Otherwise its closed form is not known.

When the underlying random variable X is endowed with unimodal pdf, its cdf is necessarily strictly increasing and continuous. Hence Property 3.2.5 implies that cdf of any tempered stable random distribution is invertible. The inverse of cdf will be denoted as $F_X^{-1}(y)$.

Figures 3.1–3.3 depict tempered stable densities for parameter $\beta \in \{0.25, 0.5, 0.75, 1\}$. On each graph parameter α amounts to 0.5, 1 and 1.5 on, respectively, top, middle on bottom charts. For the sake of comparison, Figure 3.1 was drawn for mild ($\theta = 0.2$), Figure 3.2 – for moderate ($\theta = 1$), while Figure 3.3 – for strong ($\theta = 5$) tempering. Figures 3.1–3.3 may be compared with stable densities, plotted in Samorodnitsky and Taqqu (2000) on p. 36. Note that presenting a collection of (potentially) heavy–tailed pdfs always entails a trade–off. Either all the charts are plotted on the same interval, and thus it is possible to compare the shapes of distributions, or each density is drawn in the domain that reflects the gravity of its tails, and thus it is possible to compare probabilities of large deviations. As one of the next sections is dedicated to measuring risk, associated with large losses, the latter approach is taken. For the set of parameters given above the intervals that support pdfs are extended and pdf are again evaluated, until densities, obtained at each interval end-points, fall below 10^{-3}. The outputs of this procedure are depicted in Figures 3.1–3.3. All the distributions are centred and endowed with unit variance (standardized).

Figure 3.4 depicts tempered stable distribution with $\beta = 0$ and parameter $\alpha \in \{0.5, 1.0, 1.5, 1.9, 2\}$. Parameter θ amounts to 0.4, 1.0 and 5 on, respectively, top, middle on bottom chart. Value 0.4 was used instead of $\theta = 0.2$, as the latter for $\alpha = 0.5$ produced a density that was so spiky that the remaining four pdfs could no longer be read from the picture. Again all the distributions are standardized.

Property 3.2.6 (Reflection). *If $Y = -X$ then $Y \sim TS_\alpha(-\beta, \delta, -\mu, \theta)$.*

In consequence pdf of X satisfies $f_X(x;\beta,\mu) = f_X(-x;-\beta,-\mu)$ while its cdf fulfils $F_X(x;\beta,\mu) = 1 - F_X(-x;-\beta,-\mu)$.

Property 3.2.7 (Linear transformations). *If $Y = aX + b$ where $a \neq 0$ and b are real constants, then*

$$Y \sim TS_\alpha((\operatorname{sgn} a)\beta, |a|\delta, a\mu + b, \theta/|a|). \tag{3.3}$$

Property 3.2.8 (Additivity). *If $X_k \sim TS_\alpha(\beta_k, \delta_k, \mu_k, \theta)$ are independent r.v.'s for $k = 1, 2$ and $Y = X_1 + X_2$ then $Y \sim TS_\alpha(\beta, \delta, \mu, \theta)$ with*

$$\beta = \frac{\beta_1 \delta_1^\alpha + \beta_2 \delta_2^\alpha}{\delta_1^\alpha + \delta_2^\alpha}, \ \ \delta = (\delta_1^\alpha + \delta_2^\alpha)^{\frac{1}{\alpha}}, \ \ \mu = \mu_1 + \mu_2.$$

Hence the outcome of summation of independent tempered stable random variates remains tempered stable only if both α's and θ's match. This result is identical to Property 2.4.5 of α–stable distributions given in chapter 2.

Property 3.2.9 (Scaling). *Let $f_X(x;\mu,\theta)$ stand for a pdf of X, then*

$$f_X^{*n}(x;\mu,\theta) = n^{-1/\alpha} f_X(n^{-1/\alpha}x; n^{1-1/\alpha}\mu, n^{1/\alpha}\theta),$$

where f_X^{*n} stands for convolution power of order n.

If Y is a sum of n independent copies of X, then f_X^{*n} is pdf of Y.

Note that if X_1, X_2 are independent random variables with tempered stable distributions, then the expression $bX_1 + cX_2$ does not have a tempered stable distribution for all $b, c \in \mathbb{R}$. However, the outcome of the convolution remains tempered stable in the sense of Rosiński (2007). Furthermore, what follows from Properties 3.2.7 and 3.2.8, tempered stable distributions (in the sense of Definition 3.2.1) are no longer *self–decomposable*.

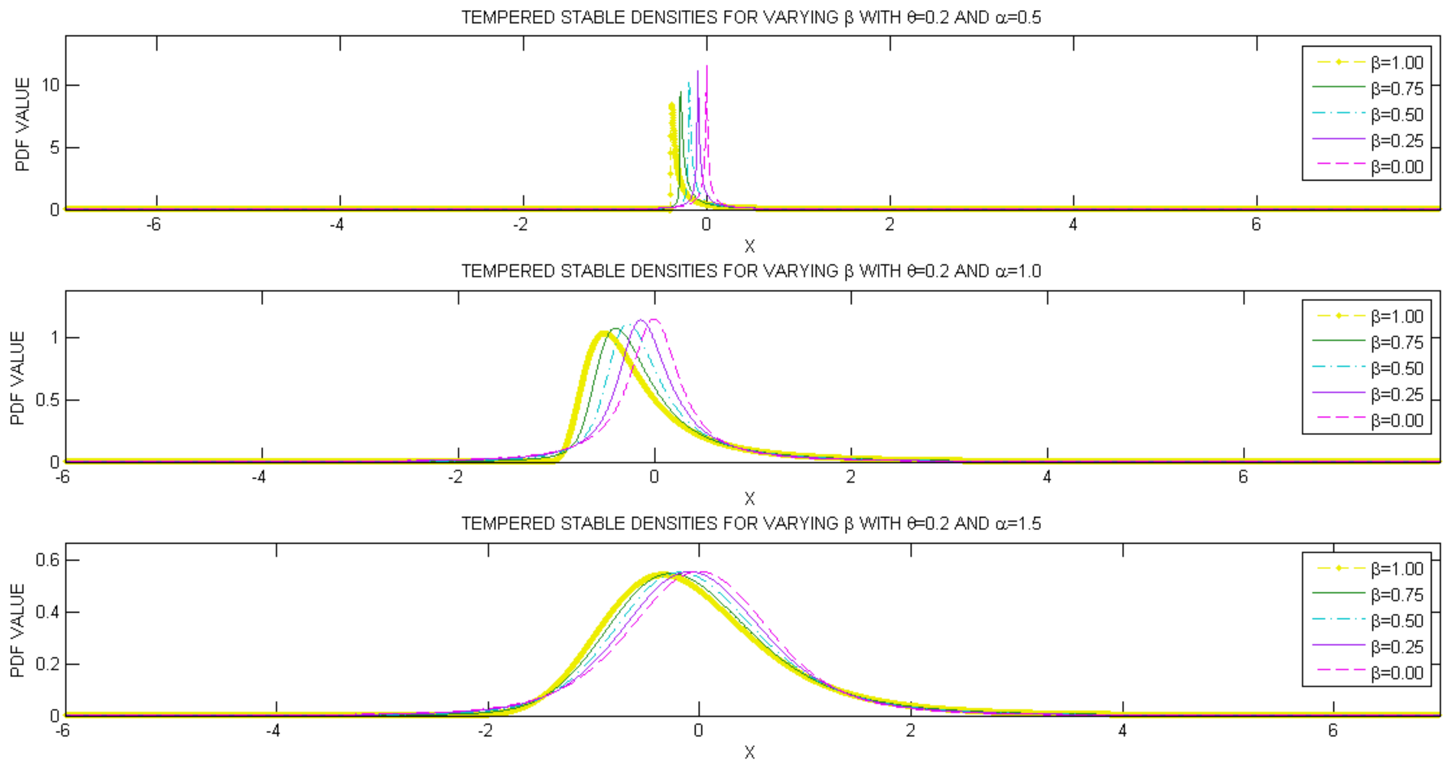


Figure 3.1: Tempered stable densities with $\theta = 0.2$ and varying β's. Parameter α amounts to, respectively, 0.5, 1.0 and 1.5 on top, middle and bottom chart. All three distributions are standardized.

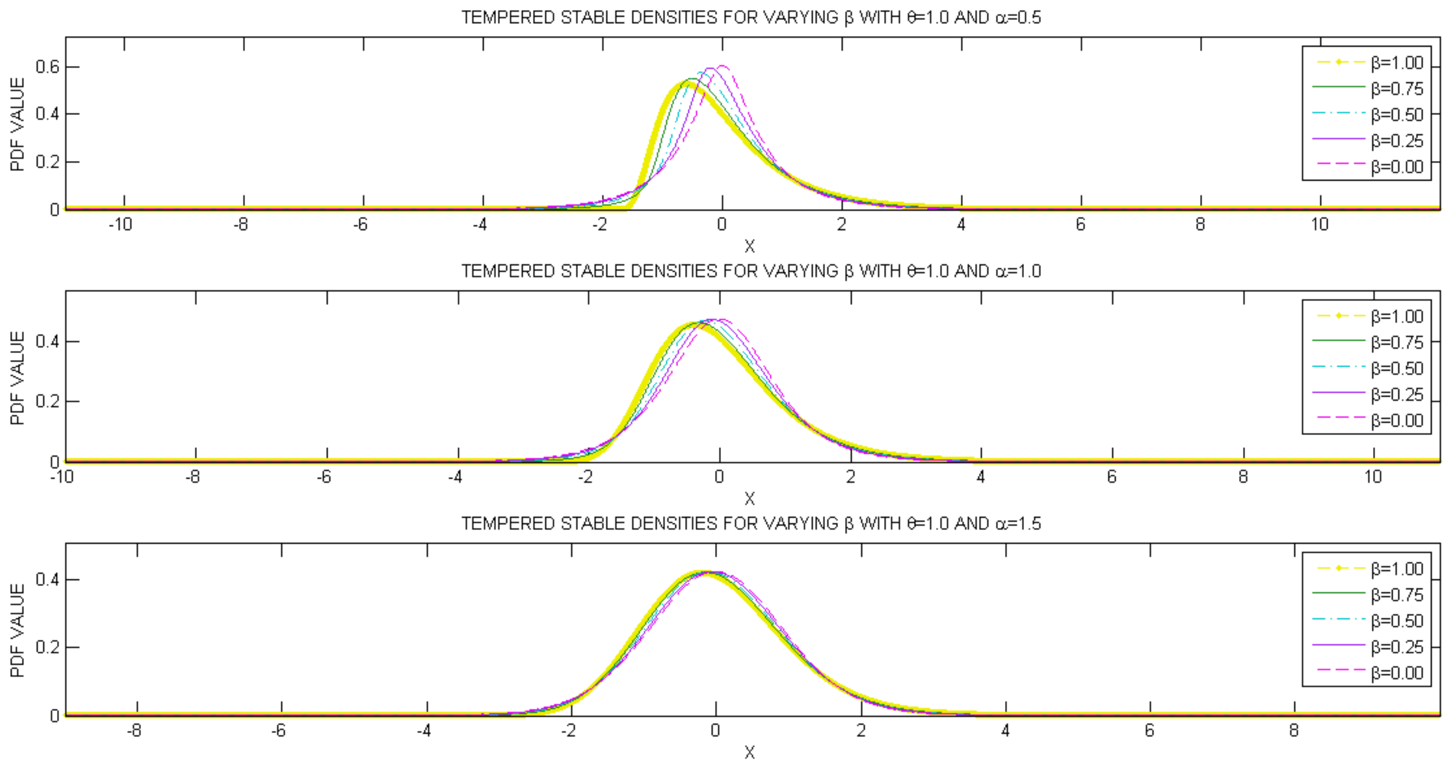


Figure 3.2: Tempered stable densities with $\theta = 1.0$ and varying β's. Parameter α amounts to, respectively, 0.5, 1.0 and 1.5 on top, middle and bottom chart. All three distributions are standardized.

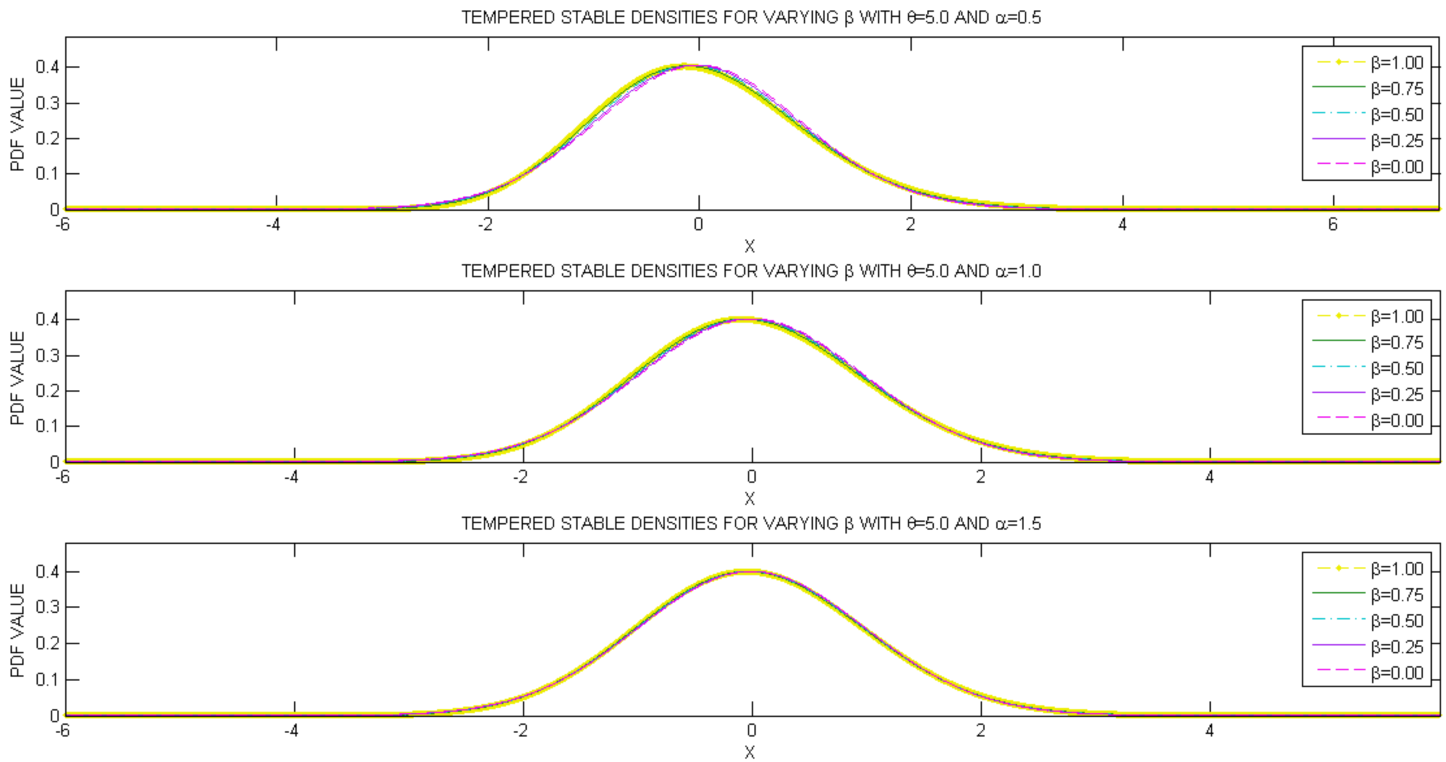


Figure 3.3: Tempered stable densities with $\theta = 5.0$ varying β's. Parameter α amounts to, respectively, 0.5, 1.0 and 1.5 on top, middle and bottom chart. All three distributions are standardized.

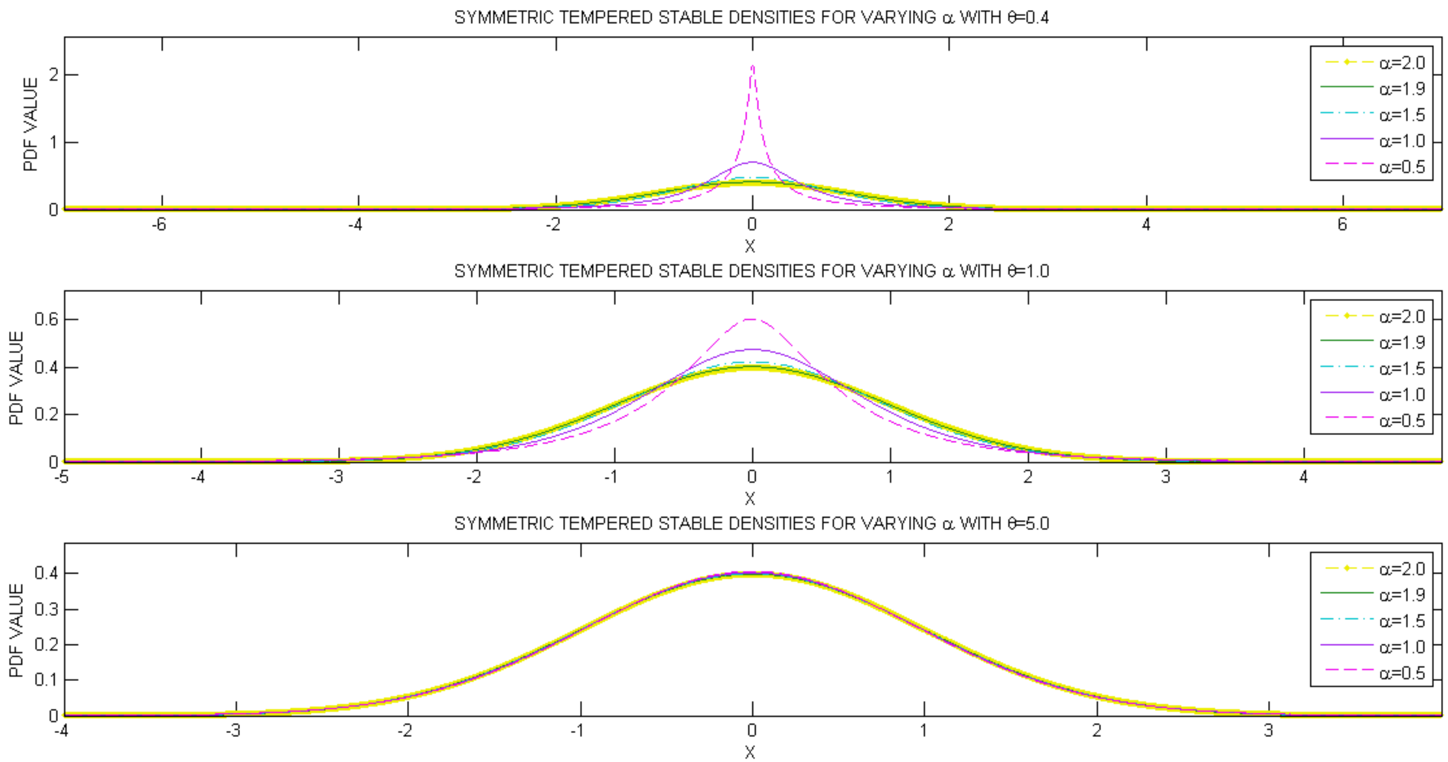


Figure 3.4: symmetric ($\beta = 0$) tempered stable densities with varying α's. Parameter θ amounts to, respectively, 0.4, 1.0 and 5.0 on top, middle and bottom chart. All three distributions are standardized.

While multivariate tempered stable distribution might be defined through non–probabilistic spectral measures as in Rosiński (2007), such characterization would not be practical. Here a simple but more tractable multivariate extension of tempered stable densities is proposed, which might be defined as follows.

Definition 3.2.2 (Multivariate tempered stable distribution)**.** *Let X_j be independent random variables such that $X_j \sim TS_\alpha(\beta_j, \delta_j, 0, \theta)$, given $\beta_j \in [-1, 1]$ and $\delta_j > 0$ for $j \in \{1, \cdots, m\}$ while $\alpha \in (0, 2)$ and $\theta > 0$. Assume A is an invertible $(m \times m)$ matrix, set $\mu_j \in \mathbb{R}$ for $j \in \{1, \cdots, m\}$ and*

$$\boldsymbol{\beta} = (\beta_1, \cdots, \beta_m)^T \ , \ \boldsymbol{\delta} = (\delta_1, \cdots, \delta_m)^T \ , \ \boldsymbol{\mu} = (\mu_1, \cdots, \mu_m)^T \ , \ \boldsymbol{X} = (X_1, \cdots, X_m)^T.$$

Then a random vector

$$\boldsymbol{Y} = \mu + A\boldsymbol{X}$$

is said to have a multivariate tempered stable (mTS) distribution with parameters α, $\boldsymbol{\beta}$, $\boldsymbol{\delta}$, $\boldsymbol{\mu}$, θ and A. This relation will be denoted as $\boldsymbol{Y} \sim mTS_\alpha(\boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{\mu}, \theta, A)$.

The following fact justifies the appropriateness of Definition 3.2.2. It follows from the remark, communicated to the author by Mark M. Meerschaert.

Fact 3.2.10. *If $\boldsymbol{Y} \sim mTS_\alpha(\boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{\mu}, \theta, A)$ then the distribution of $\boldsymbol{Y}$ lies in Rosiński (2007) class.*

In particular, the distribution of Y is well defined and infinitely divisible. It also has all the moments finite. Hence the structure of dependence between different random variates may be captured in mTS density via correlation matrix.

The properties presented above allow for comparison of α–stable and tempered stable distributions. The both classes share Properties 3.2.3, 3.2.5 and 3.2.6. Both α–stable and tempered stable densities display additivity, the formulations of which are almost identical (respectively, Propositions 2.4.5 and 3.2.8). The only difference is that in order to conveniently add two independent α–stable random variates the underlying densities are required to have the same stability index α, while in the tempered stable case both α and θ are required to match. In contrast to α–stable distributions which possess no moments of the order higher than α (Property 2.4.3), tempered stable densities have all moments finite (Property 3.2.1). The exact formulas for these moments follow from the characterization, provided in the next section (Corollary 3.3.1). As all the moments exist, parameters of tempered stable distribution may be estimated by a variant of Method of Moments (Proposition 3.4.1 from section 3.3). Furthermore, it is also possible characterize the structure of dependence between the marginals of the multivariate extension of tempered stable distributions (Definition 3.2.2) with correlation matrix. As this extension relies on Fourier transformation of probability measure, it is markedly more tractable than the definition of stable random vector (Definition 2.4.1). However, it also does not possess the convenient Property 2.4.8, which is unique to stable densities. As Definition 3.2.1 implies that for $\mu = 0$ the corresponding density is centred, Property 3.2.7 (contrary to Property 2.4.4) no longer requires two separate cases. By a similar token, the mixture representation formulated as Proposition 3.5.1 in section 3.5 takes a more compact form than Property 2.4.6.

3.3 Cumulants and moments

It follows from Property 3.2.1 from previous section that all moments of $TS_\alpha(\beta, \delta, \mu, \theta)$ exist regardless of parameters. Section 3.3 provides the exact formulas for both theoretical moments and cumulants of this distribution. For a given probability density the comparison of sample and theoretic moments is a straightforward way to evaluate the quality of any random number generator. In case of tempered stable densities it is the only available approach to assess the accuracy of randomization algorithm without tedious numerical approximations. The cumulants are an auxiliary result, necessary to obtain the moments.

After Stuart and Ord (1994) define cumulants[1] of integer order p as

$$\kappa_p = \frac{1}{i^p}\left(\frac{d^p}{du^p}\ln\Phi_X(u)\right)\Big|_{u=0}.$$

The moments of tempered α–stable random variates are all finite, which implies existence of all the cumulants, but do not take any convenient form. These cumulants are highly tractable.

When $\alpha \neq 1$ the cumulants may be elicited from Terdik and Woyczyński (2006). For $\alpha = 1$ their formulas are no longer valid and the cumulants need to be found directly. Combining both sets of results yields

[1] Having different application in mind, Barndorff-Nielsen and Shephard (2012) rely on another definition.

Corollary 3.3.1 (Cumulants). *If random variable $X \sim TS_\alpha(\beta, \delta, \mu, \theta)$ then its cumulants fulfil*

$$\kappa_p = \begin{cases} \mu & \text{if } p = 1, \\ \frac{2\delta}{\pi}\theta^{1-p}(p-2)!(I_p + \beta I_{p+1}) & \text{if } p \neq 1,\ \alpha = 1, \\ \alpha \prod_{j=1}^{p-1}(j-\alpha)(\cos\frac{\pi\alpha}{2})^{-1}\delta^\alpha\theta^{\alpha-p}(I_p + \beta I_{p+1}) & \text{if } p \neq 1,\ \alpha \neq 1, \end{cases} \tag{3.4}$$

where $I_p = 2^{-1}(1 + (-1)^p)$.

Let $\mathbb{S}kwX$ stand for skewness, $\mathbb{K}urX$ denote excess kurtosis of random variable X. For $\alpha \neq 1$ formulas (3.89–3.90) from Stuart and Ord (1994) combined with Corollary 3.3.1 imply

$$\mathbb{E}X = \mu,\ \mathbb{V}arX = \alpha(1-\alpha)\Big(\cos\frac{\pi\alpha}{2}\Big)^{-1}\delta^\alpha\theta^{\alpha-2},$$

$$\mathbb{S}kwX = (2-\alpha)\beta\sqrt{\frac{\cos\frac{\pi\alpha}{2}}{\alpha(1-\alpha)(\delta\theta)^\alpha}},\ \mathbb{K}urX = \frac{(2-\alpha)(3-\alpha)\cos\frac{\pi\alpha}{2}}{\alpha(1-\alpha)(\delta\theta)^\alpha}.$$

If $\alpha = 1$ it holds that

$$\mathbb{E}X = \mu,\ \mathbb{V}arX = \frac{2\delta}{\pi\theta},\ \mathbb{S}kwX = \frac{\beta}{\sqrt{\delta\theta}}\sqrt{\frac{\pi}{2}},\ \mathbb{K}urX = \frac{\pi}{\delta\theta}.$$

It is possible to guarantee that the resulting distribution has unit variance by setting appropriate $\delta > 0$. If the density is standardised, skewness and excess kurtosis are its third and fourth cumulant. Note that as $\mathbb{K}urX > 0$ for all $\alpha \in (0, 2)$, TS distribution is leptokurtic.

By μ_p' and μ_p denote, respectively, moments about the origin and about the mean. Given the cumulants, recursion formula from p. 88–91 in Stuart and Ord (1994) yields moments

$$\mu_p' = \kappa_p + \sum_{j=1}^{p-1}\binom{p-1}{j-1}\kappa_j\mu_{p-j}'.$$

Moments about the origin of order p are polynomials of order p of the first p cumulants

$$\mu_1' = \kappa_1,\ \mu_2' = \kappa_2 + \kappa_1^2,\ \mu_3' = \kappa_3 + 3\kappa_2\kappa_1 + \kappa_1^3,$$

$$\mu_4' = \kappa_4 + 4\kappa_3\kappa_1 + 3\kappa_2^2 + 6\kappa_2\kappa_1^2 + \kappa_1^4,$$

$$\mu_5' = \kappa_5 + 5\kappa_4\kappa_1 + 10\kappa_3\kappa_2 + 10\kappa_3\kappa_1^2 + 15\kappa_2^2\kappa_1 + 10\kappa_2\kappa_1^3 + \kappa_1^5,\ \dots\ .$$

The central moments fulfil the similar set of equations with $\kappa_1 = 0$, i.e.

$$\mu_1 = 0,\ \mu_2 = \kappa_2,\ \mu_3 = \kappa_3,$$

$$\mu_4 = \kappa_4 + 3\kappa_2^2,\ \mu_5 = \kappa_5 + 10\kappa_3\kappa_2,\ \dots\ .$$

We can always choose $\mu = 0$ and $\delta > 0$ such that the resulting distribution is standardised. If convenient, we may change the parametrization and drop the term μ_X in (3.1). As it affects only the first cumulant, the expected value of the resulting distribution becomes $\kappa_1 = \mu_X + \mu$ while all the remaining results still hold.

3.4 Outline of estimation

This section is dedicated to estimation of parameters of tempered stable distributions. It covers a number of feasible approaches, ordered from analytically tractable to numerically intensive. These methods are: Cumulant Matching, Numerical Method of Cumulants, Numerical Method of Moments, Numerical Maximum Likelihood and Empirical Characteristic Function estimator. It also treats the issue of dimension reduction and applying different numerical solvers.

The problem of parameter estimation for TS distributions has been barely touched in the literature. The paper by Terdik and Gyires (2009) provides the only method that was 1) described along with implementation details in the original paper and 2) applied to real data. This paper relied on analytical and numerical cumulant matching. What is typically given in the literature are just numerical estimates of Maximum Likelihood (ML) with no implementation details. The works of Scherer et al. (2010) and Bianchi et al. (2013) do contain significant amount of details, but only regarding characteristic function inversion. More complex route is taken by Li et al. (2012), who perform a spectral estimation of CTS parameters via the Generalized Method of Moments (GMM) with continuum of moment conditions. The formulas for the weights that attain the efficiency of ML estimators are not known in CTS case, hence their approach is less efficient than ML. As this technique was tailored to stochastic volatility models, it may be difficult to apply in different setting. Contrary to what they claim, Küchler and Tappe (2011) did not solve the problem of parameter estimation for CTS distribution. In their paper the estimates that solve equations 3.5 are required first to obtain the estimators. Hence the problem they did solve is inverse to estimation problem.

Throughout this section it is always assumed that $x_1, \ldots, x_n$ are independent realizations of $X \sim TS_\alpha(\beta, \delta, \mu, \theta)$. Data vector $\mathbf{x}$ is defined as $(x_1, \ldots, x_n)$. Whenever dependence on the parameters is important, moments about the origin of order p are denoted as $\mu_p^{'}(\alpha, \beta, \delta, \mu, \theta)$, cumulants of order p with $\kappa_p(\alpha, \beta, \delta, \mu, \theta)$, pdfs with $f_X(x; \alpha, \beta, \delta, \mu, \theta)$ and characteristic functions as $\Phi_X(u; \alpha, \beta, \delta, \mu, \theta)$.

Perhaps the easiest approach to parameter estimation is to obtain the estimates via Cumulant Matching (CM). For integer k denote sample cumulants as $\hat{\kappa}_p$. Note that if $\hat{\mu}_p^{'}$ stands for sample moment about the origin of integer order p, i.e.

$$\hat{\mu}_p^{'} = \frac{1}{n} \sum_{k=1}^{n} x_k^p,$$

then the formulas from previous section may be inverted to yield

$$\hat{\kappa}_1 = \hat{\mu}_1^{'}, \;\; \hat{\kappa}_2 = \hat{\mu}_2^{'} - \hat{\kappa}_1^2, \;\; \hat{\kappa}_3 = \hat{\mu}_3^{'} - 3\hat{\kappa}_2\hat{\kappa}_1 - \hat{\kappa}_1^3,$$

$$\hat{\kappa}_4 = \hat{\mu}_4^{'} - 4\hat{\kappa}_3\hat{\kappa}_1 - 3\hat{\kappa}_2^2 - 6\hat{\kappa}_2\hat{\kappa}_1^2 - \hat{\kappa}_1^4,$$

$$\hat{\kappa}_5 = \hat{\mu}_5^{'} - 5\hat{\kappa}_4\hat{\kappa}_1 - 10\hat{\kappa}_3\hat{\kappa}_2 - 10\hat{\kappa}_3\hat{\kappa}_1^2 - 15\hat{\kappa}_2^2\hat{\kappa}_1 - 10\hat{\kappa}_2\hat{\kappa}_1^3 - \hat{\kappa}_1^5, \;\; \ldots \; .$$

Cumulant Matching relies on the observation that, when sample size increases, sample cumulants by the Strong Law of Large Numbers converge almost surely to their analytical counterparts. Hence in sufficiently large samples they should be good approximations of theoretic cumulants from Corollary 3.3.1. If we assume they are exactly equal for as many low–order (non–zero) sample cumulants as there are (non–zero) parameters, the solution to the system of equations

$$\hat{\kappa}_p = \kappa_p \text{ for } p \in \{1, \ldots, m\} \tag{3.5}$$

with $m \geq 5$ yields the estimates.

Cumulant Matching is a variant of Method of Moments (MM). As noted by Press (1967), the estimates thus obtained are consistent, but not necessary sufficient or efficient. Furthermore, as they are functions of sample moments, the large sample distribution of the cumulant estimators will be normal.

Note that given Corollary 3.3.1, the main difficulty in eliciting the estimators lies in solving a system of five non–linear equations. Contrary to a linear case, in non–linear systems it may be difficult not only to obtain the desired solution, but even to demonstrate it actually exists. However, in this case the cumulants of TS distribution may be directly mapped to parameters. The estimators disentangled from the initial equations are given in the proposition below, along with the corresponding existence conditions.

Proposition 3.4.1 (Cumulant matching estimators)**.** *Assume that* $\hat{\kappa}_4 > 0$*.*

i) If $\hat{\kappa}_3 \neq 0$*,* $\hat{\kappa}_5 \neq 0$ *and* $\hat{\kappa}_5\hat{\kappa}_2 \neq \hat{\kappa}_4\hat{\kappa}_3$ *with* $(\hat{\kappa}_5\hat{\kappa}_2 - \hat{\kappa}_4\hat{\kappa}_3)/\hat{\kappa}_4\hat{\kappa}_3 \neq 2$*, then CM estimators of all parameters exist if and only if*

$$\frac{\hat{\kappa}_5\hat{\kappa}_2}{\hat{\kappa}_4\hat{\kappa}_3} \leq \frac{2\hat{\kappa}_4\hat{\kappa}_2}{\hat{\kappa}_3^2} - 1,$$

these estimators are given by

$$\hat{\alpha} = 2\Big(1 - \frac{\hat{\kappa}_4\hat{\kappa}_3}{\hat{\kappa}_5\hat{\kappa}_2 + \hat{\kappa}_4\hat{\kappa}_3}\Big),\ \hat{\beta} = \frac{\hat{\kappa}_5\hat{\kappa}_2 - \hat{\kappa}_4\hat{\kappa}_3}{2\hat{\kappa}_4\hat{\kappa}_2}\hat{\theta},\ \hat{\delta} = \Big(\frac{\hat{\kappa}_2 \cos \pi\hat{\alpha}/2}{\hat{\alpha}(1-\hat{\alpha})}\Big)^{1/\hat{\alpha}}\hat{\theta}^{\,2/\hat{\alpha}-1},$$

$$\hat{\mu} = \hat{\kappa}_1,\ \hat{\theta} = \sqrt{\frac{2\hat{\kappa}_3\hat{\kappa}_2(\hat{\kappa}_5\hat{\kappa}_2 + \hat{\kappa}_4\hat{\kappa}_3)}{(\hat{\kappa}_5\hat{\kappa}_2 - \hat{\kappa}_4\hat{\kappa}_3)^2}}. \tag{3.6}$$

ii) If $\hat{\kappa}_3 \neq 0$*,* $\hat{\kappa}_5 \neq 0$ *and* $\hat{\kappa}_5\hat{\kappa}_2 \neq \hat{\kappa}_4\hat{\kappa}_3$ *but* $(\hat{\kappa}_5\hat{\kappa}_2 - \hat{\kappa}_4\hat{\kappa}_3)/\hat{\kappa}_4\hat{\kappa}_3 = 2$*, then CM estimators of all parameters exist if and only if*

$$\frac{\hat{\kappa}_3^2}{\hat{\kappa}_4\hat{\kappa}_2} \leq \frac{1}{2},$$

these estimators are given by

$$\hat{\alpha} = 1,\ \hat{\beta} = \frac{\hat{\kappa}_3}{\hat{\kappa}_2}\hat{\delta},\ \hat{\delta} = \frac{\pi}{2}\hat{\kappa}_2\hat{\theta},\ \hat{\mu} = \hat{\kappa}_1,\ \hat{\theta} = \sqrt{\frac{2\hat{\kappa}_2}{\hat{\kappa}_4}}. \tag{3.7}$$

iii) If $\hat{\kappa}_3 = 0$*,* $\hat{\kappa}_5 = 0$ *and* $\hat{\kappa}_6\hat{\kappa}_2 \neq \hat{\kappa}_4^2$ *with*

$$(4\hat{\kappa}_4^2 + \sqrt{\hat{\kappa}_6^2\hat{\kappa}_4^2 + 14\hat{\kappa}_6\hat{\kappa}_4^2\hat{\kappa}_2 + \hat{\kappa}_4^4})/(\hat{\kappa}_6\hat{\kappa}_2 - \hat{\kappa}_4^2) \neq 3,$$

then CM estimators of all parameters exist if and only if

$$3\hat{\kappa}_6\hat{\kappa}_2 > 10\hat{\kappa}_4^2,$$

these estimators are given by

$$\hat{\alpha} = \frac{5}{2} - \frac{1}{2(\hat{\kappa}_6\hat{\kappa}_2 - \hat{\kappa}_4^2)}\Big(4\hat{\kappa}_4^2 + \sqrt{\hat{\kappa}_6^2\hat{\kappa}_4^2 + 14\hat{\kappa}_6\hat{\kappa}_4^2\hat{\kappa}_2 + \hat{\kappa}_4^4}\Big),\ \hat{\beta} = 0,$$

$$\hat{\delta} = \frac{\pi}{2}\hat{\kappa}_2\hat{\theta},\ \ \hat{\mu} = \hat{\kappa}_1,\ \ \hat{\theta} = \sqrt{(14 - 4\hat{\alpha})\frac{\hat{\kappa}_4\hat{\kappa}_2}{\hat{\kappa}_6\hat{\kappa}_2 - \hat{\kappa}_4^2}}. \tag{3.8}$$

iv) If $\hat{\kappa}_3 = 0$, $\hat{\kappa}_5 = 0$ and $\hat{\kappa}_6\hat{\kappa}_2 \neq \hat{\kappa}_4^2$ but

$$(4\hat{\kappa}_4^2 + \sqrt{\hat{\kappa}_6^2\hat{\kappa}_4^2 + 14\hat{\kappa}_6\hat{\kappa}_4^2\hat{\kappa}_2 + \hat{\kappa}_4^4})/(\hat{\kappa}_6\hat{\kappa}_2 - \hat{\kappa}_4^2) = 3,$$

then CM estimators of all parameters are

$$\hat{\alpha} = 1,\ \ \hat{\beta} = 0,\ \ \hat{\delta} = \frac{\pi}{2}\hat{\kappa}_2\hat{\theta},\ \ \hat{\mu} = \hat{\kappa}_1,\ \ \hat{\theta} = \sqrt{\frac{2\hat{\kappa}_2}{\hat{\kappa}_4}}. \tag{3.9}$$

In all the remaining cases CM estimator of at least one parameter does not exist.

Note that for $\beta = 0$ all sample cumulants of odd orders are equal to zero and thus can not be used to estimate the remaining parameters. This is why the formulas, presented in point iii) in the statement above, depend only on sample cumulants of even orders. It is reasonable to use part iii) of Proposition 3.4.1 only if $\hat{\kappa}_3 \approx \hat{\kappa}_5 \approx 0$ and part ii) or iv) only if we obtain $\hat{\alpha} \approx 1$ from either part i) or iii). Thus in a typical situation part i) is sufficient.

Albeit technically convenient, estimators given above are of little practical use. The reason is that CM estimators rely on asymptotic results that do not work well in *small*[1] samples. The main culprits are: non–linearity of the estimator formulas, interdependence of estimates and heavy–tails of the underlying distribution.

It is clear that the fitted estimates will always be distorted versions of true estimates. This is because sample cumulants in finite samples will not be equal to theoretic cumulants, hence the set of true estimates will not fulfil equation (3.5). When sample cumulants are treated as approximations of theoretic cumulants and inserted into CM estimators, approximation errors may be amplified due to non–linearity of equations (3.6)–(3.9). As estimator formulas are interdependent, these errors may add up across the estimates of different parameters. Furthermore, as the underlying distribution is heavy–tailed, higher order cumulants may be dominated by just a few largest observations and thus vary wildly even across large samples. Therefore ES estimators tend to be fragile. It is also not clear whether the estimates that solve the distorted system of non–linear conditions are good approximation of true parameter values. Finally, as the four cases listed in Proposition 3.4.1 are not exhaustive, CS estimators are not guaranteed to exist. In general, it seems that this fully tractable method entails serious problems that can not be easily solved.

[1] Here *small* may mean *well above* $10,000$.

When analytical approaches fail, it is frequently reasonable to resort to computationally intensive methods. These are numerical versions of: Cumulant Matching, Method of Moments, Characteristic Function Matching and Maximum Likelihood. In all these cases the admissible parameter range is constrained with $\alpha \in (0,1)$, $\beta \in [-1,1]$, $\delta, \theta > 0$ and $\mu \in \mathbb{R}$.

The first possible approach is to follow the steps of Terdik and Gyires (2009). First assume that the system of equations, utilized to derive CM estimator, holds only approximately. Then the desired parameter estimates may be obtained numerically by minimizing the distance between two vectors – sample cumulants and theoretic cumulants, implied by the estimates. A popular way to measure distance in this type of application is Euclidean norm. Then the minimand is a sum of squares. Analytical solution to the cumulant matching problem is still useful, it may serve as a starting point for non–linear solver. If sample cumulants of order higher than 5 contain a relevent information on the shape of the underlying distribution it may often be beneficial to include them into the minimand. The resulting Numerical Method of Cumulants (NMC) estimator is given by

$$(\hat{\alpha}, \hat{\beta}, \hat{\delta}, \hat{\mu}, \hat{\theta}) = \operatorname{argmin} \sum_{p=1}^{m} \left(\hat{\kappa}_p - \kappa_p(\alpha, \beta, \delta, \mu, \theta) \right)^2, \ m \geq 5. \tag{3.10}$$

A similar estimator was applied by Terdik and Gyires (2009) in their Appendix D to elicit parameters of totally skewed ($\beta = 1$) TS distributions. In case of internet traffic data they investigated the quality of estimates was best when first eight sample cumulants ($m = 8$) were accounted for. Note that $m > 5$ may make the estimates either more robust or more fragile, depending on scaling of the data.

Another route to parameter estimation is to utilize Method of Moments. Although theoretical MM estimator could be likely derived for TS distribution, this possibility will not be explored any further. As Method of Moments estimators are asymptotically equivalent to Cumulant Matching they would display similar finite sample features. Hence any MM estimator would have all the weaknesses of the estimator derived above while it would be deprived of its main strength – tractability that stems from compact formulas, given in Corollary 3.3.1. However, an approach worth investigating is using Numerical Method of Moments (NMM) estimator, defined by

$$(\hat{\alpha}, \hat{\beta}, \hat{\delta}, \hat{\mu}, \hat{\theta}) = \operatorname{argmin} \sum_{p=1}^{m} \left(\hat{\mu}'_p - \mu'_p(\alpha, \beta, \delta, \mu, \theta) \right)^2, \ m \geq 5. \tag{3.11}$$

The reason is that, albeit NMM is asymptotically equivalent to NMC, both methods display slightly different small sample properties. A set of two alternative estimators with different minimands might also be useful on data sets where identification problems arise due to numerical reasons.

The third, less tractable method is Maximum Likelihood. As pdf (in form of an infinite series expansion) is only available (Kawai and Masuda, 2011; Palmer et al., 2008) for $\alpha < 1$ and $\beta = 1$ Maximum Likelihood estimators can not be directly used in general case. However, a numerical approximation of ML estimates may be easily obtained if we can approximate pdf values in consecutive points of the sample. Note that this problem is essentially solved if we can evaluate pdf in equally spaced points over the essential support of the distribution. Then for each $k \in \{1, \ldots, n\}$ value $f_X(x_k)$ could be approximated with a cubic spline interpolation. Numerical cost of such approximation is negligible.

Let $x_l = a + hl$ with $l = 0, \ldots, N$ and $h = (b - a)N^{-1}$ for some real $a < 0 < b$ and large integer N. Pdf values in each x_l are given by inverse Fourier transformation of the characteristic function

$$f_X(x_l) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iux_l} \cdot \Phi_X(u)\, du, \;\; l = 0, 1, \ldots, N.$$

If $\Phi_X(u)$ decays fast enough this integral can be truncated, approximated by a finite sum and conveniently computed via Fast Fourier Transform. In Appendix A.2 the modified version of Mittnik et al. (1999) FFT algorithm is presented that may be conveniently utilized to obtain the values of $f_X(x_l)$. This version evaluates pdf over asymmetric interval $[a, b]$ while the original algorithm requires $a := -b$.

After a numerical method to invert Fourier transform is selected we may define an (approximate) Numerical Maximum Likelihood estimator as

$$(\hat{\alpha}, \hat{\beta}, \hat{\delta}, \hat{\mu}, \hat{\theta}) = \operatorname{argmax} \prod_{k=1}^{n} \tilde{f}(x_k; \alpha, \beta, \delta, \mu, \theta), \tag{3.12}$$

where $\tilde{f}(x_k; \alpha, \beta, \delta, \mu, \theta)$ stands for a cubic spline interpolation described above. The outputs of Proposition 3.4.1, obtained at modest computational cost, may be set as a starting point for numerical maximization of likelihood.

The fourth and most computationally intensive approach is to use Empirical Characteristic Function (ECF) estimator described in Yu (2004). This approach is essentially a special case of GMM with continuum of moment conditions. It also relies on numerical minimization of (weighted) distance, but now the objects proximity of which is desirable is characteristic function $\Phi_X(u)$ and its sample counterpart.

Define sample characteristic function as

$$\widehat{\Phi}(u) = \frac{1}{n}\sum_{k=1}^{n} e^{iux_k},$$

then ECF estimator takes the form of

$$(\hat{\alpha}, \hat{\beta}, \hat{\delta}, \hat{\mu}, \hat{\theta}) = \operatorname{argmin} \int_{-\infty}^{+\infty} \left|\widehat{\Phi}(u) - \Phi_X(u; \alpha, \beta, \delta, \mu, \theta)\right|^2 g(u)\, du, \tag{3.13}$$

where $|z| \equiv \sqrt{z\bar{z}}$ for $z \in \mathbb{C}$ and $g(u)$ is a preselected weighting function.

The aim of all four computationally intensive methods is to obtain an ensemble of five estimates. A convenient way to decrease the number of estimated parameters (and the computation time required) is to fit the distribution to standardized data. As noted by Scherer et al. (2010), standardization of the data may be used to control the errors of numerical approximations. Assume m and s stand for, respectively, sample mean and sample standard deviation of $x_1, \ldots, x_n$, define $Y \sim s^{-1}(X - m)$. On the sample of $y_k = s^{-1}(x_k - m)$ for $k = 1, \ldots, n$ estimate (via any preferred method) parameters of $TS_\alpha(\beta, \delta_{\text{std}}(\alpha, \theta), 0, \theta)$ where

$$\delta_{\text{std}}(\alpha, \theta) = \begin{cases} \left(\frac{\cos \pi\alpha/2}{\alpha(1-\alpha)}\right)^{1/\alpha} \theta^{\,2/\alpha - 1} & \alpha \neq 1, \\ \frac{\pi}{2}\theta^* & \alpha = 1, \end{cases} \tag{3.14}$$

The constraint above implies that $\mathbb{V}arY \equiv 1$. As $X \sim sY + m$, random variable X is a linear transformation of Y. By Property 3.2.7 the distribution of X is $TS_\alpha(\beta, s \cdot \delta_{\text{std}}(\alpha, \theta), m, \theta/s)$. Thus the estimates $(\hat{\alpha}, \hat{\beta}, \delta_{\text{std}}(\hat{\alpha}, \hat{\theta}), 0, \hat{\theta})$ obtained for the standardized data are consistent with $(\hat{\alpha}, \hat{\beta}, s \cdot \delta_{\text{std}}(\hat{\alpha}, \hat{\theta}), m, \hat{\theta}/s)$ being appropriate for the non–standardized data.

The last issue treated in this section is the choice of numerical minimization routine. MATLAB offers two different solvers that could deal with this type of tasks – *fmincon* in case of constrained minimization problems and *fminunc* for non–constrained problems. As parameter space of TS distribution is constrained – $\alpha \in (0, 1)$, $\beta \in [-1, 1]$, $\delta, \theta > 0$ – *fmincon* routine seems to be appropriate. However, when this procedure hits parameter boundary it terminates, instead of going one step back and changing the search direction. This feature makes *fmincon* routine of limited use in case of data sets where parameter identification becomes a problem due to numerical reasons. Therefore, due to constrained solver implementation issues, it is often convenient to perform numerical minimization via an unconstrained algorithm. Then

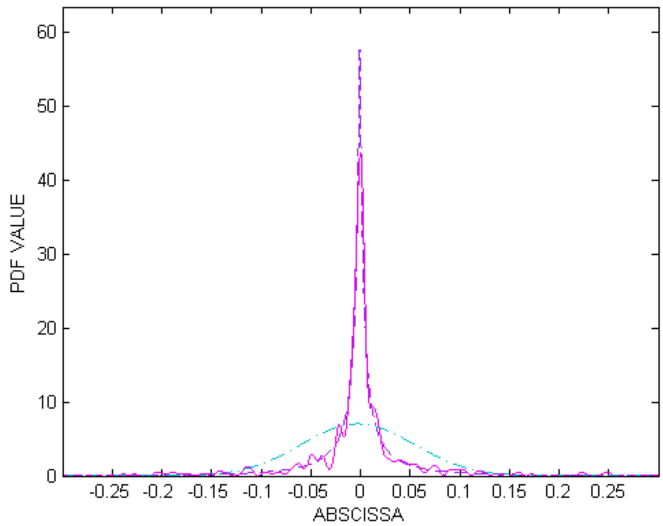


Figure 3.5: Spline–smoothed, normalised sample histogram of 1 week EURIBOR (ACT/365) offered rate (magenta solid line) versus fitted distributions: Gaussian (cyan dash–dotted line) and TS (violet dashed line). Numerical (approximate) Maximum Likelihood parameter estimates obtained for TS were: $\hat{\alpha} = 0.478$, $\hat{\beta} = -0.009$, $\hat{\delta} = 0.018$, $\hat{\mu} = -0.001$ and $\hat{\theta} = 5.704$.

what is passed to the solver is an unconstrained real vector that needs to be converted to an appropriate parameter range in the body of the minimized function. What is required to follow this route is an invertible mapping that transforms unconstrained real vector into the desired parameter space. A mapping that fulfils all this requirements is given below. Assume F is defined as

$$F(y_1, y_2, y_3, y_4, y_5) = \Big(\frac{2}{\pi}\arctan dy_1 + 1, \frac{2}{\pi}\arctan dy_2, e^{y_3/d}, y_4/d, e^{y_5/d}\Big) \tag{3.15}$$

for arbitrary real $d > 0$. Then mapping $F : \mathbb{R}^5 \to (0, 2) \times [-1, 1] \times \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}_+$ is invertible and

$$F^{-1}(x_1, x_2, x_3, x_4, x_5) = \Big(\frac{1}{d}\tan\frac{\pi(x_1 - 1)}{2}, \frac{1}{d}\tan\frac{\pi x_2}{2}, d\ln x_3, dx_4, d\ln x_5\Big). \tag{3.16}$$

Constant d above is a scaling factor which either squeezes the chart of the function to avoid identification problems and speed up convergence, or stretches it to increase accuracy of the estimates.

To assess the quality of the proposed estimation procedures TS distribution was fitted to the increments of one–week EURIBOR (EURo Inter–Bank Offered Rate) offered rate quoted from 30–th December 1998 to 30–th December 2011. While in an actual model of the inter–bank offer rate the parameters of this density would have to be estimated on model residuals, this exercise has a more modest aim of verifying the capability of TS density to depict the observed data. Thus, although the recorded values are most likely not independent and identically distributed, the distribution was fitted directly to the raw observations.

In order to obtain the estimates of parameters a feasible starting point for a numerical solver was first found by Proposition 3.4.1. Next numerical maximization of likelihood approximated with cubic spline was performed by an inverse FFT algorithm from Appendix A.2. The spline smoothed sample histogram and the resulting TS pdf are depicted in Figure 3.5. This picture may be compared with the chart on p. x. in Rachev (2003), which represents the pdf of α–stable distribution fitted to the increments of the same financial variable. The fitted value of characteristic exponent is equal to 0.478. If the data were quoted at even higher frequency, this estimate would typically decrease.

3.5 Random numbers generation

This section is dedicated to generating random draws from a tempered stable distribution. It contains a new random number generation algorithm that is easy to implement and valid for the entire parameter range. This procedure relies on the novel result, presented below as Proposition 3.5.1. Next, the sixth section outlines the two alternative approaches applicable in the parameter range plausible for macroeconomic data. The following two sections extend results presented in Jelonek (2012).

An important property of tempered stable distributions utilized below is that every TS random variable might be expressed as weighted average of two independent TS random variates with $\beta = 1$. The similar result for α–stable distributions may be found in the previous chapter as Property 2.4.6. This result is attributed to Zolotarev, but no direct reference has been traced. The following representation stems from Definition 3.2.1 and is employed further on to generate random numbers from TS distribution.

Proposition 3.5.1 (Mixture representation)**.** *Let Y^+, Y^- be independent, set $Y^\pm \sim TS_\alpha(1,1,0,\theta^\pm)$, set $V^\pm = \delta(1\pm\beta)^{1/\alpha}2^{-1/\alpha}$, $\theta^\pm = \theta V^\pm$, then $X = V^+Y^+ - V^-Y^- + \mu \sim TS_\alpha(\beta,\delta,\mu,\theta)$.*

This result provides foundation for Algorithm 2. It may be extended to CGMY by altering $\theta^\pm$.

Mixture representation formulated above takes a particularly convenient form due to the choice of parameterisation. If the term μ_X is omitted in Definition 3.2.1, then this equation needs to be corrected for $\alpha = 1$. The exact form of this correction depends on β.

The issue of random number generation for TS distribution has not been solved yet in a satisfactory way. Methods that are fast and easy to implement are only available for certain parameter values. Four algorithms – rejection by Brix (1999), generalized Kanter method by Devroye (2009), Laplace transform inversion by Ridout (2009) and approximate exponential rejection by Baeumer and Meerschaert (2010) – where dedicated to generating random draws from exponentially tempered stable distributions. However, the first is valid only for $\alpha < 1$, the latter requires that $\beta = 1$ while the remaining two are applicable if both conditions hold. In the general case only generic methods – shot–noise representation of Cohen and Rosiński (2007), compound Poisson approximation algorithms proposed in Kawai and Masuda (2011) or rejection–squeeze technique by Devroye (1981) – remain viable. The expectation of shot–noise infinite series representation matches the expectation of tempered stable random variable only in the limit. Hence every finite truncation of shot–noise series is by construction biased. As it was designed to generate an entire trajectory of stochastic process, shot–noise is extremely inefficient method to generate random numbers. Compound Poisson approximation algorithms are by construction inexact, they also require prior generation of multiple auxiliary random variates as a prerequisite to obtain a single pseudo–random draw. As it is reasonable to expect the first two approaches would be either imprecise or slow, rejection–squeeze algorithm remains the only promising option. No results for any of these methods have been reported for $\alpha \geq 1$ and $|\beta| \neq 1$.

Intuition suggests that TS distributions, particularly relevant in modelling macroeconomic data, display moderate departure from Gaussianity and mild skewness, which translates to $\alpha \geq 1$ and $|\beta| \neq 1$. However, as Palmer et al. (2008) and Kawai and Masuda (2011) in their numerical experiment consider only $\beta = 1$ (the former also assumes $\alpha < 1$), there is no literature treating this case. The aim of this section is to bridge this gap.

The problem investigated further is formulation of random number generator valid for all admissible values of α and β. The proposed method is easy to implement and much faster than the alternative approach of Devroye (1981) for moderately tempered distributions. The new algorithm relies on mixture representation of TS random variables (given as Proposition 3.5.1) that is parallel to decomposition property of α–stable random variates (Property 2.4.6 in chapter 2). It is valid for all parameter values and may be written in just one line of code, provided that random numbers generation for TS distribution with $\beta, \delta = 1$ is readily available. Hence it is particularly easy to implement.

In order to utilize the mixture representation a method to generate $Y^{\pm}$ from $TS_\alpha(1,1,0,\theta^{\pm})$ distribution is first required. Although a number of different algorithms might be used to generate TS random variates $Y^{\pm}$ endowed with $\beta = 1$, solely the procedure proposed by Baeumer and Meerschaert (2010) will be utilized. Out of all the methods investigated by Kawai and Masuda (2011) this algorithm performed best in terms of accuracy and computation time. Assume $S_\alpha(1,1,0)$ stands for α–stable distribution with unit skewness β and scale δ and naught location μ, defined as in Samorodnitsky and Taqqu (2000). Equate c to sufficiently low percentile of this distribution. Baeumer and Meerschaert algorithm is the following.

Algorithm 1 (Baeumer & Meerschaert, 2010)
Step 0. Determine constant c.
Step 1. Generate $U \sim U(0,1)$, $V \sim S_\alpha(1,1,0)$.
Step 2. If $U \leq e^{-\theta(V+c)}$, return $Y = V - \alpha\theta^{\alpha-1}/\cos\frac{\pi\alpha}{2}$ for $\alpha \neq 1$ or $Y = V + 2(\ln\theta + 1)/\pi$ for $\alpha = 1$, otherwise go to Step 1.
Algorithm 1 returns pseudo–random number Y drawn from $TS_\alpha(1,1,0,\theta)$.

Note that the scope of this algorithm is limited as it is viable only if $\beta = 1$.

If $\alpha \geq 1$ constant c in Algorithm 1 depicts truncation threshold of α–stable distribution supported on the entire real line. Hence the resulting procedure is approximate. As demonstrated in Brix (1999), for $\alpha < 1$ and $c = 0$ this rejection becomes exact. Random draws from $S_\alpha(1,1,0)$ may be generated with *rdnsta* procedure[1] by McCulloch, based on Chambers et al. (1976).

The following procedure stems directly from Proposition 3.5.1.

Algorithm 2 (Mixture representation)
Step 0. Set $V^{\pm} = \delta(1 \pm \beta)^{1/\alpha}\, 2^{-1/\alpha}$, $\theta^{\pm} = \theta V^{\pm}$.
Step 1. Generate independent $Y^+ \sim TS_\alpha(1,1,0,\theta^+)$, $Y^- \sim TS_\alpha(1,1,0,\theta^-)$.
Step 2. Return $X = V^+Y^+ - V^-Y^- + \mu$.
Algorithm 2 returns pseudo–random number X obtained for $TS_\alpha(\beta,\delta,\mu,\theta)$.

Unlike most available methods Algorithm 2 is viable for all parameter values, including $\alpha \geq 1$ and $|\beta| \neq 1$. Note that its output is endowed with arbitrary values of both β and δ.

[1] Available at: econ.ohio-state.edu/jhm/programs/RNDSSTA.

The remaining part of this section contains the description of two additional randomization algorithms valid for the entire parameter range of TS distribution. These methods will be utilized further on to draw a comparison with Algorithm 2 that builds on Proposition 3.5.1. Algorithm 3 is a benchmark where pseudo–random draws are generated by an inverse of piecewise linear cdf approximation obtained via Fast Fourier Transform (FFT). Algorithm 4 relies on rejection-squeeze technique proposed by Devroye (1981). The results obtained for the algorithm proposed in the previous section are compared with the outcomes of these two procedures.

Perhaps the easiest way to obtain alternative (benchmark) random draws from TS distribution is to invert approximated cdf. Let $F(x)$ be the cdf of $TS_\alpha(\beta, \delta, 0, \theta)$ with essential support $[a, b]$ and pdf $f(x)$. Assume N is large integer, for $k = 0, \ldots, N-1$ denote $x_k = a + hk$ with $h = (b-a)/N$. The numerical inversion of linearly approximated cdf may be implemented as follows.

Algorithm 3 (Cdf inversion)
Step 0. Evaluate $F(x_0), \ldots, F(x_{N-1})$.
Step 1. Generate $U \sim U(0,1)$. Find n such that $F(x_n) \leq U < F(x_{n+1})$.
Step 2. Return $X = x_n + h(U - F(x_n))(F(x_{n+1}) - F(x_n))^{-1} + \mu$.
Algorithm 3 returns pseudo–random number X obtained for $TS_\alpha(\beta, \delta, \mu, \theta)$.

In Algorithm 3 cdf is approximated by a piecewise linear function.

To implement Algorithm 3 pointwise values of $F(x_k)$ need to be first evaluated. If pdf proxies are initially obtained by FFT, the sought quantities may be found from $F(x_{k+1}) = F(x_k) + hf(x_k)$ under boundary condition $F(x_0) = 0$. In Appendix A.2 the modified version of Mittnik et al. (1999) FFT algorithm is presented that may be conveniently utilized to obtain the values of $f(x_k)$. This version evaluates pdf over asymmetric interval $[a, b]$ while the original algorithm requires $a = -b$.

The final algorithm considered relies on the rejection–squeeze technique. Given

$$d_1 = \frac{1}{2\pi} \int_{\mathbb{R}} |\Phi_X(u)| \, du, \ d_2 = \frac{1}{2\pi} \int_{\mathbb{R}} |\Phi_X^{(2)}(u)| \, du$$

the result formulated by Devroye (1981) states that pdf $f(x)$ of $TS_\alpha(\beta, \delta, 0, \theta)$ fulfils

$$\forall \ x \in \mathbb{R} : f(x) \leq \min \{d_1, d_2/x^2\}. \tag{3.17}$$

This inequality was originally utilized to derive the following rejection–squeeze algorithm.

Algorithm 4 (Devroye, 1981)
Step 0. Evaluate d_1 and d_2.
Step 1. Generate independent $U \sim U(0,1)$, $V, W \sim U(-1,1)$.
Set $Y = \sqrt{d_2/d_1} \cdot V/W$. If $|V| < |W|$, then go to Step 3.
Step 2. If $U < f(Y)Y^2/d_2$, then return $X = Y + \mu$. Otherwise, go to Step 1.
Step 3. If $U < f(Y)/d_1$, then return $X = Y + \mu$. Otherwise, go to Step 1.
Algorithm 4 returns pseudo–random number X drawn from $TS_\alpha(\beta, \delta, \mu, \theta)$.

The expected number of times Step 1 is executed to generate one random number is $4\sqrt{d_1 d_2}$.

In order to run Algorithm 4 some preliminary work is required. First of all, the formula for second order derivative of $\Phi_X(u)$ has to be found. Define $C_{\alpha,\delta} = \alpha\delta^\alpha(\cos\frac{\pi\alpha}{2})^{-1}/2$, then for $\alpha \neq 1$ this derivative is

$$\Phi_X^{(2)}(u) = -C_{\alpha,\delta}(C_{\alpha,\delta}[(1+\beta)(\theta - iu)^{\alpha-1} - (1-\beta)(\theta + iu)^{\alpha-1} - 2\beta\theta^{\alpha-1}]^2 +$$

$$+(1-\alpha)[(1+\beta)(\theta - iu)^{\alpha-2} + (1-\beta)(\theta + iu)^{\alpha-2}]) \cdot \Phi_X(u),$$

while for $\alpha = 1$ it holds that

$$\Phi_X^{(2)}(u) = -\frac{\delta}{\pi}(\frac{\delta}{\pi}[(1+\beta)\ln(\theta - iu) - (1-\beta)\ln(\theta + iu) - 2\beta\ln\theta]^2 + 2\frac{\theta + i\beta u}{\theta^2 + u^2}) \cdot \Phi_X(u).$$

Secondly, the integrals have to be determined. As analytic results for d_1 and d_2 seem difficult to compute it is probably necessary to approximate both quantities numerically. Finally, the pdf of TS distribution needs to be evaluated in arbitrary points of its domain. If the pdf is first approximated in the points equally spaced over its essential support (as in Algorithm 3), this last step may be done via cubic spline interpolation performed on each subinterval.

Note that in case of $\alpha \geq 1$ there are no known formulas for pdf of TS distribution. Furthermore, it is no longer possible to follow the route of Kawai and Masuda (2011) and elicit the pdf from the relation, binding densities of TS distribution and α–stable distribution being tempered. The reason is that for $|\beta| \neq 1$ this identity does not hold. Therefore in Algorithm 3 and 4 Fourier inversion of characteristic functions is utilized. As Devroye (1981) procedure is exact and the underlying pdf approximation involves (in addition) cubic splines, it is reasonable to conjecture that Algorithm 4 would be more precise than Algorithm 3.

3.6 Evaluation of random number generation algorithms

The purpose of section seven is to evaluate the quality of random numbers, generated by Algorithm 2, 3 and 4 from section six.

In the numerical exercise conducted below the main criterion used to evaluate the competing procedures was precision of mean sample moments. There are three reasons for it. First, in order to compute minimum distance measure, such as Kolmogorov–Smirnov metric[1], theoretical results for cdfs are required. As formulas for cdf of TS are only known in the form of integral (see Property 3.2.4), it would have to be approximated numerically. This approach favours randomization methods that rely on the same approximation. Next, formulas for cumulants and moments of TS distribution constitute the only available theoretic result that can be used to assess the quality of random numbers generator. Any other approach would necessarily resort to numerics, the same procedures that were once used to produce the output would have to be rerun to establish validity of these results. This could obfuscate potential errors if any of these routines were faulty. Finally, sample moments are easier to interpret and allow for more intuitive assessment of the results thus obtained.

The objective of the experiment presented in this section is to confirm the general validity of the proposed randomization procedures and to compare their quality. Thus, in contrast to section 3.7, we focus on samples that would be typically regarded as *large*. Another, more extensive expercise would be required in order to assess the small sample performance of all the investigated routines. While such experiment lies outside of the scope of this thesis, its results could possibly vary with the size of simulated samples.

In order to compare the quality of random numbers generated with Algorithm 2 and by the remaining two approaches the following exercise was undertaken. Each procedure was run with 64 different sets of parameters, defined as all possible combinations of $\alpha \in \{1.2, 1.4, 1.6, 1.8\}$, $\beta \in \{0, 0.25, 0.5, 0.75\}$ and $\theta \in \{0.25, 0.5, 0.75, 1\}$, to obtain 100 samples of 10^6 pseudo–random numbers. This parameter choice corresponds to moderately skewed TS distributions, relatively close to the underlying α–stable densities. Selected range of stability index α is typical for low frequency (i.e monthly, quarterly) data. All distributions were standardised (with $\mu = 0$ and δ implying unit variance) and thus parametrized with just α, β and θ. Both mean computation time and first five mean sample moments about the origin were recorded

[1] Extensive review of possible approaches may be found in Basu et al. (2011).

in each case. The results are depicted in Appendix C, Tables 1–16 and might be readily compared with values, obtained with the theoretical formulas from section 3.3. Figures in brackets denote unbiased estimates of standard deviation. As computation time does not vary much across replications the corresponding estimates of standard deviation were not reported. Emphasized numbers indicate either the smallest mean execution time (in seconds), or the mean sample moment most similar to the relevant theoretic value.

Implementation details were as follows. In all FFT procedures $N = 2^{13}$ was utilized as powers of 2 are computationally most efficient. In Algorithm 2 constant c was set to the bottom 0.1 percentile of the sample of 10^6 random draws coming from the relevant α–stable distribution. Results of all the calculations presented in this work were performed in MATLAB©7.11.0 (R2010b) on a PC with Inter®Core™i7 2630QM CPU (2.00 GHz, 8.0 GB RAM) under 64–bit Windows®7 Home Premium operating system. The code is available upon request.

In the undertaken exercise mixture representation algorithm produced most precise sample moments in 48.75% (156 out of 320) times. Rejection–squeeze method delivered most accurate moments in 37.19% of instances (119 cases out of 320) while cdf inversion ranked first for the remaining 14.06% (45 out of 320) of cases. Therefore for the chosen parameter values Algorithm 2 is the best in terms of quality of pseudo–random numbers generated, it is also markedly better than Algorithm 4. Algorithms 3 is clearly the least precise. If we consider only $\alpha \in \{1.2, 1.4\}$, rejection–squeeze technique is more precise then the mixture representation algorithm, while for $\alpha \in \{1.6, 1.8\}$ the superiority of the latter is striking. Hence Algorithm 2 is particularly suited for generating random draws from TS distribution close to Gaussian. Devroye procedure always requires most computation time. For the investigated parameter selection algorithms 2 and 3 need, respectively, only 8.49%–57.27% and 5.17%–14.76% of computation time required by Algorithm 4. While for $\theta = 1$ cdf inversion is the fastest out of the three algorithms, in case of $\theta = 0.25$ it performs marginally worse than the mixture representation in terms of execution speed. This difference would become more evident for smaller θ, when Baeumer and Meerschaert (2010) algorithm accepts candidate draws more often. Note that by Property 3.2.7 rescaling the data generated from TS distribution by a sufficiently large factor guarantees that the resulting parameter θ is small. All procedures are less reliable in capturing higher order moments. In case of Algorithms 3 and 4 the reason is that in order to perform numerical approximation of pdf its support needs to be constrained. Hence pseudo–random numbers above (or below) certain values will not be generated. In the case of mixture representation the culprit is auxiliary Algorithm 1 where the left tail of the distribution is trimmed. Therefore extreme values are returned less often.

Out of the procedures whose precision was investigated in this section, Algorithm 3 has not been used yet to generate random numbers from TS distribution. Results for random number generation from Algorithm 4 in case of TS density with $\alpha \geq 1$ and $|\beta| \neq 1$ have not been reported in the literature.

3.7 Application to joint currency crises

The problem investigated in this section is evaluation of the probability of joint currency crises. Currency crisis is regarded as *joint* (see e.g. Hartmann et al. (2010)) if it involves relative depreciation or appreciation of all the investigated currencies that exceeds certain threshold. What is especially interesting are the crisis of a magnitude larger than the one historically observed, which are untractable from historical data. They may both be treated as indicators of financial stability and approximate the scale of yet unseen macroeconomic events which nonetheless may occur.

Currency crises are devastating to economy (see e.g. Desai (2000) for a brief account of the Russian crisis). They are generated by complex events dependent on a large number of factors. These factors are often either not properly measured, or remain unknown until foreign exchange–rates crash. It is reasonable to assume they remain constant within the sample and are not likely to change in the immediate future. Thus the quoted exchange–rate returns may be treated as random and coming from the same probability distribution. While in this exercise we are agnostic about the model which generates the data, the results presented in this section could be further enhanced if we could quantify the factors that affect short and mid term dynamics of exchange–rates. In such instant the subject of modelling would not be the rates themselves, but rather the residuals of a given foreign exchange model.

In this section it is further assumed that the uncertainty associated with foreign exchange–rates may be represented with (univariate or multivariate) tempered stable distributions. Probabilities of large yet unseen joint exchange–rate movements may be then evaluated in a two stage procedure. First, the corresponding theoretic densities have to be fitted to empirical data. Next, probabilities of the event of interest may be approximated with the relative frequencies, evaluated on a sample of pseudo–random draws generated from the fitted distributions.

The data analysed in this example consists of monthly and daily returns of Russian ruble to euro and British pound to euro exchange–rates. While monthly data is recorded from January 1999 till October 2012, the daily quotations span the period from 13th January 1999 till 1st November 2012. What makes this pair of currencies interesting is that the both countries have markedly different characteristics. While Russia is a developing country that exports commodities, the economy of United Kingdom is classified as developed and its main exports are services. Hence both series of data display different statistical properties. In particular, the distribution of Russian ruble is much more skewed and volatile. This observation remains in line with the results of Ibragimov et al. (2013) who demonstrate that *"[...]while moments of order* $p \in (2.6, 2.8)$ *are finite for most of the developed country exchange–rates, they may be (or are) infinite for most of the emerging country exchange rates"*. Both monthly exchange–rates may be treated as random. They also represent aggregate quantities which may be useful in macroeconomic analysis and are quoted at low frequency.

There are two feasible probabilistic approaches to depict the risk, underlying the selected pair of currencies. The first option is to rely on Definition 3.2.1 and represent both distributions as separate univariate densities. The second possibility is to estimate a joint distribution, using Definition 3.2.2. The latter has a number of significant advantages. First, it allows to take into account the partial dependence of both variates. Next, it gives us the opportunity to model events that involve common movement of both exchange–rates, such as joint currency crisis, investigated by Hartmann et al. (2010). Third, joint estimation ameliorates the problem of short sample sizes in Numerical Maximum Likelihood. Hence it is possible to obtain more precise parameter estimates after shorter computation time. Finally, as marginals of multivariate distribution may be employed in all the situations where univariate densities are required, joint density has broader applicability. While results of univariate estimation are also reported, this section is focused on a more general multivariate case.

When estimating parameters of a distribution, it is often convenient to standardize the sample. In univariate case standardization involves two steps. First, sample mean is subtracted from all the observations. Next, the remaining terms are divided by an unbiased estimator of sample standard deviation. The prerequisite for this second step is the existence of the variance of the underlying theoretic distribution. Once both sample mean and standard deviation are obtained, we may fit the selected distribution to the standardized sample. As standardization is a linear operation, its inverse is also linear. Hence the parameters of the original density may by obtained by applying this linear operation to random variable, parameters of which were estimated from the standardized data. In case of univariate tempered stable distribution the outcome of such transformation is described by Proposition 3.2.7.

A similar approach is feasible in the multivariate case. Let $\boldsymbol{y}$ stand for $(n \times m)$ matrix, where series of the data are placed in consecutive columns. Assume $\hat{\boldsymbol{\mu}}$ is $(m \times 1)$ vector of sample means while $(m \times m)$ matrix $\widehat{\Sigma}$ represents the unbiased estimator of sample covariance. If $\widehat{\Sigma}$ is positive–definite, its Cholesky decomposition yields a unique lower–triangular matrix $\widehat{A}$ such that $\widehat{A}\widehat{A}^T = \widehat{\Sigma}$. The operation which transforms $\widehat{\Sigma}$ into $\widehat{A}$ is a matrix counterpart of a square root, thus $\widehat{A}$ is multivariate equivalent of sample standard deviation. Let $1_{(n\times 1)}$ stand for $(n \times 1)$ vector with unit entries. In multivariate case standardization involves two steps. First, the consecutive entries of $\hat{\boldsymbol{\mu}}$ are subtracted column by column from $\boldsymbol{y}$, which yields $\boldsymbol{y} - 1_{(n\times 1)} \times \hat{\boldsymbol{\mu}}^T$. Next, the remainder is divided from the right hand side by $\widehat{A}^T$, which produces $(\boldsymbol{y} - 1_{(n\times 1)} \times \hat{\boldsymbol{\mu}}^T)(\widehat{A}^T)^{-1}$. This operation is feasible if covariances exist. It will be further termed a *standardization* of multivariate data.

In our example the standardized data represents either monthly or daily quotations of returns from Russian ruble to euro and British pound to euro exchange–rates. In the data matrix $\boldsymbol{y}$ the first and second column constitute, respectively, RUB/EUR and GBP/EUR exchange–rate returns. While the meaning of the raw data is obvious, the interpretation of the marginals of the standardized series is no longer clear. Cholesky matrix $\hat{A}$ is lower–triangular by construction. Its transposition is upper–triangular, and so does the inverse of this transposition. The first column of the standardized bivariate data represents centred RUB/EUR returns scaled by a constant, the second column constitutes a weighted average of centred RUB/EUR and GBP /EUR returns. Assume that both exchange–rates constitute realizations of random vector $\boldsymbol{Y} \sim mTS_\alpha(\boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{\mu}, \theta, A)$, where all the parameters are set as in Definition 3.2.2. Then the marginal distribution of the first column of the standardized data is univariate tempered stable by Property 3.2.7. The marginal distribution of the second column is univariate tempered stable only if the second column of $(\widehat{A}^T)^{-1}$ contains only ones and thus Property 3.2.8 applies. Therefore the interpretation of the marginals of standardized data depends on the sequence in which different data series are inserted into $\boldsymbol{y}$. The only exception are the fitted densities which, once the data is standardized, have no degrees of freedom left. A prominent example of such a distribution is multivariate Gaussian density. Standardization is geometrically equivalent to change of coordinates. Next paragraph reports preliminary estimation results.

Table 3.7 contains parameter estimates of univariate TS densities, fitted independently to the standardized bivariate data. These estimates were obtained by Numerical Maximum Likelihood for monthly data and via much faster Numerical Method

Data	Frequency	Method	Parameter estimates $\hat{\alpha}$	$\hat{\beta}$	$\hat{\delta}$	$\hat{\mu}$	$\hat{\theta}$
RUB/EUR	Monthly	NML	1.5210	0.6706	0.5704	0.0938	0.0998
	Daily	NMM	1.6317	0.5039	0.6845	0.0194	0.3282
GBP/EUR	Monthly	NML	1.5832	0.4875	0.5571	0.1465	0.0000
	Daily	NMM	1.6293	0.3257	0.6791	0.0100	0.3177

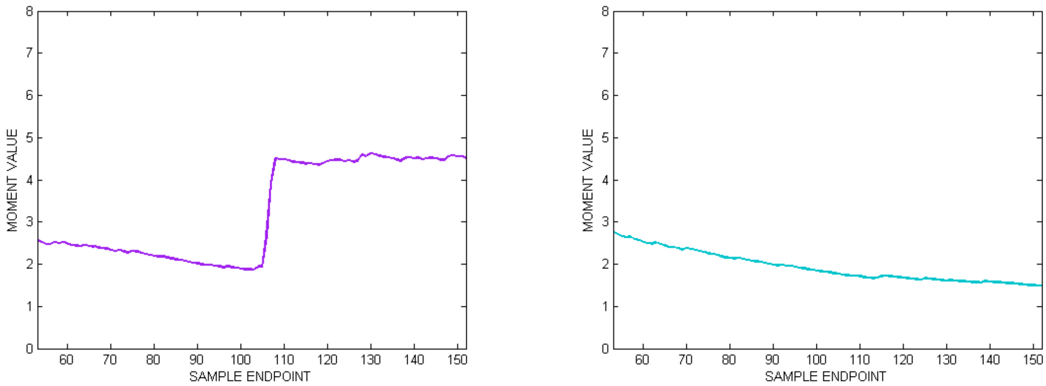
Table 3.1: Parameter estimates of univariate TS densities, fitted to the standardized bivariate data.

of Moments in the case of the daily quotations. For the two exchange–rates and the both investigated data frequencies the fitted values of skewness parameter β were markedly different from zero while the estimates of stability index α and tempering parameter θ were, respectively, larger than one and smaller than one. This range of values is typical for macroeconomic densities. The value of θ estimated on GBP/EUR exchange–rates transformed via bivariate standardization is approximately 10^{-6}, which may suggest the corresponding marginal is α–stable rather than tempered stable. The estimates of stability indexes for the both monthly exchange–rates amount are relatively close. Hence it might be justified to use one common stability index for both data series. The estimates obtained for daily frequency are not typical for financial data, in case of which both $\hat{\alpha}$ and $\hat{\beta}$ usually decline as data sampling frequency increases. Furthermore, financial series are often characterized by $\hat{\theta} > 1$.

The estimates from Table 3.7 may suggest that returns of British pound to euro exchange–rate after bivariate standardization has an α–stable distribution ($\hat{\theta} \approx 0$). Figure 3.6 is intended to verify this claim. It depicts second sample moments about the origin of the second column of our standardized, bivariate data. The moments are evaluated on the samples of increasing sizes, encompassing first n available observations. Thus they depend on the index of the last available observation and for $n \in \{53, \cdots, 153\}$ are given by

$$\hat{\mu}_2'(n) = \frac{1}{n} \sum_{k=1}^{n} x_{k,2}^2.$$

The chart may be compared with Fig. 2 from Mandelbrot (1963a). In his example the approximations of second moment about the origin were both erratic and explosive. This suggests that cotton prices he investigated may be best depicted with a density variance of which does not exist. In particular, this distribution



(a) Real standardized data. (b) Synthetic data.

Figure 3.6: Second sample moment computed for: (a) returns from GBP/EUR monthly exchange–rates transformed via bivariate standardization, (b) synthetic data, generated from Gaussian distribution with the same mean and variance. Sample size increases from 53 to 152 observations.

could be α–stable (i.e. fat–tailed). In the case of our distribution the trajectories obtained are also erratic, but – contrary to Mandelbrot's example – bounded and thus non–explosive. This dynamics implies that while the underlying distribution is certainly not Gaussian (trajectory displays jumps), it does have variance. Hence the density obtained after standardization is best depicted with either heavy–tailed, or semi heavy–tailed distribution, just as tempered stable. From the point of view of modelling macroeconomic uncertainty, this is a good news. While tempered stable distributions have all the moments finite (Property 3.3), they never fit to the data worse than α–stable densities (which is a direct consequence of Definition 3.2.1).

Multivariate tempered stable distribution is formalized in a way parallel to the definition of *elliptical*[1] distributions with finite covariance matrix Σ and a well defined density. Let $\boldsymbol{X}$ be a random vector, marginals distributions of which are independent and endowed with unit variance. Assume A is a lower–triangular (and thus invertible) matrix, $\boldsymbol{\mu}$ is a real while $\boldsymbol{Y}$ is a random vector defined by the formula $\boldsymbol{Y} = \boldsymbol{\mu} + A\boldsymbol{X}$. If $\boldsymbol{X}$ has univariate tempered stable marginals (Definition 3.2.1) with unit variances, null means and common parameters α and θ, then the distribution of $\boldsymbol{Y}$ is multivariate tempered stable (Definition 3.2.2). By Property 3.2.5 and invertibility of A mTS distribution always has a density, by Property 3.2.1 its covariance matrix exists. If $\boldsymbol{X} \sim R\boldsymbol{X}$ for every matrix R such that $RR^T = I$ where I is the identity

[1] Definitions of elliptical and spherical distributions may be found in Cambanis et al. (1981) and Eaton (1986).

matrix (density of $\boldsymbol{X}$ is *spherical*), then the distributions of $\boldsymbol{Y}$ is elliptical. Note that in the latter case the marginals of $\boldsymbol{X}$ are, in particular, centred, symmetric and identical. Given the identifying assumptions made above the covariance matrices of both elliptical and tempered stable multivariate distributions fulfil $AA^T = \Sigma$. The fact that both classes are defined in a similar way makes comparison between in sample fit of mTS and any selected elliptical distribution (such as Gaussian or t) particularly convenient. It also simplifies evaluation of joint likelihood.

If matrix A is known, then random vector $A^{-1}(\boldsymbol{Y} - \boldsymbol{\mu}) = \boldsymbol{X}$ has in both cases independent marginals centred around zero. The operation on left hand side of this formula is equivalent to changing coordinates. In the new coordinates both series may be treated as realizations of independent random variates. If matrix A is not known, then its estimator may be elicited from the sample covariance matrix. If unbiased estimate $\widehat{\Sigma}$ of covariance matrix is positive–definite, Cholesky decomposition yields a unique[1] matrix $\hat{A}$ with strictly positive diagonal entries such that $\hat{A}\hat{A}^T = \widehat{\Sigma}$. Columns of $\hat{A}$ constitute a single coordinate system in which marginals of the joint distribution may be treated as independent. In this system the pdf of $\boldsymbol{X}$ is a product of its marginal distributions f_{X_j} for $j \in \{1, \cdots, m\}$. The joint pdf of $\boldsymbol{Y}$ by change of variables factorizes as

$$f_{\boldsymbol{Y}}(y_1, \cdots, y_m) = (\det A)^{-1} \prod_{j=1}^{m} f_{X_j}(x_j).$$

This observation implies that negative log–likelihood for the standardized sample

$$\boldsymbol{x} \equiv (x_{11}, x_{12}; \cdots; x_{n1}, x_{n2})$$

may be obtained as

$$\mathbb{L}(\boldsymbol{x}) = -n \cdot \ln|\det \hat{A}| - \sum_{k=1}^{n} \ln \tilde{f}_{X_1}(x_{k,1}) - \sum_{k=1}^{n} \ln \tilde{f}_{X_2}(x_{k,2})$$

where $\tilde{f}_{X_1}$ and $\tilde{f}_{X_2}$ represent the marginal densities fitted to the standardized data. This formula is further utilized in numerical likelihood maximization.

In this section the results obtained for mTS distribution are compared with the outputs produced for two benchmarks – multivariate Gaussian and elliptical t densities[2]. The latter is obtained when the marginals of spherical distribution that

[1] See: Trefethen and Bau (1997), p. 174.

[2] For the applications of elliptical distributions in economics and finance see Chamberlain (1983) and Owen and Rabinovitch (1983).

Frequency	Data	Method	Parameter estimates $\hat{\alpha}$	$\hat{\beta}$	$\hat{\delta}$	$\hat{\mu}$	$\hat{\theta}$
Monthly	RUB/EUR	CNML	1.3240	0.7055	0.6999	0.1068	0.4117
	GBP/EUR	CNML	1.3240	0.4109	0.6999	0.0682	0.4117
Daily	RUB/EUR	CNML	0.5187	0.0330	7.3791	0.0194	1.0177
	GBP/EUR	CNML	0.5187	0.0729	7.3791	0.0100	1.0177

Table 3.2: Joint parameter estimates of bivariate TS densities.

underlies elliptical distribution are independent and t distributed. The motivation behind the choice of benchmarks is simple. Multivariate Gaussian is the most seminal distribution in both financial and economic applications. Elliptical t distribution is one of the most popular multivariate distributions and probably the most tractable multivariate density that displays heavy tails. There also seems to be a consensus in the heavy–tails literature (see for instance section 2.3) that elliptical t density is the best candidate for a benchmark heavy–tailed distribution. Another interesting candidate for a heavy–tailed reference distribution, which is not considered here, is a multivariate Normal Inverse Gaussian (mNIG) density[1]. However, even if it was used as a benchmark, the qualitative results obtained would not change. The reason is that all three multivariate distributions – NIG, t and Gaussian – are elliptical. Before the framework presented in this text is applied to the data, it is useful to discuss how the characterization, provided in Definition 3.2.2, may be applied in practice.

The estimation procedure for all three distributions – multivariate Gaussian, elliptical t and mTS – proceeds in three stages. First, unbiased estimates of sample mean $\hat{\mu}$ and matrix $\hat{A}^T$ are calculated. Second, bivariate data is standardized. Third, the appropriate bivariate distribution is fitted to the standardized data. For Gaussian density this distribution has no degrees of freedom left (multivariate Gaussian with identity covariance matrix), so no further calculation is required. As t distribution with unit variance does not exist, a similar approach for elliptical t density is not feasible. In this case the estimate of a number of degrees of freedom $\hat{\nu}$ that best describes the joint sample distribution of transformed data is first obtained by Maximum Likelihood. This fitted value implies that $\mathbb{V}ar X_1 = \mathbb{V}ar X_2 = \hat{\nu}/(\hat{\nu}-2)$. In order to obtain elliptical t distribution with covariance matrix that matches sample covariances, the estimate of matrix A is rescaled to $\hat{A} := \hat{A}\sqrt{(\hat{\nu}-2)/\hat{\nu}}$. This simple operation concludes the estimation. In case of mTS density the two independent marginals that need to be fitted to the standardized data are $TS_\alpha(\beta_1, \delta_1, 0, \theta)$ and $TS_\alpha(\beta_2, \delta_2, 0, \theta)$.

[1]See e.g. Barndorff-Nielsen (1997) and Øigård et al. (2005).

By Definition 3.2.2 both are endowed with the same parameters α and θ. We also have by construction $\mathbb{V}arX_1 = \mathbb{V}arX_2 = 1$, so δ_1 and δ_2 are given by equation (3.14) and thus are equal. Hence both marginal densities differ only by the corresponding values of skewness parameters β_1 and β_2. This additional degree of freedom gives flexibility which is required to depict different shapes of the marginals of standardized distributions. It also signifies that mTS density is non–elliptical.

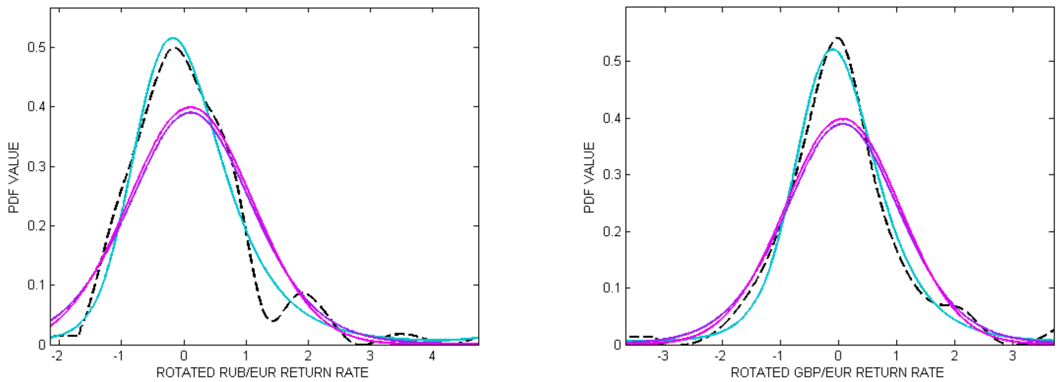
The results of a joint parameter estimation are as follows. Vectors of sample means and estimates of Cholesky matrices evaluated on monthly and daily quotations of RUB/EUR and GBP/EUR forex return rates are

$$\hat{\boldsymbol{\mu}}_m = (0.2636, 0.1950)^T, \qquad \hat{\boldsymbol{\mu}}_d = (0.0113, 0.0088)^T,$$

$$\hat{A}_m = \begin{pmatrix} 2.4678 & 0 \\ 0.7825 & 1.6347 \end{pmatrix}, \qquad \hat{A}_d = \begin{pmatrix} 0.5832 & 0 \\ 0.2154 & 0.4592, \end{pmatrix}.$$

Standardized monthly data is given by $(\boldsymbol{y}_m - 1_{[n\times 1]} \times \hat{\boldsymbol{\mu}}_m^T)(\hat{A}_m^T)^{-1}$, standardized daily data is $(\boldsymbol{y}_d - 1_{[n\times 1]} \times \hat{\boldsymbol{\mu}}_d^T)(\hat{A}_d^T)^{-1}$. The estimates of parameter ν for elliptical t density are fitted jointly and amount to, respectively, for 10 and 11 for the daily and monthly data. The sought parameters of multivariate TS distribution are obtained from the standardized sample via Constrained Numerical Maximum Likelihood (CNML). The utilized constraint implies that both vector of means and covariance matrix of the fitted distribution match their sample counterparts. The implementation of the estimation procedure relies on cubic spline interpolation and the FFT algorithm from Appendix A.2. The estimates from Proposition 3.4.1 may be used as the starting point for numerical solvers. All the estimated parameters are presented in Table 3.7.

It is worth pointing out that CNML is not equivalent to Numerical Maximum Likelihood. In particular, if we estimated the marginals of the standardized sample without constraints, the estimates thus obtained would imply neither nil mean, nor unit variance. Furthermore, the correlation rate implied by NML would not match the value, elicited on the sample. The CNML estimates not only take less time to obtain, but they are also much less susceptible to numerical problem, resulting from dimensionality of parameter space. As they exactly replicate some basic characteristics (means, variances, correlation rate) of the sample, these estimates are particularly useful in the applications where low order moments strongly affect the produced results. An example of such applications are risk management and portfolio analysis.



(a) RUB/EUR standardized return rate.

(b) GBP/EUR standardized return rate.

Figure 3.7: Spline–smoothed sample histogram (dashed black line) and pdfs fitted to the standardized bivariate returns from RUB/EUR and GBP/EUR monthly exchange–rates: tempered stable (cyan), elliptical t (magenta) and Gaussian (violet).

The results for the fitted multivariate TS distribution are as follows. Negative log–likelihood for mTS density fitted to monthly data amounts to 188.95 versus 211.54 and 218.35 obtained, respectively, for elliptical t distribution and multivariate Gaussian distribution. This substantial gain implies that, given sample size, each data point is (on average) 21.34% more likely to come from the estimated tempered stable then from the fitted Gaussian distribution and 16.02% more likely to come from the estimated TS then from the fitted t distribution. In order to compare the fitness of multivariate TS distribution against the better performing benchmark (elliptical t density) we may conduct a likelihood ratio test for non–nested models[1], proposed by Vuong (1989). In this case the sample distribution of his test statistic is standard normal. For the monthly data Vuong's statistic amounts to 2.16 when the number of parameters is penalized as in the Akaike Information Criterion (AIC). It is equal to 1.46 when we account for extra parameters using the more stringent Bayesian Information Criterion (BIC). We reject the null hypothesis that the both distributions are equivalent in favour of the hypothesis that mTS distribution is closer to the data generating density at either 1.5% (BIC) or 7.2% (BIC) significance level. Negative log–likelihood for mTS density fitted to daily data amounts to 13352.52 versus 13664 and 13877.87 obtained, respectively, for elliptical t distribution and multivariate Gaussian distribution. This moderate gain implies that, given sample size, each data point is (on average) 17.03% more likely to come from the estimated tempered

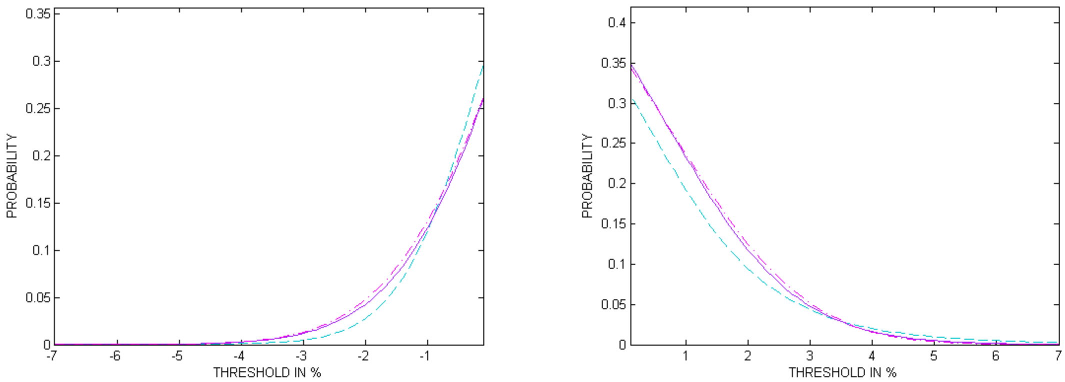
[1]All the three probability models considered here are non–nested in the sense of Vuong (1989), for more details see p. 317 therein.

stable than from the fitted Gaussian distribution and 9.77% more likely to come from the estimated tempered stable than from the fitted t distribution. The sample values of Vuong's statistic amount to 7.88 (AIC) and 7.57 (BIC), p–value of this test is nil up to machine precision. The null hypothesis is rejected with certainty, at every significance level.

Hence mTS distribution better depicts empirical densities for both monthly and daily data. These results still indicate that mTS is more capable to replicate the observed data after we take into account the number of parameters of each distribution. What is typical, the evidence in favour of mTS density is stronger for the data of higher frequency.

Figure 3.7 (a) and (b) depict the marginals of the three distributions – tempered stable, Gaussian and t – fitted to the monthly data after its bivariate standardization. The chart implies that the two marginal distributions are markedly different. Figure 3.7 (a) is strongly positively skewed while Figure 3.7 (b) appears to be almost symmetric. Multivariate tempered stable distribution is able to capture this clear diversity. The fitted multivariate t and Gaussian distributions, on the contrary, deliver two identical marginal densities that poorly befit the data. The reason is that both multivariate t and Gaussian densities are elliptical. In elliptical distributions the only admissible source of skewness is covariance. Once the data is standardized, all the fitted marginal densities are necessarily identical. Hence the both benchmark distributions have severe problems with capturing the joint distribution of the investigated series. In a standardized mTS distribution each marginal has still one remaining skewness parameter that can be tweaked to capture the shape of empirical density. Due to this property non–elliptical mTS can depict the joint historical data in a much more faithful way.

Once all the parameters of the mTS distribution have been estimated, it possible to use Algorithm 2 (derived from Proposition 3.5.1) to simulate the joint dynamics of both exchange–rates. These pseudo–random draws may be utilized in any Monte Carlo experiment, designed to deliver the answers for interesting macroeconomic questions. An example of such a task is the evaluation of probabilities of joint currency crisis. Figure 3.8 (a) depicts probabilities that both monthly returns on RUB/EUR and GBP/EUR exchange–rates fall below thresholds, indicated on horizontal axis. Figure 3.8 (b) presents probabilities that both monthly returns on RUB/EUR and GBP/EUR exchange–rates exceed thresholds, denoted on horizontal axis. As in case of mTS distribution the exact form of its multivariate cdf is not known, the only way to obtain this type of result leads via simulation. The probability of joint decline exceeding 7% in both currencies evaluated for the fitted mTS distribution amounts



(a) Joint decline.   (b) Joint upsurge.

Figure 3.8: Sample probabilities of joint: (a) crash or (b) upsurge in return rates obtained for mTS (cyan dashed line), elliptical t (violet dash–dotted line) and multivariate Gaussian (magenta solid line) fitted distributions.

to 5.3×10^{-5}. This figures is 2.12 times larger than the one obtained for elliptical t density and 13.15 times larger than the estimate for multivariate Gaussian distribution. The probability of upsurge in returns exceeding 7% in both currencies obtained for the mTS distribution amounts to 2.6×10^{-4}. This figures is 4.92 times larger than the one elicited for elliptical t density and 29.47 times larger than the estimate form multivariate Gaussian distribution. Despite the shapes of pdfs, presented on Figure 3.7, the tails of mTS density along the diagonal are heavier than the tails of either of the two elliptical distributions. As the proposed multivariate extension of TS distribution (Definition 3.2.2) provides a superior fit to the empirical data, it also delivers more more realistic estimates of the probabilities of extreme events. In particular, it provides better estimates of the probabilities of the joint currency crises, attributing much more probability mass to the events that have not been observed within the sample. These probabilities evaluated on historical data imply that joint sudden appreciation of the both currencies against euro is five times more likely then their simultaneous decline. Given the recent financial crisis this result is hardly surprising.

3.8 Conclusions

The main contributions of this chapter are as follows.

First, this chapter introduces a tractable, non–standard definition of TS distributions (Definition 3.2.1). This definition directly relates tempered stable density to the underlying α–stable distribution, properly handles the case of $\alpha = 1$ and yields centred distributions for $\mu = 0$. The chapter also surveys the main properties of TS densities, some of which were only known in less general setting ($\alpha < 1$, $\alpha \neq 1$, $\beta = 1$). Furthermore, it presents the complete set formulas for cumulants and moments of TS distributions, previously known only for $\alpha \neq 1$. This chapter also proposes a tractable multivariate extension of TS distribution (Definition 3.2.2). The corresponding proofs may be found in Appendix B. It also proposes a new Fourier Transform discretization scheme, which is efficient for asymmetric densities. This algorithm generalizes the work of Mittnik et al. (1999) and is provided in Appendix A. In all the cases the chapter provides ready to use formulas, valid for the entire parameter range of TS densities, which can be readily used by practitioners. The results listed above are further applied to fit TS densities to real data and to assess the quality of random number generation procedures.

Next, this chapter introduces (as Proposition 3.4.1) a new Cumulant Matching estimator for all parameters of TS density along with the necessary existence conditions. It also contains an overview of feasible, but more computationally intensive approaches to parameter estimation. A novel Proposition 3.5.1 demonstrates that any tempered stable random variate may be represented as a mixture. The chapter formulates a random number generator (Algorithm 2) that relies on mixture representation and generates random draws form TS distribution via rejection. The algorithm is valid for all tempered stable distributions regardless parameter range. It may also be written in just one line of code provided that Baeumer and Meerschaert (2010) procedure is already implemented.

Third, speed and accuracy of mixture representation algorithm are compared with two other techniques that yield random draws from tempered stable density. Algorithm 2 is the most accurate method for the considered set of parameters, which seems appropriate for low frequency macroeconomic distributions. These distributions are typically close to Gaussian ($\alpha > 1$), moderately skewed ($|\beta| \neq 1$) and moderately tempered ($\theta < 1$). The mixture representation algorithm is much faster than the alternative approach of Devroye (1981). It is also faster than the benchmark for mildly tempered distributions (sufficiently small θ). Thus it may be a useful tool for Monte Carlo simulations that involve macroeconomic densities. It is worth noting that the performance of the approximate cdf inversion is surprisingly good.

Fourth, the chapter applies these new methods to evaluate probabilities of joint crashes of Russian ruble and British pound against euro. First, mTS distribution is fitted to both monthly and daily quotations of ruble to euro and pound to euro exchange–rate growth rates. The quality of the obtained fit is compared against two benchmark multivariate distributions – elliptical t and Gaussian. While in these two densities the only possible source of skewness is covariance matrix (both distributions are elliptical), non–elliptical mTS takes the skewness in the marginals of the standardized multivariate data directly into account (Figure 3.7). Hence it is capable to faithfully depict the investigated empirical distribution. Finally, the probabilities of joint crisis in the two currencies are estimated via Monte Carlo. The mTS density fitted to historical data implies markedly higher probabilities of joint currency crisis than the two benchmark distributions. It also indicates that the probability of a sudden decline of euro against the other two currencies by more than 7% is five times more likely than its unexpected upsurge by the same amount. This result is hardly surprising given the recent financial crisis.

Chapter 4

Inter–bank Network Formation – From Heterogeneity to Systemic Risk

This chapter develops an endogenous network formation model for the overnight inter–bank lending network.

4.1 Introduction

Preventing the meltdown of financial system was one of the crucial issues after the collapse of Lehman Brothers in September 2008. A realistic model of banking system is a key component, necessary to verify both the efficiency of different financial regulations and the robustness of alternative market structures. Banks in this model should have assets, liabilities, take into account investment risk and be diversified enough (heterogeneous) to depict a real world diversity of the system. Simultaneously, the banks need to make make their borrowing and lending decisions in a way which has economic meaning (is in some sense optimal) and depends just on their own characteristics and the characteristics of their counterparts (is endogenous). While diversity of market participants is necessary to make the model realistic, endogenous network formation is crucial if we intend to investigate stability of the entire system. If the inter–bank linkages are not endogenous but random or semi–random, the outcome of any stability analysis depends on network configurations that may never arise in practice. Endogenous network formation is also vital in quantifying the implications of endogenous bankruptcies in a distressed banking system. The reason is that the systemic impact of such event to large extent results from the characteristics of both creditors and debtors of the insolvent institution.

This chapter proposes a computational model of endogenous network formation for the inter–bank overnight lending market. The market structure emerges from interactions of heterogeneous agents who are endowed with assets, liabilities and take into account investment risk. As all the banks are large and their trading affects the prices of risky assets, the costs of price slippage breaks the symmetry of portfolio problem, making inter–bank borrowing and lending more desirable. The model takes into account three channels of contagion – bankruptcy cascades, common component of risky asset returns and erosion of liquidity. The network formation algorithm outputs the ensemble of optimal transactions, the outcome of the corresponding link formation process is pairwise stable. This framework is next employed to investigate the stability of the endogenously generated banking systems.

The structure of this chapter is as follows. Section two reviews the literature. The third section presents the overview of the model. Section four describes the choices of borrower and lender from an institutional perspective where the agents know their trade affects market prices. The solutions to the corresponding portfolio problems is given as Appendix D. The fifth section presents the endogenous network formation algorithm. Section six is dedicated to calibration of the model. The seventh section presents the results. The final section concludes.

4.2 Literature review

In the literature of systemic risk and banking system stability two different approaches to deliver a model of a banking system are considered. The theoretical papers allow for strategic (endogenous) network formation, but in stylized settings where diversity of the agents is strongly constrained. In these papers banks typically face limited investment risks and are endowed only with rudimentary assets and liabilities. The computational papers depict more realistic pictures of banking system, but at the expense of higher model complexity. In these works sophisticated models of individual bank behaviour are typically embedded into random network formation schemes. This works bridges both strands of literature, providing a computational protocol for endogenous network formation on the inter–bank overnight market.

Allen and Gale (2001) demonstrate on the system of four identical banks that the extent of contagion depends crucially on the patter of interconnectedness in the network. Freixas et al. (2000) show that a network where banks on periphery are connected to money–centre banks, but not to each other, is also susceptible to contagion. Babus (2009) proposes a model where banks form costly links in order to insure against liquidity risk. These links expose the system to a small risk of contagion. Gai and Kapadia (2010) investigate phase transition in an inter–bank network with arbitrary vertex degree distribution, independent of agents' characteristics. Allen et al. (2012) show that unclustered asset structure of individual banks entails lower funding costs and thus lower bankruptcy costs and higher welfare. Elliott et al. (2013) describe the non–monotonic effect of diversification and integration on cascades of defaults. Acemoglu et al. (2013) demonstrates in a theoretic framework that no network structure yields maximum system resilience regardless the circumstances, a result that was earlier obtained via simulation by Ladley (2013). Finally, Zawadowski (2013) investigates a system with large and peripheral banks, where a default of insurance provider may trigger run on the insuring banks.

One of the earliest computational approaches to modelling financial systems is the work of Eisenberg and Noe (2001). The authors provide an algorithm that cleared mutual claims in a (possibly cyclic) network, a metric of vertex systemic exposure is a by–product of their procedure. Elsinger et al. (2006a) combined standard risk management tools with a (computational) network model of interbank loans and identified correlations in banks' portfolios as the main source of systemic risk. The authors found that while insolvency cascades are rare, they might nonetheless wipe out the major part of the Austrian banking system. In the subsequent work Elsinger et al. (2006b) simulate asset distribution implied by defaults of individual institutions. Thus the results of stress tests they conduct are not conditional on all UK banks remaining solvent.

The first simulation of insolvency cascades in a system of heterogeneous (in size or liquidity) banks endowed with assets and liabilities was undertaken by Iori et al. (2006). While in their model banks do not optimize investments decisions (assets were stochastic), the authors demonstrate that in heterogeneous system incomplete inter–bank network structures are more robust than complete geometries, what reverses[1] a classical result of Allen and Gale (2001). The role of both inter–bank network connectivity and capital requirements on system resilience was further investigated by Nier et al. (2007), who found that the corresponding relationships were non–monotone and non–linear. The impact of a number of inter–bank network geometries on banking system stability was inspected by Georg (2013), the author identifies money–centre networks as the most robust. He also estimates that stabilizing effect of a central bank is non–linear with respect to fraction of banks's asset

[1]Haldane and May (2011) on p. 353 points out that while excessive homogeneity minimizes the risk for each individual bank, it maximizes the probability of collapse of the entire system.

acceptable as collateral. No (random) network structure yields maximum system resilience under all feasible conditions for a banking system in the (static) evolutionary equilibrium, simulated by Ladley (2013) . The impact of heterogeneous asset volumes on erosion of confidence and long term inter–bank lending relationships was established via simulation by Arinaminpathy et al. (2012), extreme heterogeneity (in terms of market concentration) of the US CDS market also plays a pivotal role in the design of super–spreader tax, proposed by Markose et al. (2012). Vallascas and Keasey (2012) demonstrated that a cap imposed on a bank size may be the most efficient tool to reduce its default risk given a systemic event. Martínez-Jaramillo et al. (2010) found that fragility was determined by probabilities of default of individual institutions, their correlations and the number of banks that are instantaneously insolvent if their debtors default. Krause and Giansante (2012) in a recent extension of the work of Iori et al. (2006) identified network geometry and tiering[1] as the two most important factors that determine probability of contagion. Gai et al. (2011) verified how concentration of linkages under geometric and Poisson network affects liquidity hoarding and systemic crisies, he also investigated the inerdependence between market liquidity and haircuts. Bluhm et al. (2013) simulate measures of systemic risk in a complex system of interacting banks. However, their network formation heuristics[2] does not arise from optimal agents' behaviour and thus is not endogenous. Cohen-Cole et al. (2013) depict inter–bank overnight market via Cournot quantity competition model embedded into scale–free networks, they also provide a measure of contribution of individual vertices to systemic risk. Their network formation protocol is not endogenous as it involves an arbitrary probability of forming a link. In all the quoted papers inter–bank market is formed by equating supply and demand and next matching potential creditors and debtors in a random or semi–random manner.

The deficiency of computational model for endogenous network formation in the models of banking system has serious implications. If the inter–bank linkages are random, the fact that two banks are in a lending relationship does not depend on their characteristics other than liquidity demand. The network configurations that thus come into being might never arise had the agents been allowed to chose their counterparts. As the effect of insolvencies on the entire system crucially depends on the geometry of the inter–bank lending network, lack of endogenous network for-

[1]Neglected by all the quoted theoretical papers, with an exception of Zawadowski (2013).

[2]In their paper the agents look for the *closest matching partner* in terms of trading volume, see p. 9 in Bluhm et al. (2013). The only mechanism which could justify heuristics which minimizes number of connections is fixed transaction costs. However, under fixed transaction costs the agents using this rule would in some circumstances prefer a cost of nc over $2c$ for arbitrarily large n. A numerical counterexample is available upon request.

mation affects the outcomes of any systemic stability analysis. This distortion is most severe when we measure the impact of endogenous bankruptcies on a distressed banking system. The reason is that the systemic implications of such event result from the characteristics of both creditors and debtors of the insolvent institution. Another important characteristics of systems in crisis is change of the behaviour of the agents. As the banks learn they past decisions were not optimal and decide to reallocate their resources, they often change their linking patterns. If the overnight lending network is exogenous, then to capture the dynamics of this process the underlying vertex distribution needs to be re–calibrated in an attempt to catch up with the market. If this network is instead endogenous, the agents learn about the change of environment and take it into account in their (optimal) lending decisions. Thus the models with exogenous lending networks might be inadequate to investigate banking systems during crisis.

The model presented here is the first computational framework that allows for endogenous network formation. It takes into account the mutual dependence of market and funding liquidity (Brunnermeier, 2009) that can plunge the system into downward spiral of fire sales and credit denials. As the banks dynamically update their assessment of investment risk and probability of counterparty default, they may shorten their positions in the asset they consider too risky, triggering a flight to quality episode (Caballero and Krishnamurthy, 2008). Just as in the work of Arinaminpathy et al. (2012) our model admits three different channels of contagion, the importance of which was highlighted by Haldane and May (2011): (i) erosion of liquidity, where banks constrain lending in fear of counterparty default (Brunnermeier, 2009; Gai et al., 2011), (ii) fire sales deteriorating market liquidity (Adrian and Shin, 2010; Battiston et al., 2012; Cifuentes et al., 2005; Coval and Strafford, 2007; Shleifer and Vishny, 2011), (iii) bankruptcy cascades due to counterparty credit risk (Elsinger et al., 2006a; Gai et al., 2011; Nier et al., 2007; Upper, 2011). The model allows for the study of endogenous bankruptcies and takes into account different possible behaviour of the banks under varying condition, thus addressing the criticism of the literature on systemic stability, voiced by Upper (2011). While many papers consider either one (Allen and Gale, 2001; Gai et al., 2011; Georg, 2013; Ladley, 2013) or two (Babus, 2009; Zawadowski, 2013) types of banks with respect their total assets, bank sizes in this paper are fully heterogeneous (Iori et al., 2006; Krause and Giansante, 2012). The model allows returns on risky investment of different banks to be correlated, but this dependence is implemented neither via elasticities of demand (Bluhm et al., 2013; Cifuentes et al., 2005), asset commonality (Allen et al., 2012) or common shocks (Georg, 2013), but rather through the demand related component in asset prices.

There are two possible approaches to evaluating computational models. The first criterion is to take into account variety of interacting mechanisms that the model is able to depict and the plausibility of the assumptions behind it. In this instance the proposed model covers three different channels of contagion, its foundations are far less restrictive than the assumptions of theoretic models featuring endogenous network formation. The second criterion is the ability of the proposed model to reproduce empirical characteristics of both the inter–bank lending networks and the asset structure of individual banks. The networks generated by the model are endowed with degree distribution that is close to scale–free, which matches the empirical findings of Soramäki et al. (2007) and Cohen-Cole et al. (2013). The density of the simulated networks lies in the plausible range (Becher et al., 2008; Müller, 2006), small banks are (on average) creditors of a larger institutions (Müller, 2006). The lending relationship is also disassortative – in the simulation banks tend to lend from the institutions of different size to themselves (Cocco et al., 2009). Furthermore, the size of inter–bank market, the distribution of bank sizes and the strength of mutual dependence of asset returns in the system all closely match the corresponding empirical values.

4.3 Model overview

The economy consists of N regions, indexed with $k \in \{1, \cdots, N\}$. Each region harbours n identical consumers whose total mass[1] amounts to h_k. Each region is endowed with a single local bank k. There are T periods, indexed with $t \in \{1, \cdots, T\}$. In every period all local consumers approach bank k to place their deposits of h_k/n. Each consumer, after placing her deposit at time t, liquidates it at $t + 1 + S$ where S is a Poisson distributed random variate with intensity $\lambda - 1 > 0$. The deposits are being held for λ periods on average, expected net value of deposits placed at k amounts to λh_k. The volume of deposits placed in every region is in the long run stationary.

The banking system constitutes of N regional banks. Banks accept consumer deposits and do not compete with each other on deposit market. All the banks are required to keep a fraction ρ of deposits held as obligatory reserves. The remaining part is invested in the two assets – inter–bank loans and *risky* asset. Inter–bank loans cover transient liquidity shortages and are perceived as almost riskless. They

[1]The weight of consumers in all regions and adds up to N.

last only one period and can not be extended, but the debtor may obtain a new loan next period on competitive basis. The risky asset replicates portfolio structure of the aggregate banking system. It consists of the components that are regarded as far more risky than inter–bank loans. The expected volume of assets held by each bank is initially (at $t = 1$) proportional to the size of the corresponding region. Banks vary in lending needs, which arise from variability of net deposits. They also display different attitudes towards risk and different risk perception. Hence the banking system is heterogeneous, with bank characteristics varying along four dimensions.

Each bank has its own portfolio which represents long term risky investments. All the portfolios share common shocks and liquidity effects. Let $P_{k,t}$ stand for a value of a unit of long term investment of bank k in period t. Given $P_{k,-1}, P_{k,0} := 1$, the dynamics of $\Delta \ln P_{k,t+1} := \ln P_{k,t+1} - \ln P_{k,t}$ is driven for every integer $k > 1$ by the following iterative equations

$$\Delta \ln P_{k,t+1} = \eta(\mu - \Delta \ln P_{k,t}) + I_t + \sigma Z_{k,t}, \quad Z_{k,t} \sim \text{NID(0,1)}, \tag{4.1}$$

$$I_t = \nu(D_t - S_t)/(D_t + S_t).$$

Above η represents the autoregression coefficient, while μ stands for the mean reversion parameter. D_t and S_t denote, respectively, the aggregate demand for and supply of the risky assets in the entire banking system. Constant ν maps excess relative demand or supply into prices. The corresponding liquidity effect is captured by I_t. This common component of asset returns may be motivated by three mechanisms. First, in a competitive environment all the large banks face similar investment opportunities. Next, the fact that all the market participants use the same risk evaluation and portfolio optimization tools may cause spontaneous dependence of investment outcomes. Finally, as all the banks are large, the aggregate quantities of risky assets they trade always affect the market. Parameter σ reflects conditional standard deviation of asset return rates. $Z_{k,t+1}$ are independent random variables with standard normal density. They represent idiosyncratic components of price dynamics. Each bank is allowed to trade only the units of its own portfolio. The banks are oblivious of the form of risky asset pricing formula, but they record the realized sample mean and standard deviation. Diverse price trajectories represent the effects of individual portfolio composition on the bank overall performance.

Assets of each bank consist of investment portfolio $a_{k,t}$, reserves $r_{k,t}$, net cash $c_{k,t}$ and loans to other banks $l_{k,t}$. While long term investment and loans to other institutions are both risky, the term *risky asset* will be further reserved for the units of portfolio of given bank. Liabilities of each bank constitute of equity $e_{k,t}$, customer deposits $d_{k,t}$ and loans from other banks $b_{k,t}$. Hence for each bank k at every point in time $t \in \{1, \cdots, t\}$ it holds that

$$a_{k,t} + r_{k,t} + c_{k,t} + l_{k,t} = d_{k,t} + e_{k,t} + b_{k,t}.$$

If a ratio of equity to risk weighted assets for any bank falls below 4%, this bank becomes bankrupt. Banks learn unconditional distribution of risky asset prices from past data. They also form expectations on probabilities of counterparty default.

Banks are managed by CEO's who each period receive bonuses[1] proportional to net profits of their bank. The management of each bank is concerned with variability of their income, but only as long as their own institution remains solvent. In order to capture this two effects it is assumed that the banks maximize expected conditional utility derived from profit. As deposits placed are exogenous, in order to do so it is enough to maximize expected conditional utility derived from a unit portfolio of capital. To preserve the tractability of the model it is assumed that banks are endowed with constant absolute risk aversion (CARA).

Given the selected utility function, bank k needs to solve two problems in order to optimize its investment. First, it needs to decide how much money would it like to lend (or borrow) and at what interest rate. Next, it identifies a willing counterpart with whom the desired transaction could be concluded.

4.4 Portfolio problem

Throughout this section it is assumed that the volume of risky asset traded by every bank is always considered *large* by the market. Banks know that trading risky asset adversely affects transaction prices due to price slippage. When they sell risky asset, they face instantaneous loss and relinquish future stochastic profits. When they buy risky asset, they enjoy uncertain profits but at the expense of instantaneous losses. Hence a bank with transient liquidity shortage might prefer taking loan on the inter–bank market to selling risky asset. By a similar token, a bank with transient liquidity surplus might prefer lending money on inter–bank market to buying risky asset.

In order to depict investment choices of large banks this section approaches portfolio selection problem from institutional perspective, in which market participants are aware of price distortions, caused by their size. In the first subsection basic notation is introduced. Next the problems of borrower and lender are formally stated. Then

[1]Exact values of these bonuses affect only *numèraire* and thus may be neglected.

we assume that utility of each individual bank displays Constant Absolute Risk Aversion and, given this assumption, investigate the first order conditions of borrower and lender. In the fourth subsection we assume in addition that the distribution returns of risky asset is Gaussian. The equations for credit supply and demand of individual bank that follow are derived given Gaussian returns. The final passage explains how this solution of individual portfolio problem may affect endogenous network formation.

4.4.1 Notation

Index $k \in \{1, \cdots, N\}$ is reserved to denote any bank. Index b denotes a bank that is a borrower while l stands for a bank which is a lender. As the interest rate is always agreed between both parties, for the sake of brevity we shall denote it with $i \equiv i_{bl}$. All the banks dynamically update their joint beliefs p_{t+1} on the probability of counterpart default. Every bank k is endowed with risk aversion coefficient γ_k and learns the mean $\mu_{k,t+1}$ and standard deviation $\sigma_{k,t+1}$ of risky asset growth rates from its own past data. Again for the sake of brevity we shall introduce the following shortcut notation: $p \equiv p_{t+1}$, $\mu_k \equiv \mu_{k,t+1}$ and $\sigma_k \equiv \sigma_{k,t+1}$. Demand and supply characterize the behaviour of, respectively, borrower and lender, thus we may write $d \equiv d_b$, $s \equiv s_l$.

Assume $B_{b,t+1}$ is equal to one if b defaults at $t+1$ and is equal to zero otherwise. Bank l expects that given b's default only $(1-\theta)$ of the amount lent will be recovered. Multiplicative *loss given default* is equal to θ. If b does not default, it pays back the entire amount plus agreed interest rate i. Return on risky asset at $t+1$ is denoted by R_{t+1} and banks treat it as independent of $B_{b,t+1}$. Let c stand for multiplicative cost of trading either in terms of price slippage and transaction costs, or incurred due to unfavourable timing of liquidating or purchasing the risky asset. These concern buyers and sellers equally[1], as each of the banks is large enough to affect market prices.

Let w stand for the desired fraction of stocks held in portfolio at the end of current period, let $\underline{w} \geq 0$ be the share of stocks in portfolio at the beginning of current period. Assume v denotes net cash surplus after the bank paid out deposits due and satisfied obligatory reserve requirements. If bank holds a surplus cash, $w = \underline{w}$ implies no necessity to trade shares, and hence no extra costs. If the bank is not able to either meet reserve requirements or pay out the deposits at the beginning of the current period, v becomes negative.

[1] See i.e. Coval and Strafford (2007).

Assume bank k has in its portfolio m loans to/from $j_1, \dots, j_m$ whose volumes amount to $w_{kj_1}, \cdots, w_{kj_m}$. Define w_{kj_i} as positive only if in the i–th transaction bank k is a borrower and j is a lender. If in i–th transaction bank k is a lender then w_{kj_i} is negative. Let $\hat{w}$ stand for the aggregate net loans granted to k by other members of the banking system prior current transaction. Let $\hat{i}$ be the aggregate interest rate on these loans. Thus we have

$$\hat{w} = \sum_{i=1}^{m} w_{kj_i},\ \hat{w}(1+\hat{i}) = \sum_{i=1}^{m} w_{kj_i}[(1+i_{kj_i})\mathbb{1}_{\{w_{kj_i}>0\}}(w_{kj_i}) + (1+i_{j_ik})\mathbb{1}_{\{w_{kj_i}<0\}}(w_{kj_i})].$$

Above $\mathbb{1}_{\{A\}}(w) = 1$ iff $w \in A$ and zero otherwise. The quantities v, $\hat{w}$ and $\hat{i}$ that appear in the borrower problem describe characteristics of the borrower, the same notation in the lender problem pertains to lender, in both cases the indexes were omitted for easier display.

For the sake of convenience set the following notation

$$\chi_b := c^{-1}\mathbb{1}_{\{w\geq \underline{w}\}}(w) + c\mathbb{1}_{\{w<\underline{w}\}}(w),\ \chi_l := c^{-1}\mathbb{1}_{\{w>\underline{w}\}}(w) + c\mathbb{1}_{\{w\leq\underline{w}\}}(w).$$

Above χ_k is a step functions of w for $k \in \{b, l\}$. Note that both threshold functions differ in $\underline{w}$. Define the following constants

$$A_k := [(1+i)(v+\hat{w}+\chi_k\underline{w}) - \hat{w}(1+\hat{i})],\ B_k := [1-(1+i)\chi_k],$$

$$E_k := [(1-\theta)(v+\hat{w}+\chi_k\underline{w}) - \hat{w}(1+\hat{i})],\ F_k := [1-(1-\theta)\chi_k],$$

Let $V_{t+1} \equiv V_{k,t+1}$ stand for capital of bank k at $t+1$.

Bank b is a prospective borrower if the amount of risky asset it intends to hold can not be financed with its net cash. For $w \geq \underline{w}$ this condition may be written as $(v+\hat{w}) - c^{-1}(w-\underline{w}) < 0$, which is equivalent to $w > \underline{w} + c(v+\hat{w})$. In case of $w < \underline{w}$ it may be represented as $(v+\hat{w}) + c(\underline{w}-w) < 0$ and thus $w > \underline{w} + c^{-1}(v+\hat{w})$. Both formulas may be expressed as $w > \underline{w} + [c\mathbb{1}_{\{w\geq\underline{w}\}}(w) + c^{-1}\mathbb{1}_{\{w<\underline{w}\}}(w)](v+\hat{w})$. Define a set of portfolio choices where b is a borrower as I_B. We have

$$I_B = \{w:\ w \geq \underline{w} + [c\mathbb{1}_{\{w\geq\underline{w}\}}(w) + c^{-1}\mathbb{1}_{\{w<\underline{w}\}}(w)](v+\hat{w})\} = \{w:\ w \geq \underline{w} + \chi_b^{-1}(v+\hat{w})\}.$$

Let I_L stand for a set of portfolio choices where l is a lender, then

$$I_L = \{w:\ w \leq \underline{w} + \chi_l^{-1}(v+\hat{w})\}.$$

Borrower and lender problem are solved for, respectively, $w \in I_B$ and $w \in I_L$. Banks are not allowed to keep cash between consecutive periods for reasons other than precautionary.

4.4.2 Borrower and lender problems

Assume $w \in I_B$. For $w > \underline{w}$ the capital of b at $t+1$ amounts to

$$\underline{w}(1+R_{t+1})+\mathbb{1}_{\{w>\underline{w}\}}(w)(w-\underline{w})(1+R_{t+1})-(1+i)(c^{-1}\mathbb{1}_{\{w>\underline{w}\}}(w)(w-\underline{w})-v-\hat{w})-\hat{w}(1+\hat{i}),$$

in case of $w = \underline{w}$ we have

$$\underline{w}(1+R_{t+1}) - (1+i)(-v-\hat{w}) - \hat{w}(1+\hat{i}),$$

while for $w < \underline{w}$ we obtain

$$\underline{w}(1+R_{t+1})+\mathbb{1}_{\{w<\underline{w}\}}(w)(w-\underline{w})(1+R_{t+1})-(1+i)(c\mathbb{1}_{\{w<\underline{w}\}}(w)(w-\underline{w})-v-\hat{w})-\hat{w}(1+\hat{i}).$$

Combining the three cases above yields for borrower b the following formula

$$V_{t+1}(w) = w(1+R_{t+1}) - (1+i)(\chi_b(w-\underline{w}) - v - \hat{w}) - \hat{w}(1+\hat{i}).$$

The conditional utility that borrower b expects to obtain at $t+1$ by investing in a portfolio, consisting of $w \in I_B$ units of shares and a loan from l, is

$$\mathbb{E}u(V_{t+1}(w)|B_{b,t+1}=0) = \mathbb{E}u(A_b + B_b w + wR_{t+1}).$$

Now assume $w \in I_L$. For $w > \underline{w}$ the capital of l at $t+1$ amounts to

$$\underline{w}(1+R_{t+1}) + \mathbb{1}_{\{w>\underline{w}\}}(w)(w-\underline{w})(1+R_{t+1})+$$

$$+((1-\theta)B_{b,t+1} + (1+i)(1-B_{b,t+1}))(v+\hat{w}-c^{-1}\mathbb{1}_{\{w>\underline{w}\}}(w)(w-\underline{w})) - \hat{w}(1+\hat{i}),$$

in case of $w = \underline{w}$ we have

$$\underline{w}(1+R_{t+1}) + ((1-\theta)B_{b,t+1} + (1+i)(1-B_{b,t+1}))(v+\hat{w}) - \hat{w}(1+\hat{i}),$$

while for $w < \underline{w}$ we obtain

$$\underline{w}(1+R_{t+1}) + \mathbb{1}_{\{w<\underline{w}\}}(w)(w-\underline{w})(1+R_{t+1})+$$

$$+((1-\theta)B_{b,t+1}+(1+i)(1-B_{b,t+1}))(v+\hat{w}-c\mathbb{1}_{\{w<\underline{w}\}}(w)(w-\underline{w}))-\hat{w}(1+\hat{i}).$$

After combining the three cases above and substituting χ_l for lender l we obtain

$$V_{t+1}(w)=(w(1+R_{t+1})+((1-\theta)B_{b,t+1}+(1+i)(1-B_{b,t+1}))(v+\hat{w}-\chi_l(w-\underline{w}))-\hat{w}(1+\hat{i}).$$

Therefore by the rule of iterated expectation

$$\mathbb{E}u(V_{t+1}(w))=\mathbb{E}[\mathbb{E}u(V_{t+1}(w)|B_{b,t+1})]=$$

$$=p_{t+1}\cdot\mathbb{E}u(E_l+F_lw+wR_{t+1})+(1-p_{t+1})\cdot\mathbb{E}u(A_l+B_lw+wR_{t+1}).$$

4.4.3 First order conditions under CARA

Assume that each bank k is endowed with constant absolute risk aversion with risk aversion parameter $\gamma_k>0$. Let $f_{R_t}(x)$ be a pdf of random variable R_t. Assume R_t has a moment generating function $M_t(q)$, defined for $q\in I\subset\mathbb{R}$.

Under these assumptions made for the borrower problem we have

$$\mathbb{E}u(V_{t+1}(w)|B_{b,t+1}=0)=\mathbb{E}u(A_b+B_bw+wR_{t+1})=1-e^{-\gamma_b(A_b+B_bw)}\cdot M_{t+1}(-\gamma_bw).$$

Thus the optimal portfolio is given by

$$w_b^*=\underset{w\in I_B}{\operatorname{argmax}}\left\{1-e^{-\gamma_b(A_b+B_bw)}\cdot M_{t+1}(-\gamma_bw)\right\}=\underset{w\in I_B}{\operatorname{argmin}}\left\{e^{-\gamma_b(A_b+B_bw)}\cdot M_{t+1}(-\gamma_bw)\right\}.$$

The borrower's f.o.c. may be written as

$$\frac{d\log M_{t+1}(q)}{dq}\Big|_{q=-\gamma_bw}+B_b=0. \tag{4.2}$$

To validate the consistency of assumptions we need to check if $-\gamma_bw_b^*\in I$.

For the lender problem we obtain

$$\mathbb{E}u(V_{t+1}(w))=p_{t+1}\cdot\mathbb{E}u(E_l+F_lw+wR_{t+1})+(1-p_{t+1})\cdot\mathbb{E}u(A_l+B_lw+wR_{t+1})=$$

$$=1-p_{t+1}\cdot e^{-\gamma_l(E_l+F_lw)}\cdot M_{t+1}(-\gamma_lw)-(1-p_{t+1})\cdot e^{-\gamma_l(A_l+B_lw)}\cdot M_{t+1}(-\gamma_lw).$$

Thus the optimal unit portfolio is given by

$$w_l^*=\underset{w\in I_L}{\operatorname{argmax}}\left\{1-\left(p_{t+1}\cdot e^{-\gamma_l(E_l+F_lw)}+(1-p_{t+1})\cdot e^{-\gamma_l(A_l+B_lw)}\right)\cdot M_{t+1}(-\gamma_lw)\right\}=$$

$$= \operatorname*{argmin}_{w \in I_L} \left\{ \left(p_{t+1} \cdot e^{-\gamma_l (E_l + F_l w)} + (1 - p_{t+1}) \cdot e^{-\gamma_l (A_l + B_l w)} \right) \cdot M_{t+1}(-\gamma_l w) \right\}.$$

The lender's f.o.c. may be written as

$$\frac{d \log M_{t+1}(q)}{dq} \Big|_{q=-\gamma_b w} + \frac{p_{t+1} F_l \cdot e^{-\gamma_l (E_l + F_l w)} + (1 - p_{t+1}) B_l \cdot e^{-\gamma_l (A_l + B_l w)}}{p_{t+1} \cdot e^{-\gamma_l (E_l + F_l w)} + (1 - p_{t+1}) \cdot e^{-\gamma_l (A_l + B_l w)}} = 0. \quad (4.3)$$

To validate the consistency of assumptions we need to check if $-\gamma_l w_l^* \in I$.

Fact D.5.1 in Appendix D demonstrates that every internal, admissible solution of either borrower or lender f.o.c. yields maximum utility level.

4.4.4 Credit demand and supply under Gaussian returns

From this point assume, in addition, that density of risky asset returns for each bank k is Gaussian with mean μ_k and variance σ_k^2. Under this assumption logarithmic differentials of moment generating functions are linear, thus it is possible to derive analytic formulas for the largest interest rate acceptable to a borrower and the smallest interest rate acceptable to a lender. From these two equations we may obtain demand and inverse demand, supply and inverse supply formulas of individual banks that are consistent with the assumptions, introduced for the network formation algorithm in section 4.5. These four formulas are the main result of the following section. For the sake of exposition all the intermediate steps, required to derive the sought results, are relegated to Appendix D.

The monetary demand for credit d depends on the difference between the optimal portfolio choice and the largest position that b could finance without additional founding. For sufficiently small $d > 0$ the optimal portfolio choice w_b^* is an internal solution of the borrower problem and lies above $e_b := \underline{w} + \chi_b^{-1}(v + \hat{w})$. There are two distinct cases. If $v + \hat{w} \geq 0$ then $w_b^* \geq \underline{w}$. Bank b utilizes credit to buy more shares than it had at the end of previous period. As buying large quantities of shares causes price slippage, for every extra unit of shares it has to pay c^{-1} units of money. Monetary demand for credit is given by $c^{-1}(w_b^* - e_b)$. By a similar token, if $v + \hat{w} < 0$ and if $d > 0$ is sufficiently small, then we have $w_b^* < \underline{w}$. Bank b takes credit in order to avoid selling shares it already has in its portfolio. As disposing of large quantities of assets incurs cost, for every c units of money b can refrain from selling one unit of shares. So the monetary demand for credit is given by $c(w_b^* - e_b)$. In this case credit demand is sufficiently small if only $d < -v - \hat{w}$. As threshold function χ_b depends on w_b^*, this additional restriction guarantees that borrower's actions are indeed optimal. In our notation monetary credit demand may be in both cases expressed as

$$d(i) = \chi_b(w_b^* - e_b) = \chi_b(w_b^* - \underline{w}) - (v + \hat{w}).$$

Given previous results the solution of borrower's problem w_b^* may be treated as given for each offered interest rate i. In both cases we may determine $\bar{i}_b$ as a limit interest rate for $d \to 0^+$.

Now we may find the largest interest rate $\bar{i}_b$ that borrower b would be willing to accept. Monetary demand for credit as a function of the offered interest rate is

$$d(i) = \chi_b\Big(\frac{1}{\gamma_b\sigma_b^2}[(1+\mu_b) - \chi_b(1+i)] - \underline{w}\Big) - (v + \hat{w}) \equiv$$

$$\equiv i(d) = \chi_b^{-1}\Big((1+\mu_b) - \gamma_b\sigma_b^2[\underline{w} + \chi_b^{-1}(v + \hat{w} + d)]\Big) - 1.$$

Hence we have

$$\bar{i}_b = \lim_{d\to 0^+} i(d) = \chi_b^{-1}\Big((1+\mu_b) - \gamma_b\sigma_b^2[\underline{w} + \chi_b^{-1}(v + \hat{w})]\Big) - 1. \tag{4.4}$$

The monetary supply of credit s depends on the difference between the optimal portfolio choice and the largest position that l could finance without additional founding. For sufficiently small credit supply $s > 0$ the optimal portfolio choice w_l^* is an internal solution of the lender problem and lies below $e_l := \underline{w} + \chi_l^{-1}(v + \hat{w})$. There are two distinct cases. If $v + \hat{w} \geq 0$ and if in addition credit supply $s < v + \hat{w}$ then $w_l^* \geq \underline{w}$. Bank l finances current loan from surplus cash. As buying large quantities of shares causes price slippage, for every extra unit of shares that l does not to buy it can grant a loan of c^{-1} units of money. Monetary supply of credit is given by $c^{-1}(e_l - w_l^*)$. By a similar token, if $v + \hat{w} < 0$ then we have $w_l^* < \underline{w}$. Bank l finances the loan by selling speculative asset. As disposing of large quantities of assets incurs cost, for every unit of shares l sells it obtains c units of money that can be lent on the inter–bank market. So the monetary supply of credit is $c(e_l - w_l^*)$. Using previous notation, monetary credit supply may be in both cases expressed as

$$s(i) = \chi_l(e_l - w_l^*) = \chi_l(\underline{w} - w_l^*) + (v + \hat{w}).$$

Again we may determine $\underline{i}_l$ as a limit interest rate implied by $s \to 0^+$.

The solution of lender's problem satisfies f.o.c at $w_l^* := \underline{w} + \chi_l^{-1}(v + \hat{w} - s(i))$. To find $\underline{i}_l$ it is sufficient to substitute w_l^* to lender's f.o.c. and obtain i in the limit.

Thus the smallest interest rate that the lender would be able to accept is

$$\underline{i}_l = \bar{i}_l \frac{1}{1-p_{t+1}} + \theta \frac{p_{t+1}}{1-p_{t+1}}. \tag{4.5}$$

The maximal size of a loan that b is eager to accept at interest rate $\tilde{i}$ is the volume $\tilde{w}$ that makes $\tilde{i}$ the largest acceptable interest rate. By a similar token, the maximal size of a loan that l is eager to grant at interest rate $\underset{\sim}{i}$ is the volume $\underset{\sim}{w}$ that equates $\underset{\sim}{i}$ to $\underline{i}_l$. Given the interior solution of borrower problem his maximal loan volume may be obtained as

$$\tilde{i} = \chi_b^{-1}\Big((1+\mu_b) - \gamma_b\sigma_b^2[\underline{w} + \chi_b^{-1}(v+\hat{w}+\tilde{w})]\Big) - 1 \equiv \tag{4.6}$$

$$\equiv \tilde{w} = \gamma_b^{-1}\chi_b^2\sigma_b^{-2}(\bar{i}_b - \tilde{i}),$$

while for the lender it is given by

$$\underset{\sim}{i} \approx (\bar{i}_l - \gamma_l\chi_l^{-2}\sigma_l^2\underset{\sim}{w})\frac{1}{1-p_{t+1}} + \theta\frac{p_{t+1}}{1-p_{t+1}} \equiv \tag{4.7}$$

$$\equiv \underset{\sim}{w} \approx \frac{1}{\gamma_l\chi_l^{-2}\sigma_l^2}[(\bar{i}_l - \underset{\sim}{i})(1-p_{t+1}) + (\bar{i}_l+\theta)p_{t+1}].$$

The pairs of equations (4.6) and (4.7) stand for, respectively, demand and inverse demand for, supply and inverse supply of inter–bank loans, obtained for an individual bank k. These formulas are recovered from reservation interest rates of borrower and lender which are slightly perturbed and inverted in the limit. In case of borrower b this approach approximates the solution of borrower's f.o.c. This approximation is exact if an offer interest rate $\tilde{i}$ is sufficiently close to $\bar{i}_b$ or, alternatively, $\tilde{w} > 0$ is sufficiently small. The behaviour of a debtor who borrows $\tilde{w}$ at an interest rate $\tilde{i}$ is optimal. No such reverse engineering is possible in case of lender l. This is because non–linearity of credit supply equation vanishes in the limit for $s(i) \to 0^+$. While $\tilde{w}$ is approximate solution to (borrower) f.o.c., $\underset{\sim}{w}$ is just a solution to approximate (lender) f.o.c. These two concepts do not need to coincide, lender's behaviour is not optimal by itself. However, it becomes optimal due to assumption 2) in the next section.

Note that formulas (4.4) and (4.5) imply satiation effect as both parties become more reluctant to trade when their volume of either borrowing or lending increases. Furthermore, no lender ever lends at an interest rate it would accept as a borrower. No bank $k \in \{b, l\}$ would ever trade with itself as it always holds that $\underline{i}_k > \bar{i}_k$.

4.4.5 From portfolio selection to network formation

Given the solution of portfolio problem, there are two main reasons why inter–bank market exists. First, there are costs of trading. Banks who finance a deficit of net deposits by selling risky asset suffer the costs of portfolio adjustment and opportunity lost. Thus they may prefer to take a loan on the inter–bank market and repay it with interest. Simultaneously, banks who invest a surplus of net deposits in the risky asset face instant loss due to price slippage, while their future profits remain uncertain. Hence they may prefer to grant a loan on the inter–bank market and collect a (seemingly) certain interest. Financing a new loan by selling a risky asset is more expensive than financing it with net deposits. Therefore banks who enjoy a surplus of net deposits have stronger incentives to become lenders, bunks who run a deficit have motivation to become borrowers. In this model transaction costs play the role of *symmetry–breaking* mechanism, as market participants assume that buying and then selling a unit of risky asset is not neutral to their financial situation. Next, banks are heterogeneous. Different degree of risk aversion, diverse risk perception and liquidity needs of individual banks map to reservation interest rates. If these rates are diversified enough, the banks find it desirable to trade.

4.5 Network formation algorithm

In this section the assumptions behind the proposed model are being first discussed. Next, a network formation algorithm that utilizes the solution to the portfolio problem derived in section 4.4 is presented. The basic idea is that the two groups of banks who would surely trade with each other are best borrowers and best lenders. Hence the optimal network of lending relations may be approximated by, sequentially, allowing most desirable counterparts to trade by borrowing or lending sufficiently small amounts and next updating weak preference relation.

4.5.1 Assumptions

The model presented in this chapter builds on the following assumptions. 1) Banks joint beliefs on the probability of counterparts bankruptcy is given by p. The value of p is time dependent and is a subject of collective learning. 2) All inter–bank lending is concluded at the midpoints of reservation prices of borrowers and lenders. 3) Banks are able to foresee all the interim stages of the proposed network formation protocol.

Out of the three assumptions introduced above the first one is probably the most problematic. It implies that probability of bankruptcy does not vary with counterpart level of equity or reserves. Thus assumption one is clearly false. The main motivation behind it are the limitations of computational approach, taken in this paper. If the agents are supposed to evaluate risk of lending money to given counterpart, they need to form beliefs on a probability of counterpart bankruptcy in a dynamic way. One way to implement this learning process which is optimal from the point of view of information theory is to employ bayesian probit model. However, inter–bank bankruptcies are rare events. Hence for all practical purposes the output of such a model would be equivalent to prior information. Furthermore, it is reasonable to expect that the outcome of network formation process would be sensitive to how this prior is calibrated. Heterogeneous probabilities of counterpart default do not only make model calibration arbitrary, they also result in causality of the model being more blurred as the impact of different variables is harder to disentangle. A fringe benefit of this assumption is that matching algorithm, presented in further part of this chapter, needs to keep track of just a single bid and ask price for each market participant.

There are two main reasons behind the second assumption. First, it provides an incentive for the proponents of the best bid and ask offers to trade together. Any inter–bank lending which happens in the interim stages of the algorithm presented here takes place between parties who are each other's first choices, either in terms of asked or offered interest rate. The aggregate outcome of this trade is necessary pairwise stable. Next, there is no reason to *a priori* assume that any side of the market – either borrowers or lenders – has a bargaining power over the other. This assumption may be at ease relaxed, as it affects neither link formation process, nor transaction volumes. It is therefore irrelevant to geometry of the emergent network.

Assumption three implies that the banks possess full information on characteristics of other agents and the network formation protocol. It performs a dual role of both rationality and consistency condition – each bank attempts to optimize its actions given the information available at a time of forming links. After this preliminary discussion we can introduce the network formation algorithm.

4.5.2 Algorithm

The network formation protocol, introduced in this section, constitutes of two phases. In the *interim* phase the four following stages are consecutively repeated. First, the algorithm identifies the sets of best prospective borrowers and lenders. Only the members of these two sets are allowed to trade. Next it determines the largest admissible volume of trade which obeys certain conditions. In step three this quantity is split among all the bilateral transactions and the common interest rate is set. Finally banks reservation rates are updated as if the transactions agreed upon were already concluded. The proposed algorithm iterates through the four stages listed above until there are no more creditors or debtors who are willing to trade. This ends the interim phase. The second part of the algorithm is termed the *update* phase. In the update phase all the transactions, agreed upon in all the stages of the first phase, are finally concluded. Furthermore, all the bank characteristics are properly updated. The detailed account of each stage of the interim phase is given below.

First stage of the interim phase consists of the following steps. 1) For every non–bankrupt bank we compute the largest interest rate it is able to accept as a borrower and the lowest interest rate it is able to accept as a lender. 2) All the solvent prospective borrowers are listed in descending order according to the largest interest rate they are willing to accept. All the solvent prospective lenders are listed in ascending order according to the lowest interest rate they are able to accept. Hence each bank is listed twice, once on every side of the market. These two hierarchies introduce a natural ordering among both debtors and creditors – it is clear whose offer on given side of the market may be termed *best*. In consequence we can distinguish the two cliques, consisting of the best prospective borrowers and the best prospective lenders. 3) The prospective debtors and prospective creditors with, respectively, the highest bid and the lowest ask price, are allowed to declare trade with each other. In given stage of the interim phase these two groups are termed *active* borrowers and *active* lenders.

Second stage of interim phase determines the volume of trade between active debtors and creditors, selected in stage one. This quantity is defined as the largest value which preserves the following conditions.

i) The volume of aggregate trade can not be larger than the volume which equates aggregate supply and demand. The latter corresponds to the market clearing interest rate, derived as equation D.10.

ii) If there is a second best bid interest rate, then total volume of loans granted to any prospective debtor can not be larger than the amount that makes his bid interest rate equal to the second best bid rate. By a similar token, if there is a second best ask interest rate, then total volume of loans granted by any prospective creditor can not be larger than the amount that makes his ask interest rate equal to the second best ask rate.

iii) If any active borrower experiences a deficit of net cash, then the maximum total volume of the loans he is allowed to accept during current interim stage exactly offsets his deficit. If any active lender runs a surplus of net cash, then the maximum total volume of the loans he is allowed to grant during current interim stage exactly offsets his surplus.

iv) The amount of lending any active creditor may grant has to be financed either with its reserves or by selling risky asset. After all the other constraints are applied, the volume of lending has to be equal to the volume of borrowing.

The short discussion of the four postulated conditions is as follows.

Condition i) above is simply a market clearing condition. If no other restriction binds, all the borrowers and lenders active during current interim stage are satiated and should exit the market.

Condition ii) reflects the constraint, imposed on either borrower or lender by his counterpart. It may be triggered if in the course of trading the highest (lowest) interest rate, acceptable to active borrowers (lenders), starts equating the second best interest rate. Then active lenders (borrowers) would also want to trade with all the non–active borrowers (lenders) who display the best reservation rate. Trading with more counterparts implies either the same amount of credit at more favourable price, or more credit with no deterioration of price. Hence no rational bank would trade beyond this point.

Condition iii) is merely a consistency condition. If surplus of net cash of any borrower of lender is exactly equal to zero, his reservation interest rates need to be recalculated. As this update takes place at the end of the fourth stage of the interim phase, no further trade is possible in current iteration.

Condition iv) is balancing condition. No lender can breach its budget constraint. Each borrower (lender) needs to have a counterpart willing to lend him (borrow from him) the agreed amount of money. This last condition is necessary as, after all the other conditions are fulfilled, supply may exceed demand or demand may exceed supply. In the first case we need to find the aggregate creditor reservation rate which generates supply equal to demand. In the second case the aggregate debtor reservation rate is required which corresponds to demand equal to supply. Both aggregate reservation rates may be found as formulas (D.11) and (D.12) in Appendix D.

A simple way to implement restrictions i)–iv) is to map these conditions into reservation interest rates. For each of these restrictions we can find the largest (or the smallest) interest rate acceptable for active debtors (creditors) that guarantees the given condition is fulfilled. The maximal reservation rate for all active borrowers constitutes a single restriction which binds all the active debtors. The minimal reservation rate for all active lenders constitutes a single restriction which binds all the active creditors. Hence the proposed implementation relies on formulas (4.6) and (4.7).

Note that conditions i)–iv) do not rule out cycles. Thus an additional precaution is required in order to guarantee that the network formation protocol actually stops. In the simulation, presented in this chapter, it was simply assumed that any active borrower who in the next iteration of the algorithm becomes active lender is no longer allowed to become active borrower. By a similar token, any active lender who becomes active borrower is no longer allowed to become active lender. This completes the description of the second stage of the interim phase.

In the third stage of interim phase the total volume that each debtor (creditor) has committed to borrow (lend) is distributed among all his active counterparts. It follows from assumption 1) that banks are indifferent with whom they trade, as long as their partner on the opposite side of the market offers them the most favourable interest rate. Hence if all the banks were of the same size (in terms of the total assets), all trade would be symmetric. In general case the agreed volume of bilateral trade is proportional to the counterpart's share in either aggregate supply or aggregate demand multiplied by the share of counterpart's assets in bank system assets. By assumption 2) the price the banks agree upon is a midpoint of reservation prices, which characterize each side of the market.

We might consider networks of lending relations which arise in interim phase. The vertices in such interim network are active borrowers and active lenders. Borrower b is connected to lender l if and only if both parties agreed on a loan during four consecutive stages of given interim phase. All such networks are necessary *complete*. Here completeness means that every borrower is connected to every lender, and every lender is connected to every borrower. This feature follows from the fact that during every interim stage all active borrowers offer the same (maximal available) interest rate to all active lenders. Simultaneously, all active lenders ask the same (minimal available) interest rate from all active borrowers. As transactions are concluded at the midpoint of reservation prices (assumption 2), each active borrower is strictly better off by borrowing from all the active lenders, which implies either larger loan at the same interest rate, or more favourable rate on the loan of the same size. By the same token each active lender is strictly better off by lending to all the active borrowers. As it is demonstrated in section 4.7, aggregation of these simple interim phase networks leads to complicated network geometry.

Finally, in stage four the reservation rates of active borrowers an lenders are updated with formulas (4.4) and (4.5) as if the transactions agreed upon were already concluded. Next the algorithm iterates through the four stages of the interim phase until there are no more creditors or debtors who are willing to trade. When this condition is met, all the agreed trade takes place. This ends the interim phase.

The second stage of the algorithm is the update phase. Here the deposits and risky assets prices are updated for each bank. Solvent debtors pay back inter–bank loans with an interest. For each insolvent debtor the creditor receives $100(1-\theta)\%$ of the principal. Finally equity, reserves, net cash and reservation interest rates are also recalculated. The next section describes calibration of this network formation model.

4.6 Calibration

Twelve different parameters are required to simulate market formation with the proposed protocol. The following section describes how their values were selected. Unless stated otherwise, the sources quoted below were accessed in July 2013. For the sake of brevity, the numbers obtained during calibration were often rounded.

The banking system simulated in this paper comprises of $N := 35$ banks whose relative assets match those of the top US–chartered commercial banks having total assets of at least 300M USD. This choice is motivated by a fact that, according to the Board of Governors of the Federal Reserve System large commercial banks data (accessed on 31–st December 2012), the share of 35 largest institutions in the aggregate assets of all 1700 banks amounts to 92.6%. High concentration of banking system is also typical for many European countries. In example, Blåvarg and Nimander (2002) report that four large Swedish banks constitute at least 80% of the entire domestic banking sector. A network of 35 banks for a consolidated market should be sufficient to model the bulk of the industry without excessive computational burden.

Loss Given Default (LGD) is calibrated to $\theta := 0.05$. Kaufman (1994) reports that LGD amounted to only 5% in the case of Continental Illinois, which is also the value used by Georg (2013) in his simulation of German inter–bank market. Given that banks become insolvent if their equity to risk–weighted asset ratio falls below 4%, this level corresponds to an overnight loss of borrower's total assets of approximately 7%. As risky assets display high persistence and low volatility, this event seems highly unlikely. Thus the assumed values may be regarded as conservative.

The multiplicative *cost of trading* is set to $c := 0.997$. Thus risk–averse banks assume that each transaction that involves risky asset entails a small additional cost due to price slippage. I assume that prior the simulation takes place the bankruptcy of commercial bank in a population of 100 banks happens on average once every five

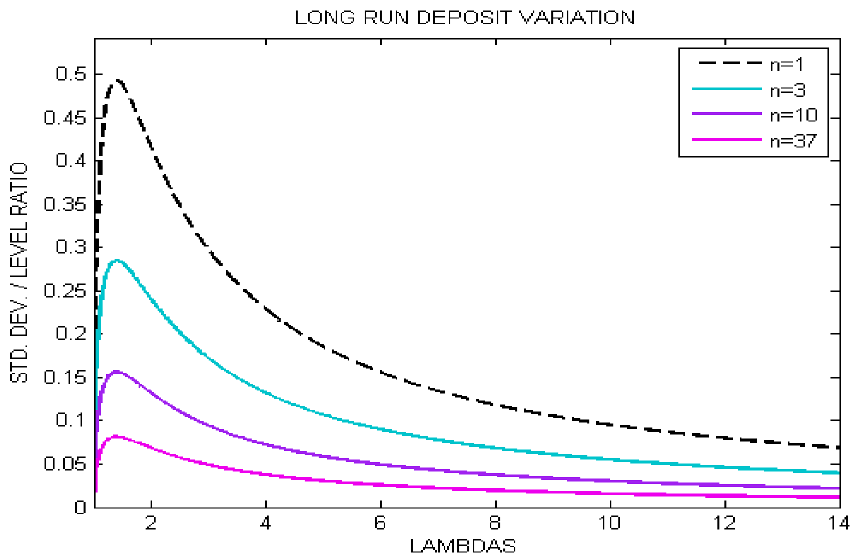


Figure 4.1: Long run relative variation of deposits as a function of parameter λ, plotted for $n \in \{1, 3, 10, 37\}$ agents in the region.

years, which (given 260 working days per year) implies $p := 10^{-6}$. The later indicates that inter–bank loans are perceived as (virtually) riskless.

According to the Federal Reserve Requirements (Regulation D) reserve rate of 10% applies to all deposits above 71 M USD threshold. As all the members of the depicted system are *large*, the deposits in the simulation are all subject to 10% reserve rate. Starting equity to assets ratio of each bank amounts to 11.5%. This figure corresponds to the share of asset minus liabilities residual in total assets of the US banking system according to H.8 statement of the Board of Governors of the Federal Reserve System (31–st December 2012).

At the beginning of every simulation *risk aversion* parameter γ_k for every bank $k \in \{1, \cdots, N\}$ is drawn from a uniform distribution $U(\gamma_{\min}, \gamma_{\max})$. The values of $\gamma_{\min}$ and $\gamma_{\max}$ were set to, respectively, 55 and 60, as this range delivers plausible magnitude of inter–bank overnight interest rates. While they may seem unusually high, this range of parameters ascertains that banks displaying constant absolute risk aversion are able to make economically meaningful choices over daily inter–bank interest rates that are typically very small.

Parameter λ represents the expected duration of deposits. The average volume of deposits held by bank k is given by λh_k, where h_k stands for the size of corresponding region. Let H_k be a random variable, depicting net deposits placed in bank k at given point in time. Standard deviation of H_k depends both on parameter λ and on the number of bank customers in given region, denoted by n. In the long run it may be approximated by

$$\mathbb{V}ar H_k = \frac{h_k^2}{n}\Big(1 - e^{-2(\lambda-1)}\sum_{j=0}^{+\infty}\frac{(\lambda-1)^{2j}}{(j!)^2}\Big).$$

Figure 4.1 depicts the ratio of standard deviation of net deposits to their volume as a function of λ.

Average deposit duration of $\lambda := 11$ days may be obtained from savings rate. Assume bank customers spend money with constant intensities and the only available mean of saving are current accounts. Given 21 working day per month the customers who are expected to collect their deposits after exactly 10.5 working days have on average no savings. Those who are expected to collect their deposits half a day later run a surplus of $0.5/21 \approx 4.76\%$. According to the most recent OECD household survey data[1] the saving rate of US and UK households amount to, respectively, 4.0% and 5.4%. This range of values corresponds approximately to the calibrated parameter value. If each region hosts $n := 10$ bank clients, then for given λ the long term deposit variance attains a moderate value of 2.52%. Note that as the model does not cover consumer behaviour, the only role of deposits is to cause net cash variability which aids the existence of inter–bank lending. For a deposit intensity of $\lambda = 12$ monthly saving rate increases on average to $1.5/21 \approx 14.28\%$, which is more appropriate in case of France (household saving rate of 15.8%) or Spain (13.6%).

Price process parameters are set to $\mu := 0.0037$, $\eta := 0.0369$ and $\sigma := 0.0004$. These values are estimated by fitting formula (4.1) to synthetic index that replicates the relative aggregate asset composition of US–charted banks, revealed in statement H.8 of the Board of Governors of the Federal Reserve System. Mortgages, mortgage backed treasury bills, commercial and commercial real estate loans constitute 85% of the index. They are assumed to be risk–free and yield an annual interest rate of 4.5%. US treasury bonds contribute 6% to the weight of the index. Their prices are substituted with the prices of 10–year benchmark US bonds. Foreign bonds drive the remaining 9% of synthetic index dynamics. Their prices are replaced with Markowitz risk–minimizing portfolio of 10–year benchmark bonds obtained for EU, Germany, UK, France, Italy and Japan. The index does not account for inter–bank loans. While only the bond components are allowed to display volatility, the index represents a portfolio which is far less diversified, and thus most likely more volatile, than the actual portfolios of US banks.

[1]Source: OECD Economic Outlook No. 91, June 2012.

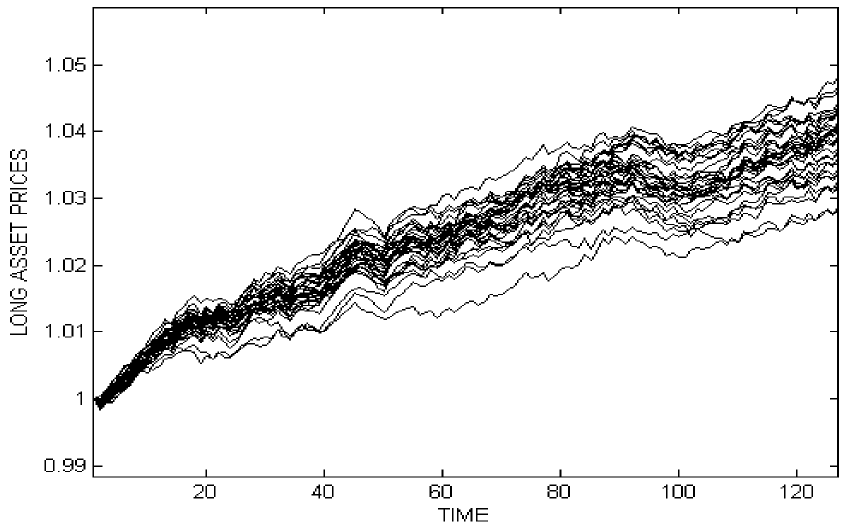


Figure 4.2: Simulated risky asset values recorded by $N = 35$ banks during 126 working days (6 months). In this instance no bank became insolvent.

Any bank becomes insolvent if its ratio of equity to risk weighted assets falls below 4%. Given the construction of the synthetic index that replicates the asset composition of US–chartered banks, risk weighted assets are computed as $0.35 \times 0.85 \times a_{k,t} + 0.2 \times l_{k,t}$. Here $a_{k,t}$ and $l_{k,t}$ are the volumes of risky asset and loans to other banks (held by bank k at time t), 0.2 is the mandatory weight of short–term unsecured loans, 0.35 is the risk weight that applies to mortgages[1], 85% is the share of mortgages and mortgage backed securities in the risky asset. The remaining 15% of the volume of the risky asset consists of sovereign debt, treated as risk–free.

To obtain $\nu := 0.0027$ assume for a moment that banks hold only risky asset. In case where all banks have the same equity to assets ratio there is a constant which, multiplied by risky asset returns of given bank, yields equity growth rate of this bank. If financial markets are efficient, then equity growth rates are equal to stock price growth rates. In equation (4.1) the term I_t stands for a component of risky asset returns common to all banks. Under the assumptions made above to replicate the average correlation rate of individual banks' portfolios it is enough to select ν for which average correlations of banks' stock growth rates and simulated risky asset returns are equal. The average 5–year correlation rate of daily stock returns, obtained for 10 largest publicly listed US–chartered banks, amounts to 68.5%. The selected parameter value for the banks sizes calibrated to the US market yields average correlation of returns equal to 70.2%. The dynamics of individual risky asset

[1]Provided they account for less than 60% of the value of the associated property.

prices driven by equation (4.1) with all parameters equal to their calibrated values is depicted on Figure 4.2.

The calibrated tuple $(\mu, \eta, \sigma; \nu)$ implies that the prices of risky assets display two important empirical features. They are both persistent[1] and dependent on the excess demand or supply generated by the other market participants. The latter is particularly important as it enables positive feedback mechanism where excess demand for or supply of risky assets constitutes an externality, which affects financial position of given bank. What formula (4.1) does not take into account is moderate volatility clustering, displayed by the synthetic index. However, as to derive the model we only need unconditional distribution of asset returns to be Gaussian, the framework proposed here could be easily extended.

The next section presents simulation results.

4.7 Results

This section describes the geometry of the inter–bank market, obtained with the proposed network formation protocol. It also investigates the stability of the entire banking system in two scenarios. The first scenario assumes that all the banks are initially endowed with the same expected amount of assets. In the second scenario the expected assets of consecutive banks are calibrated to reflect heterogeneity, similar to the US banking system. In both cases all the observed bankruptcies arise endogenously. All the figures depicting networks were plotted with Pajek software package, described in de Nooy et al. (2011). Unless stated otherwise, all the simulation parameters are calibrated with values from section 4.6.

The following network terminology is employed in the remaining part of this section. An *inter–bank network* consists of set of vertices (banks) connected by *links* or *connections* (overnight loans). In every transaction there is a borrower and a lender, hence all the links are *directed*. A convention followed here is that links run from creditor to debtor, just as the corresponding money flow. The *weight* of a link (volume of a loan) indicates significance of given link. *Degree* of a vertex is the number of links that either originate (*in degree*) or terminate (*out degree*) in the given node. As the network formation protocol delivers net lending between two parties, there can be no more than one link between each pair of banks. The ratio of all links formed in a network to all potential connections is termed network *connectivity*. *Assortativity* coefficients indicate how probable vertices of different types are to form

[1] AR(1) coefficient estimated for the constructed synthetic index is significant at 5% level.

links with each other.

To investigate the geometry of inter–bank lending 600 networks were generated, each depicting the development of financial linkages after one month (21 working days) of trading. As all the relevant quantities are normalized each turn (the model is stationary), inter–bank network characteristics stabilize after a short period in which all the banks are allowed to adjust their portfolios. A limited time span of this simulation rules out insolvencies that affect[1] the maximum feasible number of connections. To obtain sample correlates of systemic risk the simulation was run 200 times, the duration of each experiment amounted to 6 months (126 working days).

4.7.1 Homogeneous bank sizes

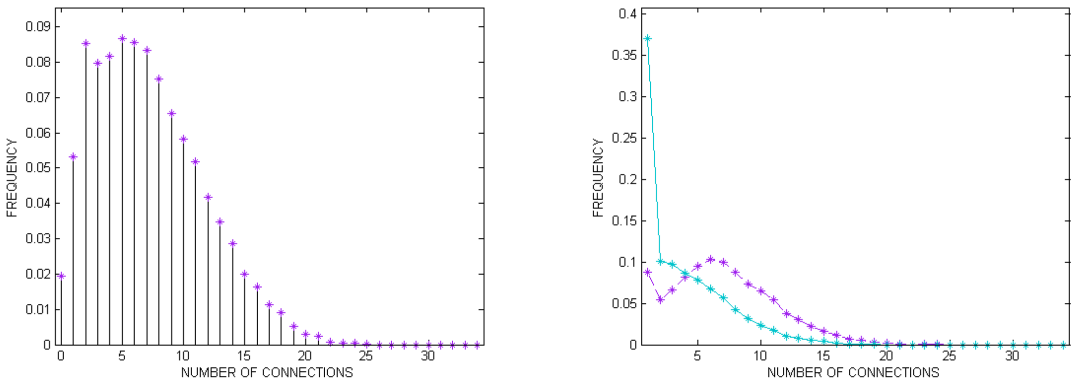
The first investigated case is the benchmark where the expected assets of each bank are equal and amount to a single unit.

The simulated degree distribution is depicted by Figure 4.6 (a). While this chart takes into account the total number of connections, it does not distinguish between lenders and borrowers. The fraction of unconnected vertices amounts to 1.93%. Mean vertex degree (including non–connected nodes whose degree is zero) is equal to 7.18, average ratio of formed links to all feasible links is 10.56%. If sample degree distribution followed Erdös and Rény (1959) random attachment model, then for given connectivity the mean vertex degree should be less then 3.6. Conversely, connectivity which corresponds in this model to the sample mean vertex degree is 21.12%. Hence the simulated networks entail a degree distribution with heavier tails than the one implied by purely random networks. In result few banks trade with many counterparts while majority trades with few.

Connectivities of complete, ring, and star–shaped networks of N vertices amount to, respectively 1, $1/(N-1)$ and $1/N$. In a network of 35 banks these are: 100, 2.9 and 2.8 percent. The simulated lending networks are therefore less interconnected than complete, but markedly more than ring or star–shaped networks. The degree distribution generated by network formation protocol is hump–shaped and appears to be unimodal. Each simulation returns one connected graph.

The aggregate ratio of loans to assets in the simulated banking system amounts to 2.89%. This is more than twice the volume of inter–bank lending expressed as a fraction of total consolidated assets of US–chartered banks, which amounts to approximately 1.36% (31–st December 2012, only asset categories present in the model are being included).

[1]Bankrupted banks do not trade.



(a) Simulated degree frequency.

(b) Borrowers (solid cyan line) vs. lenders (dashed violet line).

Figure 4.3: Degree frequencies simulated for homogeneous bank sizes. Plots obtained: (a) including non connected vertices, (b) conditional on bank k being either a borrower or a lender.

Figure 4.6 (b) presents the differences in simulated degree distributions between borrowers and lenders. In the sample 4.74% of the vertices are exclusively lenders, 46.63% are only borrowers while 46.70% of all the banks are simultaneously creditors and debtors. As this last group is particularly large, instantaneous bankruptcy cascades are both possible and likely. The considerable amount of lenders who are also borrowers may arise when numerous banks in the course of trading either cover their entire net cash deficit or exhaust all their cash surplus. As hitting a boundary implies their reservation interest rates are updated (in a non–linear fashion), these banks may become most desirable counterparts on the opposite side of the market in one of the following interim stages. These figures also suggest that the depicted inter–bank market is a lender market with large excess credit demand. Mean borrower degree amounts to 3.85 while maximum borrower degree is 19. Mean lender degree is 6.98, maximum lender degree is 24. Hence in the entire banking system there is relatively more debtors with smaller number of connections and relatively fewer more interconnected creditors. Within the sample the lenders are relatively more likely to have 1–4 connections while the borrowers more often have 5–20 links.

A typical characterization of debtor–creditor pair is one of network characteristics that may affect stability of the entire system. As expected asset volumes of all the banks are initially equal, the main factors that distinguish different market participants from each other are their risk aversion and risk perception. While for given bank the former may be expressed with CARA coefficient, that latter is captured

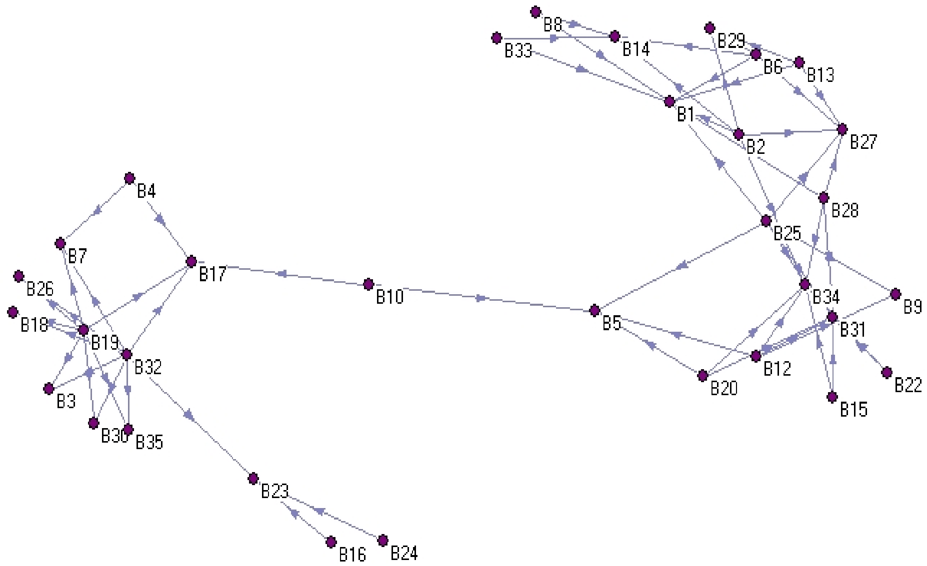


Figure 4.4: Inter–bank overnight lending network plotted with Pajek. All banks are solvent, non–connected banks are omitted. The layout of the vertices was determined by Fruchterman–Reingold (2D) energy minimization algorithm.

with sample variance of returns. Each of those characteristics may be termed either *low* or *high*, depending on whether its value falls below or above sample median. The entries of matrices A_g and A_s below are empirical frequencies, computed form sample of $71,189$ links formed during the simulation. They describe how often pairs of banks that differ with respect to risk aversion and risk perception form links with each other

$$A_g = \begin{pmatrix} 0.195 & 0.375 \\ 0.164 & 0.267 \end{pmatrix}, \qquad\qquad A_s = \begin{pmatrix} 0.236 & 0.266 \\ 0.249 & 0.245 \end{pmatrix}.$$

In A_g the intersection of first row and first column contains probability that a pair of connected banks consists of a borrower and lender who both have low (i.e. below sample median) gamma parameter. The intersection of first row and second column conveys the probability that borrower has low while lender has high (above sample median) level of gamma. The intersection of second row and first column contains probability that borrower has high while lender has low gamma parameter. Finally, the intersection of second row and second column conveys the probability that both borrower and lender have hight level of gamma. Matrix A_s carries a similar information for risk perception.

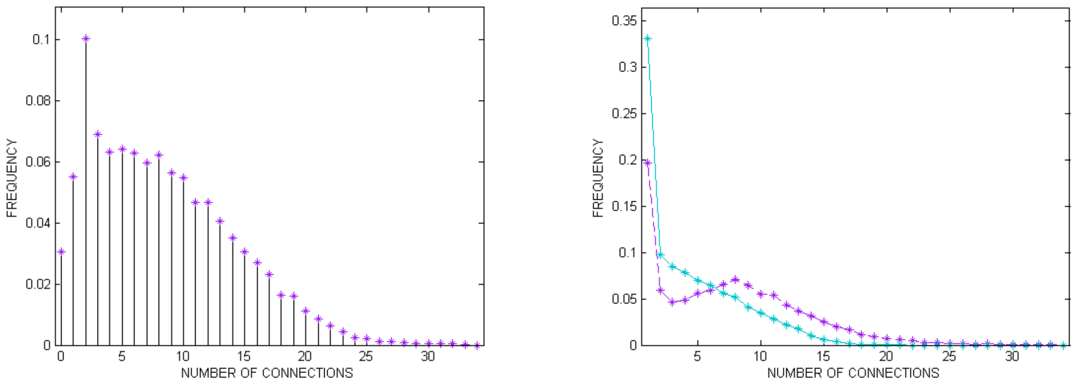
Assortativity reveals the propensity of the agents' to conclude transactions with the counterparts whose characteristics are different to their own. It captures the important feature of real world trade networks where involved parties differ in a much more fundamental way than just with the respective value of demand, supply or initial endowment. This is because the latter three quantities arise endogenously and stem from the fact that the agents are heterogeneous.

In order to assess significance of sample assortativity matrices, their largest entries may be compared against the simulated quantiles of matrix supremum norms. Critical values for the largest matrix entry simulated for a population of 10,000 links at 10, 5, 1 and 0.1% levels amount to 25.85, 25.97, 26.22 and 26.52%. As the simulated critical levels decrease with a number of links in each sample, these values may be considered as conservative. Both A_g and A_s matrices are significant at 0.1% level. Note that as the entries are dependent (sum up to one), significance of empirical frequencies has to be tested jointly.

In case of both risk aversion and risk perception it is more likely that borrower and lender are of a different than of same types. The most common pairing consists of a borrower who is less risk averse and perceives lower level of risk than his counterpart. This result is consistent with basic economic intuition: the risk averse banks with conservative assessment of investment risk regard interbank loans as desirable asset and are thus more likely to grant loans on the inter–bank market. Simultaneously, the risk–prone institutions who believe that the risk is low are more inclined to purchase the risky asset. This also means that the risk–prone debtor expects higher benefits in favourable scenarios. The least probable pair consists of highly risk averse borrower and less risk averse lender who both expect low variance of unconditional risky asset returns.

The fact that agent's linking decisions are assortative has two important implications. First, it implies that financial linkages are dependent on characteristics of both involved parties, and thus are not random. Second, if a risk–prone bank which perceives investment risk as low suddenly defaults, then the institutions hit most by its collapse are likely to be risk–averse and to perceive investment risk as high. This is precisely the group of banks that is worth saving by financial authorities in case of major financial crisis.

Figure 4.4 depicts an example of a network of inter–bank connections, created endogenously by the proposed algorithm. Network formation protocol delivers a complex system of financial linkages in which creditors typically reside in the centre of star–shaped formations and lend to banks, located on the periphery. Furthermore, the borrowers form fewer links than the lenders.



(a) Simulated degree frequency.

(b) Borrowers (solid cyan line) vs. lenders (dashed violet line).

Figure 4.5: Degree frequencies simulated for bank sizes calibrated to US market. Plots obtained: a) including non connected vertices, b) conditional on bank k being either a borrower or a lender.

4.7.2 Calibrated bank sizes

Next investigated case is when sizes (in terms of total assets) of the banks in the system where calibrated to match top 35 banks, active on the US market.

The simulated degree distribution is depicted in Figure 4.5 (a). While this chart takes into account the total number of connections, it does not distinguish between lenders and borrowers. The fraction of unconnected vertices amounts to 3.05%. Mean vertex degree (including non–connected nodes degree of which is zero) is equal to 8.25, average ratio of formed links to all feasible links is 12.13%. Connectivities of complete, ring, and star–shaped networks for 35 vertices amount to, approximately, 100%, 2.9% and 2.8%. The networks obtained in the heterogeneous case are again far less interconnected than complete networks, but markedly more than ring or star–shaped networks. Each simulation returns one connected graph.

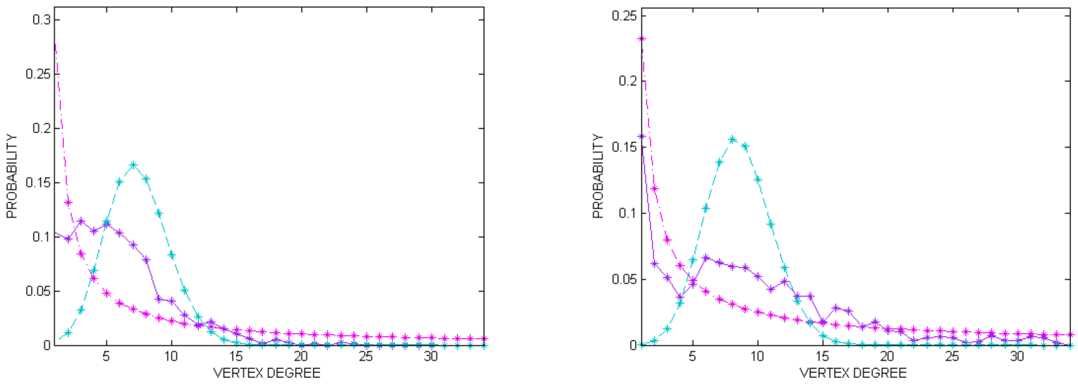
Although the magnitude of the simulated networks connectivity is too large for entire banking systems, it lies in the range typical for subnetworks or clusters of these systems. In example, Müller (2006) reports that while Swiss inter–bank network connectivity amounts to only 3%, it soars to 27% for the subnetwork of cantonal banks. Average connectivities of CHAPS Sterling and *Giant Strongly Connected Components*[1] of CHAPS Sterling and Fedwire overnight inter–bank networks provided by Becher et al. (2008) amount, respectively, 88%, 5.1% and 0.3%.

[1]See Becher et al. (2008) for more details.

Figure 4.6 presents the simulated degree distributions for the both investigated cases, conditional on vertex having at least one connection. This empirical density is plotted against *scale–free* and conditional *purely random* degree distributions, fitted to match the simulated sample mean. Estimated values of characteristic exponent for scale–free density amount to 1.11 for homogeneous and 0.97 for calibrated bank assets. Both numbers are low and imply fat–tails. As scale–free distribution does not allow for vertices with no connections, purely random density is conditioned on vertex having at least one counterpart.

The degree distribution generated by the network formation protocol is not scale–free. Empirical densities depicted on Figure 4.6 (a) and (b) are not monotonously decreasing, they also have lighter tails than the scale–free distributions with the same average vertex degrees. Sample density obtained for homogeneous banks sizes appears to be hump—shaped, while the distribution simulated for the banks sizes calibrated to the US data seems to be multimodal. Furthermore, the densities generated by the network formation algorithm do not comply with Erdös and Rény (1959) random attachment model. Both simulated distributions are not symmetric around mean vertex degree, their tails are also heavier than in binomial density. Connectivity of purely random network is identical to the estimated probability of forming a link. Hence if the network simulated for the calibrated bank assets was purely random, its connectivity would be 24.26%, which is exactly two times more than the actual sample value. The chart indicates that both simulated inter–bank networks are somewhere in between purely random and scale–free networks. Sample density obtained for the case with the calibrated bank assets oscillates around scale–free pdf and intersects with it four times. Hence it may be regarded as approximately scale–free.

The shape of degree distribution presented above conforms with the findings of Iori et al. (2008), who observed that vertex degree distribution on Italian segment European e–MID market is neither scale–free, nor purely random. It contrasts with a more recent picture of the entire e–MID market (250 banks) provided by Cohen-Cole et al. (2013), who found that degree density of European inter–bank market is approximately scale–free and thus (approximately) monotonously decreasing. The similar results were obtained by Soramäki et al. (2007) who demonstrated that vertex distribution for US Fedwire market was close to scale–free with characteristic coefficient equal to approximately 2.2. The difference between these empirical results and the shape of the simulated densities may be attributed to two factors. First, the investigated system consists of only 35 major banks. As noted by Blåvarg and Nimander (2002), in more concentrated systems large banks, who typically trade significant



(a) Homogeneous assets. (b) Calibrated assets.

Figure 4.6: Simulated degree distribution, conditional on vertex having at least one connection (violet solid line). Degree distributions: scale–free (magenta dash–dotted line) and random (cyan dashed line).

volumes of assets, have fewer alternative counterparties. Thus a network, consisting only of large banks, is necessarily much more interconnected than the system, where majority of banks is either small (eMID with 250 banks) or negligible (Fedwire, 1700 banks). When small banks who typically form only few links are being omitted, low degree vertices are under–represented and relative frequency of vertices with large number of links is weighted up. Next, in the proposed algorithm all the market participants are endowed with full information about the other members of the banking system. In result they find it optimal to conclude more transactions of smaller volume. These two factors lead to endogenous formation of highly interconnected networks with distribution that may be approximately humped–shaped (just as in homogeneous case).

The simulated average aggregate volume of loans as a fraction of total borrower assets amounts to 4.46%. The same quantity expressed as a percentage of total lender assets is equal to 8.02%. Hence in the investigate system it is the smaller banks who are crediting their larger counterparts. Empirically this is demonstrated by Müller (2006). The average total share of loans in the aggregate assets equates to 3.44%, which is 2.5 times more than the share of inter–bank loans in consolidated assets of US–chartered banks (1.36% on 31–st December 2012). However, while inter–bank market plays secondary role in the US, its importance is much more pronounced in some European countries. In example, Degryse and Nguyen (2007) report that in their data inter–bank loans amount to 20–30% of assets and 30–40% of Belgian bank liabilities.

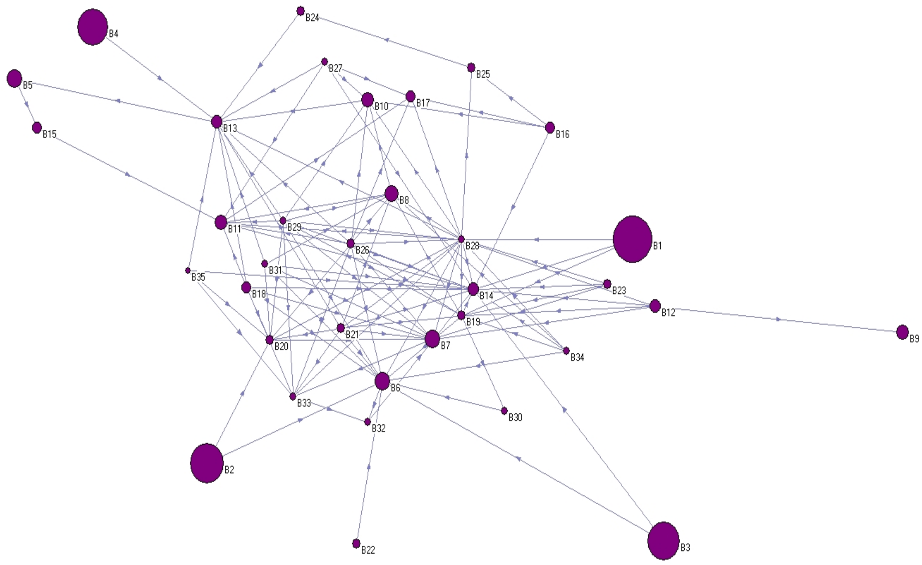


Figure 4.7: Example of inter–bank overnight lending network plotted with Pajek. Non–connected and insolvent banks are omitted.

Figure 4.5 (b) presents the differences between borrowers and lenders in simulated degree distributions. In the sample 5.09% of the vertices are exclusively lenders, 42.29% are only borrowers, while 49.58% of all the banks are simultaneously creditors and debtors. As this last group is particularly large, instantaneous bankruptcy cascades are both possible and likely. Mean borrower degree amounts to 4.49 while maximum borrower degree is 21. Mean lender degree is 7.55, maximum lender degree is 33. Hence in the entire banking system there is relatively more debtors with smaller number of connections and relatively fewer more interconnected creditors. Within the sample the borrowers are relatively more likely to have 1–5 connections while the borrowers more often have 6–24 links.

A typical characterization of debtor–creditor pair is one of network characteristics that may affect stability of the entire system. In this model the main factors that distinguish different market participants from each other are their: size (in terms of total assets), risk aversion (CARA coefficient) and risk perception (sample variance of returns). Each of those characteristics may be termed either *low* or *high*, depending on whether its value falls below or lies above sample median. Empirical frequencies in the matrices

$$A_h = \begin{pmatrix} 0.226 & 0.303 \\ 0.212 & 0.258 \end{pmatrix}, \qquad A_g = \begin{pmatrix} 0.199 & 0.335 \\ 0.191 & 0.270 \end{pmatrix}, \qquad A_s = \begin{pmatrix} 0.231 & 0.267 \\ 0.246 & 0.249 \end{pmatrix},$$

describe how often in the generated networks pairs of banks that differ with respect to size, risk aversion and risk perception form links with each other. The frequencies summarize the characteristics of $81,115$ inter–bank links, formed during the simulation. In A_h the intersection of first row and first column contains probability that a pair of connected banks consists of a borrower and lender who both have low level of assets. The intersection of first row and second column conveys the probability that borrower has low while lender has high level of assets. The intersection of second row and first column contains probability that borrower has high while lender has low level of assets. Finally, the intersection of second row and second column conveys the probability that both borrower and lender have hight level of assets. Matrices A_h and A_s carry a similar information for risk aversion and risk perception.

The largest entries of all three matrices are significant at 0.1% level. The most important factor that differentiates banks' linking patterns is their risk aversion and next volume of their assets, risk perception parameter is somehow less important.

In the simulated networks the lenders are more likely to perceive investment risk as *high*. The borrowers who participate in the inter–bank trade more often are *small* and display *high* risk aversion while the lenders are *large* and display *high* risk aversion. In line with the empirical results by Cocco et al. (2009) banks tend to lend or borrow from the institutions of different size to themselves. A typical debtor–creditor pair consists of a borrower with low level of assets, low risk aversion and low risk perception and a lender with high level of assets, high risk aversion and risk perception. The least probable pairing consists of small risk–loving lender who perceives risk as low and large risk–averse borrower who perceives investment risk as low.

Just as in the homogeneous case, financial linkages are dependent on characteristics of both involved parties and thus are not random. Again if a risk–prone bank which perceives investment risk as low suddenly defaults, then the institutions hit most by its collapse are likely to be risk–averse and to perceive investment risk as high. By tempering the exuberance of the first group, a prudent banking system supervisor would protect the latter.

Figure 4.7 depicts an example of a network of inter–bank connections, created endogenously by the proposed algorithm. Network formation protocol delivers a complex system of financial linkages in which banks with more assets (typically of lower indices) reside on periphery of star–shaped formations and borrow from smaller banks, located in the centre. The relatively less numerous lenders form more links and are endowed with a smaller volume of assets. They are more exposed to counterparty default, as the volume of credit they grant constitutes more significant portion of their total assets.

The networks simulated for relative quantities of assets calibrated to the US market have a degree density that displays heavier tails than the networks obtained for homogeneous case. Their sample vertex distribution is also much closer to scale–free density and thus more realistic. Both instances yield similar market structures, with few smaller lenders granting numerous loans and numerous larger borrowers taking fewer loans. As the networks generated in heterogeneous case are more interconnected, for a sufficiently mild crisis they are expected to be more robust than the market configurations, obtained in homogeneous case. However, heterogeneous system would be much more vulnerable to cascades of insolvencies if the crisis was sufficiently severe. It could be also conjectured it would display much more vehement transition from the state where all the banks are solvent to the state where a significant fraction of the inter–bank market is bankrupt.

4.7.3 Correlates of systemic risk

This section investigates the correlates of systemic risk. Contrary to other the bulk of literature on inter–bank network simulation, the bankruptcies investigated here are all endogenous events. A single insolvency signals that the entire system is distressed and thus it is more likely to be followed by a cascade of bankruptcies. In order to generate sufficient number of insolvencies the equity to asset ratio was set to 4% of all the banks. This figure corresponds to the average relative equity of US banks in the eve of 2007 financial crisis. Furthermore, the parameter μ in formula (4.1) was reset to -0.0037. Hence what is modeled in this exercise is a *moderate* crisis where the banks in the short term (simulation time span) on average loose money on their investment, but the returns on investment itself do not become more volatile.

Table 4.7.3 contains sample correlation rates of system aggregates with three indicators of banking system stability. These indicators are: aggregate equity (denoted as *Eq.*), number of insolvent banks (*Ins.*) and a share of assets of bankrupted banks in total assets of the banking system (*Sh.*). The components of both the assets and the liabilities are normalized to one. Funding liquidity is approximated in the model by an average volume of inter–bank loans, expressed as a fraction of banking system assets. Market liquidity is approximated with an average relative contribution to the risky asset price of the excess supply or demand for this asset, generated by the banks. Note that probability of default may be estimated on the sample as a proportion of bankrupted banks to all banks. Hence correlation rate between any quantity and a number of insolvent banks is (for sufficiently large sample) identical to correlation rate between given quantity and the estimated probability of de-

	Bank assets					
	Homogeneous			Calibrated		
Variable	Eq.	Ins.	Sh.	Eq.	Ins.	Sh.
Deposits	-0.527****	-0.334****	-0.333****	0.565****	-0.925****	-0.937****
Risky asset	0.737****	-0.975****	-0.975****	0.852****	-0.994****	-0.975****
Reserves	-0.527****	-0.334****	-0.334****	0.565****	-0.925****	-0.937****
Loans	-0.685****	0.985****	0.985****	-0.840****	0.995****	0.977****
Cash	-0.297****	0.501****	0.501****	-0.297****	0.369****	0.373****
Borrower share	-0.943****	0.695****	0.696****	-0.908****	0.972****	0.950****
Lender share	-0.924****	0.805****	0.806****	-0.924****	0.888****	0.869****
Interest rate	0.814****	-0.886****	-0.886****	0.900****	-0.884****	-0.829****
Leverage	-0.973****	0.756****	0.756****	-0.977****	0.926****	0.884****
Funding liquidity	-0.067	-0.021	-0.021	-0.226****	-0.090*	-0.116**
Market liquidity	-0.394****	0.602****	0.602****	-0.668****	0.914****	0.918****
Prob. of def.	-0.614****	1.000****	1.000****	-0.826****	1.000****	0.989****
Equity	1.000****	-0.614****	-0.615****	1.000****	-0.826****	-0.771****

Table 4.1: Simulated correlation rates of aggregates. *Eq.* stands for equity, *Ins.* for total number of observed insolvencies, *Sh.* for insolvencies as a share of assets of the banking system. A number of 1–4 stars denote significance levels of, respectively, 10, 5, 1 and 0.1%.

fault. One to four stars denote rejection of null hypothesis that sample correlation rate is insignificant at significance levels of, respectively, 10, 5, 1 and 0.1 percent. The table indicates that most of the estimated correlation coefficients is significant at 0.1% level. As normality assumptions underlying Pearson's test are not fulfilled, these values may be only treated as reference. The results obtained were as follows.

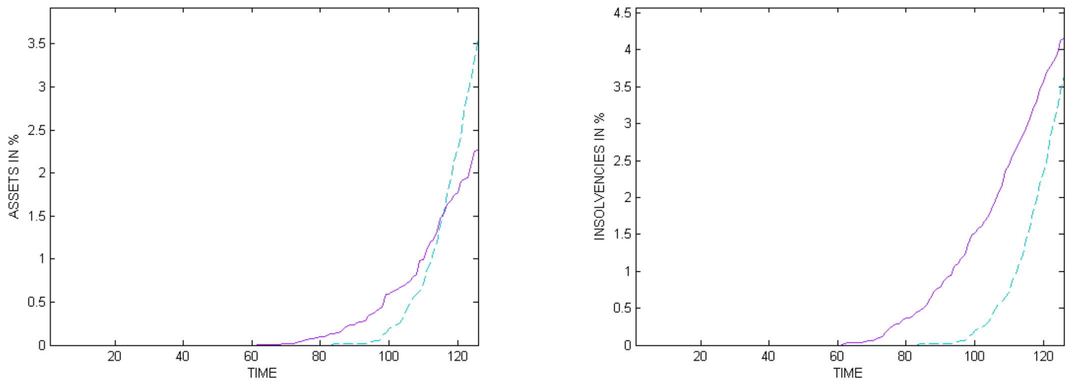
The aggregate results allow us to identify two important subsets of aggregates. The quantities that are positively correlated with equity and negatively with a number of insolvencies and share of bankrupted banks in total assets are the variables whose high values characterize a system that is more robust. These aggregates are: share of equity and deposits in liabilities, share of risky asset and reserves in assets and inter–bank interest rate. Deposits converted to risky asset constitute the main source of profits. Retained profits contribute to equity and thus make the banks less susceptible to insolvency. Reserves are by definition proportional to deposits. High levels of the aggregate interest rate make inter–bank lending more profitable and thus increase equity of the lenders. Both reserves and equity provide a safety buffer, high levels of which make the bank more resilient to crisis. The quantities that are negatively correlated with equity and positively with a number of insolvencies and share of bankrupted banks' assets in total assets are the variables whose high levels characterize a system that is more fragile. These aggregates are: share of loans in assets

	Bank assets			
	Homogeneous		Calibrated	
Variable	Equity	Solvent	Equity	Solvent
In degree	0.076****	0.163****	0.022**	0.183****
Out degree	-0.350****	0.107****	-0.324****	0.128****
Gamma	0.013	0.028***	-0.047****	-0.074****
Sigma	-0.204****	-0.054****	-0.007	0.023**
Deposits	-0.213****	-0.137****	0.125****	-0.163****
Equity	1.000****	0.360****	1.000****	0.389****
Loans to	0.222****	0.129****	-0.158****	0.151****
Risky asset	0.485****	0.004	0.579****	-0.021**
Reserves	-0.339****	-0.173****	0.085****	-0.182****
Loans from	-0.432****	0.091****	-0.527****	0.119****
Net cash	-0.134****	-0.160****	-0.179****	-0.345****
Equity ratio	0.993****	0.363****	0.991****	0.396****
Leverage	-0.952****	-0.499****	-0.929****	-0.516****
Size	0.040****	0.005	0.070****	0.059****

Table 4.2: Correlation rates simulated for individual banks. A number of 1–4 stars denote significance levels of, respectively, 10, 5, 1 and 0.1.

of borrowers and lenders, the ratio of loans to aggregate assets, net cash, leverage, market liquidity and the estimate of the probability of counterparty default common to all the banks in the system. If risky asset is profitable, large amount of loans in result to the costs of lost opportunity is detrimental to equity of the entire system. Net cash is invested in the risky assets at the end of current period and it decreases equity due to slippage costs. High levels of leverage are associated with high risk and larger losses if the investment will not be profitable. Significant market liquidity is associated with higher volumes of trade and may decrease equity due to trade frictions. High assessment of the probability that the counterpart defaults is characteristic for the systems where large number of bankruptcies has already occurred and thus the aggregate equity is low.

Table 4.7.3 contains results simulated for 7000 individual banks. For each of the two investigated cases the first column contains correlates of consecutive indicators with bank equity. The second column presents correlation of given quantity with a binary variable which denotes that bank is solvent. The latter is equivalent to negative correlation rate with probability of default estimated within the sample. Both the asset and the liability size of each bank's balance sheet is normalized to one. A size of a bank is defined as the amount of assets the bank holds and is treated as *numéraire*. Again the majority of the estimated correlation coefficients is significant at 0.1% level.



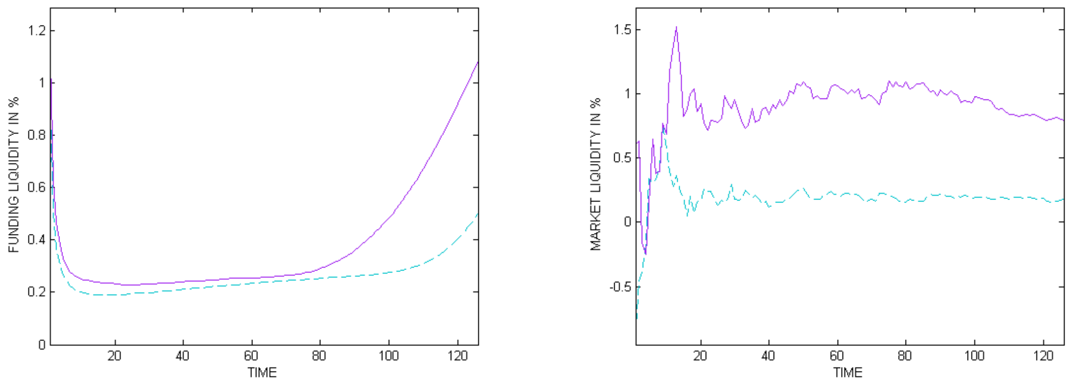
(a) Assets of the system.

(b) Number of insolvencies.

Figure 4.8: Expected: (a) loss of banking system assets, (b) number of insolvent banks for homogeneous (dashed cyan line) and calibrated (solid violet line) bank sizes.

Vertex in degree, bank size (total volume of assets), equity and equity ratio are positively correlated with both equity and the binary index which denotes that the bank remains solvent. Thus a larger bank, especially a one borrowing from numerous sources, is more resilient to crisis. The high relative amount of risky asset, accumulated reserves or deposits held (in heterogeneous case) coincide with high level of equity and lower probability of remaining solvent. As both assets and liabilities are normalized to one, the higher volume of deposits (and thus the higher obligatory reserves) typically implies the lower equity. Volume of loans both to and from given bank, vertex out degree and the risk perception parameter correspond to higher level of equity and lower level of the binary variable which indicates that the bank remains solvent. Thus while inter–bank lending during the simulated crisis decreases equity, it also makes the banks less likely to go bankrupt. Leverage, net cash and risk aversion coefficient gamma are negatively correlated with equity and positively correlated with probability of default.

At the level of individual banks we may distinguish two subsets of variables. The quantities that are positively correlated both with equity and the binary variable equal to one when bank remains solvent contribute to resilience of given bank. These indicators are: vertex in degree, bank size (total volume of assets), equity and equity ratio. The quantities negatively correlated both with equity and the binary variable equal to one when bank remains solvent are the ones whose high levels characterize banks prone to failure. These variables are: leverage, amount of net cash and risk aversion coefficient.



(a) Funding liquidity.

(b) Market liquidity.

Figure 4.9: The moderate crisis dynamics of the indicators of: (a) funding liquidity and (b) market liquidity, plotted for homogeneous (cyan dashed line) and calibrated (solid violet line) bank sizes.

Figure 4.8 depicts as functions of time: sample mean number of insolvent banks and mean share of insolvent banks' assets in the total assets of the banking system. It reveals that after quiet spell of 80 working days the system with homogeneous bank assets enters a much more turbulent phase when a fraction the total number of banks becomes bankrupt. These insolvent institutions represent 3.5% of both the total number of banks and the aggregate assets of the banking system. In the case of calibrated bank assets the crisis starts taking its toll earlier (60 working days), but it is also less harmful. While more banks become bankrupt (4.1%), their total assets account for only 2.2% of the aggregate assets of the entire system. Hence heterogeneity of bank assets makes the entire system more stable and resilient to ruptures.

Figures 4.9 (a) and (b) represent the mean dynamics of the indicators of funding liquidity and market liquidity, recorder during the simulated crisis. Figure 4.9 (a) reveals that after a short period of decline in the amount of inter–bank loans caused by initial portfolio adjustments, the volume of the overnight loan market stabilizes and then slowly increases. This gradual process is abruptly interrupted some time after the first insolvency is observed. Once the banks learn that their portfolios incur losses, this sudden realization causes a flight to quality episode (Caballero and Krishnamurthy, 2008) where they rapidly increase involvement in the less risky inter–bank market. This effect in the case of homogeneous bank assets is triggered somehow later, it is also less pronounced. Figure 4.9 (b) indicates that the contribution of the banking system to the risky asset price peaks in between 10th (homogeneous assets) or 15th (calibrated assets) working day. This effect decays over time and may be also attributed to the cumulated portfolio adjustments. Next the indicators of market liquidity in both investigated cases start to oscillate around a certain fixed level. It amounts to approximately 0.8% for the calibrated bank assets and only 0.2% when bank sizes are homogeneous. Thus when the banks substantially differ in the amount of assets they hold, risky asset market is much more liquid. This fact may contribute to the higher overall systemic resilience.

4.7.4 Parameter sensitivity

The results presented in this chapter were obtained for one combination of model parameters. The objective of the following two series of exercises is to verify the impact of parameter perturbation on both the generated networks and the correlates of systemic risk. In the tests summarized below it is always a single parameter that is being altered, while all the remaining model characteristics are calibrated as in section 4.6.

In order to obtain results comparable with the ones, presented in subsection 4.7.2, the first series of tests was run for 21 periods and repeated 600 times. The qualitative results thus obtained were as follows. Setting $c := 0.999$ slightly decreases connectivity of generated networks. The share of loans in the assets of, respectively, borrowers and lenders falls by 0.29% and 0.9%. Setting $c := 0.995$ has exactly the opposite effect. It marginally increases the connectivity, while boosting the share of the loans in the assets of debtors and creditors by 0.31% and 1.1%. In this second scenario borrowers and lenders not only conclude transactions of larger volume, but they also form more (by 1% and 2%, respectively) links. The source of this effect are larger incentives to become involved in the inter–bank market, which stem from higher transaction costs when the agents invest in the risky asset. When lenders' assessment of the probability of their counterparty overnight default goes up to $p := 0.00001$, the size of the inter–bank market declines by 11% while the network connectivity falls by almost 15%. Further growth of this probability to $p := 0.0001$ has even more devastating effects as the inter–bank market dwindles by total 48% and the connectivity drops by 66%. Simultaneously, mean network degree falls by two thirds while the average vertex degrees of borrower or lender decline by, respectively 15% and 35%. When p is high, the inter–bank loans are perceived as a more risky (and thus less desirable) asset. Assuming $\lambda := 12$ marginally decreases the total inter–bank loan volume and has negligible effect on all the other model characteristics. Larger λ implies lower volatility of the deposits, and thus smaller demand for loans, triggered by transient liquidity shortages. If $\gamma \sim U(53, 54)$ then the connectivity of the generated networks drops by 90.1%. The banks no longer differ with respect to risk aversion, their size and risk perception become the two main factors that determine their linking decisions. If the dispersion of risk aversion parameters is low, the banks are more similar to each other and thus lack incentives to trade. Finally, setting $\gamma \sim U(20, 25)$ increases the size of the inter–bank loans market by 216%, while the connectivity of the generated networks increases by 37%. As the utility functions in this parameter region are more steep, the banks much more fundamentally differ in their reservation interest rates, and thus find it optimal to lend or borrow more on the inter–bank market.

To deliver results comparable with subsection 4.7.3, the second series of tests was run for 126 periods and repeated 200 times. In the seven investigated exercises one model parameter always assumed either different value or range of values ($c := 0.999$, $c =: 0.995$, $p := 0.00001$, $p := 0.0001$, $\nu := 0.05$, $\gamma \sim U(53, 54)$, $\gamma \sim U(20, 25)$), while all the remaining parameters were set as in section 4.6. The qualitative results thus obtained for the banking system calibrated to the US market were the following. At the level of the entire banking system the correlates of: loans, cash, share of the loans in either borrowers' or lenders' assets, aggregate interest rate, systemic leverage estimated probability of default and aggregate system equity were qualitatively stable. Stable was also the correlation rate of market liquidity and the share of risky asset in banks' portfolio with a number of insolvent banks. At the level of individual institutions the correlates of: in degree, out degree, net cash equity ratio and leverage were qualitatively stable. Stable were also the correlates with bank equity of: the share of risky asset in total assets and risk aversion parameter gamma (the latter with the exception of the case when $\nu := 0.05$) and the correlates with a binary variable which indicates that the bank remains solvent of: equity, loans to given bank and reserves (again with the exception of $\nu := 0.05$). The fact that not all correlates of systemic risk prove to be stable is not surprising, as some of the considered exercises are rather extreme (e.g. parameter p is multiplied by a hundred).

4.8 Conclusions

This chapter made the following contributions.

First, the paper proposes a computational model of endogenous network formation designed for the inter–bank market. The algorithm relies on the solution of portfolio problem where banks displaying constant absolute risk aversion maximize their expected utility while simultaneously taking into account price slippage, costs of financing and investment risk. The banks who differ with respect to size, risk aversion and risk perception, form links with their most preferred counterparts. The emergent market structure arises due to banks being heterogeneous. The network formation protocol yields simultaneously the optimal: choice of a counterpart, volume of a loan and agreed interest rate. The generated system may be analysed as a whole, at the level of individual banks or separate transactions. The outcome of network formation procedure is deterministic and pairwise–stable. According to our best knowledge, this is the only fully endogenous (depending solely on agent's characteristics) computational model featuring link formation that simultaneously incorporates investment risk and takes into accounts assets and liabilities of commercial banks.

Second, the proposed model is calibrated to the subnetwork of 35 largest US banks and run to simulate network geometries. In the generated networks banks with more assets typically reside on the periphery of star–shaped formations and borrow from smaller banks, located in the centre. The lenders form more links than the borrowers. Thus they are more exposed to counterparty default as the volume of credit they grant constitutes more significant portion of their total assets. Just as in real world financial networks, degree distribution in the generated networks displays a tail gravity of which is between that of purely random and scale–free network. If the expected initial bank assets are homogeneous, the factor which affect linking decisions of the agents most is risk aversion. If the assets are calibrated to the US market, the factors that affects bank's linking behaviour are first risk aversion and next volume of assets while risk perception seems least important. As there is a fraction of banks who are simultaneously borrowers and lenders, the geometry of the resulting networks allows for instantaneous bankruptcy cascades.

Third, the correlates of systemic stability are being investigated. The quantities that on aggregate level characterize more robust system are: share of equity and deposits in liabilities, share of risky asset and reserves in assets and inter–bank interest rate. Thus one of the ways to make system more resilient to crisis is to incite the banks to rise additional capital. As high inter–bank interest rates may contribute to equity via the retained profits, keeping interest rates artificially high might be detrimental to systemic stability. At the level of individual banks the factors which contribute to resilience of given bank are: vertex in degree, bank size (total volume of assets), equity and equity ratio. These results indicate that the lender of the last resort could mitigate the investigated crisis, by providing loans to the distressed banks. They also suggest one of the reasons behind high banking system concentration might be that the concentrated systems are more robust. The simulation also shows that size matters – for the considered crisis scenario the case where bank assets are calibrated to US market generates a more robust system with less common bankruptcies than the homogeneous case.

The model presented in this work has some interesting extensions. First, the proposed framework could be used to analyse robustness of the banking system under different crisis scenarios. Next, it could be utilized to investigate the impact of fire sales and asset price erosion on systemic stability. Third, the model maybe employed as a simulation tool to quantify the impact of different banking system regulations (such as Basel III accord) on the emergent market structure.

Chapter 5

Summary and Contributions

The contributions of chapters 2-4 may be summarized as follows.

Chapter 2 demonstrates that 1) heavy–tails in macroeconomic data are not an exception, 2) if not properly taken into account, heavy–tails may cause serious problem in both estimation of and inference from econometric models. It also provides an original comparison of different classes of densities, shedding a new light on their applicability in economics. It is argued that out of the diverse approaches to modelling large deviations, the one that is best suited to modelling macroeconomic uncertainties is the approach which employs tempered stable distributions.

Chapter 3 introduces a tractable, non–standard definition of TS distributions (Definition 3.2.1) and its multivariate extension (Definition 3.2.2), which allows for capturing multivariate dependences with covariance matrix. The chapter presents the complete set of formulas for cumulants and moments of TS densities, previously known only for $\alpha \neq 1$. It investigates a number of properties of TS distributions which were known in less general setting ($\alpha < 1$, $\alpha \neq 1$, $\beta = 1$). It introduces the new Cumulant Matching estimator (Proposition 3.4.1) and the novel mixture representation for tempered stable random variates (Proposition 3.5.1). The random number generation algorithm that relies on the latter (Algorithm 2) is valid for the entire admissible parameter range and may be implemented in just one line of code. This algorithm is demonstrated to be the most accurate and the fastest randomization method in the range of parameters which seems appropriate for low frequency macroeconomic distributions ($\alpha > 1$, $|\beta| \neq 1$ and $\theta < 1$). Finally, the chapter applies these new methods to simulate the probabilities of joint currency crises on Russian ruble and British pound. It is demonstrated that the mTS distribution much better represents the standardized multivariate data than the two investigated benchmark multivariate densities (Gaussian and elliptical t). It thus delivers much more realistic estimates of the probabilities of extreme events. All the proofs are relegated to Appendix B.

Chapter 4 presents a computational model of endogenous network formation designed for the overnight inter–bank market. The algorithm relies on the solution of portfolio problem (Appendix D) where the banks display constant absolute risk aversion and maximize their expected utility, while taking into account price slippage, costs of financing and investment risk. The network formation protocol yields simultaneously the optimal: choice of a counterpart, volume of a loan and agreed interest rate. Next the proposed model is calibrated to the subnetwork of 35 largest US banks and run to simulate network geometries. In the generated networks banks with more assets typically reside on the periphery of star–shaped formations and borrow from smaller banks, located in the centre. The lenders also form more links than the borrowers. The simulated degree distribution displays a tail gravity of which lies between that of purely random and scale–free network. The outcome of network formation procedure is deterministic and pairwise–stable. The chapter also investigates the correlates of systemic risk both at the level of individual banks and the entire banking systems. According to our best knowledge, the proposed novel framework is the only available fully endogenous (depending solely on agent's characteristics) computational model of network formation in banking systems.

The proposed new methods have certain limitations.

Tempered stable distribution still lack a reliable estimation technique that would not resort to numerical approximations. This problem is particularly pressing in the case of their multivariate extension. Thus a multivariate approach that would rely on direct estimation of spectral measures could be here especially useful.

The main drawback of the proposed network formation protocol is that it employs the estimate of the probability of counterparty default that is common to all the banks in the system. Hence, in particular, it can not depend on any characteristics of the debtor. While it is possible to relax this assumption, it would result in a much more complicated model, a solution of which through the interim stage iterations would be highly computationally demanding. Such a modified algorithm would generate in its interim phases disjoint inter–bank lending networks, which could possibly lead to the final overnight inter–bank lending network also being disjoint. This last feature would not necessarily be desirable as it is most likely counterfactual. Another assumption which might cause concerns is the full information of the agents on the characteristics of their counterparts in the system. While it seems to be a sufficiently good first approximation, a more realistic incomplete information framework would be more than welcome. A serious disadvantage of the model is that it allows for only one maturity of the loans. Thus it can not depict a build-up of maturity mismatch, which is one of the potential sources of fragility of banking systems. Any framework, attempting to overcome this obstacle, would have to combine optimization over multiple periods with modelling banks' intertemporal choices. At the moment it is neither clear, how to construct such an extension, nor what would be its impact an the stability and dynamics of the entire system. Finally, the model is computational, and thus not fully analytically tractable. Furthermore, the numerical cost of analysing large inter–bank systems may often prove to be prohibitive.

Appendix A

Algorithms

A.1 Random number generation for stable distributions

The algorithm below follows the style of McCuloch's *rndsta* routine[1]. This procedure is an implementation of Chambers et al. (1976) random number generation algorithm, valid for skewed α–stable random variates.

Algorithm 0 (Chambers, Mallows and Stuck, 1976)

Step 0. Generate independent $Z_1, Z_2 \sim N(0,1)$, set ε to sufficiently small number.

Step 1. Set $V := \pi(Z_1 - 0.5)$, $W := -\ln Z_2$, $\zeta := \beta \tan \frac{\pi\alpha}{2}$.

Step 3. **If** $|1-\alpha| < \varepsilon$ **then**

$$Y := \frac{\sin \alpha V + \zeta \cos \alpha V}{\cos V} \left(\frac{\cos(1-\alpha)V + \zeta \sin(1-\alpha)V}{W \cos V} \right)^{\frac{1-\alpha}{\alpha}};$$

else

$\eta := 0.5\pi + \beta V;$

$Y := \frac{2}{\pi}(\eta \tan V - \beta \ln \frac{\pi W \cos V}{\eta});$

If $\alpha \neq 1$ **then** $Y := Y + \eta;$

end.

Step 4. Set $X := \delta Y$.

If $|1-\alpha| < \varepsilon$ **then** $X := X + \frac{2}{\pi}\beta\delta \cdot \ln \delta$.

$X := X + \mu$.

Step 5. Return X.

[1]Available at: http://www.econ.ohio-state.edu/jhm/programs/RNDSSTA.

Algorithm 0 returns pseudo–random number X drawn from $S_\alpha(\beta, \delta, \mu)$. Auxiliary random variable Y defined above is generated from $S_\alpha(\beta, 1, 0)$.

A.2 Inverse FT on asymmetric domain

Assume set $[a, b]$ is divided it into N disjoint sections of equal length. The aim procedure derived here is to evaluate pdf of random variable with known characteristic function $\Phi_X(u)$ in the lower bounds of these sections.

For $l = 0, \ldots, N-1$ set $x_l = a + hl$ with $h = (b-a)N^{-1}$. For N sufficiently large (h sufficiently small) constant $c = \pi/h$ is also large and

$$f_X(x_l) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iux_l}\cdot\Phi_X(u)\,du \approx \frac{1}{2\pi}\int_{-c}^{c} e^{-iux_l}\cdot\Phi_X(u)\,du =$$

$$= \{u = 2\pi\omega, \frac{1}{2\pi}du = d\omega\} = \int_{-N/2(b-a)}^{N/2(b-a)} e^{-2\pi i\cdot\omega x_l}\cdot\Phi_X(2\pi\omega)\,d\omega.$$

Set $\omega_n = (n - N/2)s$ for $n = 0, \ldots, N-1$ with $s = (hN)^{-1} = (b-a)^{-1}$ to obtain

$$\int_{-N/2(b-a)}^{N/2(b-a)} e^{-2\pi i\cdot\omega x_l}\cdot\Phi_X(2\pi\omega)\,d\omega = \int_{-Ns/2}^{Ns/2} e^{-2\pi i\cdot\omega x_l}\cdot\Phi_X(2\pi\omega)\,d\omega \approx$$

$$\approx s\sum_{n=0}^{N-1}\Phi_X(2\pi\omega_n)\cdot e^{-2\pi i\cdot\omega_n x_l} = s\sum_{n=0}^{N-1}\Phi_X\Big(2\pi s(n - \frac{N}{2})\Big)\cdot e^{-2\pi i\cdot(\frac{a}{h}+l)(n-\frac{N}{2})hs}.$$

As $e^{\pi i} = -1$, it follows that

$$e^{-2\pi i\cdot(\frac{a}{h}+l)(n-\frac{N}{2})hs} = e^{-2\pi i\cdot(\frac{aN}{b-a}+l)(\frac{n}{N}-\frac{1}{2})} = e^{\pi i\cdot(\frac{aN}{b-a}+l)}\cdot e^{-2\pi i\cdot(\frac{an}{b-a}+l\frac{n}{N})} =$$

$$= (-1)^{\frac{a}{b-a}N+l}\cdot(-1)^{-\frac{2a}{b-a}n}\cdot e^{-2\pi i\cdot l\frac{n}{N}}.$$

Finally

$$f_X(x_l) \approx \frac{1}{b-a}(-1)^{\frac{a}{b-a}N+l}\sum_{n=0}^{N-1}(-1)^{-\frac{2a}{b-a}n}\cdot\Phi_X\Big(\frac{2\pi}{b-a}(n - \frac{N}{2})\Big)\cdot e^{-2\pi i\cdot l\frac{n}{N}}.$$

The sought result may be computed by evaluation of Inverse Fourier Transformation

$$\sum_{n=0}^{N-1} y_n \cdot e^{-2\pi i \cdot l \frac{n}{N}},\ l = 0, 1, \dots, N-1$$

via Fast Fourier Transform (FFT) algorithm applied to the sequence

$$y_n = (-1)^{-\frac{2a}{b-a}n} \cdot \Phi_X\Big(\frac{2\pi}{b-a}(n - \frac{N}{2})\Big),\ n = 0, 1, \dots, N-1.$$

An output of FFT procedure is a vector. In order to obtain it in MATLAB run *fft* procedure on $(y_0, \dots, y_{N-1})$. To get valid pdf values multiply the entries thus obtained by $\frac{1}{b-a}(-1)^{\frac{a}{b-a}N+l}$.

Appendix B

Proofs

Property 3.2.1 The corresponding Rosiński measure

$$R(dx) = C_\alpha(\theta\delta)^\alpha[(1+\beta)\delta(x-1/\theta) + (1-\beta)\delta(x+1/\theta)]\,dx,$$

where $\delta(x \pm 1/\theta)$ stands for Dirac's delta and constant C_α is equal to

$$C_\alpha = \begin{cases} \alpha(1-\alpha)[2\cos\left(\frac{\pi\alpha}{2}\right)\Gamma(2-\alpha)]^{-1} & \alpha \neq 1 \\ \frac{1}{\pi} & \alpha = 1 \end{cases}$$

satisfies $R(\{x : |x| > 1/\theta\}) = 0$. The statement follows from Corollary 2.13 in Rosiński (2007).

Property 3.2.2 For $V \sim Gamma(1/2, \theta)$ the corresponding Fourier transformation is

$$\Phi_V(u) \equiv \mathbb{E}e^{iuV} = (1-\theta iu)^{-1/2}.$$

If V_1, V_2 are independent copies of V, then for $X \sim (V_1 - V_2)/\theta^2$ we have

$$\Phi_X(u) \equiv \mathbb{E}e^{iu(V_1-V_2)/\theta^2} = \mathbb{E}e^{i(u/\theta^2)V_1 + i(-u/\theta^2)V_2} = \Phi_V(u/\theta^2) \cdot \Phi_V(-u/\theta^2) =$$

$$= (1-iu/\theta)^{1/2}(1+iu/\theta)^{1/2}$$

what matches the formula, provided by Boyarchenko and Levendorskii (2000) for $\alpha = 0$.

Property 3.2.3 Distribution of $X \sim S_\alpha(\beta, \delta, 0)$ is concentrated on $\mathbb{R}_+$ for $\alpha < 1$,

$\beta = 1$ and on $\mathbb{R}$ otherwise. As distribution in Definition 3.2.1 is first obtained from $S_\alpha(\beta, \delta, 0)$ via exponential tempering (which does not affect the support) and next shifted by μ_X to obtain centred density for $\mu = 0$, the statement follows.

Property 3.2.4 Kim et al. (2010a) give the proof for $\alpha \neq 1$. Their reasoning also holds for $\alpha = 1$ and $\rho \in (0, \theta)$ if only $|\Phi_X(u + i\rho)| < +\infty$ for all $u \in \mathbb{R}$. We shall demonstrate this inequality. First note that the logarithm in Definition 3.2.1 is a principal branch of complex valued natural logarithm, hence $\ln z = \ln |z| + i \arg z$. This equality implies that

$$\ln\left(\theta \pm i(u + i\rho)\right) = \ln\sqrt{(\theta \pm \rho)^2 + u^2} \pm i \arcsin\frac{u}{\sqrt{(\theta \pm \rho)^2 + u^2}}.$$

Next define real functions $A(u)$ and $B(u)$ so that

$$A(u) + iB(u) := \ln \Phi_X(u + i\rho) = \psi_X(u + i\rho) + i(\mu - \mu_X)(u + i\rho) =$$

$$= \frac{1}{\pi}\delta[(1+\beta)(\theta - \rho - iu)(\ln\sqrt{(\theta-\rho)^2 + u^2} - i \arcsin\frac{u}{\sqrt{(\theta-\rho)^2+u^2}})]+$$

$$+\frac{1}{\pi}\delta[(1-\beta)(\theta + \rho + iu)(\ln\sqrt{(\theta+\rho)^2 + u^2} + i \arcsin\frac{u}{\sqrt{(\theta+\rho)^2+u^2}})]+$$

$$+i(\mu-\mu_X)(u+i\rho) = \frac{1}{\pi}\delta\Big[(1+\beta)\Big((\theta-\rho)\ln\sqrt{(\theta-\rho)^2+u^2} - u\arcsin\frac{u}{\sqrt{(\theta-\rho)^2+u^2}}\Big)+$$

$$+(1-\beta)\Big((\theta+\rho)\ln\sqrt{(\theta+\rho)^2+u^2} - u\arcsin\frac{u}{\sqrt{(\theta-\rho)^2+u^2}}\Big) + (\mu-\mu_X)\rho\Big]+$$

$$+i\frac{1}{\pi}\delta\Big[(1+\beta)\Big(-u\ln\sqrt{(\theta-\rho)^2+u^2} - (\theta-\rho)\arcsin\frac{u}{\sqrt{(\theta-\rho)^2+u^2}}\Big)+$$

$$+(1-\beta)\Big(u\ln\sqrt{(\theta+\rho)^2+u^2} + (\theta+\rho)\arcsin\frac{u}{\sqrt{(\theta+\rho)^2+u^2}}\Big) + (\mu-\mu_X)u\Big]$$

Note that for sufficiently large $|u|$ we have

$$(\theta \pm \rho)\ln\sqrt{(\theta\pm\rho)^2+u^2} - u\arcsin\frac{u}{\sqrt{(\theta\pm\rho)^2+u^2}} \approx (\theta\pm\rho)\ln|u| - \frac{\pi}{2}|u|$$

and thus

$$\lim_{|u|\to+\infty} A(u) = -\infty, \quad \lim_{|u|\to+\infty} e^{A(u)} = 0. \tag{B.1}$$

This yields

$$|\Phi_X(u+i\rho)| = |e^{A(u)+iB(u)}| = \sqrt{e^{A(u)+iB(u)} \cdot e^{A(u)-iB(u)}} =$$

$$= \sqrt{e^{2A(u)}(\cos B(u) + i\sin B(u))(\cos B(u) - i\sin B(u))} =$$

$$= e^{A(u)}\sqrt{\cos^2 B(u) + \sin^2 B(u)} = e^{A(u)}$$

Now for sufficiently large real R and $u \in (-R, R)$ function $|\Phi_X(u+i\rho)|$ is arbitrarily small on $(-\infty, -R) \cup (R, +\infty)$ by equation (B.1) and, as a continuous complex function defined on a compact, bounded on $(-R, R)$. The statement holds by Proposition 1 in Kim et al. (2010a).

Property 3.2.5 Treat r as complex, note that $|(-r)^{-1-\alpha}| = |r^{-1-\alpha}||(-1)^{-1-\alpha}| = r^{-1-\alpha}$. Spectral density λ defined in Wolfe (1971) on p. 2069 fulfils

$$\lambda(r) - |\lambda(-r)| = r\Big(r^{-1-\alpha}e^{-\theta r}(1+\beta)\delta^\alpha C_\alpha/2 - \Big|(-r)^{-1-\alpha}e^{\theta r}(1+\beta)\delta^\alpha C_\alpha/2\Big|\Big) =$$

$$= r^{-\alpha}\Big(e^{-\theta r}(1+\beta) - e^{\theta r}(1-\beta)\Big)\delta^\alpha C_\alpha/2 = \frac{(e^{-\theta r} - e^{\theta r}) + \beta(e^{-\theta r}) + e^{\theta r})}{r^\alpha}\delta^\alpha C_\alpha/2.$$

The term $\delta^\alpha C_\alpha/2$ is strictly positive and does not depend on r. From L'Hospital rule

$$\lambda(0^+) + |\lambda(0^-)| = \lim_{r\to 0^+}(\lambda(r) - |\lambda(-r)|) =$$

$$= \delta^\alpha C_\alpha/2 \cdot \lim_{r\to 0^+} \frac{(e^{-\theta r} - e^{\theta r}) + \beta(e^{-\theta r}) + e^{\theta r})}{r^\alpha} = +\infty.$$

By Theorem 4 in Wolfe (1971) function $f_X(x)$ exists and is smooth. This calculation also implies that the corresponding cdf belongs either to class I_6 or to class I_7 defined in Sato and Yamazoto (1978). By either point (vii) or (xi) of Theorem 1.3 therein $f_X(x)$ is unimodal.

Property 3.2.6 As $\mu_Y = -\mu_X$, in consequence $\psi_X(-u) + i(\mu - \mu_X)(-u) = \psi_Y(u) + i(-\mu - \mu_Y)u$ and we have $\Phi_{-X}(u) = \Phi_Y(u)$.

Property 3.2.7 Assume for simplicity that $a > 0$. For $\alpha \neq 1$ we have $\mu_Y = a\mu_X$ and $\psi_X(au) = \psi_Y(u)$, while in case of $\alpha = 1$ it holds that $\mu_Y = a\mu_X + \frac{2}{\pi}\beta\delta a \ln a$ and $\psi_X(au) = \psi_Y(u) - i(\frac{2}{\pi}\beta\delta a \ln a)u$. In both cases $\psi_X(au) + i(\mu - \mu_X)au + ibu = \psi_Y(u) + i(a\mu + b - \mu_Y)u$ and thus $\Phi_{aX+b}(u) = \Phi_Y(u)$. If $a < 0$ then $\Phi_{aX+b}(u) = \Phi_{-(|a|X-b)}(u)$ and the statement holds by Property 3.2.6.

Property 3.2.8 Follows from multiplication of characteristic functions and collection of terms.

Property 3.2.9 By Property 3.2.5 pdf of X exists. By Property 3.2.8 we have

$$f_X^{*n}(x;\mu,\theta) = f_X(x; n^{1/\alpha}\delta, n\mu).$$

Let $Y \sim TS_\alpha(\beta,\delta,n^{1-1/\alpha}\mu, n^{1/\alpha}\theta)$. We obtain

$$\ln\Phi_X(u; n^{1/\alpha}, n\mu) = \psi_X(u; n^{1/\alpha}\delta) + i(n\mu - n\mu_X)u =$$

$$= \begin{cases} \psi_Y(n^{1/\alpha}u; n^{1/\alpha}\theta) + i(n^{1-1/\alpha}\mu - \mu_Y)(n^{1/\alpha}u) & \alpha \neq 1, \\ \psi_Y(nu; n\theta) + i(\frac{2}{\pi}\beta\delta\ln n)(nu) + i(\mu - \mu_Y + \frac{2}{\pi}\beta\delta\ln n)(nu) & \alpha = 1, \end{cases}$$

$$= \psi_Y(n^{1/\alpha}u; n^{1/\alpha}\theta) + i(n^{1-1/\alpha}\mu - \mu_Y)(n^{1/\alpha}u).$$

Express pdf via Inverse Fourier Transform to obtain

$$f_X(x; n^{1/\alpha}, n\mu) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iux}\cdot\Phi_X(u)\,du =$$

$$= \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-i(n^{1/\alpha}u)(n^{-1/\alpha}x)}\cdot e^{\psi_Y(n^{1/\alpha}u;n^{1/\alpha}\theta)+i(n^{1-1/\alpha}\mu-\mu_Y)(n^{1/\alpha}u)}\,du = \{v = n^{1/\alpha}u;$$

$$du = n^{-1/\alpha}dv\} = n^{-1/\alpha}\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iv(n^{-1/\alpha}x)}\cdot e^{\psi_Y(v;n^{1/\alpha}\theta)+i(n^{1-1/\alpha}\mu-\mu_Y)v}\,dv =$$

$$= n^{-1/\alpha}f_X(n^{-1/\alpha}x; n^{1-1/\alpha}\mu, n^{1/\alpha}\theta).$$

Fact 3.2.10 First introduce the following notation. Define v_j as $(n\times 1)$ vector with one on j–th place and all the remaining entries being equal to zero. Define X_j^0 as $(n\times 1)$ random vector with X_j on j–th place and the remaining entries being equal to zero. Assume $\alpha \in (0,2)$ and the constant C_α is defined as in chapter 2, set $\sigma_{j,\pm} := \delta_j^\alpha C_\alpha(1\pm\beta_j)/2$, $\delta_j > 0$ and $\beta_j \in [-1,1]$ for $j \in \{1,\cdots,n\}$.

Next, characterize the distribution of $\boldsymbol{X}$ from Fact 3.2.10 with the underlying spectral measure (expressed in polar coordinates). If $X_j \sim TS_\alpha(\beta,\delta,\mu,\theta)$ then the spectral measure corresponding to X_j for each $j \in \{1,\cdots,n\}$ may be written as

$$r^{-1-\alpha}e^{-\theta r}[\sigma_{j,+}\delta(v-v_1) + \sigma_{j,-}\delta(v+v_1)]\,dr\,dv$$

where δ stands for Dirac's delta. Hence the spectral measure of random vector X_j^0 is

$$r^{-1-\alpha}e^{-\theta r}[\delta(v-v_j)+\delta(v+v_j)][\sigma_{j,+}\delta(v-v_j)+\sigma_{j,-}\delta(v+v_j)]\,dr\,dv =$$

$$= r^{-1-\alpha}e^{-\theta r}\,dr\,[\sigma_{j,+}\delta(v-v_j)+\sigma_{j,-}\delta(v+v_j)]\,dv.$$

The last equality follows from the fact that this measure is non–zero only on the intersection of j–th axis with the unit sphere. If random variates X_j are independent, random vectors X_j^0 are also independent. Hence Fourier transformation of their sum is the product of (univariate) Fourier transformations, and thus spectral measure of their sum is the sum of (univariate) spectral measures. Note that random vector $\boldsymbol{X}$ may be expressed as $\boldsymbol{X} = \sum_{j=1}^n X_j^0$, so in polar coordinates its spectral measure is give as

$$r^{-1-\alpha}e^{-\theta r}\,dr\sum_{j=1}^n[\sigma_{j,+}\delta(v-v_j)+\sigma_{j,-}\delta(v+v_j)]\,dv =: r^{-1-\alpha}e^{-\theta r}\,dr\,\sigma(v)\,dv.$$

The assignment above (from right to left) defines spherical measure σ.

In the third step note that, by Example 2.3.5 in Samorodnitsky and Taqqu (2000), the expression $r^{-1-\alpha}\,dr\,\sigma(v)\,dv$ stands for spectral measure of the multivariate stable vector with independent entries. Thus the multivariate distribution, obtained via the uniform, exponential tempering of the spectral measure (expressed in polar coordinates) of multivariate stable vector with independent entries via $e^{-\theta||x||} = e^{-\theta r}$ may be characterized with

$$T(dr,dv) = e^{-\theta r}S(dr,dv) = r^{-1-\alpha}e^{-\theta r}\,dr\,\sigma(v)\,dv.$$

The last equality implies that random vector $\boldsymbol{X}$ has a tempered stable distribution where the tempering is defined as in Rosiński (2007).

Finally, observe that by Corollary 2.12 proved therein the class of tempered stable distributions defined in Rosiński (2007) is closed with respect to linear transformations. Hence the distribution of $\boldsymbol{Y}$ given as $\boldsymbol{Y} = \mu + A^T\boldsymbol{X}$ is a member of Rosiński (2007) class and thus a multivariate tempered stable distribution. Note that this result is established irrespective of the form of A^T. In particular, it is required to be neither an invertible nor a full–rank matrix. The distribution of $\boldsymbol{Y}$ is therefore a well-defined multivariate tempered stable distribution which may be further denoted as $mTS_\alpha(\boldsymbol{\beta},\boldsymbol{\delta},\boldsymbol{\mu},\theta,A)$.

Corollary 3.3.1 To obtain cumulants when $\alpha \neq 1$ integrate eq. (14) in Terdik and Woyczyński (2006) with respect to Rosiński measure

$$R(dx) = C_\alpha(\theta\delta)^\alpha[(1+\beta)\delta(x-1/\theta)+(1-\beta)\delta(x+1/\theta)]\,dx.$$

Definition 3.2.1 relies on different parametrization than the one used in Rosiński (2007), it implies that $\kappa_1 = \mu$ while all higher order cumulants remain intact. In case of $\alpha = 1$ it holds that

$$\frac{d}{du}\psi_X(u) = i\frac{2\delta}{\pi}\Big[\sum_{j=1}^{+\infty}(I_{j+1}+\beta I_j)\theta^{-j}(j-1)!\frac{(iu)^j}{j!} - \beta(1+\ln\theta)\Big],$$

so $\kappa_1 = (\mu - \mu_X) - i\psi_X'(0) = \mu$ and $\kappa_p = (-i)^p\psi_X^{(p)}(0) = \frac{2\delta}{\pi}(I_p + \beta I_{p+1})\theta^{1-p}(p-2)!$ for $p \geq 2$.

Proposition 3.4.1 i) We have $\hat{\kappa}_3 \neq 0$ and $\hat{\kappa}_5 \neq 0$. From Corollary 3.3.1 it holds iff $\beta \neq 0$. The CM estimator may be obtained by matching $\hat{\kappa}_p = \kappa_p$ for $p \in \{1,\dots,5\}$ given $\alpha \neq 1$. Derive the expression for $\hat{\alpha}$ using ratios

$$\frac{\hat{\kappa}_5/\hat{\kappa}_3}{\hat{\kappa}_4/\hat{\kappa}_2} = \frac{4-\alpha}{2-\alpha}.$$

This yields

$$\hat{\alpha} = 2\Big(1 - \frac{\hat{\kappa}_4\hat{\kappa}_3}{\hat{\kappa}_5\hat{\kappa}_2 + \hat{\kappa}_4\hat{\kappa}_3}\Big).$$

As $\hat{\kappa}_5\hat{\kappa}_2 + \hat{\kappa}_4\hat{\kappa}_3 \neq 0$ this estimate exists. It is assumed that $\hat{\kappa}_4\hat{\kappa}_3/(\hat{\kappa}_5\hat{\kappa}_2 + \hat{\kappa}_4\hat{\kappa}_3) \neq 1/2$, so the estimate $\hat{\alpha} \neq 1$ is consistent with the initial set of conditions. It holds that $\hat{\alpha} \in (0,2)$ iff $\hat{\kappa}_5\hat{\kappa}_2/\hat{\kappa}_4\hat{\kappa}_3 > 2$. Substitute the formula for $\hat{\alpha}$ to the ratio

$$\hat{\kappa}_4\hat{\kappa}_2 = (2-\alpha)(3-\alpha)\theta^{-2}$$

to derive the formula for θ^2. As $\hat{\kappa}_2 = \widehat{\mathbb{V}ar}X > 0$, this expression is positive iff $\hat{\kappa}_4 > 0$. We thus obtain

$$\hat{\theta} = \sqrt{\frac{2\hat{\kappa}_3\hat{\kappa}_2(\hat{\kappa}_5\hat{\kappa}_2 + \hat{\kappa}_4\hat{\kappa}_3)}{(\hat{\kappa}_5\hat{\kappa}_2 - \hat{\kappa}_4\hat{\kappa}_3)^2}}.$$

Next note that $\hat{\kappa}_5\hat{\kappa}_4 = (4-\alpha)\beta/\theta$ and $\hat{\kappa}_3\hat{\kappa}_2 = (2-\alpha)\beta/\theta$. Subtract the latter from the previous and use the estimates $\hat{\alpha}$, $\hat{\theta}$ to derive

$$\hat{\beta} = \frac{\hat{\kappa}_5\hat{\kappa}_2 - \hat{\kappa}_4\hat{\kappa}_3}{2\hat{\kappa}_4\hat{\kappa}_2}\hat{\theta}.$$

We have $\hat{\beta} \in [-1, 1]$ iff $\hat{\beta}^2 \leq 1$. The last condition is equivalent to $\hat{\kappa}_5\hat{\kappa}_2/\hat{\kappa}_4\hat{\kappa}_3 \leq 2\hat{\kappa}_4\hat{\kappa}_2/\hat{\kappa}_3^2 - 1$. Inverting formula for $\hat{\delta}$ and inserting all the estimates above yields

$$\hat{\delta} = \left(\frac{\hat{\kappa}_2 \cos \pi\hat{\alpha}/2}{\hat{\alpha}(1-\hat{\alpha})}\right)^{1/\hat{\alpha}} \hat{\theta}^{\,2/\hat{\alpha}-1}.$$

Given the assumptions already made, this estimator is always positive and thus does not require any further existence conditions. Finally set $\hat{\mu} = \kappa_1$. Point i) is completed.
ii) We have $\hat{\kappa}_3 \neq 0$ and $\hat{\kappa}_5 \neq 0$. From Corollary 3.3.1 it holds iff $\beta \neq 0$. As $\hat{\kappa}_5\hat{\kappa}_2 + \hat{\kappa}_4\hat{\kappa}_3 \neq 0$, the estimator $\hat{\alpha}$ defined in point i) exists, but now $\hat{\kappa}_4\hat{\kappa}_3/(\hat{\kappa}_5\hat{\kappa}_2+\hat{\kappa}_4\hat{\kappa}_3) = 1/2$, hence $\hat{\alpha} = 1$. The CM estimator may be obtained from $\hat{\kappa}_p = \kappa_p$ for $p \in \{1, \ldots, 4\}$ given $\alpha = 1$. Use the ratio $\hat{\kappa}_4/\hat{\kappa}_2 = 2/\theta^2$ to find

$$\hat{\theta} = \sqrt{\frac{2\hat{\kappa}_2}{\hat{\kappa}_4}},$$

this estimator exists iff $\hat{\kappa}_4 > 0$. Inverting the formula for second cumulant yields

$$\hat{\delta} = \frac{\pi}{2}\hat{\kappa}_2\hat{\theta}.$$

Use the ratio $\hat{\kappa}_3/\hat{\kappa}_2 = \beta/\theta$ to obtain

$$\hat{\beta} = \frac{\hat{\kappa}_3}{\hat{\kappa}_2}\hat{\delta}.$$

We have $\hat{\beta} \in [-1, 1]$ iff $\hat{\beta}^2 \leq 1$. The latter condition is equivalent to $2\hat{\kappa}_3^2/\hat{\kappa}_4\hat{\kappa}_2 \leq 1$. Set $\hat{\mu} = \kappa_1$. This concludes point ii).
iii) We have $\hat{\kappa}_3 = 0$ and $\hat{\kappa}_5 = 0$. From Corollary 3.3.1 it holds iff $\beta = 0$. Set $\hat{\beta} = 0$. The CM estimator may be obtained by matching $\hat{\kappa}_p = \kappa_p$ for $p \in \{1, 2, 4, 6\}$ given $\alpha \neq 1$. Use the ratios

$$\frac{\hat{\kappa}_6}{\hat{\kappa}_4} = (4-\alpha)(5-\alpha)\theta^{-2}, \ \frac{\hat{\kappa}_4}{\hat{\kappa}_2} = (2-\alpha)(3-\alpha)\theta^{-2}, \ \frac{\hat{\kappa}_6}{\hat{\kappa}_4} - \frac{\hat{\kappa}_4}{\hat{\kappa}_2} = (14-4\alpha)\theta^{-2}$$

to obtain

$$\theta^{-2} = \frac{1}{14-4\alpha}\frac{\hat{\kappa}_6\hat{\kappa}_2 - \hat{\kappa}_4^2}{\hat{\kappa}_4\hat{\kappa}_2}.$$

Set $a := \hat{\kappa}_4^2/(\hat{\kappa}_4^2 - \hat{\kappa}_6\hat{\kappa}_2)$, insert the formula above to the expression for $\hat{\kappa}_4/\hat{\kappa}_2$. It follows that α fulfils quadratic equation $\alpha^2-(5+4a)\alpha+(6+14a) = 0$ with $\Delta = 16a^2-16a+1$. Only smaller of the two feasible solutions could ever satisfy $0 < \alpha < 2$. This solution lies in the desired region iff $a \in (-3/7, 0)$. The last condition is equivalent to $3\hat{\kappa}_6\hat{\kappa}_2 > 10\hat{\kappa}_4^2$. When it holds, we have

$$\alpha = \frac{5}{2} + 2a - \sqrt{4a^2 - 4a + \frac{1}{4}},$$

which in turn yields

$$\hat{\alpha} = \frac{5}{2} - \frac{1}{2(\hat{\kappa}_6\hat{\kappa}_2 - \hat{\kappa}_4^2)}\Big(4\hat{\kappa}_4^2 + \sqrt{\hat{\kappa}_6^2\hat{\kappa}_4^2 + 14\hat{\kappa}_6\hat{\kappa}_4^2\hat{\kappa}_2 + \hat{\kappa}_4^4}\Big).$$

As $\hat{\kappa}_6\hat{\kappa}_2 \neq \hat{\kappa}_4^2$, this estimate exists. We assumed $(4\hat{\kappa}_4^2 + \sqrt{\hat{\kappa}_6^2\hat{\kappa}_4^2 + 14\hat{\kappa}_6\hat{\kappa}_4^2\hat{\kappa}_2 + \hat{\kappa}_4^4})/$ $/(\hat{\kappa}_6\hat{\kappa}_2 - \hat{\kappa}_4^2) \neq 3$, so the estimate $\hat{\alpha} \neq 1$ is consistent with the initial set of conditions. Use the expression for $\hat{\alpha}$ and the formula for θ^{-2} given above to obtain

$$\hat{\theta} = \sqrt{(14-4\hat{\alpha})\frac{\hat{\kappa}_4\hat{\kappa}_2}{\hat{\kappa}_6\hat{\kappa}_2 - \hat{\kappa}_4^2}}.$$

If $3\hat{\kappa}_6\hat{\kappa}_2 > 10\hat{\kappa}_4^2$ then $\hat{\kappa}_6\hat{\kappa}_2 - \hat{\kappa}_4^2 > 0$, we have also assumed $\hat{\kappa}_4 > 0$, hence this estimator is well defined. Inverting formula for $\hat{\delta}$ and using the remaining estimates yields

$$\hat{\delta} = \Big(\frac{\hat{\kappa}_2 \cos \pi\hat{\alpha}/2}{\hat{\alpha}(1-\hat{\alpha})}\Big)^{1/\hat{\alpha}} \hat{\theta}^{\,2/\hat{\alpha}-1}.$$

Given the assumptions already made this estimator is always positive and thus does not require any further existence conditions. Finally set $\hat{\mu} = \kappa_1$. Point iii) is completed.

iv) We have $\hat{\kappa}_3 = 0$ and $\hat{\kappa}_5 = 0$. From Corollary 3.3.1 it holds iff $\beta = 0$. Set $\hat{\beta} = 0$. As $\hat{\kappa}_6\hat{\kappa}_2 \neq \hat{\kappa}_4^2$, the estimator $\hat{\alpha}$ defined in point iii) exists, but now $(4\hat{\kappa}_4^2 + \sqrt{\hat{\kappa}_6^2\hat{\kappa}_4^2 + 14\hat{\kappa}_6\hat{\kappa}_4^2\hat{\kappa}_2 + \hat{\kappa}_4^4})/(\hat{\kappa}_6\hat{\kappa}_2 - \hat{\kappa}_4^2) = 3$, hence $\hat{\alpha} = 1$. The CM estimator may be obtained by matching $\hat{\kappa}_p = \kappa_p$ for $p \in \{1, 2, 4\}$ given $\alpha = 1$. Use the ratio

$\hat{\kappa}_4/\hat{\kappa}_2 = 2/\theta^2$ to find

$$\hat{\theta} = \sqrt{\frac{2\hat{\kappa}_2}{\hat{\kappa}_4}},$$

this estimator exists iff $\hat{\kappa}_4 > 0$. Inverting the formula for second cumulant yields

$$\hat{\delta} = \frac{\pi}{2}\hat{\kappa}_2\hat{\theta}.$$

Set $\hat{\mu} = \kappa_1$. This concludes point iv).

Proposition 3.5.1 It is enough to show that $\ln \Phi_X(u) = \ln \Phi_{Y^+}(V^+u) + \ln \Phi_{Y^-}(-V^-u) + i\mu u$. In case of $\alpha \neq 1$ it holds that

$$\ln \Phi_{Y^\pm}(\pm V^\pm u) = -\frac{\delta^\alpha}{2\cos\frac{\pi\alpha}{2}}(1 \pm \beta)[(\theta \mp iu)^\alpha - \theta^\alpha] \pm i\frac{\delta^\alpha}{2\cos\frac{\pi\alpha}{2}}\alpha(1 \pm \beta)\theta^{\alpha-1}u,$$

while $\alpha = 1$ yields

$$\ln \Phi_{Y^\pm}(\pm V^\pm u) = \frac{1}{\pi}\delta[(1\pm\beta)(\theta\mp iu)\ln(\theta\mp iu) - (1\pm\beta)\theta\ln\theta] \pm i\frac{1}{\pi}\delta(1\pm\beta)(\ln\theta + 1)u,$$

so in both cases $\ln \Phi_{Y^+}(V^+u) + \ln \Phi_{Y^-}(-V^-u) = \psi_X(u) - i\mu_X u$.

Appendix C

Tables

Parameters	Method	Mean sample moments (1)	(2)	(3)	(4)	(5)	Mean time (sec)
$\alpha = 1.20$, $\beta = 0.00$, $\theta = 0.25$	Theoretic	0.0000	1.0000	0.0000	26.0400	0.0000	NA
	Mixture	-0.0002	0.9997	-0.0057	25.8809	2.6837	0.662
		(0.0010)	(0.0047)	(0.0724)	(1.2546)	(36.2074)	
	Inversion	0.0060	1.0010	0.0148	**26.0838**	**1.3157**	0.435
		(0.0010)	(0.0050)	(0.0682)	(1.2923)	(31.7305)	
	Devroye	**0.0000**	**0.9998**	**-0.0045**	25.9491	-1.7717	6.291
		(0.0010)	(0.0050)	(0.0594)	(1.2472)	(35.4470)	
$\alpha = 1.20$, $\beta = 0.00$, $\theta = 0.50$	Theoretic	0.0000	1.0000	0.0000	8.7600	0.0000	NA
	Mixture	**-0.0001**	1.0001	-0.0037	8.7548	-0.2700	0.930
		(0.0011)	(0.0028)	(0.0179)	(0.1952)	(2.6853)	
	Inversion	0.0035	**1.0000**	0.0096	**8.7607**	**0.0944**	0.422
		(0.0010)	(0.0032)	(0.0175)	(0.2149)	(2.9491)	
	Devroye	-0.0002	1.0006	**-0.0006**	8.8004	0.4099	6.284
		(0.0010)	(0.0029)	(0.0169)	(0.1959)	(2.8851)	
$\alpha = 1.20$, $\beta = 0.00$, $\theta = 0.75$	Theoretic	0.0000	1.0000	0.0000	5.5600	0.0000	NA
	Mixture	**0.0000**	0.9998	**0.0001**	5.5631	**0.0338**	1.410
		(0.0011)	(0.0022)	(0.0111)	(0.0730)	(0.6436)	
	Inversion	0.0024	**1.0000**	0.0072	5.5558	0.0394	0.404
		(0.0011)	(0.0020)	(0.0114)	(0.0699)	(0.6463)	
	Devroye	-0.0001	0.9999	-0.0021	**5.5577**	-0.1155	6.282
		(0.0010)	(0.0020)	(0.0095)	(0.0600)	(0.6113)	
$\alpha = 1.20$, $\beta = 0.00$, $\theta = 1.00$	Theoretic	0.0000	1.0000	0.0000	4.4400	0.0000	NA
	Mixture	0.0001	0.9999	0.0012	4.4394	0.0284	2.340
		(0.0009)	(0.0017)	(0.0070)	(0.0367)	(0.2460)	
	Inversion	0.0022	**0.9999**	0.0068	4.4386	0.0585	0.398
		(0.0011)	(0.0019)	(0.0076)	(0.0334)	(0.2340)	
	Devroye	**-0.0001**	0.9999	**-0.0003**	**4.4404**	**0.0198**	6.265
		(0.0009)	(0.0016)	(0.0076)	(0.0335)	(0.2687)	

Table C.1: Theoretic vs. mean sample moments about the origin for standardised TS distribution. Emphasized numbers indicate either the smallest (mean) execution time, or the (mean) sample moment most similar to corresponding theoretic value.

Parameters	Method	Mean sample moments (1)	(2)	(3)	(4)	(5)	Mean time (sec)
$\alpha = 1.20$, $\beta = 0.25$, $\theta = 0.25$	Theoretic	0.0000	1.0000	0.8000	26.0400	72.5120	NA
	Mixture	**-0.0000** (0.0010)	0.9993 (0.0051)	0.7963 (0.0731)	**26.0122** (1.4045)	69.4444 (41.4620)	0.657
	Inversion	0.0063 (0.0010)	1.0006 (0.0050)	0.8223 (0.0672)	26.2182 (1.3281)	75.9623 (36.4537)	0.818
	Devroye	-0.0001 (0.0010)	**0.9999** (0.0051)	**0.7999** (0.0682)	26.0814 (1.3458)	**74.3815** (35.4963)	6.780
$\alpha = 1.20$, $\beta = 0.25$, $\theta = 0.50$	Theoretic	0.0000	1.0000	0.4000	8.7600	12.0640	NA
	Mixture	-0.0002 (0.0011)	**1.0000** (0.0027)	**0.3989** (0.0195)	8.7471 (0.1789)	12.3754 (2.7282)	0.944
	Inversion	0.0034 (0.0009)	0.9998 (0.0031)	0.4085 (0.0169)	**8.7614** (0.2127)	12.1297 (2.6439)	0.877
	Devroye	**-0.0001** (0.0010)	1.0002 (0.0028)	0.4021 (0.0183)	8.7583 (0.1701)	**12.0874** (2.3579)	6.809
$\alpha = 1.20$, $\beta = 0.25$, $\theta = 0.75$	Theoretic	0.0000	1.0000	0.2667	5.5600	5.0560	NA
	Mixture	-0.0001 (0.0010)	0.9997 (0.0021)	**0.2662** (0.0098)	5.5552 (0.0653)	**5.0050** (0.6186)	1.431
	Inversion	0.0029 (0.0010)	0.9997 (0.0021)	0.2750 (0.0102)	5.5548 (0.0749)	5.1201 (0.6274)	0.911
	Devroye	**0.0001** (0.0010)	**0.9999** (0.0018)	0.2680 (0.0087)	**5.5590** (0.0570)	5.1245 (0.5732)	6.837
$\alpha = 1.20$, $\beta = 0.25$, $\theta = 1.00$	Theoretic	0.0000	1.0000	0.2000	4.4400	3.0080	NA
	Mixture	**0.0001** (0.0010)	0.9997 (0.0018)	**0.2005** (0.0072)	4.4388 (0.0354)	**2.9951** (0.2429)	2.425
	Inversion	0.0021 (0.0011)	**1.0000** (0.0015)	0.2073 (0.0074)	**4.4406** (0.0355)	3.0922 (0.2545)	0.921
	Devroye	0.0001 (0.0012)	1.0001 (0.0020)	0.2015 (0.0078)	4.4411 (0.0352)	3.0222 (0.2493)	6.824

Table C.2: Theoretic vs. mean sample moments about the origin for standardised TS distribution. Emphasized numbers indicate either the smallest (mean) execution time, or the (mean) sample moment most similar to corresponding theoretic value.

Parameters	Method	Mean sample moments (1)	(2)	(3)	(4)	(5)	Mean time (sec)
$\alpha = 1.20$, $\beta = 0.00$, $\theta = 0.25$	Theoretic	0.0000	1.0000	1.6000	26.0400	145.0240	NA
	Mixture	**0.0000** (0.0010)	1.0003 (0.0048)	**1.5991** (0.0669)	**26.0772** (1.4646)	141.9092 (34.7801)	0.656
	Inversion	0.0062 (0.0011)	1.0004 (0.0055)	1.6176 (0.0678)	26.3081 (1.4909)	**145.4225** (34.3135)	0.819
	Devroye	0.0001 (0.0011)	**1.0002** (0.0053)	1.6069 (0.0659)	26.2189 (1.3830)	151.5961 (48.3729)	7.121
$\alpha = 1.20$, $\beta = 0.00$, $\theta = 0.50$	Theoretic	0.0000	1.0000	0.8000	8.7600	24.1280	NA
	Mixture	-0.0001 (0.0010)	**0.9999** (0.0026)	0.7993 (0.0182)	**8.7562** (0.1879)	24.2283 (2.8353)	0.922
	Inversion	0.0035 (0.0010)	0.9999 (0.0030)	0.8091 (0.0188)	8.7761 (0.2073)	23.9692 (2.8306)	0.879
	Devroye	**0.0000** (0.0010)	0.9997 (0.0027)	**0.7999** (0.0213)	8.7686 (0.2211)	**24.0839** (3.1010)	6.975
$\alpha = 1.20$, $\beta = 0.00$, $\theta = 0.75$	Theoretic	0.0000	1.0000	0.5333	5.5600	10.1120	NA
	Mixture	0.0001 (0.0009)	0.9997 (0.0020)	**0.5329** (0.0112)	5.5546 (0.0644)	10.0294 (0.6911)	1.469
	Inversion	0.0027 (0.0009)	0.9999 (0.0020)	0.5418 (0.0107)	5.5621 (0.0713)	10.1472 (0.7168)	0.911
	Devroye	**-0.0001** (0.0010)	**1.0000** (0.0022)	0.5322 (0.0108)	**5.5614** (0.0700)	**10.1135** (0.6121)	6.931
$\alpha = 1.20$, $\beta = 0.00$, $\theta = 1.00$	Theoretic	0.0000	1.0000	0.4000	4.4400	6.0160	NA
	Mixture	0.0003 (0.0010)	0.9994 (0.0019)	0.4021 (0.0075)	4.4365 (0.0384)	6.0813 (0.2711)	2.607
	Inversion	0.0023 (0.0011)	**1.0001** (0.0017)	0.4064 (0.0078)	4.4452 (0.0358)	**6.0420** (0.2673)	0.924
	Devroye	**-0.0001** (0.0011)	0.9999 (0.0016)	**0.3985** (0.0069)	**4.4348** (0.0330)	5.9691 (0.2416)	6.892

Table C.3: Theoretic vs. mean sample moments about the origin for standardised TS distribution. Emphasized numbers indicate either the smallest (mean) execution time, or the (mean) sample moment most similar to corresponding theoretic value.

Parameters	Method	Mean sample moments (1)	(2)	(3)	(4)	(5)	Mean time (sec)
$\alpha = 1.20$, $\beta = 0.00$, $\theta = 0.25$	Theoretic	0.0000	1.0000	2.4000	26.0400	217.5360	NA
	Mixture	0.0000 (0.0009)	**1.0001** (0.0052)	**2.3986** (0.0639)	26.1151 (1.4795)	**218.7404** (40.5140)	0.653
	Inversion	0.0060 (0.0009)	1.0003 (0.0048)	2.4154 (0.0695)	26.0701 (1.4754)	214.3631 (34.4277)	0.819
	Devroye	**0.0000** (0.0010)	1.0004 (0.0055)	2.3951 (0.0610)	**26.0043** (1.2247)	215.2783 (31.1859)	7.691
$\alpha = 1.20$, $\beta = 0.00$, $\theta = 0.50$	Theoretic	0.0000	1.0000	1.2000	8.7600	36.1920	NA
	Mixture	-0.0001 (0.0010)	0.9997 (0.0027)	1.1989 (0.0194)	8.7292 (0.1677)	35.7832 (2.7124)	0.926
	Inversion	0.0036 (0.0010)	1.0005 (0.0028)	1.2102 (0.0205)	8.7896 (0.1908)	36.2584 (2.8720)	0.879
	Devroye	**-0.0001** (0.0009)	**0.9998** (0.0031)	**1.1994** (0.0195)	**8.7548** (0.2169)	**36.2549** (3.9570)	7.273
$\alpha = 1.20$, $\beta = 0.00$, $\theta = 0.75$	Theoretic	0.0000	1.0000	0.8000	5.5600	15.1680	NA
	Mixture	**0.0001** (0.0010)	**0.9999** (0.0024)	**0.8009** (0.0114)	**5.5638** (0.0796)	**15.2235** (0.7378)	1.554
	Inversion	0.0026 (0.0011)	0.9997 (0.0020)	0.8073 (0.0091)	5.5649 (0.0658)	15.2699 (0.5627)	0.910
	Devroye	0.0001 (0.0008)	1.0002 (0.0022)	0.8018 (0.0097)	5.5676 (0.0621)	15.2245 (0.6085)	7.112
$\alpha = 1.20$, $\beta = 0.00$, $\theta = 1.00$	Theoretic	0.0000	1.0000	0.6000	4.4400	9.0240	NA
	Mixture	0.0005 (0.0010)	0.9988 (0.0018)	0.6031 (0.0082)	4.4344 (0.0428)	9.0573 (0.3384)	2.955
	Inversion	0.0019 (0.0011)	0.9999 (0.0019)	0.6049 (0.0080)	4.4418 (0.0432)	9.0432 (0.2733)	0.922
	Devroye	**0.0002** (0.0010)	**1.0000** (0.0020)	**0.6004** (0.0085)	**4.4401** (0.0458)	**9.0408** (0.3480)	7.017

Table C.4: Theoretic vs. mean sample moments about the origin for standardised TS distribution. Emphasized numbers indicate either the smallest (mean) execution time, or the (mean) sample moment most similar to corresponding theoretic value.

Parameters	Method	Mean sample moments (1)	(2)	(3)	(4)	(5)	Mean time (sec)
$\alpha = 1.20$, $\beta = 0.00$, $\theta = 0.25$	Theoretic	0.0000	1.0000	0.0000	18.3600	0.0000	NA
	Mixture	0.0000	**0.9999**	**0.0033**	**18.3849**	**0.3150**	0.701
		(0.0009)	(0.0040)	(0.0555)	(0.8778)	(23.5989)	
	Inversion	0.0059	1.0001	0.0156	18.4215	-0.8777	0.434
		(0.0011)	(0.0045)	(0.0565)	(0.9974)	(26.4343)	
	Devroye	**0.0000**	1.0003	-0.0044	18.4302	0.6460	6.219
		(0.0009)	(0.0044)	(0.0482)	(0.9463)	(23.4834)	
$\alpha = 1.20$, $\beta = 0.00$, $\theta = 0.50$	Theoretic	0.0000	1.0000	0.0000	6.8400	0.0000	NA
	Mixture	**0.0001**	1.0001	**0.0001**	6.8433	0.0794	1.029
		(0.0010)	(0.0025)	(0.0137)	(0.1546)	(1.6497)	
	Inversion	0.0035	1.0002	0.0117	**6.8396**	**0.0369**	0.413
		(0.0011)	(0.0021)	(0.0144)	(0.1467)	(1.8643)	
	Devroye	0.0001	**1.0000**	-0.0033	6.8491	-0.5212	6.247
		(0.0009)	(0.0026)	(0.0141)	(0.1351)	(2.0512)	
$\alpha = 1.20$, $\beta = 0.00$, $\theta = 0.75$	Theoretic	0.0000	1.0000	0.0000	4.7067	0.0000	NA
	Mixture	0.0001	0.9998	0.0014	4.7091	0.0406	1.578
		(0.0010)	(0.0021)	(0.0080)	(0.0589)	(0.4515)	
	Inversion	0.0026	**0.9999**	0.0073	**4.7077**	0.0711	0.404
		(0.0011)	(0.0017)	(0.0097)	(0.0486)	(0.4488)	
	Devroye	**-0.0001**	0.9996	**-0.0005**	4.6981	**-0.0104**	6.253
		(0.0009)	(0.0019)	(0.0083)	(0.0502)	(0.4652)	
$\alpha = 1.20$, $\beta = 0.00$, $\theta = 1.00$	Theoretic	0.0000	1.0000	0.0000	3.9600	0.0000	NA
	Mixture	**0.0000**	0.9996	**0.0000**	3.9582	**-0.0035**	2.717
		(0.0008)	(0.0016)	(0.0061)	(0.0259)	(0.1678)	
	Inversion	0.0022	**1.0003**	0.0070	**3.9611**	0.0583	0.405
		(0.0011)	(0.0018)	(0.0065)	(0.0269)	(0.1918)	
	Devroye	0.0000	1.0004	0.0002	3.9656	-0.0146	6.391
		(0.0008)	(0.0017)	(0.0054)	(0.0275)	(0.1798)	

Table C.5: Theoretic vs. mean sample moments about the origin for standardised TS distribution. Emphasized numbers indicate either the smallest (mean) execution time, or the (mean) sample moment most similar to corresponding theoretic value.

Parameters	Method	Mean sample moments (1)	(2)	(3)	(4)	(5)	Mean time (sec)
$\alpha = 1.20$, $\beta = 0.00$, $\theta = 0.25$	Theoretic	0.0000	1.0000	0.6000	18.3600	45.9360	NA
	Mixture	**0.0000** (0.0010)	**1.0000** (0.0039)	**0.6005** (0.0575)	18.2698 (0.9310)	44.1277 (26.3777)	0.734
	Inversion	0.0057 (0.0011)	0.9998 (0.0034)	0.6128 (0.0492)	**18.3982** (0.9655)	**45.7336** (27.8289)	0.997
	Devroye	-0.0001 (0.0010)	1.0002 (0.0041)	0.5935 (0.0503)	18.2433 (0.8961)	43.8200 (19.9383)	6.958
$\alpha = 1.20$, $\beta = 0.00$, $\theta = 0.50$	Theoretic	0.0000	1.0000	0.3000	6.8400	7.9920	NA
	Mixture	**0.0000** (0.0009)	1.0002 (0.0021)	**0.3022** (0.0140)	6.8373 (0.1395)	8.2052 (1.9206)	1.035
	Inversion	0.0035 (0.0009)	**0.9999** (0.0024)	0.3084 (0.0145)	6.8470 (0.1536)	**8.1545** (2.1488)	1.027
	Devroye	0.0002 (0.0011)	1.0003 (0.0026)	0.3034 (0.0144)	**6.8402** (0.1467)	8.1974 (1.8822)	6.957
$\alpha = 1.20$, $\beta = 0.00$, $\theta = 0.75$	Theoretic	0.0000	1.0000	0.2000	4.7067	3.4791	NA
	Mixture	0.0002 (0.0010)	0.9997 (0.0020)	**0.2006** (0.0089)	**4.7061** (0.0519)	**3.4880** (0.5747)	1.612
	Inversion	0.0026 (0.0010)	1.0001 (0.0021)	0.2083 (0.0084)	4.7240 (0.0492)	3.5865 (0.4492)	1.009
	Devroye	**-0.0001** (0.0009)	**1.0000** (0.0020)	0.1991 (0.0090)	4.6986 (0.0589)	3.4312 (0.4758)	6.951
$\alpha = 1.20$, $\beta = 0.00$, $\theta = 1.00$	Theoretic	0.0000	1.0000	0.1500	3.9600	2.1240	NA
	Mixture	0.0003 (0.0010)	0.9994 (0.0018)	0.1513 (0.0062)	3.9549 (0.0278)	2.1149 (0.1677)	2.711
	Inversion	0.0021 (0.0010)	**0.9999** (0.0017)	0.1569 (0.0068)	**3.9608** (0.0324)	2.2049 (0.2546)	0.987
	Devroye	**0.0000** (0.0011)	0.9997 (0.0017)	**0.1500** (0.0065)	3.9526 (0.0290)	**2.1232** (0.1844)	6.909

Table C.6: Theoretic vs. mean sample moments about the origin for standardised TS distribution. Emphasized numbers indicate either the smallest (mean) execution time, or the (mean) sample moment most similar to corresponding theoretic value.

Parameters	Method	Mean sample moments (1)	(2)	(3)	(4)	(5)	Mean time (sec)
	Theoretic	0.0000	1.0000	1.2000	18.3600	91.8720	NA
$\alpha = 1.20$,	Mixture	-0.0001	**1.0001**	1.1969	**18.3333**	88.4030	0.704
		(0.0010)	(0.0042)	(0.0520)	(1.0720)	(26.4375)	
$\beta = 0.00$,	Inversion	0.0058	1.0005	1.2125	18.3868	88.4468	0.986
		(0.0011)	(0.0038)	(0.0451)	(1.0207)	(24.2300)	
$\theta = 0.25$	Devroye	**0.0001**	1.0003	**1.1990**	18.3274	**90.6778**	7.084
		(0.0010)	(0.0045)	(0.0471)	(0.8857)	(25.0640)	
	Theoretic	0.0000	1.0000	0.6000	6.8400	15.9840	NA
$\alpha = 1.20$,	Mixture	0.0001	0.9998	0.6021	6.8588	16.5354	1.024
		(0.0010)	(0.0027)	(0.0153)	(0.1812)	(2.9643)	
$\beta = 0.00$,	Inversion	0.0035	1.0002	0.6112	6.8597	16.3037	1.015
		(0.0010)	(0.0021)	(0.0175)	(0.1736)	(2.6321)	
$\theta = 0.50$	Devroye	**0.0001**	**1.0000**	**0.6005**	**6.8310**	**16.0191**	6.974
		(0.0010)	(0.0024)	(0.0130)	(0.1362)	(1.6097)	
	Theoretic	0.0000	1.0000	0.4000	4.7067	6.9582	NA
$\alpha = 1.20$,	Mixture	**0.0001**	0.9999	**0.4002**	**4.7065**	6.9089	1.648
		(0.0012)	(0.0017)	(0.0089)	(0.0573)	(0.4536)	
$\beta = 0.00$,	Inversion	0.0025	**1.0000**	0.4078	4.7145	6.9861	0.998
		(0.0011)	(0.0018)	(0.0082)	(0.0505)	(0.5604)	
$\theta = 0.75$	Devroye	-0.0001	1.0002	0.3995	4.7157	**6.9544**	6.935
		(0.0010)	(0.0021)	(0.0084)	(0.0458)	(0.4381)	
	Theoretic	0.0000	1.0000	0.3000	3.9600	4.2480	NA
$\alpha = 1.20$,	Mixture	0.0004	0.9991	0.3031	3.9490	4.2772	2.860
		(0.0010)	(0.0018)	(0.0059)	(0.0256)	(0.1879)	
$\beta = 0.00$,	Inversion	0.0019	0.9999	0.3061	**3.9602**	4.3046	0.981
		(0.0010)	(0.0018)	(0.0067)	(0.0325)	(0.2116)	
$\theta = 1.00$	Devroye	**0.0000**	**1.0000**	**0.2998**	3.9586	**4.2471**	6.891
		(0.0010)	(0.0017)	(0.0065)	(0.0278)	(0.1979)	

Table C.7: Theoretic vs. mean sample moments about the origin for standardised TS distribution. Emphasized numbers indicate either the smallest (mean) execution time, or the (mean) sample moment most similar to corresponding theoretic value.

Parameters	Method	Mean sample moments (1)	(2)	(3)	(4)	(5)	Mean time (sec)
$\alpha = 1.20$, $\beta = 0.00$, $\theta = 0.25$	Theoretic	0.0000	1.0000	1.8000	18.3600	137.8080	NA
	Mixture	0.0002 (0.0010)	1.0010 (0.0037)	1.8114 (0.0485)	18.5616 (1.0119)	141.8032 (29.2527)	0.695
	Inversion	0.0058 (0.0011)	**0.9997** (0.0040)	1.8114 (0.0462)	18.2798 (1.0921)	139.2773 (35.7956)	0.975
	Devroye	**0.0000** (0.0011)	0.9995 (0.0043)	**1.8002** (0.0538)	**18.3133** (1.0302)	**136.4863** (28.3968)	7.298
$\alpha = 1.20$, $\beta = 0.00$, $\theta = 0.50$	Theoretic	0.0000	1.0000	0.9000	6.8400	23.9760	NA
	Mixture	0.0002 (0.0008)	**1.0001** (0.0023)	0.9035 (0.0144)	6.8498 (0.1392)	**24.0893** (1.7887)	1.036
	Inversion	0.0034 (0.0010)	1.0004 (0.0024)	0.9129 (0.0131)	6.8819 (0.1523)	24.3171 (2.0236)	1.014
	Devroye	**0.0000** (0.0010)	0.9997 (0.0026)	**0.9003** (0.0147)	**6.8361** (0.1420)	24.1728 (1.7990)	7.140
$\alpha = 1.20$, $\beta = 0.00$, $\theta = 0.75$	Theoretic	0.0000	1.0000	0.6000	4.7067	10.4373	NA
	Mixture	0.0002 (0.0010)	0.9998 (0.0019)	0.6019 (0.0089)	4.7017 (0.0519)	**10.4224** (0.4838)	1.707
	Inversion	0.0024 (0.0010)	**1.0000** (0.0018)	0.6068 (0.0079)	4.7105 (0.0481)	10.4835 (0.4410)	0.999
	Devroye	**-0.0001** (0.0009)	1.0001 (0.0020)	**0.5994** (0.0080)	**4.7056** (0.0514)	10.4010 (0.5098)	7.050
$\alpha = 1.20$, $\beta = 0.00$, $\theta = 1.00$	Theoretic	0.0000	1.0000	0.4500	3.9600	6.3720	NA
	Mixture	0.0006 (0.0008)	0.9984 (0.0015)	0.4550 (0.0057)	3.9522 (0.0262)	6.4225 (0.1966)	3.154
	Inversion	0.0021 (0.0009)	**0.9999** (0.0019)	0.4565 (0.0063)	3.9624 (0.0305)	6.4160 (0.2189)	0.981
	Devroye	**0.0000** (0.0009)	0.9998 (0.0016)	**0.4499** (0.0057)	**3.9571** (0.0296)	**6.3651** (0.1941)	6.970

Table C.8: Theoretic vs. mean sample moments about the origin for standardised TS distribution. Emphasized numbers indicate either the smallest (mean) execution time, or the (mean) sample moment most similar to corresponding theoretic value.

Parameters	Method	Mean sample moments (1)	(2)	(3)	(4)	(5)	Mean time (sec)
$\alpha = 1.60$, $\beta = 0.00$, $\theta = 0.25$	Theoretic	0.0000	1.0000	0.0000	11.9600	0.0000	NA
	Mixture	**0.0000** (0.0010)	**1.0007** (0.0035)	0.0010 (0.0361)	**12.0570** (0.7073)	1.4137 (17.8088)	0.805
	Inversion	0.0011 (0.0010)	0.9884 (0.0029)	**0.0002** (0.0176)	8.7298 (0.1449)	-0.1658 (1.0945)	0.333
	Devroye	0.0000 (0.0010)	0.9881 (0.0027)	-0.0010 (0.0174)	8.7220 (0.1310)	**-0.0694** (1.1528)	6.131
$\alpha = 1.60$, $\beta = 0.00$, $\theta = 0.50$	Theoretic	0.0000	1.0000	0.0000	5.2400	0.0000	NA
	Mixture	**0.0000** (0.0011)	**1.0001** (0.0021)	-0.0010 (0.0127)	**5.2386** (0.1049)	-0.2791 (2.0371)	1.176
	Inversion	0.0011 (0.0008)	0.9978 (0.0020)	0.0051 (0.0090)	4.9156 (0.0496)	0.0784 (0.3032)	0.323
	Devroye	0.0000 (0.0012)	0.9980 (0.0019)	**0.0001** (0.0094)	4.9156 (0.0437)	**-0.0050** (0.3069)	6.169
$\alpha = 1.60$, $\beta = 0.00$, $\theta = 0.75$	Theoretic	0.0000	1.0000	0.0000	3.9956	0.0000	NA
	Mixture	**-0.0001** (0.0010)	**0.9996** (0.0017)	**-0.0002** (0.0065)	**3.9909** (0.0376)	-0.0165 (0.2689)	1.901
	Inversion	0.0009 (0.0010)	0.9992 (0.0017)	0.0026 (0.0055)	3.9230 (0.0213)	0.0199 (0.1312)	0.320
	Devroye	-0.0001 (0.0011)	0.9995 (0.0019)	-0.0008 (0.0070)	3.9276 (0.0254)	**-0.0062** (0.1564)	6.188
$\alpha = 1.60$, $\beta = 0.00$, $\theta = 1.00$	Theoretic	0.0000	1.0000	0.0000	3.5600	0.0000	NA
	Mixture	0.0002 (0.0010)	0.9990 (0.0016)	0.0005 (0.0057)	**3.5488** (0.0223)	-0.0011 (0.1164)	3.103
	Inversion	0.0008 (0.0010)	**0.9998** (0.0014)	0.0020 (0.0047)	3.5304 (0.0172)	0.0046 (0.0750)	0.319
	Devroye	**-0.0001** (0.0011)	0.9993 (0.0017)	**-0.0002** (0.0052)	3.5232 (0.0190)	**-0.0004** (0.0792)	6.165

Table C.9: Theoretic vs. mean sample moments about the origin for standardised TS distribution. Emphasized numbers indicate either the smallest (mean) execution time, or the (mean) sample moment most similar to corresponding theoretic value.

Parameters	Method	Mean sample moments (1)	(2)	(3)	(4)	(5)	Mean time (sec)
$\alpha = 1.60$, $\beta = 0.25$, $\theta = 0.25$	Theoretic	0.0000	1.0000	0.4000	11.9600	25.5040	NA
	Mixture	**0.0000** (0.0010)	**0.9991** (0.0031)	**0.4013** (0.0356)	**11.8545** (0.6430)	**28.1014** (18.4283)	0.802
	Inversion	0.0012 (0.0011)	0.9861 (0.0029)	0.3723 (0.0182)	8.4574 (0.1341)	13.8200 (1.1095)	0.897
	Devroye	-0.0002 (0.0011)	0.9862 (0.0027)	0.3651 (0.0185)	8.4535 (0.1278)	13.5220 (1.1236)	6.730
$\alpha = 1.60$, $\beta = 0.25$, $\theta = 0.50$	Theoretic	0.0000	1.0000	0.2000	5.2400	4.6880	NA
	Mixture	0.0002 (0.0010)	**0.9998** (0.0019)	**0.2003** (0.0118)	**5.2279** (0.1038)	**4.7655** (1.2274)	1.188
	Inversion	0.0010 (0.0008)	0.9971 (0.0020)	0.2076 (0.0088)	4.8643 (0.0467)	4.5192 (0.2989)	0.863
	Devroye	**0.0001** (0.0009)	0.9975 (0.0020)	0.2062 (0.0073)	4.8705 (0.0473)	4.5228 (0.2635)	6.709
$\alpha = 1.60$, $\beta = 0.25$, $\theta = 0.75$	Theoretic	0.0000	1.0000	0.1333	3.9956	2.1298	NA
	Mixture	**0.0001** (0.0010)	0.9990 (0.0018)	**0.1332** (0.0071)	**3.9864** (0.0371)	**2.0949** (0.3906)	1.876
	Inversion	0.0011 (0.0010)	**0.9993** (0.0016)	0.1420 (0.0063)	3.9116 (0.0259)	2.3267 (0.1343)	0.825
	Devroye	0.0002 (0.0010)	0.9991 (0.0021)	0.1402 (0.0061)	3.9101 (0.0303)	2.3234 (0.1405)	6.676
$\alpha = 1.60$, $\beta = 0.25$, $\theta = 1.00$	Theoretic	0.0000	1.0000	0.1000	3.5600	1.3360	NA
	Mixture	0.0004 (0.0010)	0.9990 (0.0017)	0.1027 (0.0054)	**3.5511** (0.0213)	**1.3571** (0.1442)	3.138
	Inversion	0.0005 (0.0011)	**0.9998** (0.0015)	**0.0984** (0.0058)	3.5290 (0.0179)	1.1879 (0.0820)	0.782
	Devroye	**-0.0001** (0.0009)	0.9996 (0.0016)	0.0963 (0.0051)	3.5277 (0.0165)	1.1767 (0.0789)	6.645

Table C.10: Theoretic vs. mean sample moments about the origin for standardised TS distribution. Emphasized numbers indicate either the smallest (mean) execution time, or the (mean) sample moment most similar to corresponding theoretic value.

Parameters	Method	Mean sample moments (1)	(2)	(3)	(4)	(5)	Mean time (sec)
$\alpha = 1.60$, $\beta = 0.50$, $\theta = 0.25$	Theoretic	0.0000	1.0000	0.8000	11.9600	51.0080	NA
	Mixture	**0.0001** (0.0010)	**0.9996** (0.0035)	**0.8020** (0.0372)	**11.9443** (0.7070)	**52.8672** (18.0740)	0.790
	Inversion	0.0009 (0.0010)	0.9849 (0.0027)	0.7274 (0.0156)	8.3453 (0.1268)	25.9464 (0.9759)	0.897
	Devroye	-0.0003 (0.0011)	0.9855 (0.0026)	0.7289 (0.0162)	8.3742 (0.1257)	26.2184 (1.0264)	6.788
$\alpha = 1.60$, $\beta = 0.50$, $\theta = 0.50$	Theoretic	0.0000	1.0000	0.4000	5.2400	9.3760	NA
	Mixture	**0.0000** (0.0012)	**0.9998** (0.0022)	**0.4012** (0.0129)	**5.2425** (0.0939)	**9.3886** (1.1789)	1.191
	Inversion	0.0007 (0.0009)	0.9971 (0.0019)	0.3861 (0.0082)	4.8544 (0.0431)	7.4010 (0.2862)	0.863
	Devroye	-0.0003 (0.0010)	0.9974 (0.0020)	0.3832 (0.0085)	4.8621 (0.0448)	7.4038 (0.2913)	6.748
$\alpha = 1.60$, $\beta = 0.50$, $\theta = 0.75$	Theoretic	0.0000	1.0000	0.2667	3.9956	4.2596	NA
	Mixture	0.0004 (0.0009)	**0.9995** (0.0018)	0.2692 (0.0066)	**3.9924** (0.0407)	**4.3012** (0.3516)	1.930
	Inversion	0.0008 (0.0010)	0.9993 (0.0017)	**0.2664** (0.0059)	3.9088 (0.0250)	3.9788 (0.1279)	0.825
	Devroye	**0.0001** (0.0009)	0.9992 (0.0018)	0.2648 (0.0058)	3.9072 (0.0274)	3.9707 (0.1393)	6.706
$\alpha = 1.60$, $\beta = 0.50$, $\theta = 1.00$	Theoretic	0.0000	1.0000	0.2000	3.5600	2.6720	NA
	Mixture	0.0007 (0.0011)	0.9982 (0.0017)	0.2051 (0.0054)	**3.5449** (0.0221)	2.7207 (0.1168)	3.249
	Inversion	0.0007 (0.0009)	0.9993 (0.0017)	0.2030 (0.0049)	3.5159 (0.0188)	**2.6345** (0.0846)	0.780
	Devroye	**0.0000** (0.0010)	**0.9995** (0.0018)	**0.2005** (0.0049)	3.5172 (0.0192)	2.6162 (0.0759)	6.605

Table C.11: Theoretic vs. mean sample moments about the origin for standardised TS distribution. Emphasized numbers indicate either the smallest (mean) execution time, or the (mean) sample moment most similar to corresponding theoretic value.

Parameters	Method	Mean sample moments (1)	(2)	(3)	(4)	(5)	Mean time (sec)
$\alpha = 1.60$, $\beta = 0.75$, $\theta = 0.25$	Theoretic	0.0000	1.0000	1.2000	11.9600	76.5120	NA
	Mixture	**-0.0001** (0.0011)	**0.9995** (0.0040)	**1.1960** (0.0379)	**11.8614** (0.6827)	**74.0155** (16.4289)	0.749
	Inversion	0.0013 (0.0010)	0.9888 (0.0030)	1.1284 (0.0206)	8.9725 (0.1763)	43.6179 (1.6690)	0.885
	Devroye	0.0001 (0.0011)	0.9885 (0.0026)	1.1225 (0.0197)	8.9439 (0.1541)	43.3129 (1.5476)	6.855
$\alpha = 1.60$, $\beta = 0.75$, $\theta = 0.50$	Theoretic	0.0000	1.0000	0.6000	5.2400	14.0640	NA
	Mixture	**0.0001** (0.0010)	**0.9996** (0.0021)	**0.6007** (0.0105)	**5.2368** (0.0996)	**14.0554** (1.2328)	1.123
	Inversion	0.0005 (0.0009)	0.9966 (0.0019)	0.5806 (0.0094)	4.8376 (0.0504)	11.1898 (0.3315)	0.852
	Devroye	-0.0001 (0.0010)	0.9970 (0.0019)	0.5792 (0.0079)	4.8419 (0.0490)	11.2072 (0.2758)	6.742
$\alpha = 1.60$, $\beta = 0.75$, $\theta = 0.75$	Theoretic	0.0000	1.0000	0.4000	3.9956	6.3893	NA
	Mixture	0.0003 (0.0009)	**0.9993** (0.0018)	0.4030 (0.0071)	**3.9920** (0.0435)	**6.4052** (0.3430)	1.939
	Inversion	0.0007 (0.0010)	0.9988 (0.0019)	0.4017 (0.0064)	3.9063 (0.0287)	6.0654 (0.1417)	0.823
	Devroye	**0.0000** (0.0010)	0.9991 (0.0017)	**0.4002** (0.0059)	3.9067 (0.0267)	6.0661 (0.1399)	6.732
$\alpha = 1.60$, $\beta = 0.75$, $\theta = 1.00$	Theoretic	0.0000	1.0000	0.3000	3.5600	4.0080	NA
	Mixture	0.0011 (0.0011)	0.9975 (0.0017)	0.3088 (0.0054)	**3.5379** (0.0231)	**4.0786** (0.1276)	3.553
	Inversion	0.0005 (0.0009)	**0.9995** (0.0013)	0.2957 (0.0042)	3.5144 (0.0165)	3.7128 (0.0708)	0.788
	Devroye	**-0.0004** (0.0011)	0.9993 (0.0015)	**0.2927** (0.0050)	3.5099 (0.0160)	3.6918 (0.0705)	6.735

Table C.12: Theoretic vs. mean sample moments about the origin for standardised TS distribution. Emphasized numbers indicate either the smallest (mean) execution time, or the (mean) sample moment most similar to corresponding theoretic value.

Parameters	Method	Mean sample moments (1)	(2)	(3)	(4)	(5)	Mean time (sec)
$\alpha = 1.80$, $\beta = 0.00$, $\theta = 0.25$	Theoretic	0.0000	1.0000	0.0000	6.8400	0.0000	NA
	Mixture	**0.0000** (0.0008)	**0.9997** (0.0025)	-0.0006 (0.0212)	**6.8254** (0.4086)	0.9242 (8.1849)	0.852
	Inversion	0.0008 (0.0010)	0.9910 (0.0024)	0.0028 (0.0092)	5.0994 (0.0671)	0.0397 (0.4069)	0.345
	Devroye	0.0002 (0.0009)	0.9909 (0.0018)	**0.0003** (0.0092)	5.0945 (0.0553)	**0.0027** (0.3653)	6.237
$\alpha = 1.80$, $\beta = 0.00$, $\theta = 0.50$	Theoretic	0.0000	1.0000	0.0000	3.9600	0.0000	NA
	Mixture	**-0.0001** (0.0010)	**0.9998** (0.0016)	**-0.0008** (0.0077)	**3.9459** (0.0497)	-0.0109 (0.5335)	1.262
	Inversion	0.0007 (0.0010)	0.9985 (0.0016)	0.0017 (0.0060)	3.7838 (0.0266)	0.0053 (0.1559)	0.330
	Devroye	-0.0002 (0.0011)	0.9983 (0.0016)	-0.0010 (0.0058)	3.7839 (0.0252)	**-0.0030** (0.1406)	6.130
$\alpha = 1.80$, $\beta = 0.00$, $\theta = 0.75$	Theoretic	0.0000	1.0000	0.0000	3.4267	0.0000	NA
	Mixture	-0.0001 (0.0010)	0.9995 (0.0016)	0.0004 (0.0054)	**3.4176** (0.0240)	0.0406 (0.2208)	2.029
	Inversion	0.0006 (0.0010)	0.9994 (0.0016)	0.0016 (0.0050)	3.3789 (0.0148)	0.0038 (0.0750)	0.323
	Devroye	**-0.0001** (0.0011)	**0.9996** (0.0014)	**-0.0003** (0.0054)	3.3805 (0.0166)	**-0.0014** (0.0744)	6.107
$\alpha = 1.80$, $\beta = 0.00$, $\theta = 1.00$	Theoretic	0.0000	1.0000	0.0000	3.2400	0.0000	NA
	Mixture	0.0001 (0.0011)	0.9984 (0.0015)	**-0.0001** (0.0049)	**3.2278** (0.0156)	**0.0000** (0.0768)	3.336
	Inversion	0.0008 (0.0010)	**0.9994** (0.0014)	0.0025 (0.0046)	3.2090 (0.0126)	0.0128 (0.0509)	0.325
	Devroye	**-0.0001** (0.0010)	0.9994 (0.0015)	-0.0009 (0.0043)	3.2093 (0.0129)	-0.0108 (0.0471)	6.093

Table C.13: Theoretic vs. mean sample moments about the origin for standardised TS distribution. Emphasized numbers indicate either the smallest (mean) execution time, or the (mean) sample moment most similar to corresponding theoretic value.

Parameters	Method	Mean sample moments (1)	(2)	(3)	(4)	(5)	Mean time (sec)
$\alpha = 1.80$, $\beta = 0.25$, $\theta = 0.25$	Theoretic	0.0000	1.0000	0.2000	6.8400	10.4480	NA
	Mixture	**0.0000** (0.0010)	**1.0003** (0.0024)	**0.2015** (0.0256)	**6.9020** (0.4019)	**11.2557** (10.7739)	0.824
	Inversion	0.0008 (0.0010)	0.9923 (0.0018)	0.1910 (0.0101)	5.2755 (0.0671)	5.9576 (0.5152)	0.863
	Devroye	0.0001 (0.0011)	0.9928 (0.0020)	0.1884 (0.0109)	5.2793 (0.0690)	5.8653 (0.5222)	6.654
$\alpha = 1.80$, $\beta = 0.25$, $\theta = 0.50$	Theoretic	0.0000	1.0000	0.1000	3.9600	2.0560	NA
	Mixture	**0.0001** (0.0010)	**0.9998** (0.0015)	**0.1000** (0.0069)	**3.9535** (0.0522)	1.8741 (0.9295)	1.268
	Inversion	0.0011 (0.0011)	0.9983 (0.0015)	0.1083 (0.0062)	3.7617 (0.0239)	**2.0549** (0.1478)	0.821
	Devroye	0.0002 (0.0010)	0.9985 (0.0017)	0.1044 (0.0056)	3.7627 (0.0249)	2.0043 (0.1294)	6.612
$\alpha = 1.80$, $\beta = 0.25$, $\theta = 0.75$	Theoretic	0.0000	1.0000	0.0667	3.4267	0.9796	NA
	Mixture	0.0001 (0.0011)	0.9991 (0.0016)	**0.0675** (0.0055)	**3.4166** (0.0239)	**0.9547** (0.1915)	2.019
	Inversion	0.0006 (0.0010)	0.9994 (0.0015)	0.0651 (0.0044)	3.3772 (0.0143)	0.8113 (0.0675)	0.778
	Devroye	**-0.0001** (0.0010)	**0.9994** (0.0015)	0.0628 (0.0050)	3.3782 (0.0161)	0.7955 (0.0716)	6.570
$\alpha = 1.80$, $\beta = 0.25$, $\theta = 1.00$	Theoretic	0.0000	1.0000	0.0500	3.2400	0.6320	NA
	Mixture	0.0005 (0.0009)	0.9980 (0.0014)	0.0541 (0.0041)	3.2220 (0.0143)	**0.6770** (0.0876)	3.349
	Inversion	0.0006 (0.0011)	**0.9998** (0.0014)	0.0550 (0.0050)	**3.2286** (0.0125)	0.7408 (0.0595)	0.778
	Devroye	**0.0001** (0.0010)	0.9997 (0.0014)	**0.0535** (0.0045)	3.2268 (0.0125)	0.7324 (0.0549)	6.563

Table C.14: Theoretic vs. mean sample moments about the origin for standardised TS distribution. Emphasized numbers indicate either the smallest (mean) execution time, or the (mean) sample moment most similar to corresponding theoretic value.

Parameters	Method	Mean sample moments (1)	(2)	(3)	(4)	(5)	Mean time (sec)
$\alpha = 1.80$, $\beta = 0.50$, $\theta = 0.25$	Theoretic	0.0000	1.0000	0.4000	6.8400	20.8960	NA
	Mixture	**0.0000**	**0.9996**	**0.3925**	**6.8336**	**18.3487**	0.817
		(0.0009)	(0.0025)	(0.0204)	(0.4108)	(9.8022)	
	Inversion	0.0010	0.9916	0.3744	5.2340	11.1214	0.863
		(0.0011)	(0.0021)	(0.0113)	(0.0838)	(0.6140)	
	Devroye	0.0001	0.9916	0.3701	5.2205	10.9838	6.653
		(0.0010)	(0.0020)	(0.0109)	(0.0665)	(0.5142)	
$\alpha = 1.80$, $\beta = 0.50$, $\theta = 0.50$	Theoretic	0.0000	1.0000	0.2000	3.9600	4.1120	NA
	Mixture	**0.0000**	**0.9995**	**0.2012**	**3.9574**	**4.4035**	1.263
		(0.0009)	(0.0017)	(0.0077)	(0.0860)	(1.8661)	
	Inversion	0.0007	0.9979	0.1941	3.7545	3.2538	0.821
		(0.0010)	(0.0018)	(0.0060)	(0.0295)	(0.1473)	
	Devroye	-0.0002	0.9982	0.1925	3.7588	3.2701	6.602
		(0.0009)	(0.0016)	(0.0052)	(0.0248)	(0.1223)	
$\alpha = 1.80$, $\beta = 0.50$, $\theta = 0.75$	Theoretic	0.0000	1.0000	0.1333	3.4267	1.9591	NA
	Mixture	0.0003	0.9988	0.1361	**3.4155**	**1.9743**	2.062
		(0.0010)	(0.0014)	(0.0056)	(0.0226)	(0.1801)	
	Inversion	0.0006	**0.9993**	0.1350	3.3690	1.8519	0.780
		(0.0010)	(0.0016)	(0.0048)	(0.0167)	(0.0717)	
	Devroye	**-0.0001**	0.9992	**0.1326**	3.3670	1.8395	6.554
		(0.0010)	(0.0016)	(0.0042)	(0.0166)	(0.0645)	
$\alpha = 1.80$, $\beta = 0.50$, $\theta = 1.00$	Theoretic	0.0000	1.0000	0.1000	3.2400	1.2640	NA
	Mixture	0.0009	0.9972	0.1079	3.2123	1.3368	3.489
		(0.0010)	(0.0017)	(0.0046)	(0.0158)	(0.0787)	
	Inversion	0.0008	0.9998	0.1031	3.2265	1.2802	0.780
		(0.0010)	(0.0016)	(0.0043)	(0.0154)	(0.0544)	
	Devroye	**0.0000**	**1.0000**	**0.1010**	**3.2283**	**1.2709**	6.560
		(0.0009)	(0.0014)	(0.0045)	(0.0127)	(0.0549)	

Table C.15: Theoretic vs. mean sample moments about the origin for standardised TS distribution. Emphasized numbers indicate either the smallest (mean) execution time, or the (mean) sample moment most similar to corresponding theoretic value.

Parameters	Method	Mean sample moments (1)	(2)	(3)	(4)	(5)	Mean time (sec)
$\alpha = 1.80$, $\beta = 0.75$, $\theta = 0.25$	Theoretic	0.0000	1.0000	0.6000	6.8400	31.3440	NA
	Mixture	**0.0000**	**0.9999**	**0.6012**	**6.8453**	**30.4744**	0.802
		(0.0010)	(0.0026)	(0.0233)	(0.4204)	(10.4944)	
	Inversion	0.0008	0.9913	0.5397	5.2227	15.0194	0.865
		(0.0009)	(0.0019)	(0.0101)	(0.0716)	(0.5467)	
	Devroye	-0.0002	0.9919	0.5392	5.2362	15.0942	6.688
		(0.0010)	(0.0020)	(0.0108)	(0.0704)	(0.5430)	
$\alpha = 1.80$, $\beta = 0.75$, $\theta = 0.50$	Theoretic	0.0000	1.0000	0.3000	3.9600	6.1680	NA
	Mixture	0.0001	**0.9996**	**0.3012**	**3.9509**	**6.1390**	1.272
		(0.0009)	(0.0015)	(0.0067)	(0.0587)	(0.8326)	
	Inversion	0.0007	0.9978	0.2934	3.7521	4.9577	0.822
		(0.0010)	(0.0016)	(0.0053)	(0.0242)	(0.1247)	
	Devroye	**0.0000**	0.9979	0.2914	3.7503	4.9346	6.621
		(0.0009)	(0.0018)	(0.0060)	(0.0287)	(0.1508)	
$\alpha = 1.80$, $\beta = 0.75$, $\theta = 0.75$	Theoretic	0.0000	1.0000	0.2000	3.4267	2.9387	NA
	Mixture	0.0005	0.9988	**0.2047**	**3.4146**	**2.9862**	2.143
		(0.0010)	(0.0017)	(0.0051)	(0.0237)	(0.1896)	
	Inversion	0.0004	0.9990	0.1951	3.3639	2.5665	0.778
		(0.0009)	(0.0015)	(0.0047)	(0.0151)	(0.0679)	
	Devroye	**-0.0001**	**0.9992**	0.1930	3.3638	2.5537	6.568
		(0.0011)	(0.0018)	(0.0048)	(0.0173)	(0.0754)	
$\alpha = 1.80$, $\beta = 0.75$, $\theta = 1.00$	Theoretic	0.0000	1.0000	0.1500	3.2400	1.8960	NA
	Mixture	0.0014	0.9963	0.1624	3.2017	2.0169	3.764
		(0.0010)	(0.0017)	(0.0044)	(0.0179)	(0.0808)	
	Inversion	0.0007	0.9997	**0.1503**	3.2232	**1.8192**	0.779
		(0.0010)	(0.0016)	(0.0043)	(0.0139)	(0.0515)	
	Devroye	**-0.0001**	**0.9999**	0.1481	**3.2248**	1.8083	6.572
		(0.0009)	(0.0015)	(0.0041)	(0.0134)	(0.0517)	

Table C.16: Theoretic vs. mean sample moments about the origin for standardised TS distribution. Emphasized numbers indicate either the smallest (mean) execution time, or the (mean) sample moment most similar to corresponding theoretic value.

Appendix D

Portfolio problem

Let b denote a bank that is a borrower while l stands for a lender. For the sake of brevity we shall denote $i \equiv i_{bl}$ as the interest rate is always an interbank interest rate agreed between both parties. Every bank k learns the mean and standard deviation of risky asset growth rates from its own past data. All the banks dynamically update their joint beliefs on the probability of counterpart default. Again for the sake of brevity we shall introduce the following shortcut notation: $p \equiv p_{t+1}$, $\mu \equiv \mu_{k,t+1}$ and $\sigma \equiv \sigma_{k,t+1}$, $A \equiv A_k$, $B \equiv B_k$, $E \equiv E_k$ and $F \equiv F_k$. Demand and supply characterize the behaviour of, respectively, borrower and lender, thus we may write $d \equiv d_b$, $s \equiv s_l$. The quantities v, $\hat{w}$ and $\hat{i}$ that appear in the borrower problem describe characteristics of the borrower, the same notation in the lender problem pertains to lender, in both cases the indexes were omitted for easier display. Denote $\chi_b := c^{-1}\mathbb{1}_{\{w \geq \underline{w}\}}(w) + c\mathbb{1}_{\{w < \underline{w}\}}(w)$ for borrower and $\chi_l := c^{-1}\mathbb{1}_{\{w > \underline{w}\}}(w) + c\mathbb{1}_{\{w \leq \underline{w}\}}(w)$ in case of lender. Note that both threshold functions differ in $\underline{w}$.

D.1 Borrower problem

Assume $w \in I_B$. For $w > \underline{w}$ the capital of b at $t+1$ amounts to

$$\underline{w}(1+R_{t+1}) + \mathbb{1}_{\{w>\underline{w}\}}(w)(w-\underline{w})(1+R_{t+1}) - (1+i)(c^{-1}\mathbb{1}_{\{w>\underline{w}\}}(w)(w-\underline{w}) - v - \hat{w}) - \hat{w}(1+\hat{i}),$$

in case of $w = \underline{w}$ we have

$$\underline{w}(1+R_{t+1}) - (1+i)(-v-\hat{w}) - \hat{w}(1+\hat{i}),$$

while for $w < \underline{w}$ we obtain

$$\underline{w}(1+R_{t+1})+\mathbb{1}_{\{w<\underline{w}\}}(w)(w-\underline{w})(1+R_{t+1})-(1+i)(c\mathbb{1}_{\{w<\underline{w}\}}(w)(w-\underline{w})-v-\hat{w})-\hat{w}(1+\hat{i}).$$

Combining the three cases above yields the formula for $V_{t+1}(w)$ which is

$$V_{t+1}(w) = w(1+R_{t+1}) - (1+i)(\chi_b(w-\underline{w}) - v - \hat{w}) - \hat{w}(1+\hat{i}).$$

The conditional utility that borrower b expects to obtain at $t+1$ by investing in a portfolio, consisting of $w \in I_B$ units of shares and a loan from l, is

$$\mathbb{E}u(V_{t+1}(w)|B_{b,t+1}=0) = \frac{\mathbb{E}u(V_{t+1}(w), B_{b,t+1}=0)}{\mathbb{P}(B_{b,t+1}=0)} =$$

$$= \mathbb{E}u([(1+i)(v+\hat{w}+\chi_b\underline{w}) - \hat{w}(1+\hat{i})] + [1-(1+i)\chi_b]w + wR_{t+1}) =$$

$$= \mathbb{E}u(A + Bw + wR_{t+1}).$$

D.2 Lender problem

Assume $w \in I_L$. For $w > \underline{w}$ the capital of l at $t+1$ amounts to

$$\underline{w}(1+R_{t+1}) + \mathbb{1}_{\{w>\underline{w}\}}(w)(w-\underline{w})(1+R_{t+1})+$$

$$+((1-\theta)B_{b,t+1} + (1+i)(1-B_{b,t+1}))(v+\hat{w}-c^{-1}\mathbb{1}_{\{w>\underline{w}\}}(w)(w-\underline{w})) - \hat{w}(1+\hat{i}),$$

in case of $w = \underline{w}$ we have

$$\underline{w}(1+R_{t+1}) + ((1-\theta)B_{b,t+1} + (1+i)(1-B_{b,t+1}))(v+\hat{w}) - \hat{w}(1+\hat{i}),$$

while for $w < \underline{w}$ we obtain

$$\underline{w}(1+R_{t+1}) + \mathbb{1}_{\{w<\underline{w}\}}(w)(w-\underline{w})(1+R_{t+1})+$$

$$+((1-\theta)B_{b,t+1} + (1+i)(1-B_{b,t+1}))(v+\hat{w}-c\mathbb{1}_{\{w<\underline{w}\}}(w)(w-\underline{w})) - \hat{w}(1+\hat{i}).$$

After combining the three cases above and substituting χ_l the formula for $V_{t+1}(w)$ is

$$V_{t+1}(w) = (w(1+R_{t+1})+((1-\theta)B_{b,t+1}+(1+i)(1-B_{b,t+1}))(v+\hat{w}-\chi_l(w-\underline{w}))-\hat{w}(1+\hat{i}).$$

Therefore by the rule of iterated expectation

$$\mathbb{E}u(V_{t+1}(w)) = \mathbb{E}[\mathbb{E}u(V_{t+1}(w)|B_{b,t+1})] =$$

$$= \mathbb{P}(B_{b,t+1} = 1) \cdot \mathbb{E}u(V_{t+1}(w)|B_{b,t+1} = 1) + \mathbb{P}(B_{b,t+1} = 0) \cdot \mathbb{E}u(V_{t+1}(w)|B_{b,t+1} = 0) =$$

$$= p \cdot \mathbb{E}u([(1-\theta)(v + \hat{w} + \chi_l \underline{w}) - \hat{w}(1+\hat{i})] + [1-(1-\theta)\chi_l]w + wR_{t+1})+$$

$$+(1-p) \cdot \mathbb{E}u([(1+i)(v + \hat{w} + \chi_l \underline{w}) - \hat{w}(1+\hat{i})] + [1-(1+i)\chi_l]w + wR_{t+1}) =$$

$$= p \cdot \mathbb{E}u(E + Fw + wR_{t+1}) + (1-p) \cdot \mathbb{E}u(A + Bw + wR_{t+1}).$$

D.3 Utility under CARA

Assume that for each bank k displays constant absolute risk aversion with parameter $\gamma_k > 0$. Let $f_{R_t}(x)$ be a pdf of random variable R_t. Assume R_t has a moment generating function $M_t(q)$, defined for $q \in I \subset \mathbb{R}$.

Under these assumptions made for the **borrower** problem we have

$$\mathbb{E}u(V_{t+1}(w)|B_{b,t+1} = 0) = \mathbb{E}u(A + Bw + wR_{t+1}) = 1 - \mathbb{E}e^{-\gamma_b(A+Bw+wR_{t+1})} =$$

$$= 1 - e^{-\gamma_b(A+Bw)} \cdot M_{t+1}(-\gamma_b w).$$

Thus the optimal portfolio is given by

$$w_b^* = \underset{w \in I_B}{\operatorname{argmax}} \left\{1 - e^{-\gamma_b(A+Bw)} \cdot M_{t+1}(-\gamma_b w)\right\} = \underset{w \in I_B}{\operatorname{argmin}} \left\{e^{-\gamma_b(A+Bw)} \cdot M_{t+1}(-\gamma_b w)\right\},$$

where the interval I_B is given by

$$I_B = \{w : \ w \geq \underline{w} + \chi_b^{-1}(v + \hat{w})\}.$$

The borrower's f.o.c. may be written as

$$\frac{d\log M_{t+1}(q)}{dq}\Big|_{q=-\gamma_b w} + B = 0. \tag{D.1}$$

To validate the consistency of assumptions we need to check if $-\gamma_b w_b^* \in I$.

For the **lender** problem we obtain

$$\mathbb{E}u(V_{t+1}(w)) = p\cdot\mathbb{E}u(E+Fw+wR_{t+1}) + (1-p)\cdot\mathbb{E}u(A+Bw+wR_{t+1}) =$$

$$= p(1-\mathbb{E}e^{-\gamma_l(E+Fw+wR_{t+1})}) + (1-p)(1-\mathbb{E}e^{-\gamma_l(A+Bw+wR_{t+1})}) =$$

$$= 1 - p\cdot e^{-\gamma_l(E+Fw)}\cdot M_{t+1}(-\gamma_l w) - (1-p)\cdot e^{-\gamma_l(A+Bw)}\cdot M_{t+1}(-\gamma_l w).$$

Thus the optimal unit portfolio is given by

$$w_l^* = \operatorname*{argmax}_{w\in I_L}\left\{1-\left(p\cdot e^{-\gamma_l(E+Fw)} + (1-p)\cdot e^{-\gamma_l(A+Bw)}\right)\cdot M_{t+1}(-\gamma_l w)\right\} =$$

$$= \operatorname*{argmin}_{w\in I_L}\left\{\left(p\cdot e^{-\gamma_l(E+Fw)} + (1-p)\cdot e^{-\gamma_l(A+Bw)}\right)\cdot M_{t+1}(-\gamma_l w)\right\},$$

where the interval I_L is defined by

$$I_L = \{w:\ w \le \underline{w} + \chi_l^{-1}(v+\hat{w})\}$$

The lender's f.o.c. may be written as

$$\frac{d\log M_{t+1}(q)}{dq}\Big|_{q=-\gamma_b w} + \frac{Fp\cdot e^{-\gamma_l(E+Fw)} + B(1-p)\cdot e^{-\gamma_l(A+Bw)}}{p\cdot e^{-\gamma_l(E+Fw)} + (1-p)\cdot e^{-\gamma_l(A+Bw)}} = 0. \tag{D.2}$$

To validate the consistency of assumptions we need to check if $-\gamma_l w_l^* \in I$.

D.4 Towards analytical expression for *i* underline

Start by rewriting **lender's** f.o.c.

$$\frac{d\ln M_{t+1}(q)}{dq}\Big|_{q=-\gamma_l w} + \frac{Fp\cdot e^{-\gamma_l(E+Fw)} + B(1-p)\cdot e^{-\gamma_l(A+Bw)}}{p\cdot e^{-\gamma_l(E+Fw)} + (1-p)\cdot e^{-\gamma_l(A+Bw)}} = 0 \equiv$$

$$\equiv \Big(\frac{d\ln M_{t+1}(q)}{dq}\Big|_{q=-\gamma_l w} + F\Big) p \cdot e^{-\gamma_l(E+Fw)} + \Big(\frac{d\ln M_{t+1}(q)}{dq}\Big|_{q=-\gamma_l w} + B\Big)\cdot$$

$$\cdot(1-p)\cdot e^{-\gamma_l(A+Bw)} = 0 \equiv \Big(\frac{d\ln M_{t+1}(q)}{dq}\Big|_{q=-\gamma_l w} +$$

$$+B-(B-F)\Big) + \Big(\frac{d\ln M_{t+1}(q)}{dq}\Big|_{q=-\gamma_l w} + B\Big)\cdot e^{-\gamma_l[(A-E)+(B-F)w]-\ln\frac{p}{1-p}} = 0.$$

Under Gaussian stock returns we have

$$A-E = (\theta+i)(v+\hat{w}+\chi_l\underline{w}),\ B-F = -(\theta+i)\chi_l,\ B = 1-(1+i)\chi_l,$$

$$\frac{d\ln M_{t+1}(q)}{dq}\Big|_{q=-\gamma_l w} + B = (1+\mu)-(1+i)\chi_l - \gamma_l\sigma^2 w,$$

Define the following constants

$$C_1 := -\gamma_l(A-E) - \ln\frac{p}{1-p} = -\gamma_l(\theta+i)(v+\hat{w}+\chi_l\underline{w}) - \ln\frac{p}{1-p},$$

$$C_2 := -\gamma_l(B-F) = \gamma_l(\theta+i)\chi_l,\ C_3 := (1+\mu)-(1+i)\chi_l,$$

$$C_4 := -\gamma_l\sigma^2,\ C_5 := B-F = -(\theta+i)\chi_l.$$

Assume in addition $w \neq -C_3/C_4$. Then the lender's f.o.c. may be written as

$$e^{C_1+C_2 w} = -\frac{C_3 - C_5 + C_4 w}{C_3 + C_4 w} = \frac{C_5}{C_3+C_4 w} - 1.$$

Note that if $C_5/(C_3+C_4w) \leq 1$ then the equality above has no solution as its right hand side is non–positive. Thus we need to have $C_5/(C_3+C_4w) > 1$ which is equivalent to $w < (C_5-C_3)/C_4$. For plausible parameter values C_5 is always negative, hence the latter also implies that $w \neq -C_3/C_4$. Therefore given $C_5 < 0$ the lender's f.o.c. is equivalent to

$$C_1 + C_2 w - \ln\Big(\frac{C_5}{C_3+C_4 w} - 1\Big) = 0. \tag{D.3}$$

if and only if $w < (C_5-C_3)/C_4$. Otherwise it has no interior solution.

D.5 Utility maximization under Gaussian returns

If $R_{t+1} \sim N(\mu, \sigma^2)$, we have:

$$M_{t+1}(q) = e^{\mu q + \frac{1}{2}\sigma^2 q^2}, \ \frac{d \log M_{t+1}(q)}{dq} = \mu + \sigma^2 q, \ I \equiv \mathbb{R}.$$

Then the f.o.c for **borrower** problem is equivalent to

$$\mu + \sigma^2 w + B = 0 \ \equiv \ \mu - \gamma_b \sigma^2 w + [1 - \chi_b(1+i)] = 0 \ \equiv$$

$$\equiv \ w = \frac{1}{\gamma_b \sigma^2}[(1+\mu) - \chi_b(1+i)]. \tag{D.4}$$

Before we proceed, it is convenient to identify the set of potential corners. Denote the set of admissible corners of the borrower problem as E_B, and the set of admissible corners of the lender problem as E_L, set $e_b = \underline{w} + \chi_b^{-1}(v + \hat{w})$ and $e_l = \underline{w} + \chi_l^{-1}(v + \hat{w})$. If $v + \hat{w} > 0$ then $E_B = \{\max\{0, e_b\}\}$, otherwise $E_B = \{\max\{0, e_b\}, \underline{w}\}$. If $v + \hat{w} < 0$ then $E_L = \{0, \max\{0, e_l\}\}$, otherwise $E_L = \{0, \max\{0, e_l\}, \underline{w}\}$. Hence the borrower problem has at most two while the lender problem has at most three corner solutions.

The following proposition verifies that any internal solution that satisfies first order condition of either the borrower or the lender problem is a strict (local) maximum.

Fact D.5.1 (Local strict concavity)**.** *Under Gaussian stock returns:*

i) *borrower's conditional expected utility $\mathbb{E}u(V_{t+1}(w)|B_{b,t+1} = 0)$ is strictly concave on any real interval $I \subset (e, +\infty) \setminus E_B$,*

ii) *lender's expected utility $\mathbb{E}u(V_{t+1}(w))$ is strictly concave on any real interval $I \subset (0, e) \setminus E_L$.*

Proof: i) First consider an auxiliary function $f(x) = e^{a+bx+cx^2}$ with real parameters a, b, c. We have

$$f''(x) = e^{a+bx+cx^2}[(b+2cx)^2 + 2c], \ f''(x) > 0 \equiv (b+2cx)^2 + 2c > 0.$$

To demonstrate $f(x)$ is strictly convex on any interval $I \in \mathbb{R}$ it is enough to show that $c > 0$. To show that $\mathbb{E}u(V_{t+1}(w)|B_{k,t+1} = 0)$ is strictly concave on any interval

$I \subset (e, +\infty) \setminus E$ it is enough to prove that

$$e^{-\gamma_b(A+Bw)} \cdot M_{t+1}(-\gamma_b w)$$

is strictly convex on given interval. Under Gaussian stock returns this expression reads as

$$e^{-\gamma_b(A+Bw)} \cdot e^{-\gamma_b \mu w+\frac{1}{2}\gamma_b^2\sigma^2 w^2} = e^{\frac{1}{2}(\gamma_b\sigma)^2 w^2+\ldots}$$

where all the terms of order lower than two were omitted on the right hand side. As $(\gamma_b\sigma)^2 > 0$, this function is strictly convex.

ii) Note that $f(w) = 1$ is linear, hence both weakly convex and weakly concave. By previous point the expected utility is a weighted average of one weakly concave function and two strictly concave functions. The outcome is strictly concave. □

Fact D.5.1 allows us to simplify the utility maximization problem. Define $E_B^{'}$ and $E_L^{'}$ as the set of (potential) expected utility maximizing portfolios for, respectively, borrower and lender problem. Initially set $E_B^{'} := E_B$ and $E_L^{'} := E_L$. Denote optimal portfolio choice as $w_b^* \equiv w_b^*(i_{bl})$ for the borrower and $w_l^* \equiv w_l^*(i_{bl})$ for the lender. Let w satisfy formula (D.4). If $w > \max\{\underline{w}, e_b\}$ then $E_B^{'} := E_B^{'} \cup \{w\}$, if $e_b < w < \underline{w}$ then $E_B^{'} := E_B^{'} \cup \{w\}$. Let $E_L^{'} := E_L \cup S$ where S is a set of internal solutions of lender's f.o.c's. A number of internal solutions $\#S$ is at most two.

The optimal borrower's portfolio is given by

$$w_b^* = \underset{w \in E_B^{'}}{\operatorname{argmax}} \{\mathbb{E}u(V_{t+1}(w) | B_{b,t+1} = 0)\},$$

while lender's expected utility maximizing portfolio is

$$w_l^* = \underset{w \in E_L^{'}}{\operatorname{argmax}} \{\mathbb{E}u(V_{t+1}(w))\}.$$

D.6 Credit demand and supply

Having characterized both the set of feasible corner solutions and the formulas for interior solutions we may derive demand for credit and supply of credit as a function of interest rate.

The monetary **demand** for credit results from the difference between the optimal portfolio choice and the largest position that b could finance without additional founding. Note that for sufficiently small demand for credit $d > 0$ the optimal portfolio choice w_b^* is an internal solution of the borrower problem which lies above $e_b = \underline{w} + \chi_b^{-1}(v + \hat{w})$. There are two distinct cases. If $v + \hat{w} \geq 0$ then $w_b^* \geq \underline{w}$. Bank b utilizes credit to buy more shares than it had at the end of previous period. As buying large quantities of shares causes price slippage, for every extra unit of shares it has to pay c^{-1} units of money. So monetary demand for credit is given by $c^{-1}(w_b^* - e_b)$. By a similar token, if $v + \hat{w} < 0$ and if $d > 0$ is sufficiently small, then we have $w_b^* < \underline{w}$. Bank b takes credit in order not to sell shares it already has in its portfolio. As selling large quantities of assets incurs cost, for every c units of money it borrows it can refrain from selling one unit of shares. So the monetary demand for credit is given by $c(w_b^* - e_b)$. In this case credit demand is sufficiently small if only $d < -v - \hat{w}$. As threshold function χ_b above depends on w_b^*, this additional restriction guarantees that borrowers actions are indeed optimal. In our notation monetary credit demand may be in both cases expressed as

$$d(i) = \chi_b(w_b^* - e_b) = \chi(w_b^* - \underline{w}) - (v + \hat{w}). \tag{D.5}$$

Given previous results the solution of borrower's problem w_b^* may be treated as given for each offered interest rate i. In both cases we may determine $\bar{i}_b$ as a limit interest rate implied by $d \to 0^+$.

Now we may find the largest interest rate $\bar{i}_b$ that **borrower** b would be willing to accept. The monetary demand for credit as a function of the offered interest rate is

$$d(i) = \chi_b\Big(\frac{1}{\gamma_b \sigma^2}[(1+\mu) - \chi_b(1+i)] - \underline{w}\Big) - (v + \hat{w}) \equiv \chi_b^{-1}(v + \hat{w} + d(i)) =$$

$$= \frac{1}{\gamma_b \sigma^2}[(1+\mu) - \chi_b(1+i)] - \underline{w} \equiv \gamma_b \sigma^2[\underline{w} + \chi_b^{-1}(v + \hat{w} + d(i))] =$$

$$(1+\mu) - \chi_b(1+i) \equiv 1 + i = \chi_b^{-1}\Big((1+\mu) - \gamma_b \sigma^2[\underline{w} + \chi_b^{-1}(v + \hat{w} + d(i))]\Big).$$

Hence we may write

$$i(d) = \chi_b^{-1}\Big((1+\mu) - \gamma_b \sigma^2[\underline{w} + \chi_b^{-1}(v + \hat{w} + d)]\Big) - 1.$$

Now we have

$$\bar{i}_b = \lim_{d \to 0^+} i(d) = \chi_b^{-1}\Big((1+\mu) - \gamma_b\sigma^2[\underline{w} + \chi_b^{-1}(v+\hat{w})]\Big) - 1.$$

Formula for $\bar{i}_b$ implies the more debt the borrower has, the more reluctant she is to borrow. Note that the maximal size of a loan that borrower is eager to take at interest rate $\tilde{i}$ is the volume that makes $\tilde{i}$ the largest acceptable interest rate. Given interior solution this volume may be obtained as

$$\tilde{i} = \chi_b^{-1}\Big((1+\mu) - \gamma_b\sigma^2[\underline{w} + \chi_b^{-1}(v+\hat{w}+\tilde{w})]\Big) - 1 = \bar{i}_b - \gamma_b\chi_b^{-2}\sigma^2\tilde{w} \equiv$$

$$\equiv \gamma_b\chi_b^{-2}\sigma^2\tilde{w} = \bar{i}_b - \tilde{i} \equiv \tilde{w} = \gamma_b^{-1}\chi_b^2\sigma^{-2}(\bar{i}_b - \tilde{i}). \tag{D.6}$$

The monetary **supply** of credit is the difference between the optimal portfolio choice and the largest position that l could finance without additional founding. Note that for sufficiently small credit supply $s > 0$ the optimal portfolio choice w_l^* is an internal solution of the lender problem which lies below $e_l = \underline{w} + \chi_l^{-1}(v+\hat{w})$. There are two distinct cases. If $v + \hat{w} \geq 0$ and if in addition credit supply $s < v + \hat{w}$, then $w_l^* \geq \underline{w}$. Bank l finances current loan from surplus cash. As buying large quantities of shares causes price slippage, for an extra unit of shares that l decides not to buy it can grant a loan of c^{-1} units of money. So monetary supply of credit is given by $c^{-1}(e_l - w_l^*)$. By a similar token, if $v + \hat{w} < 0$, then we have $w_l^* < \underline{w}$. Bank l finances the loan by selling speculative asset. As selling large quantities of assets incurs cost, for every unit of shares l sells it obtains c units of money that can be lent on the interbank market. So monetary supply of credit is $c(e_l - w_l^*)$. Using previous notation, monetary credit supply may be in both cases expressed as

$$s(i) = \chi_l(e_l - w_l^*) = \chi_l(\underline{w} - w_l^*) + (v + \hat{w}). \tag{D.7}$$

In both cases we may determine $\underline{i}_l$ as a limit interest rate implied by $s \to 0^+$.

The solution of lender's problem satisfies f.o.c at $w_l^* = \underline{w} + \chi_l^{-1}(v + \hat{w} - s(i))$. To find $\underline{i}_l$ it is sufficient to substitute w_l^* to lender's f.o.c. and solve for i in the limit. First compute

$$C_1 + C_2 w_l^* = -\gamma_l(\theta+i)(v+\hat{w}+\chi_l\underline{w}) - \ln\frac{p}{1-p} + \gamma_l(\theta+i)\chi_l(\underline{w} + \chi_l^{-1}(v+\hat{w}-s(i)) =$$

$$= -\gamma_l(\theta + i)s(i) - \ln \frac{p}{1-p},$$

$$C_3 + C_4 w_l^* = (1+\mu) - (1+i)\chi_l - \gamma_l \sigma^2(\underline{w} + \chi_l^{-1}(v + \hat{w} - s(i)) = (\bar{i}_l - i)\chi_l - \gamma_l \sigma^2 s(i)\chi_l^{-1}.$$

As we already have $C_5 = -(i + \theta)\chi_l$, in the limit it holds that

$$\lim_{s \to 0^+} \frac{C_5}{C_3 + C_4 w_l^*} - 1 = -\frac{(i+\theta)\chi_l}{(\bar{i}_l - i)\chi_l} - 1 = \frac{(i - \bar{i}_l) + (\theta + \bar{i}_l)}{i - \bar{i}_l} - 1 = \frac{\theta + \bar{i}_l}{i - \bar{i}_l}.$$

After inserting these results into f.o.c. we obtain

$$\lim_{s \to 0^+} \left\{ C_1 + C_2 w_l^* - \ln\left(\frac{C_5}{C_3 + C_4 w_l^*} - 1\right)\right\} = 0 \equiv -\ln \frac{p}{1-p} - \ln \frac{\theta + \bar{i}_l}{i - \bar{i}_l} = 0 \equiv$$

$$\equiv \ln \frac{p(\theta + \bar{i}_l)}{(1-p)(i - \bar{i}_l)} = 0 \equiv \frac{p(\theta + \bar{i}_l)}{(1-p)(i - \bar{i}_l)} = 1 \equiv$$

$$\equiv p\theta + \bar{i}_l = (1-p)i \equiv i = (\bar{i}_l + p\theta)/(1-p).$$

Therefore the smallest interest rate that the lender would be able to accept is

$$\underline{i}_l = \bar{i}_l \frac{1}{1-p} + \theta \frac{p}{1-p}. \tag{D.8}$$

Note that the more l lends, the less it wants to lend. The lender never lends at the interest rate it would accept as a borrower. Hence no bank l would ever trade with itself as $\underline{i}_l > \bar{i}_l$.

Now assume that the **lender** decides to lend at the interest rate $\underset{\sim}{i}$. The largest admissible volume of a loan is $\underset{\sim}{w}$ that makes $\underset{\sim}{i}$ the smallest interest rate that the lender would accept, hence

$$\underset{\sim}{i} \approx (\bar{i}_l - \gamma_l \chi_l^{-2} \sigma^2 \underset{\sim}{w}) \frac{1}{1-p} + \theta \frac{p}{1-p} \equiv \gamma_l \chi_l^{-2} \sigma^2 \underset{\sim}{w} \approx \bar{i}_l - \underset{\sim}{i}(1-p) + \theta p \equiv$$

$$\equiv \underset{\sim}{w} \approx \frac{1}{\gamma_l \chi_l^{-2} \sigma^2} [(\bar{i}_l - \underset{\sim}{i})(1-p) + (\bar{i}_l + \theta)p]. \tag{D.9}$$

D.7 Equating supply and demand

Equating supply and demand for multiple lenders and borrowers. Let B and L denote, respectively, the set of borrowers and the set of lenders. For $j \in B \cup L$ define constant $c_j := \gamma_j^{-1} \chi_j^2 \sigma^{-2}$, let f_j denote the size of portfolio of bank j. Given the agreed interest rate i, supply equates demand if

$$\sum_{b \in B} f_b w_b = -\sum_{l \in L} f_l w_l.$$

After substituting left hand side of the formula (D.6) for w_b and the approximation (D.9) for w_l we obtain

$$\sum_{b \in B} f_b c_b (\bar{i}_b - i) = -\sum_{l \in L} f_l c_l [(\bar{i}_l - i)(1-p) + (\bar{i}_l + \theta)p] \equiv$$

$$\equiv \sum_{b \in B} f_b c_b \bar{i}_b - i \sum_{b \in B} f_b c_b = -\sum_{l \in L} f_l c_l (\bar{i}_l + p\theta) + i(1-p) \sum_{l \in L} f_l c_l \equiv$$

$$\equiv \sum_{b \in B} f_b c_b \bar{i}_b + \sum_{l \in L} f_l c_l \bar{i}_l + p\theta \sum_{l \in L} f_l c_l = i \sum_{b \in B} f_b c_b + i(1-p) \sum_{l \in L} f_l c_l \equiv$$

$$\equiv i_{eq} = \frac{\sum_{b \in B} f_b c_b \bar{i}_b + \sum_{l \in L} f_l c_l \bar{i}_l + p\theta \sum_{l \in L} f_l c_l}{\sum_{b \in B} f_b c_b + (1-p) \sum_{l \in L} f_l c_l}. \tag{D.10}$$

The interest rate i_{eq} implies that the desired volumes of trade on both lender and borrower side of the market are equal.

D.8 Censoring supply or demand

We may obtain a formula for reservation interest rate which guarantees that credit demand, generated by a party of borrowers, matches any volume V. By a similar token, it is possible to derive an equation for a reservation interest rate which implies that credit supply of a group of lenders is equal to V. Both formulas are useful if the volume of trade desired by one side of the market has to be capped at a given level.

In order to derive the formula that constraints borrower side of the market observe that

$$i = (\bar{i}_b - c_b^{-1} w_b) \equiv i c_b = c_b \bar{i}_b - w_b \equiv i f_b c_b = f_b c_b \bar{i}_b - f_b w_b \equiv$$

$$\equiv i \sum_{b\in B} f_b c_b = \sum_{b\in B} f_b c_b \bar{i}_b - \sum_{b\in B} f_b w_b \equiv$$

$$\equiv i_c = \frac{\sum_{b\in B} f_b c_b \bar{i}_b - V}{\sum_{b\in B} f_b c_b} \tag{D.11}$$

where the latter equivalence stems from the fact that $\sum_{b\in B} f_b w_b$ is a volume of credit demand which corresponds to reservation interest rate i_c.

The formula that constraints the lender side of the market may be obtained from

$$i = (\bar{i}_l - c_l^{-1} w_l + \theta p)/(1-p) \equiv i(1-p)c_l = c_l \bar{i}_l - w_l + p\theta c_l \equiv$$

$$\equiv i(1-p) f_l c_l = f_l c_l \bar{i}_l - f_l w_l + p\theta f_l c_l \equiv$$

$$\equiv i(1-p) \sum_{l\in L} f_l c_l = \sum_{l\in L} f_l c_l \bar{i}_l - \sum_{l\in L} f_l w_l + p\theta \sum_{l\in L} f_l c_l \equiv$$

$$\equiv i_c = \theta \frac{p}{1-p} + \frac{\sum_{l\in L} f_l c_l \bar{i}_l + V}{\sum_{l\in L} f_l c_l} \frac{1}{1-p}, \tag{D.12}$$

where the latter equivalence stems from the fact that $-\sum_{l\in L} f_l w_l$ is a volume of credit demand which corresponds to reservation interest rate i_c.

D.9 Prices

Assume initial price P_0 is known, set $x_t = \ln P_t$. Maclaurin expansion of natural logarithm is

$$\ln(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 + \dots,$$

hence if x is sufficiently small, $x \approx \ln(1+x) + O(2)$ is its first order approximation. For real interest rate R_{t+1} such that $|R_{t+1}| \ll 1$ it holds that

$$R_{t+1} = \frac{P_{t+1} - P_t}{P_t} \approx \ln\left(1 + \frac{P_{t+1} - P_t}{P_t}\right) = \ln P_{t+1} - \ln P_t = x_{t+1} - x_t.$$

We shall first introduce a discreet price process, next it will be modified to take into account excess demand.

Assume that logarithms of prices are driven by the following **AR(1)** equation

$$x_{t+1} - x_t = \eta[\mu - (x_t - x_{t-1})] + \sigma z_t, \quad z_t \sim N(0,1).$$

Above μ is the mean growth level, parameter σ embodies volatility. Using the approximation above, the conditional distribution of growth rates fulfils

$$R_{t+1}|x_t - x_{t-1} \sim N(\eta\mu - \eta(x_t - x_{t-1}), \sigma^2),$$

while unconditional mean growth rate amounts to $\mu/(1+\eta)$.

References

Acemoglu, D., A. Ozdaglar, and A. Tahbaz-Salehi (2013). Systemic risk and stability in financial networks. NBER Working Paper No. 18727. 71

Adrian, T. and H. S. Shin (2010). Liquidity and leverage. *Journal of Financial Intermediation 19*(3), 418–437. 73

Akgiray, V. and G. G. Booth (1988). The stable–law model of stock returns. *Journal of Business & Economic Statistics 6*(1), 51–57. 14, 15

Allen, F., A. Babus, and E. Carletti (2012). Asset commonality, debt maturity and systemic risk. *Journal of Financial Economics 104*(3), 519–534. 71, 73

Allen, F. and D. Gale (2001). Financial contagion. *Journal of Political Economy 108*(1), 1–33. 70, 71, 73

Arinaminpathy, N., S. Kapadia, and R. May (2012). Size and complexity in model financial systems. Bank of England Working Paper No. 465. 72, 73

Asmussen, S. (2003). *Applied Probability and Queues*. Berlin: Springer. 12

Asmussen, S., K. Binswanger, and B. Højgaard (2000). Rare events simulation for heavy tailed distributions. *Bernoulli 6*(2), 303–322. 10, 11

Aston, J. A. D. (2006). Modeling macroeconomic time series via heavy tailed distributions. In H.-C. Ho, C.-K. Ing, and T. L. Lai (Eds.), *IMS Lecture Notes Monograph Series, Vol. 52. Time Series and Related Topics. In Memory of Ching–Zong Wei*, pp. 138–148. IMS–Cambridge University Press. 9, 16

Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics 12*, 171–178. 16

Azzalini, A. and A. Capitanio (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t–distribution. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 65*(2), 367–389. 16

Babus, A. (2009). The formation of financial networks. Tinbergen Institute Discussion Paper No. 06–093. 71, 73

Baeumer, B. and M. M. Meerschaert (2010). Tempered stable Lévy motion and transient super–diffusion. *Journal of Computational and Applied Mathematics 223*(10), 2438–2448. 50, 51, 55, 67

Balke, N. S. and T. B. Fomby (1994). Large shocks, small shocks, and economic fluctuations: Outliers in macroeconomic time series. *Journal of Applied Econometrics 9*(2), 181–200. 9, 10

Barndorff-Nielsen, O. and N. Shephard (2012). Basics of Lévy processes. Department of Economics Discussion Paper No. 610, University of Oxford. 39

Barndorff-Nielsen, O. E. (1997). Normal Inverse Gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of Statistics 24*(1), 1–13. 62

Basu, A., H. Shioya, and C. Park (2011). *Statistical Inference: The Minimum Distance Approach*. London: Chapman & Hall/CRC Press. 54

Battiston, S., D. D. Gatti, M. Gallegati, and J. E. Stiglitz (2012). Liaisons dangereuses: Increasing connectivity, risk sharing, and systemic risk. *Journal of Economic Dynamics and Control 36*(8), 1121–1141. 73

Bauwens, L. and S. Laurent (2005). A new class of multivariate skew densities, with application to generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics 23*(3), 346–354. 6, 17

Becher, C., S. Millard, and K. Soramäki (2008). The network topology of CHAPS Sterling. Bank of England Working Paper No. 355. 74, 98

Bianchi, M. L., S. T. Rachev, and F. J. Fabozzi (2013). Tempered stable Ornstein–Uhlenbeck process: A practical view. Bank of Italy Working Paper No. 912. 41

Blattberg, R. and N. Gonedes (1974). A comparison of the stable and student distributions as statistical models for stock prices. *Journal of Business 47*(2), 244–280. 6, 11, 14, 15

Blåvarg, M. and P. Nimander (2002). Interbank exposures and systemic risk. In: Risk Management and Systemic Risk. Bank for International Settlements. 89, 99

Bluhm, M., E. Faia, and J. P. Krahnen (2013). Endogenous banks' networks, cascades and systemic risk. SAFE Working Paper No. 355. 72, 73

Boyarchenko, S. I. and S. Z. Levendorskii (2000). Option pricing for truncated Lévy processes. *International Journal of Theoretical and Applied Finance 3*(3), 549–552. 10, 27, 30, 31, 118

Brix, A. (1999). Generalized gamma measure and shot–noise Cox process. *Advanced Applied Probability 31*, 929–953. 50, 51

Brooks, C. and H. M. Kat (2002). The statistical properties of hedge fund index returns and their implications for investors. *Journal of Alternative Investments 5*(2), 26–44. 6

Brunnermeier, M. K. (2009). Deciphering the liquidity and credit crunch 2007–2008. *Journal of Economic Perspectives 23*(1), 77–100. 73

Caballero, R. and A. Krishnamurthy (2008). Collective risk management in a flight to quality episode. *The Journal of Finance 63*(5), 2008. 73, 107

Cambanis, S., S. Huang, and G. Simmons (1981). On the theory of elliptically contoured distributions. *Journal of Multivariate Analysis 11*(1), 368–385. 60

Carr, P., H. Geman, D. Madan, and M. Yor (2002). The fine structure of asset returns: An empirical investigation. *Journal of Business 75*(2), 305–332. 30

Chakrabarty, A. and M. M. Meerschaert (2011). Tempered stable laws as random walk limits. *Statistics and Probability Letters 81*, 989–997. 19, 28, 30

Chamberlain, G. (1983). A characterization of the distributions that imply mean–variance utility functions. *Journal of Economic Theory 29*(1), 185–201. 61

Chambers, J. M., C. L. Mallows, and B. W. Stuck (1976). A method of simulating stable random variables. *Journal of American Statistical Association 71*, 340–344. 51, 115

Cifuentes, R., G. Ferrucci, and H. S. Shin (2005). Liquidity risk and contagion. *Journal of the European Economic Association 3*(2), 556–566. 73

Cocco, J. a. F., F. J. Gomez, and N. C. Martins (2009). Lending relationships in the interbank market. *Journal of Financial Intermediation 18*(1), 24–48. 74, 102

Cohen, S. and J. Rosiński (2007). Gaussian approximation of multivariate Lévy processes with applications to simulation of tempered stable processes. *Bernoulli 13*(1), 195–210. 50

Cohen-Cole, E., E. Patacchini, and Y. Zenou (2013). Systemic risk and network formation in the interbank market. Unpublished. 72, 74, 99

Coval, J. D. and E. Strafford (2007). Asset fire sales (and purchases) in equity markets. *Journal of Financial Economics 86*(2), 479–512. 73, 77

Crovella, M. E., M. S. Taqqu, and A. Bestavros (1998). Heavy–tailed probability distributions in the World Wide Web. In R. J. Adler, R. E. Feldman, and M. S. Taqqu (Eds.), *A Practical Guide to Heavy Tails*. Boston: Birkhäuser. 12

de Nooy, W., A. Mrvar, and V. Batagelj (2011). *Exploratory social network analysis with Pajek*. UK: Cambridge University Press. 93

Degryse, H. and G. Nguyen (2007). Interbank exposures: an empirical examination of systemic risk in the Belgian banking system. *International Journal of Central Banking 3*(2), 123–171. 100

Desai, P. (2000). Why did the ruble collapse in August 1998. *American Economic Review 90*(2), 48–52. 56

Devroye, L. (1981). On the computer generation of random variables with a given characteristic function. *International Journal of Computer Mathematics 7*(6), 547–552. 50, 52, 53, 67

Devroye, L. (2009). Random variate generation for exponentially and polynominally tilted stable distributions. *ACM Transactions on Modelling and Computer Simulation 19*. Article 18. 50

Eaton, M. L. (1986). A characterization of spherical distributions. *Journal of Multivariate Analysis 20*(1), 272–276. 60

Eberlein, E. and U. Keller (1995). Hyperbolic distributions in finance. *Bernoulli 1*(3), 281–299. 17

Eisenberg, L. and T. H. Noe (2001). Systemic risk in financial systems. *Management Science 47*(2), 236–249. 71

Elliott, M., B. Golub, and M. O. Jackson (2013). Financial networks and contagion. Unpublished. 71

Elsinger, H., A. Lehar, and M. Summer (2006a). Risk assessment for banking systems. *Management Science 52*(9), 1301–1314. 71, 73

Elsinger, H., A. Lehar, and M. Summer (2006b). Using market information for banking system risk assessment. *International Journal of Central Banking 2*(1), 137–165. 71

Erdös, P. and A. Rény (1959). On random graphs I. *Publicationes Mathematicae 6*(1), 290–297. 94, 99

Fagiolo, G., M. Napoletano, and A. Roventini (2008). Are output growth–rate distributions fat–tailed? Some evidence from OECD countries. *Journal of Applied Econometrics 23*(5), 2008. 9, 10, 11

Fama, E. F. (1963). Mandelbrot and the stable Paretian hypothesis. *Journal of Business 35*(4), 420–429. 13

Fama, E. F. (1965a). The behavior of stock–market prices. *Journal of Business 38*(1), 34–105. 6, 11, 14

Fama, E. F. (1965b). Portfolio analysis in a stable Paretian market. *Management Science 11*(3), 404–419. 14

Fama, E. F. and R. F. Roll (1965). Parameter estimates for symmetric stable distributions. *Journal of American Statistical Association 66*(334), 331–338. 11

Feller, W. (1971). *An Introduction to Probability Theory and Its Applications, Vol. II.* New York: John Wiley & Sons Inc. 13, 20, 21

Fernández, C. and M. F. J. Steel (1998). On bayesian modeling of fat–tails and skewness. *Journal of the American Statistical Association 93*(441), 359–371. 16

Fielitz, B. D. and J. P. Rozelle (1983). Stable distributions and the mixtures of distributions hypotheses for common stock returns. *Journal of American Statistical Association 78*(381), 28–36. 15

Freixas, X., B. Parigi, and J.-C. Rochet (2000). Systemic risk, interbank relations and liquidity provision by the central bank. *Journal of Money, Credit and Banking 32*(3), 611–638. 70

Friedman, B. M. and D. I. Laibson (1989). Economic implications of extraordinary movements in stock prices. *Brooking Papers on Economic Activity 2*, 137–172. 17

Gai, P., A. Haldane, and S. Kapadia (2011). Complexity, concentration and contagion. *Journal of Monetary Economics 58*(5), 453–470. 72, 73

Gai, P. and S. Kapadia (2010). Contagion in financial networks. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 466*(2120), 2401–2423. 71

Georg, C.-P. (2013). The effect of the interbank network structure on contagion and common shocks. *Journal of Banking & Finance 37*(7), 2216–2228. 71, 73, 89

Glasserman, P., P. Heidelberger, and P. Shahabuddin (2002). Portfolio value–at–risk with heavy tails risk factors. *Mathematical Finance 12*(3), 239–269. 16

Grabchak, M. (2012). On a new class of tempered stable distributions: Moments and regular variation. *Journal of Applied Probability 49*(4), 1015–1035. 28

Haas, M. (2007). Do investors dislike kurtosis? *Economics Bulettin 7*(2), 1–9. 6

Hagerman, R. (1978). More evidence on the distribution of security returns. *Journal of Finance 33*(4), 1213–1221. 6, 14, 15

Haldane, A. G. and R. M. May (2011). Systemic risk in banking ecosystems. *Nature 469*, 351–355. 71, 73

Hall, P. (1981). A comedy of errors: The canonical form for a stable characteristic function. *Bulletin of London Mathematical Society 13*(1), 23–27. 20

Hansen, B. E. (1994). Autoregressive conditional density estimation. *International Economic Review 35*(3), 705–730. 16, 17

Hartmann, P., S. Straetmans, and C. G. de Vries (2010). Heavy tails and currency crisis. *Journal of Empirical Finance 17*, 241–254. 10, 13, 16, 56, 57

Harvey, C. R. and A. Siddique (1999). Autoregressive conditional skewness. *Journal of Financial and Quantitative Analysis 34*, 465–487. 16

Harvey, C. R. and A. Siddique (2000). Conditional skewness in asset pricing tests. *Journal of Finance 55*(3), 1263–1295. 6

Hsu, D.-A., R. Miller, and D. Wichtern (1974). On the stable Paretian behavior of stock–market prices. *Journal of American Statistical Association 69*(345), 108–113. 15

Ibragimov, M., R. Ibragimov, and P. Kattuman (2013). Emerging markets and heavy tails. *Journal of Banking & Finance 37*(7), 2546–2559. 57

Ibragimov, R. (2005). On the robustness of economic models to heavy–tailedness assumptions. Deutsche Bundesbank conference on: Heavy tails and stable Paretian distributions in finance and macroeconomics. 11

Iori, G., G. De Masi, O. V. Precup, G. Gabbi, and G. Caldarelli (2008). A network analysis of the Italian overnight money market. *Journal of Economic Dynamics and Control 32*(1), 259–278. 99

Iori, G., S. Jafarey, and F. G. Padila (2006). Systemic risk on the interbank market. *Journal of Economic Behaviour & Organization 61*(4), 525–542. 71, 72, 73

Jacquier, E., N. G. Polson, and P. E. Rossi (2004). Bayesian analysis of stochastic volatility models with fat–tails and correlated errors. *Journal of Econometrics 122*(1), 185–212. 7

Jelonek, P. (2012). Generating tempered stable random variates from mixture representation. Department of Economics Working Paper No. 12/14, University of Leicester. 49

Jones, M. C. and M. J. Faddy (2003). A skew extension of the t–distribution, with applications. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 65*(1), 159–174. 16

Jorion, P. (1988). On jump processes in the foreign exchange and stock market. *The Review of Financial Studies 1*(4), 427–445. 7

Kaufman, G. G. (1994). Bank contagion: A review of the theory and evidence. *Journal of Financial Services Research 8*(2), 123–150. 89

Kawai, R. and H. Masuda (2011). On simulation of tempered stable random variates. *Journal of Computational and Applied Mathematics 235*(8), 2873–2887. 28, 32, 46, 50, 51, 53

Kim, Y. S., S. T. Rachev, M. L. Bianchi, and F. J. Fabozzi (2008). Financial market models with Lévy processes and time–varying volatility. *Journal of Banking & Finance 32*(7), 1363–1378. 7, 15

Kim, Y. S., S. T. Rachev, M. L. Bianchi, and F. J. Fabozzi (2010a). Computing VaR and AVaR in infinitely divisible distributions. *Probability and Mathematical Statistics 30*(2), 223–245. 31, 119, 120

Kim, Y. S., S. T. Rachev, M. L. Bianchi, and F. J. Fabozzi (2010b). Tempered stable and tempered infinitely divisible GARCH models. *Journal of Banking and Finance 30*(2), 223–245. 27, 30

Kolossiatis, M., J. E. Griffin, and M. F. J. Steel (2011). Modelling overdispersion with the normalized tempered stable distribution. *Computational Statistics & Data Analysis 55*(7), 2288–2301. 28

Kon, S. J. (1984). Models of stock return – A comparison. *Journal of Finance 39*(1), 147–165. 6, 15, 16, 17

Koponen, I. (1995). Analytic approach to the problem of convergence of truncated Lévy flights towards the Gaussian stochastic process. *Physical Review E: Statistical physics, Plasmas, Fluids, and Related Interdisciplinary Topics 52*, 1197–1199. 30

Krause, A. and S. Giansante (2012). Interbank lending and the spread of bank failures: A network model of systemic risk. *Journal of Economic Behaviour & Organization 83*(3), 583–608. 72, 73

Küchler, U. and S. Tappe (2011). Tempered stable distributions and applications to financial mathematics. Unpublished. 28, 31, 41

Ladley, D. (2013). Contagion and risk–sharing on the inter–bank market. *Journal of Economic Dynamics and Control 37*(7), 1384–1400. 71, 72, 73

Lévy, P. (1924). Théorie des erreurs. La loi de Gauss et les lois exceptionelles. *Bulletin de la Société Mathématique de France 52*(1), 49–85. 20

Li, J. (2007). Characteristic function based estimation of Lévy tempered stable stock price models and jump market price calibration. Unpublished. 27

Li, J., C. Favero, and F. Ortu (2012). A spectral estimation of tempered stable stochastic volatility models and option pricing. *Computational Statistics and Data Analysis 56*(11), 3645–3658. 41

Loretan, M. and P. C. B. Phillips (1994). Testing the covariance stationarity of heavy–tailed time series: An overview of the theory with applications to several financial datasets. *Journal of Empirical Finance 1*(2), 211–248. 7

Mandelbrot, B. (1963a). New methods in statistical economics. *Journal of Political Economy 71*(5), 421–440. 11, 13, 14, 59

Mandelbrot, B. (1963b). The variation of certain speculative prices. *Journal of Business 36*(4), 394–419. 13, 15, 17

Mandelbrot, B. and R. L. Hudson (2004). *The (mis)behavior of markets: A fractal view of risk, ruin, and reward*. New York: Profile Books. 12

Marinelli, C., S. T. Rachev, and R. Roll (2001). Subordinated exchange rate models: Evidence for heavy tailed distributions and long–range dependence. *Mathematical and Computer Modelling 34*(9–11), 955–1001. 7, 14

Markose, S., S. Giansante, and A. R. Shaghaghi (2012). 'Too interconnected to fail' financial networks of US CDS market: Topological fragility and systemic risk. *Journal of Economic Behaviour & Organization 83*(3), 627–546. 72

Martínez-Jaramillo, S., O. Pérez Pérez, F. Avila Embriz, and F. López Gallo Dey (2010). Systemic risk, financial contagion and financial fragility. *Journal of Economic Dynamics and Control 34*(11), 2358–2374. 72

McCulloch, H. (1986). Simple consistent estimators of stable distribution parameters. *Communications in Statistics – Simulation and Computation 15*(4), 1109–1136. 22

Mittnik, S., T. Doganoglu, and D. Chenyao (1999). Computing the probability density function of the stable Paretian distribution. *Mathematical and Computer Modelling 29*(10–12), 235–240. 29, 46, 52, 67

Müller, J. (2006). Interbank credit lines as a channel of contagion. *Journal of Financial Services Research 29*(1), 37–60. 74, 98, 100

Nakao, H. (2000). Multi–scaling properties of truncated Lévy flights. *Physics Letters A 266*(4–5), 282–289. 31

Nier, E., J. Yang, T. Yorulmazer, and A. Alertorn (2007). Network models and financial stability. *Journal of Economic Dynamics and Control 31*(6), 2033–2060. 71, 73

Nolan, J. P. (2013). *Stable Distributions – Models for Heavy Tailed Data*. Boston: Birkhäuser. In progress, Chapter 1 available online at: academic2.american.edu/∼jpnolan. 20

Officer, R. R. (1972). The distribution of stock returns. *Journal of American Statistical Association 67*(340), 807–812. 7

Øigård, T. A., A. Hanssen, R. E. Hansen, and F. Godtliebsen (2005). EM–estimation and modeling of heavy–tailed processes with the multivariate Normal Inverse Gaussian distribution. *Signal Processing 85*(8), 1655–1673. 62

Owen, J. and R. Rabinovitch (1983). On the class of elliptical distributions and their applications to the theory of portfolio choice. *The Journal of Finance 38*(3), 745–752. 61

Palmer, K. J., M. S. Ridout, and B. J. T. Morgan (2008). Modelling cell generation times by using the tempered stable distribution. *Journal of the Royal Statistical Society: Series C (Applied Statistics) 57*(4), 379–397. 20, 28, 46, 50

Pindyck, R. S. (2011). Fat tails, thin tails and climate change policy. *Review of Environmental Economics and Policy 5*(2), 258–274. 10

Praetz, P. D. (1972). The distributions of share price changes. *Journal of Business 45*(1), 49–55. 15

Press, S. J. (1967). Events model for security prices. *Journal of Business 40*(3), 317–335. 42

Rachev, S. T. (2003). Preface. In S. T. Rachev (Ed.), *Handbook of Heavy Tailed Distributions in Finance: Handbooks in Finance, Book 1*, pp. ix–xi. Elsevier/North–Holland. 49

Rachev, S. T. and S. Mittnik (2000). *Stable Paretian Models in Finance*. New York: John Wiley & Sons. 14, 15

Ridout, M. S. (2009). Generating random numbers from a distribution specified by its Laplace transform. *Statistics and Computing 19*, 439–450. 50

Rosiński, J. (2007). Tempering stable processes. *Stochastic Processes and their Applications 117*, 677–707. 18, 29, 30, 31, 33, 38, 118, 122, 123

Rozelle, J. P. and B. D. Fielitz (1980). Skewness in common stock returns. *Financial Review 15*(3), 1–19. 6

Samorodnitsky, G. and M. S. Taqqu (2000). *Stable Non–Gaussian Random Processes: Stochastic Models with Infinite Variance*. Taylor & Francis. 20, 23, 24, 32, 51, 122

Sato, K.-i. (2005). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge: Cambridge University Press. 24

Sato, K.-i. and M. Yamazoto (1978). On distribution functions of class L. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 43*(4), 273–308. 120

Scherer, M., S. T. Rachev, Y. S. Kim, and F. J. Fabozzi (2010). A FFT–based approximation of tempered stable and tempered infinitely divisible distributions. Unpublished. 28, 41, 47

Shleifer, A. and R. Vishny (2011). Fire sales in finance and macroeconomics. *Journal of Economic Perspectives 25*(1), 29–48. 73

Simkowitz, M. A. and W. L. Beedles (1980). Asymmetric stable distributed security returns. *Journal of American Statistical Association 75*(370), 306–312. 6

Soramäki, K., M. L. Bech, J. Arnold, R. J. Glass, and W. E. Beyeler (2007). The topology of interbank payment flows. *Physica A: Statistical Mechanics and its Applications 379*(1), 317–333. 74, 99

Stuart, A. and J. K. Ord (1994). *Kendall's Advanced Theory of Statistics, Vol. 1*. Halsted Press. 39, 40

Sun, W., S. T. Rachev, and F. J. Fabozzi (2007). Fractals or i.i.d.: Evidence of long–range dependence and heavy tailedness from modeling German equity market returns. *Journal of Economics and Business 59*(6), 575–595. 7

Sztonyk, P. (2010). Estimates of tempered stable densities. *Journal of Theoretical Probability 23*(1), 127–147. 28

Terdik, G. and T. Gyires (2009). Lévy flights and fractal modeling of internet traffic. *IEEE/ACM Transactions on Networking 17*(1), 120–129. 31, 41, 45

Terdik, G. and W. A. Woyczyński (2006). Rosiński measures for tempered stable and related Ohrstein–Uhlenbeck processes. *Probability and Mathematical Statistics 26*(2), 213–243. 31, 39, 123

Theodossiou, P. (1998). Financial data and the skewed generalized t distribution. *Management Science 44*(12), 1650–1661. 16

Trefethen, L. N. and D. Bau (1997). *Numerical Linear Algebra*. Philadelphia: Society for Industrial and Applied Mathematics. 61

Upper, C. (2011). Simulation methods to assess the danger of contagion in interbank markets. *Journal of Financial Stability 7*(3), 111–125. 73

Vallascas, F. and K. Keasey (2012). Bank resilience to systemic shocks and the stability of banking systems: Small is beautiful. *Journal of International Money and Finance 31*(3), 1745–1776. 72

Vlaar, P. J. C. and F. C. Palm (1993). The message in weekly exchange rates in European Monetary System: Mean reversion, conditional heteroskedasticity and jumps. *Journal of Business & Economic Statistics 11*(3), 351–360. 7, 10

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica 57*(2), 307–333. 64

Wang, D., S. T. Rachev, and F. J. Fabozzi (2009). Pricing of credit default index swap tranches with one–factor heavy–tailed copula models. *Journal of Empirical Finance 16*(2), 201–215. 15

Wolfe, S. J. (1971). On the continuity properties of L functions. *The Annals of Mathematical Statistics 42*(6), 2064–2073. 120

Yamazato, M. (1978). Unimodality of infinitely divisible distribution functions of class L. *The Annals of Probability 6*(4), 523–531. 22

Yu, J. (2004). Empirical characteristic function estimation and its applications. *Econometric Reviews 23*(2), 93–123. 46

Zawadowski, A. (2013). Entangled financial systems. *The Review of Finanacial Studies 26*(5), 1291–1323. 71, 72, 73