

Université catholique de Louvain



Louvain School of Management

Université Paris 1 Panthéon Sorbonne



Ecole de Management de la Sorbonne

PhD dissertation

Essays on Risk Management and Financial Stability

Saifeddine Ben hadj

President: **Pr. Jean-Paul Laurent** Université Paris 1 Panthéon Sorbonne

Supervisors: **Pr. Isabelle Platten** Université catholique de Louvain

Pr. Roland Gillet Université Paris 1 Panthéon Sorbonne

Jury: **Pr. Hans Dewachter** National Bank of Belgium & KULeuven

Pr. Frédéric Lobe Université de Lille 2

Pr. Mikael Petitjean Université catholique de Louvain et IESEG

July 4, 2017

Saifeddine Ben hadj

Essays on Risk Management and Financial Stability

PhD dissertation, July 4, 2017

Supervisors: Isabelle Platten and Roland Gillet

Université catholique de Louvain

Louvain School of Management

Place de l'université

1348 and Louvain-la-Neuve

Université Paris 1 Panthéon Sorbonne

Ecole de Management de la Sorbonne

Place du Panthéon

75231 and Paris

Contents

1	Introduction	3
2	A non-uniform nested simulation algorithms in portfolio risk measurement	13
2.1	Literature Review	18
2.2	General Framework	20
2.2.1	Model Framework for VaR	21
2.2.2	Simulation Framework for VaR	22
2.2.3	Estimating Value-at-Risk	24
2.2.4	Estimating Solvency II capital requirements:	25
2.3	Sampling Algorithms	28
2.3.1	Uniform Sampling	28
2.3.2	Optimal Uniform Sampling	30
2.3.3	Sequential Simulation	31
2.3.4	Stratified Sequential Simulation	41
2.4	Numerical Results	45
2.4.1	Experimental Settings	45
2.4.2	Simulation Results	52
2.4.3	Algorithm Comparison	57
A	Proof of Equations 2.18 and 2.19	59
B	Proof of equation (2.26)	61
C	Sensitivity based optimization algorithm	61
A	Proof of Equations 2.24	62
3	Financial Institutions Externalities and Systemic Risk: Tales of Tails	
	Symmetry	65
3.1	State of the Art	70
3.2	Relation between Tails' symmetry and extreme losses	73

3.2.1	A model of the banking system	73
3.2.2	Definition of tails symmetry	75
3.2.3	Implication of tail symmetry in distress probability :	77
3.2.4	System stability illustrated by simulations	81
3.3	Measure of tail asymmetry and capital provisions	89
3.3.1	Tail index	89
3.3.2	Tail imbalance factor	91
3.4	Measuring the bank externalities based on asymmetry	93
3.4.1	Data description	96
3.4.2	Results	97
3.5	The post-financial crisis fines	102
A	Proof of theorem 2	106
A	Illustration of λ as a measure of externalities	108
4	A game theory approach for systemic risk and international regulatory coordination	111
4.1	State of the Art	117
4.2	History and channels of contagion	120
4.3	Description of the model	123
4.3.1	Elementary asset value units	124
4.3.2	Description of the risk profile	126
4.3.3	Hedging and recovery mechanism	127
4.3.4	State transition dynamic	129
4.3.5	Crisis and government interventions	133
4.4	General transmission dynamics	135
4.5	Strategic decision of individual countries	145
4.6	Policy implication and discussion of the model	156
	List of Symbols	161
A	Proof of Equations (4.10)	162
4.2	Parameter used for figures (4.5), (4.6), (4.7) and (4.8)	164
5	Conclusions	165

Acknowledgement

Undertaking this PhD has been a truly life-changing experience for me and it would not have been possible to do without the guidance and support that I received from many people.

I would like to first say a very big thank you to my supervisors Prof Isabelle Platten and Prof Roland Gillet for all the support and guidance they provided during the elaboration of this PhD without their constant feedback this PhD would not have been achievable.

Besides my advisor, I would like to thank the rest of my thesis committee: Prof Mikael Petitjean, Prof Jean-Paul Laurent for their insightful comments and encouragement, but also for the hard questions which incited me to widen my research from various perspectives.

I also thank Prof Hans Dewachte and Prof Frédéric Lobez for carefully reading this thesis and providing valuable feedbacks and comments.

I would like to thank my colleagues in UCL especially in Mons campus for the stimulating discussion and all the fun we had. I also address my thanks to my colleagues in the National Bank of Belgium who encouraged me during the last miles of this journey.

Last but not least, I would like to thank my family: my parents and my sisters for supporting me and their prayers throughout writing this thesis and my life in general. My gratitude goes to my wife Mariem the woman who made all the difference in my life who supported me, encouraged me and gave me the strength that I needed when I despaired. My thanks go also to little Anass who supported me

without knowing with his joyful smile and continuously radiating happiness in our home.

Introduction

In the late 19th century, it was the failure of railroad companies that triggered major economic crises. Their role as intermediate between miners, producers, and consumers was the cornerstone and booster of economic activity. Nowadays, it is banks and more generally financial institutions that play the leading role in the economy. Some claim that the ability of banks to intermediate between those who are willing to lend and others in need of borrowing is the key determinant of growth and economic welfare. In the absence of this modern system of intermediation, it would be difficult for companies to fulfill their investment needs and for individuals to invest in durable goods and consume non-durable ones. Driven by regulation and the natural survival instinct living in each economic agent seeking profit, banks are investing billions of dollars yearly in their risk management department.

Economic historian reached no consensus about the origin of the concept of risk. However, one story seems to have more proponents than others. It traces back risk to trades between Christian and Arab merchants in the Middle Age and attributes its origins to the middle-eastern language. Italian traders qualify traveling merchandise in the middle of the sea as risk having in mind the negative outcome of its loss, while Arabs consider risk in a positive way. In fact, a risk to Arab merchant is the gain that God attributed to them however hard work is still needed to cash it. The missing link in the history of risk is how the positive connotation has transformed into a negative one after crossing the Mediterranean. Despite this little agreement about the origin of the word, it seems that the modern concept of risk in the western civilization emerged with economic development and replaced the notion of dangers and hazards in the circles of businessmen. Later, entrepreneurs stopped considering risk as a fatality and started to develop strategies to mitigate the effect of risk or what we call now risk management.

Modern risk management in the financial industry started to grow in the second half of the 20th century as investment banks ventured into derivatives. Companies saw there a real opportunity to ship uncertainty out of their balance sheet and focus on their core business activities. It is safe to consider investment bankers as professional risk managers. The task turned to be cumbersome for the industry and their supervisors. The latest financial crisis violently demonstrated the impact of incomplete risk assessment on the viability of the financial system and the continuity of financial intermediation. It also showed how governments efforts to even keel after the storm without proper crisis management plan could lead to a Pyrrhic victory.

Series of financial crises which began from the great depression of the 1930's to the great recession that peaked in 2008 has given birth to regulation that shaped the modern financial system. The reforms are generally the result of lessons learned by regulators after the storm. From that angle, the latest financial crisis was a great learning experience for economic agents. The first lesson is that the next crisis is unlikely to be the result of over-investment in A-rated Mortgage Bases Securities (MBS). The second and more important lesson is that regulation needs to be re-engineered having in mind the evolutionary characteristic of the financial actors. It is important to marry micro-prudential measures with macro oriented regulations while keeping an eye on international coordination. In fact, the responsibility of financial system stability is a burden that must be shared between individual banks, national governments, and international regulatory organizations.

No later than 1933, the US government established explicit deposit insurance to protect customers' deposits against future bank runs via the creation of the FDIC. To avoid moral hazard problems, banks agreed to allow regulators monitor their risk-taking behavior. Besides, they were also concerned about the viability of the financial system and the continuity of its services. Despite all the regulatory efforts (or because too much was done according to some) and the evolution of risk management techniques by banks, the 2008 crisis clearly demonstrated that economic agents were unprepared to cushion the negative effects of a full-scale financial crisis. More specifically, they acknowledged the existence of imperfections in the risk management techniques and banking regulations.

At the level of banks, risk managers were unable to estimate correctly the risk that was undertaken by their institutions. Consequently, most of them failed to anticipate unprecedented market downturns that endangered the viability of their banks. Even the very few which were able to foresee the wave before hitting the shores failed to reckon the scale of the future crisis. This failure is partially due to the oversimplification of risk assessment models that dealt with sizable portfolios of exotic derivatives combined with structured products. This complexity led to a lack of understanding of the real level of risk created by those products. In short, financial engineers were highly overrunning risk managers who had trouble keeping up with the increasing complexity of financial instruments.

In the meanwhile, regulators were lacking the tools and technologies to identify the financial institutions which could threaten the stability of the system. In fact, the financial system could be compared to a soccer game played by banks and refereed by regulators. The common features between both games are that players are highly skilled compared to the arbitrator. Nevertheless, a good pair of eyes is enough to detect unfair actions even by the most resourceful players in soccer games. Unfortunately, more complex observation tools are required to identify players who are breaking the rules in the financial system. In fact, the main indicator of the importance of a bank to the economy was the size of the institution. National champions were qualified as being too big to fail and benefited from the implicit guarantee that governments will step-in to bail them out in the case of financial distress.

Of course, such superficial analysis of the contribution of banks to a systemic crisis will fail to anticipate critical downturns and could only result in a massive bailout as it was the case in the US and Europe after 2008. In addition, it was also clear to governments that the lack of international coordination of financial regulation was a source of regulatory arbitrage.

In fact, the globalization of the financial markets resulted in the global spread of the financial crisis and when burst it required remedies at the international level. This dissertation is structured into three proposed essays. In each chapter, we choose to tackle the issues related to risk management from a different angle: the point of

view of individual banks, national regulators and international regulatory institutions.

While subjects may seem different, they all try to give answers to the same question: How should we improve the process of risk management to enhance financial stability at the banks level and more importantly at the national and international level? For instance, at the level of banks, we propose technical solutions to reconcile the flexibility of risk management assessment models with the feasibility and real-time implementation of those models. In the upper layer of risk management i.e. regulators, we propose to provide them with tools that can detect hazardous innovation in the financial system without the need of the costly thorough analysis of individual bank positions. Finally, given the fragmentation of regulatory bodies at international or even national levels, we propose a model of strategic interaction between different regulators. The chapter aims at studying to what extent collaboration between regulators is beneficial. The titles of each chapter are as following:

- Chapter 1: Non-uniform nested simulation algorithms in portfolio risk measurement
- Chapter 2: Financial Institutions Externalities and Systemic Risk: a Tale of Tails Symmetry
- Chapter 3: A game theory approach for systemic risk and international regulatory coordination

In the first chapter of this dissertation, we focus on the question of improving risk assessment within individual banks and more precisely for those holding portfolios of complex derivatives. In fact, the failure of correct risk assessment during the crisis showed how realistic were the standard assumption. The sacrifices that were made to accuracy in order to obtain a solution in the time limits are no longer acceptable today. To add flexibility to those models, financial engineers are bounded to use simulations that call for the use of computationally greedy algorithms. In this chapter, we question whether the pricing complexity leads inevitably to massive computational spendings in risk management applications. The focus of this chapter

is the widely used risk Value-at-Risk that requires the use of nested simulations or a two-stage simulations: the outer simulation and the inner simulation. The outer simulation is used to sample risk factors over a given time horizon. The inner simulation reprices the portfolio instruments conditional on the drawn risk factors. We focus on Value-at-Risk because it goes beyond risk management. It is also applied as a limit for managing trading desk in big investment banks and can also be used as an asset allocation criterion.

The core of the first chapter is to improve the direct nested simulation technique to allow for the use of more realistic pricing models while remaining within reasonable computational efforts. These ideas are based on the work of Gordy and Juneja (2010) and inspired from the work of Broadie et al. (2011) on evaluating probabilities of large losses. The biggest contribution of this chapter to the literature of computational finance is in the applied methodology. In fact, it is the first technique applied to quantile-based risk measures and it has no prerequisite that could prevent its implementation in practice. Later, we provide theoretical justification and numerical implementations of the proposed algorithms to show its efficiency.

The second chapter focuses on a broader concept related to the stability of the financial system. Regulators from the early 2000s knew the importance of marrying macro and micro-prudential regulations. In fact, what can be considered as a rational and desirable behavior at the individual level can have serious negative effects on the system. The objective of this chapter is to propose a theoretical and practical framework for identifying and measuring the negative externalities, or the social costs generated by the banking activities. The core of this chapter is to suggest a new framework for assessing the negative impact of the internal decision of banks on the financial system. We develop both a theoretical and an empirical framework to measure those externalities. Our biggest contribution with regard to the literature is in the concept of externalities and the way to measure it. In fact, we argue that any financial activity that creates no negative externalities should not affect the symmetry in the tails of the profit and loss distribution of the banks that take part of this transaction. The perfect example of externalities are the too-big-to-fail implicit guarantee that can distort the $P\&L$ distribution. Banks will be keeping the gains resulting from excessively risky positions while they expect governments to inter-

vene when big losses materialize. Several arguments support the concepts that we propose. First, derivatives contracts are zero-sum games which suggest that some risks are visible on the gains part of their counterparts. The second argument is a historical argument. In fact, most of the hazardous financial innovations resulted in important gains in the start and still lead to heavy losses only visible later in the future. Such outcome suggests that some inter-temporal transfer of risk and present gains are the symptoms of risk taken in a future time. To better understand crises, we argue that it is also important to have a closer look at the build-up phase and we that gains are informative as losses in this case. The last argument is related to risk management within the bank. As gains are the results of a favorable exposure to a set of risk factors, a negative outcome of the same factors will result in losses. The concept of tail symmetry that we introduce is only relevant to the tails as we tolerate skewness that results on different market anticipations.

The last chapter takes a general view of financial stability and questions the importance of international collaboration in the presence of coordination costs. We argue that international coordination is not an obvious decision and regulators should balance costs and benefits to engaging in collaborative efforts. The main incentive for international coordination is the contagion of crises that become more important due to the increasing integration of financial systems. Moreover, history showed that negotiation to share the burden of a crisis *ex post* is inefficient. This research proposes a strategic theoretical model based on the concept of contagion used in the biological environments to justify collaboration between similar financial actors in the such as regulators.

The main focus of this chapter is to propose a strategic interaction model between regulators to justify collaboration in the presence of costs. Compared to the relevant literature, we contribute both in the choice and the design of the model. First, we extend the famous SIR model (Susceptible/Infected/Recovered or Removed) to the economic context. We propose a unique design in the network literature of banking that takes into account the heterogeneity of banks and the effects that regulation could have on moral hazard. Second, by contrast to the previous works in the financial intermediation literature, we do not consider banks as the atomic unit in the network. We model the financial system as a population of unit values that

are susceptible to failure. This design allows for the possibility to study general protection measures that are not oriented toward a single financial institution. The results should encourage regulators to consider the international dimension in their expenditures related to regulatory efforts. Depending on the level of interconnect- edness, peripheral countries should help the source country in its regulatory effort beyond the optimal level if decided only by the latter. In fact, the country where the crisis starts has no incentive to stabilize more its financial system beyond its selfish optimal level despite the fact that it would have important positive effects on other countries. We show the importance of a central planner in that context.

After almost ten years since the first sign of the global financial crisis, we believe that the tentative answers provided by each chapter in this document are timely for several reasons.

First, in most advanced economies it seems that the deleveraging cycle is coming to an end. The chances are that banks will take more risk to ensure profits to their shareholders driven by an environment of very low revenue on fixed-income as- sets due to falling interest rates. In such an environment accurately measuring risk grows in importance. The first chapter is very handy in this aspect. Quantile-based risk measures are by far the most widespread technique used by the financial indus- try and it is improving the computational algorithms applicable for those measures the biggest contribution of this chapter to the literature. In fact, techniques applica- ble by practitioners to VaR which aim at improving directly nested-simulations are in rare in the literature. We show that the algorithms that we develop (called sequen- tial and stratified) for computing risk measures such as the VaR yield significant computational savings. In the simulation exercise, we show the non-optimized uni- form algorithm requires at least twice the effort in the case of the Gaussian portfolio to reach the same level of accuracy. The advantage of is even more pronounced for other more complex settings than the naive algorithm. For example, in the case of a portfolio holding a single position of a basket option, the uniform technique needs between 2^4 and $2^{7.5}$ more effort to match the performance of the sequential algorithm. The test cases are also more comprehensive than in the literature. The challenge is that the theoretical value of each quantile must be available to compute the performance metrics (MSE) In the chapter, we also provide theoretical evidence

that the sequential algorithm is superior to its uniform equivalent. In particular, the analytical findings shows that the new technique that we introduce focus the computational power around the exact value of the risk measure for the corresponding portfolio. The importance of such a gain is that it allows risk managers to stretch their model to cover risk far in the tails and tackle interconnectedness risk of which the crisis of 2008 revealed the crucial aspect. Moreover, this chapter proposes a technique that does not compete with classic variance reduction techniques. In other words, the latter simulation's optimization methods can be combined with the algorithms in this chapter. One additional advantage of the sequential and stratified algorithms is that with a small tweak they can be useful to compute other types of risk measures that also call for the use of two-steps simulations such as the regulatory stress tests.

The contribution and lessons drawn from the second chapter are also timely, especially from regulatory aspects. It is indeed critical to build up buffers that can cushion the effect of a financial crisis when the it hits the shores. Moreover, it is important to separate risks that banks can manage via traditional risk management, from those that are due to negative externalities generated by some polluting players in the system. The first type of risk is part of the economic activity of banks and is an engine of economic development. It is the latter that we believe is more dangerous as it is mostly unmonitored and hard to detect. We propose in the third chapter a new framework to detect those invisible risks. The idea is to check the probability of gains to detect the institutions behind such bubble-creating risks. This chapter, shows via three different techniques that extreme gains can be informative about the global health of the financial system. We use a theoretical model in the first part to show that regulators can impose a "fair game" by proposing a weak notion of tail's symmetry. In that case, systemic crises are limited to unpredictable, out of control events. In addition to this theoretical framework, we also establish via simulations that financial systems are safer when all banks have tail's symmetry. The model that we use for simulation cover two types of topologies advocated in the literature on banking. The first is a system where the direct connections between banks are the major contagion channel and the second topology is where a common liquidity market centralize transactions between banks. The model that we simulate cannot probably capture all the complexity of a financial system, but the results of

the simulations are overwhelming and can give hints about the importance of tails symmetry as we advocate it. Finally, we design an empirical measure of externalities based on the idea of tail symmetry. We can, via publicly available data, assess the level of externalities that a bank is creating in a period of time. The measure can be computed at the high frequency and can be updated daily. We compare the link between this measure and a proxy for externalities which is the ex-post fines paid by the US banks and show that a measure based on trials symmetry could have higher explanatory power compared to used techniques available in the traditional systemic risk literature.

The third chapter of this thesis deals with the issue of global safety net. In the aftermath of the financial crisis, it was argued that collaboration between regulators at the international level was essential to avoid and mitigate, when necessary, financial crises. A question remains unanswered, what would justify the cost of cooperation especially amid an economic crisis. In the third chapter, we enrich the literature with a special design of the financial system that not only allows answering the previous question but also considers the relationship between regulations and moral hazard. We base the strategic decision of individual regulators on a unique design of the financial system that can cover a multitude of situations where collaboration is possible. The type of situations can range from individual banks pooling up resources for times of troubles to cooperation between several regulatory agencies to collaboration between international regulator. In this chapter, we dedicate an important space to the description of the model and the dynamic of the evolution of the crisis inside our model. This is mainly because of the novelty of using this type of model in the economic context. Thanks to this design we show that international coordination is preferable when financial systems are connected enough for contagion to leave sizable effects. We also reveal that it is advisable to subsidize reforms in countries where potentially arise can start and little willingness is shown to engage in stronger regulatory oversight.

Each chapter is structured to be independent of the others. Within each of those chapters, we adopt a classic scheme. We will begin first by a brief motivation. Second, we present an overview of the most relevant literature to the research question. We finally present the methodology and the results.

A non-uniform nested simulation algorithms in portfolio risk measurement

WE investigate the computational complexity for estimating quantile-based risk measures, such as the widespread Value at Risk for banks and Solvency II capital requirements for insurance companies, via nested Monte Carlo simulations. The estimator is a conditional expectation type estimator where two stage simulations are required to evaluate the risk measure: an outer simulation is used to generate risk-factor scenarios that govern price movements and an inner simulation is used to evaluate the future portfolio value based on each of those scenarios. We propose a new set of non-uniform algorithms to evaluate risk. The algorithms place more importance upon outer scenarios which are more likely to have a direct impact on the estimator and considers the marginal changes in the risk estimator at each additional inner scenario. We demonstrate using analytical and experimental settings that our proposed heuristics outperform the uniform algorithm and result in a lower variance and bias with the same initial settings and resources. The results are also robust enough for the multidimensionality of risk factors and the non-linearity of pay-offs.

Introduction

Despite the theoretical advances made in the field of derivative pricing, a wide range of commonly used derivatives do not fall within range of a pricing formula. Therefore, practitioners have no alternative but to use Monte Carlo simulations and face the constraints of their computational cost. The closer the model is being able to substantially grasp the complexity of the financial system, the more likely is that the time budget needed for an accurate simulation will become excessive. Many simplifications of the models are, therefore, being applied and accuracy is sacrificed to lower the execution time to acceptable limits fixed by the application of the simulation. Consequently, for the purpose of computing large portfolio risk measures, the optimization of simulations is almost inevitable.

The failure of major risk assessment models to protect both big and small investors in the recent years has demonstrated the need for models which are closer to reality. The sacrifices made to accuracy to simplify the models are no longer acceptable. As flexibility should be added to the pricing models, risk practitioners will have their toolbox limited to Monte Carlo simulations for pricing complex products and nested simulations for computing risk measures.

Regardless of the risk measure, the evaluation procedure is usually divided into two stages. Risk scenarios are generated and designed either to reflect normal market conditions and the most probable evolution of the financial market or the contrary the scenarios which are the least likely but bear critical risks for the financial institution. The second stage is to evaluate the portfolio under the condition of the risk scenarios. Because any risk measure is usually to intended account for possible losses within a future time horizon, the two-levels procedure is fated unless a pricing formula exists for each position in the risky portfolio.

In spite of the fact that for risk management application, the time constraint is rather generous the portfolio's size will affect the task complexity. A second challenge which is overlaid in this particular type of application, it is the use of nested simulation, to evaluate common risk measures such as Value-At-Risk (VaR), Expected Shortfall (ES) or Solvency II capital requirement, on a given horizon. Nested

simulation calls for the use of two stages of simulations: the outer simulation and the inner simulation. The outer simulation is used to sample risk factors over a given time horizon. The number of required risk factors and the correlation between them justifies the need for Monte Carlo simulations at this level. The inner simulation reprices the portfolio instruments conditional on the drawn risk factors.

In this chapter, we question whether the pricing complexities inevitably lead to a large computational burden that may prevent accurate risk assessment in practice. Because some risk scenarios may have no direct impact on the estimator, we show that an ingenious allocation of a relatively small computation budget can yield acceptable levels of variance and bias for portfolio risk measures such as the VaR. We analyze how a fixed computational budget could be allocated across both inner and outer simulations to minimize the Mean Square Error (MSE) of the outcome estimator.

Moreover, the field of application of VaR is not limited to portfolio risk management. Large bank would usually divide its trading activities into trading desks. Management rules limit the freedom of each desk using quantifiable ceilings. Before the widespread application of VaR, these limits used to be defined in terms of notional limits that were hardly comparable between asset classes. Therefore, the use of a VaR-based trading limit is preferable for managing trading desks. For more details about VaR-based risk limits, refer to Blanco and Blomstrom (1999).

Besides, Cuoco et al. (2008) prove that if VaR is recomputed dynamically using the latest available information, then the risk exposure of trading using VaR is always lower than that computed for the unconstrained traders. Therefore, VaR could also be used as a criterion for asset allocation problems. Again, the dynamic re-evaluation of VaR will require a significant computational budget that could jeopardize the application of such a strategy in practice. It is obvious that we are working in relatively small perimeters compared to those of risk management applications, and that means we are no longer limited by the size of the problem. Nevertheless, for trading applications, the time constraint is generally very tight which creates the need for a clever simulation design.

A second major application of nested simulation in finance is the valuation of American options. It is important to remember that American options can be exercised at any time until maturity. Hence, the holder of such options is faced with an optimal stopping problem where he must choose the best execution time. The straightforward technique to resolve the issue is by nested simulation. Because a continuous-time stopping problem is burdensome in simulation, most American options are approached as Bermudan options, which have a finite set of execution times. Whenever the set of possible execution times is large enough, the American option could be treated as a Bermudan option. The simulation procedure, which is usually expensive in terms of computational resources, consists in measuring the continuation value at every step on each path via inner simulation to decide whether to exercise the option or continue to hold the option. The outer simulation will be dedicated to sampling different paths. This procedure is not privileged in practice due to its computational cost. However, when dealing with high dimensionalities, such as American style basket options, Monte Carlo repricing seems to be inevitable.

The core of this paper is to present a set of efficient heuristic algorithms to evaluate quantiles such as VaR. We are focusing our attention on improving the direct nested simulation techniques. Our main purpose is to allow the use of more realistic models in the managing of financial investments.

These ideas are based on the work of Gordy and Juneja (2010) and inspired by the work of Broadie et al. (2011) on evaluating the probabilities of large losses. The main concept behind this paper is that for a two-stage simulation, the additional budget will not have the same marginal improvement from one simulation to another. The foremost contributions of this paper to the literature of non-uniform nested estimators are:

1. We will provide a non-uniform nested simulation algorithm for estimating quantile-based risk measures such as VaR and the implementation of the internal model in the Solvency II regulations. The algorithm has no prerequisite that would prevent practitioners from adopting it. The algorithm will execute the simulation sequentially. The first set of simulations will generate preliminary results. Then extra budget will be allocated, at each step, where

one added inner simulation will have the largest impact on the desired risk measure. The numerical implementation of this estimator demonstrates that bias is reduced dramatically even for relative small additional computational budgets. We also provide a theoretical justification of the efficiency of this technique compared to the standard ones.

2. A second estimator will be proposed. In fact, given that the purpose of the non-uniform estimator is to enhance the uniform algorithm in terms of bias, a budget saving could be generated. However, where the actual level of bias is acceptable, the savings could be considered to generate additional outer scenarios and, consequently, reduce the variance. Therefore, we will present an estimator that will try to balance the bias generated by inner simulations and the variance reducible by increasing the number of outer scenarios.

This chapter will be structured as follows. Section (2.1) will provide a brief literature review. The general simulation framework and notation will be presented in the next section. The third section will be devoted to describing the optimal algorithm proposed for the quantile estimators. Finally, numerical results and conclusions will be presented in the fourth section.

2.1 Literature Review

As the measurement is the cornerstone of risk management in a financial institutions, it has received the attention of scholars and practitioners for the last few decades. For a general overview of risk measures and risk management in financial markets refer to Crouhy, 2000. Properties and requirements of risk measures were studied in Artzner et al., 1999. They proposed a set of axioms that risk measures are required to satisfy to be deemed a coherent risk measure. It should be noted that, despite the interest given to coherent risk measures in theoretical research, VaR (which is not a coherent risk measure) is, by far, the most widespread risk measure used within the banking industry. However, a number of papers including Cochran et al., 2010 have criticized the axioms of coherence and proposed a different classification for risk measures. Their newborn class of risk, called natural risk measures, includes the VaR as defined by the Basel II and Basel III regulations. They also argued that VaR is not contradictory to the principle of diversification as is claimed by Artzner et al., 1999. In fact, according to the VaR criteria, the merger of two portfolios will increase risk only in case of extremely heavy tailed-distribution in which case diversification may not be preferable. Their work seems to provide a theoretical justification for the Basel II and Basel III regulations. The determination of the VaR is also important to compute other risk measures such as the ES defined as the expected loss when the VaR is exceeded. Therefore, even if regulators will shift to the usage of ES for regulatory capitals the determination of the quantiles will remain important and its accuracy will have a direct impact on the precision of the ES. In addition, regulators of financial sectors other than the banking sector, like insurance calls for the use of a quantile-based risk measure. It is the case for example for Solvency II regulation in Europe. For a complete overview on the computational detail of Solvency II capital requirements please refer to Devineau and Loisel (2009) and Bauer et al. (2012). An alternative measure of riskiness has been developed by Bali et al., 2011, who were able to classify portfolios according to their expected return by unit of risk.

The problem of estimating risk measures using nested simulation was first introduced by Lee (1998) and Lee and Glynn (2003). Those authors have begun by

investigating the properties of the uniform nested simulation estimator. Such estimators, distribute budget over outer scenarios equally and lead to a constant number of inner simulations. Both authors demonstrated the asymptotic variation of the accuracy of such an estimator where the bias is a function of inner simulation, while the variance of the estimator is inversely proportional to the number of outer scenarios.

The work of Gordy and Juneja (2010) was able to characterize the perfect allocation of budget between inner and outer simulation for uniform nested simulations and has assessed bias for the particular case of the gaussian portfolios. For the continuous case Gordy and Juneja (2010) established that given a total computational budget of Γ the optimal asymptotic mean square error (MSE) is of order $\Gamma^{-2/3}$ for VaR and expected shortfall. Although, their work is pioneering concerning methodology, it is unclear how to compute the optimal allocation for more complex portfolios where the Profit and Losses' distribution (P&L) is not theoretically characterized. Guojun Gana (2015) also used similar techniques to value large portfolios of variable annuity (VA) products that provide downside protection from the fluctuation of financial markets in the form of a minimum guarantee.

Authors, such as Longstaff and Schwartz (2001) addressed the problem of nested simulations for pricing American options. They proposed to reduce the computational burden by using a meta-modeling methodology. More precisely, a limited number of inner simulations is generated to estimate a relationship between the price of the options and the risk factors using the least square modeling approach. For a complete overview of the latest techniques of Monte Carlo methods for valuing American options and the possible improvement, refer to Bouchard and Warin (2012).

A similar approach is that of Liu and Staum (2010) who proposed the use of stochastic kriging for estimating risk measures. Their approach is also a meta-modeling approach that takes into consideration the bias in the estimated values. Kernel-based estimators were studied in the particular case of estimation conditional expectations by Hong, 2009. They demonstrated that the estimator converges at the rate $k^{-\min(1, \frac{4}{d+2})}$ where d is the dimension of the risk factor's vector. This method beats

the nested simulation only in the case where $d \leq 3$. However, this technique will lose its competitiveness in high dimensional simulations.

The work of Broadie et al. (2011) is the closest work to our study. They established an efficient algorithm to allocate budget to inner simulation to compute the probability of large losses. Their development is based on the idea that the marginal change in the risk estimator is not uniform across scenarios when the additional budget for inner simulations is allocated. For risk measures that focus on the tail of the distribution, such as VaR and the probability of large losses, the extreme scenarios are the ones that matter the most.

In our study, we will try to make a similar extension as Broadie et al. (2011) to the work of Gordy and Juneja (2010) but to a more conventional risk measure i.e. the VaR. The main difficulty of our circumstances compared to those of Broadie et al., 2011 is that for evaluation of the probability of large losses the threshold defining extreme losses is known and is expressed in nominal terms. In other words, both the input and the output from the algorithm have the same dimension. In our work, the threshold defining the extreme losses needs to be estimated then updated after each step of the sequential simulation. Our work is also of interest to other applications that need the a conditional expectation of a quantile close to the tail such as the Solvency II capital requirements.

2.2 General Framework

In this section we will present the model framework for the two main application of nested simulation in the financial industry: Value-at-Risk and Solvency II capital requirements. It is also possible to call for the use of nested simulation for other financial application like the valuation of exposures in credit risk measurement. In this paper, we will only focus on regulatory measures. Nevertheless, only a small tweak is needed to be made to the proposed algorithm to be useful for the mentioned application.

2.2.1 Model Framework for VaR

For consistency purposes, we will follow the model and the notations proposed by Gordy and Juneja (2010).

Let X_t be the vector of v state variables that will determine the asset prices. The vector should include all the information needed for the pricing operation at time t . In the following, $\mathcal{F}_t$ is the filtration generated by X_t . In order to discount future cash flows, we denote by $B_t(s)$ the value of a unit of currency invested at time $t \leq s$ in a risk free money-market account. Given the interest rate r then:

$$B_t(s) = \exp\left(\int_t^s r(u)du\right),$$

The portfolio that we study is composed of $K + 1$ positions. The price of each position k will depend on the t , $\mathcal{F}_t$ and the legal characteristics of the instrument priced. Position p_0 , regroups the set of instruments for which an analytical pricing function is available. The portfolio is assumed to be held static over the model horizon and maturities T_k are finite for all positions $k = 1, \dots, K$. The second assumption will ensure that Monte Carlo pricing is always possible for the positions with no proposed analytical formula.

Let $C_k(t)$ be the cumulative cash flow for the position k over the time horizon $(0, t]$. Notice that conditional on $\mathcal{F}_t$, $C_k(t)$ is a deterministic function according to the instrument contractual terms. According to that, the market value of each position is the present discounted expected value of its cash flows under the risk-neutral measure $\mathcal{Q}$.

$$V_k(t) = \mathbb{E}^{\mathcal{Q}} \left[\int_t^{T_k} \frac{dC_k(s)}{B_t(s)} \middle| \mathcal{F}_t \right] \quad (2.1)$$

Whenever $t \geq T_k$, we set $V_k(t) = 0$

Given all those assumptions the portfolio loss, defined as the difference between the present portfolio value and the discounted future values adjusted to interim cash flows, Y can be written as:

$$Y(H) = \sum_{k=0}^K \left(V_k(0) - \frac{1}{B_0(H)} \left(V_k(H) + \int_0^H B_t(H) dC_k(t) \right) \right) \quad (2.2)$$

Here the present time is normalized to 0 and the model horizon is H . An implicit assumption is that the interim cash flows received at time $t, t < H$ are reinvested in the money market until time H . However, other conventions could be easily adopted. Moreover, no portfolio weights are used in this model as positions are expressed in currency units.

The combination of equations (2.1) and (2.2) illustrates the necessity of nested simulations. Observe that by construction, Monte Carlo simulation is inescapable for pricing position k where $k \neq p_0$. Therefore, repricing via equation (2.2) is only possible via simulation. Moreover, instruments value at time H and interim cash flows up to the horizon H are conditioned to the choice of the filtration $\mathcal{F}_t$. Once again, simulation is the only way through this problem, where different filtrations will be generated in order to obtain a vision of the loss distribution function.

Because the value of a position k at time H is simply a conditional expectation that does not depend on future cash flows of other positions (see equation(2.1)), it should be noted that inner simulations can be run independently across position. This could have two benefits in practice. First, it will be possible to run parallel repricing for the distinct positions in the portfolio and take advantage of the advancing parallel computing hardware. Second, this will ensure the diversification of pricing error among different positions and lead to less biased portfolio measures. We will also assume that the initial prices $V_k(0)$ are already known and can be taken as constant in our problem.

2.2.2 Simulation Framework for VaR

In this section, we will develop the notations needed to illustrate the simulation process. The simulation is nested. More precisely there is an *outer step* in which

we simulate scenarios up to the time horizon H . In each trial in the outer step, a second simulation called the *inner simulation* is needed to reprice each position (except the position p_0 by construction).

In what will follow, L will represent the number of trials in the outer simulation. In each of these trails we will execute the following steps:

1. Simulate a path X_t for $t \in (0, H]$ under the physical measure. Let ξ be the realisation of random variables $(X_t : 0 < t \leq H)$. Hence, ξ_l represents the generated information in the outer steps of trial l .
2. Evaluate the accrued value at H of the interim cash flows.
3. Evaluate the price of each position at H and this by applying the following rule
 - a) Closed-form price for position p_0 at H
 - b) Simulation with N_l inner steps trials for time period $(H, T_l]$. The inner paths are simulated under the risk neutral measure.
4. Sum both the accrued value of cash flows and the estimated value at time H and discount everything back to time 0. The estimate loss is then $\tilde{Y}(\xi_i)$

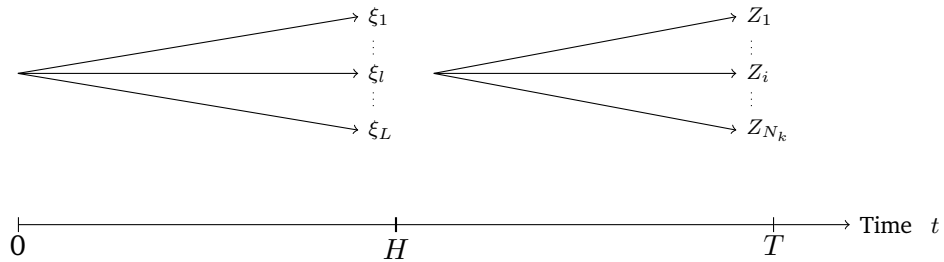


Fig. 2.1.: Illustration of the Nested Simulation sampling

Figure (2.1) illustrates the idea of the simulation and its nested aspect. Step (1) in the previous procedure is shown in the figure in the time horizon $[0, H]$. Then, starting from a generated scenario ξ_l , each position is repriced using the second simulation as specified in the step 3(b) up to the time horizon T . Using this procedure, the value of each position at time H will be the mean of the position following the generated path in the inner simulation. The inner Monte Carlo estimator is unbi-

ased. However, its inaccuracy will engender bias in the overall estimator of the risk measure.

2.2.3 Estimating Value-at-Risk

We will now go into more details regarding the problem of efficiently estimating the Value-at-Risk for the loss function Y . For a target insolvency probability α , $\alpha \in [0.1]$, VaR is the value y_α given by :

$$y_\alpha = VaR_\alpha[Y] = \inf\{y : P(Y \leq y) \geq 1 - \alpha\} \quad (2.3)$$

As specified before, the nested simulation generates samples $(\tilde{Y}(\xi_1), \dots, \tilde{Y}(\xi_L))$. We sort these draws as $\tilde{Y}_{[1]} \geq \dots \geq \tilde{Y}_{[L]}$. $\tilde{Y}_{[\alpha L]}$ provides an estimate of y_α $\lceil a \rceil$ denotes the integer ceiling of the real number a .

The focus of our work is the efficiency of the estimator. Therefore, we will begin by characterizing the mean square error (MSE). MSE is a conventional measure of the performance of the nested simulation estimator. The objective of this paper is to minimize the MSE.

The MSE $\mathbb{E}[(\tilde{Y}_{[\alpha L]} - y_\alpha)^2]$ could be decomposed as the sum of a bias and a variance :

$$\underbrace{\mathbb{E}[(\tilde{Y}_{[\alpha L]} - y_\alpha)^2]}_{MSE} = \underbrace{\mathbb{V}[\tilde{Y}_{[\alpha L]}]}_{Variance} + \underbrace{\left(\mathbb{E}[\tilde{Y}_{[\alpha L]} - y_\alpha]\right)^2}_{Bias^2} \quad (2.4)$$

For a given computational budget, the problem of estimating Value-at-Risk, beyond the pricing complexity of derivatives, is the trade-off between bias and variance. To compute the most efficient estimator of the quantile-based risk measure the budget must be allocated wisely between the two levels of simulation.

2.2.4 Estimating Solvency II capital requirements:

In Europe, Solvency II plays the role of Basel III for the insurance industry. The sets of rules in Solvency II aim at improving the solvency of insurance companies. The key element of the Solvency II approach is that the company should hold enough capital at time zero to overcome difficulties that may arise from an unforeseen event in the following year.

Economic Balance Sheet at time t	
A_t	E_t
	L_t

Where A_t is the market value of the asset at time t . E_t is the equity value at time t and L_t represents the liabilities also at time t . The objective of solvency is to ensure that insurance companies have enough capital to face a bankruptcy situation i.e $E_t = A_t - L_t < 0$. To compute the value of the economic balance sheet, we need to introduce the following notation:

- $(F_t)_{t \geq 0}$: is the filtration of the available information at time, all the elements
 : of the balance sheet are F_t measurable
 δ_t the discount factor expressed with the risk free rate r_t .
 $\delta_t = e^{-\int_0^t r_h dh}$
 P_t the cash flows of liabilities at period t
 R_t the profit of the company at time t
 (2.5)

Thus, the value of equity and liabilities at time 0 are easily computed as:

$$L_0 = \mathbb{E}_Q \left[\sum_{u \geq 1} \delta_u P_u | F_0 \right] \quad (2.6)$$

$$E_0 = \mathbb{E}_Q \left[\sum_{u \geq 1} \delta_u R_u | F_0 \right] \quad (2.7)$$

The economic capital is then evaluated using the following formula:

$$C = E_0 - P(0, 1)q_{0.5\%}(E_1) \quad (2.8)$$

With $P(0, 1)$ is the price of a zero coupon of 1-year maturity. The quantity C is the surplus of capital that need to be added to ensure that the condition that may wipe out the entire equity occurs with a probability equal to 0.5%. We should consider that the evaluation of C needs the knowledge of the quantile of the equity distribution at a future time 1. This is very similar to the computation of VaR . In this context we also have that:

$$E_1 = \mathbb{E}_Q \left[\sum_{u \geq 2} \frac{\delta_u}{\delta_1} R_u | F_1^{RW} + R_1 \right] \quad (2.9)$$

Where F_1^{RW} is the real world information of the first year.

For more details about the computation of Solvency II capital requirement please refer to Bauer et al. (2012) and Devineau and Loisel (2009). It is important to have a snapshot of the distribution of E_1 to evaluate the desired quantile. For that purpose, we need to simulate a different set of information F_1^{RW} . The valuation of the future return of the company is usually a simulation exercise. In fact, those revenues depend on several risk factor and a set of financial and non-financial variables with complex emended options in the insurance contracts. Both simulations combined is again a situation of nested simulations.

It is important to note that insurance companies are struggling to implement this technique and are resolved to use a more simplified approach that may miss situations that threaten the viability of the company. It is also worth mentioning that the particular case of Solvency II capital requirements is more complex than Basel III requirements in our opinion. In fact, the long time horizon (1 year) for Solvency II calls for the use of numerous real world scenario to ensure an acceptable level of accuracy. According to a study by Moody's (Morrison (2009)) as much as 100 000 real world scenarios are required by insurance companies. This number is almost computationally infeasible for a large portfolio of an insurance contract. Therefore,

the simulation optimization is inevitable. This paper proposes a feasible solution to this kind of problem.

2.3 Sampling Algorithms

2.3.1 Uniform Sampling

Uniform sampling is perhaps the most obvious way to proceed to a two level simulation. The estimator is a function of two variables L and N . L is the number of outer simulation and N is the number of inner simulations. The estimator is *uniform* in the sense that the number of inner stage samples is identical for each outer stage scenario. The algorithm is as follows:

Algorithm 1 VaR uniform estimator sampling

```

1: procedure UNIFORM( $L, N$ )
2:   for  $l \leftarrow 1, L$  do
3:     Generate scenario  $\xi_l$ 
4:     Evaluate the accrued value at  $H$  of the interim cash flows
5:     Estimate the closed form price for position 0
6:     Conditioned on the scenario  $\xi_l$  generate i.i.d inner samples  $\hat{Z}_{l,1}, \dots, \hat{Z}_{l,N}$ 
       of portfolio losses
7:     Compute an estimate of portfolio loss in scenario  $l$   $\tilde{Y}_l = \frac{1}{N} \sum_{i=1}^N \hat{Z}_{l,i}$ 
8:   end for
9:   Compute an estimate of the VaR,  $\hat{y}_\alpha = \tilde{Y}_{[\alpha L]}$ 
10: end procedure

```

An asymptotic characterization of the bias and variance of the uniform estimator is possible throughout a set of technical assumptions that will ensure that the higher order partial derivatives could be eliminated when proceeding with a Taylor series expansion to compute an asymptotic version of the bias¹ as follows:

Gordy and Juneja (2010) established an asymptotic characterization of the bias based upon known properties of order statistics:

Theorem 1 ² :

$$Bias = \mathbb{E}[\tilde{Y}_{[\alpha L]}] - y_\alpha = \frac{\theta_\alpha}{Nf(y_\alpha)} + o_N(N^{-1}) + O_L(L^{-1}) + o_N(1)O_L(L^{-1}) \quad (2.10)$$

¹For more information about these technical assumptions and a discussion about their implications please refer to Gordy and Juneja (2010).

²We say that a function is $O_m(h(x))$, if its absolute value is upper bounded by a constant multiplied by $h(x)$ starting from a sufficiently large m . In the same way, we say that a function is $o_m(h(x))$, if for all $\epsilon > 0$ the absolute value of the function is upper bounded by ϵ multiplied by $h(x)$ starting from a sufficiently large m

Knowing that :

$$\Theta(u) = \frac{1}{2}f(u)\mathbb{E}[\sigma_\xi^2 \mid Y(\xi) = u]$$

Where f denotes the density distribution function of Y and σ_ξ^2 is the conditional variance of the error of the portfolio inner pricing (conditioned on ξ). Finally let $\theta_\alpha = -\Theta'(y_\alpha)$.

With the aid of Theorem 1, we can visualize³ that in the case of Value-at-Risk, the bias is introduced by the uncertainty of the inner simulation. Consequently, the number of inner simulations determines of the level of bias. Allocating more budget to the inner simulation will increase the accuracy of the inner simulation and, consequently, reduce the bias. However, an increase of 1 in the number of inner scenarios will result in an increase of L in the total budget for uniform estimators.

The number of outer scenarios will be the key factor in variance reduction as increasing the number of outer scenarios sampled will decrease the variance. Theorem 2 is an asymptotic characterization of the level of variance

Theorem 2

$$\mathbb{V}[\tilde{Y}_{\lceil \alpha L \rceil}] = \frac{\alpha(1-\alpha)}{(L+2)f(y_\alpha)} + o_L(L^{-2}) + o_N(1)O_L(L^{-1}) \quad (2.11)$$

Theorems (1) and (2), besides characterizing the origin of bias and variance, also allow the convergence rate of the uniform algorithm to be determined. Depending on the application, the convergence speed of both bias and variance may oblige a practitioner to employ significant numbers of both inner and outer scenarios. Consequently, for large portfolios, uncontrollable amounts of memory and computational resources are often required to satisfy the industry standards.

³The $Bias^2$ is proportional to $\frac{1}{N^2}$, as the two parameters θ_α and $f(y_\alpha)$ are not a function of the simulation parameters. Hence, as the bias could be eliminated by dramatically increasing the number of inner simulation, the inaccuracy introduced by the inner Monte Carlo simulation is the origin of bias

2.3.2 Optimal Uniform Sampling

Gordy and Juneja (2010) demonstrated that the optimal choice of L and N could lead to a better allocation of computational budget. In other words, they established the existence of L^* and N^* that minimize the MSE. This result is conditioned by the verification of a set of assumptions. Before illustrating the value of the optimal choices, we will first assume the following notations: γ_0 is the average effort needed for generating an outer scenario. γ_1 will be the average effort needed in the inner simulation. Thus, the overall computational effort will be $\Gamma = L(N\gamma_1 + \gamma_0)$

Specifically, to find the optimal budget allocation between inner and outer simulations, we will use the approximation of MSE. It is important to remember that the MSE is the sum of $Bias^2$ and variance detailed in Theorem 1 and Theorem 2. The solution of the following optimization problem will lead to a perfect asymptotic allocation of budget between the inner and outer scenarios. It will also lead to the best possible performance using the uniform algorithm.

$$\begin{aligned} & \underset{L, N}{\text{minimize}} \quad \frac{\theta_\alpha}{Nf(y_\alpha)} + \frac{\alpha(1-\alpha)}{(L+2)f(y_\alpha)} \\ & \text{subject to} \quad L(N\gamma_1 + \gamma_0) \leq \Gamma, \\ & \quad \quad \quad L, N \geq 0. \end{aligned}$$

The solution to the problem is :

$$\begin{aligned} N^* &= \left(\frac{2\theta_\alpha^2}{\alpha(1-\alpha)\gamma_1} \right)^{1/3} \Gamma^{1/3} + o_\Gamma(\Gamma^{1/3}) \\ &\text{and} \\ L^* &= \left(\frac{\alpha(1-\alpha)}{2\gamma_1^2\theta_\alpha^2} \right)^{1/3} \Gamma^{2/3} + o_\Gamma(\Gamma^{2/3}) \end{aligned} \tag{2.12}$$

The two parameters are then injected into the uniform algorithm to compute an estimator of the VaR.⁴

⁴The proof is given in Gordy and Juneja, 2010

Despite the importance of the work of Gordy and Juneja (2010), they do not give a practical approach for implementing the simulation, and the optimal allocation solution that they present is quite difficult to use in practice as some of the parameters required are not available and must be the object of a simulation. Therefore, the optimal solution could only be used as a benchmark of other simulations designed to compute risk measures and could not be applied to industry problems.

Lee and Glynn (2003), adopted the strategy of using a two-stage simulation to compute the distribution of conditional expectations. The output of the first stage is the number of inner and outer simulations to be performed in order to attain the optimal estimator. The second step will be to perform nested simulations to compute the desired loss probability. Their algorithm delivered a small empirical improvement to the crude uniform estimator.

2.3.3 Sequential Simulation

Here, we will present a second family of algorithms studied by Broadie et al. (2011). Their work is focused on the measure of the probability of large loss. Their work is very similar to that of Lee and Glynn (2003), yet concentrates on probabilities far along the tail of the distribution.

Our development is inspired by their methodology and we tried to use similar thinking process to develop more efficient algorithms to estimate the VaR.

Because the estimator of the VaR is $\hat{Y}_{[\alpha L]}$, we can see that only the order of the estimated values and the quantity of the $[\alpha L]^{th}$ loss matters for the estimator. The basic idea behind our procedure is that an additional scenario will yield the greatest impact whenever it has a much greater probability of changing the order of the estimated losses and by consequence a greater probability of affecting the estimator. To explain and illustrate our idea, we will present a simple example. Imagine a setting where a certain number of inner and outer simulations has been performed. The quantile that we seek to estimate, is on the right tail of the distribution. Therefore, obvious allocating additional budget to the outer scenarios where losses are closer to the right tail will have a greater impact to allocate. By contrast, the additional

of an inner scenario to an outer scenario where the associated loss is located in the left tail will not impact upon the VaR estimate greatly. It is very unlikely that the added scenario will make the loss jump to the level that will alter the quantile estimator on the opposite tail of the distribution. This example illustrates both the concept behind the estimator that we are developing and, in the same occasion, demonstrated the inefficiency of the uniform estimator.

To go further into detail, let us suppose that we have already estimated L outer scenarios using N_0 inner steps. The challenge is to find the best allocation for the remaining budget $\Gamma - L_0(N_0\gamma_1 + \gamma_0)$.

Without any loss of generality, let us suppose that the set $\hat{Y}_{\xi_1}^B(k_1), \dots, \hat{Y}_{\xi_L}^B(k_L)$ is non decreasing. Where $\hat{Y}_{\xi_l}^B(k_l)$ is the loss estimated using k_l inner steps and conditioned to scenario ξ_l . The overall budget consumed for the simulation at this stage is $B = L_0(N_0\gamma_1 + \gamma_0)$.

The additional inner scenario will cause a significant if the new loss estimated affects the order of the losses around the current estimation using the available simulations. To be more specific, let us suppose that for the budget B the scenario ξ^* is the VaR scenario. In other words, $Y(\xi^*)^B(k_*) = \hat{VaR}(B)$.

$\hat{Y}_{\xi_i}^{B+1}(k_i)$ will denote the loss estimated for the scenario ξ_i using a total budget of $B + 1$ for the global simulation. The first step is to look for the outer scenario ξ_k verifying:

$$\left(\hat{Y}_{\xi_k}^B(k_k) - \hat{VaR}(B)\right) \left(\hat{Y}_{\xi_k}^{B+1}(k_k) - \hat{VaR}(B)\right) < 0 \quad (2.13)$$

Equation (2.13) ensures that the loss evaluated conditioned on scenario ξ_k will jump from the right side of the estimator when adding a single inner scenario or inversely depending on the initial order of $\hat{Y}_{\xi_k}^B(k_k)$ and $\hat{VaR}(B)$. In other words, $\hat{VaR}(B)$ should be inside the segment bounded by $\hat{Y}_{\xi_k}^B(k_k)$ and $\hat{Y}_{\xi_k}^{B+1}(k_k)$ to satisfy condition 2.13.

Regardless of the initial and final configuration, we are confident that a scenario satisfying equation (2.13) will have a direct impact on the estimator and therefore, will be more efficiently allocated. In fact, whenever a new inner simulation added to one of the outer-scenarios simulations satisfies condition (2.13), the estimator will see its value changed simply because the outer scenario that generated the VaR loss will be changed. By contrast, not every inner-scenario added can leave its footprint on the value of the estimator. In such a situation, we can say that the additional budget spent to add the corresponding inner scenario was wasted. This is often the case when allocating budget to the outer-scenarios that will result in losses in the left tail of the loss distribution.

However, the main difficulty of this setting compared to the framework of Broadie et al. (2011) and Lee and Glynn (2003), is that for measuring quantile, loss thresholds must be estimated.

It is useful to remember that, in the setting of estimating the probability of large loss, the estimator is:

$$\hat{\alpha} = \frac{1}{L} \sum_{i=1}^L \mathbb{1}_{\{\hat{Y}(\xi_i) > c\}} \quad (2.14)$$

where c is the extreme loss threshold. Both authors evaluated the probability of making an impact on the estimator according to the loss benchmark c which is an input of the estimator and is expressed in currency units. On the other hand, the input for quantile estimation, is a loss probability. Then, as specified in equation (2.13), the benchmark $VaR(B)$ is an estimated value and can bear some inaccuracy. Consequently, we should take the uncertainty of the loss estimated in the development into consideration in the sequential simulation.

The uncertainty surrounding the parameter $VaR(B)$ is illustrated in the value of $\hat{\sigma}_i$, the estimated standard deviation of loss conditioned on a scenario ξ_i . $\hat{\sigma}_i$ is considered as a commonly accepted measure of the uncertainty of the conditional expectation. Hence in the algorithm we will seek a scenario that not only change the estimator but also consider the uncertainty errors in the estimator itself. In order to incorporate the estimation error within the first scenario, our evaluation of the probability of having a direct impact upon the estimator by adding a single

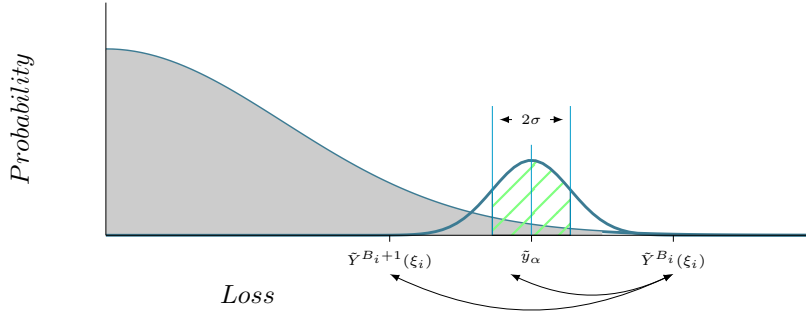


Fig. 2.2.: Illustration of the Stratified Nested Simulation sampling changing order strategy without accounting for uncertainty due to the initial configuration of the scenario and the actual loss estimate

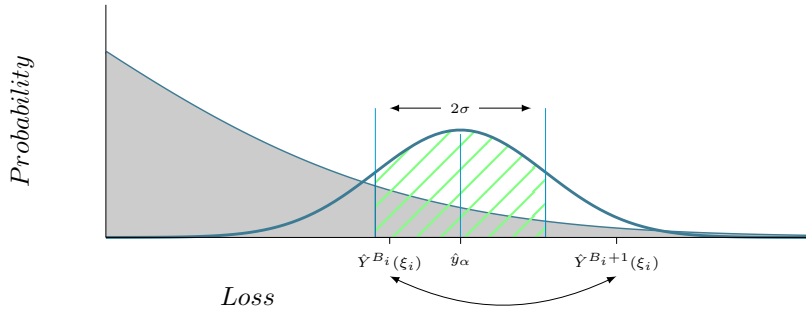


Fig. 2.3.: Illustration of the Stratified Nested Simulation sampling changing order strategy while accounting for uncertainty due to the initial configuration of the scenario and the actual loss estimate

inner scenario will depend on the initial settings. In fact, we will distinguish two different configurations:

- $|\hat{VaR}(B) - \hat{Y}_{\xi_i}^B(k_i)| > \hat{\sigma}_i$ configuration illustrated in figure 2.2.

In this case, we can say with an acceptable level of confidence that the loss generated by the outer scenario ξ is greater (or smaller) than the actual value of the quantile we are trying to estimate using the computational budget B . Our certainty is derived from the fact that the loss $\hat{Y}_{\xi_i}^B(k_i)$ is outside the uncertainty region of the estimator $\hat{VaR}(B)$. Therefore, in this setting, there is no need for special treatment to deal with the bias in the estimator.

- $|\hat{VaR}(B) - \hat{Y}_{\xi_i}^B(k_i)| \leq \hat{\sigma}_i$ configuration illustrated in figure 2.3

In fact, whenever the loss associated with a given outer scenario ξ_k is within the domain of uncertainty, the scenario is still a candidate for the $\lceil \alpha L \rceil^{th}$ scenario and consequently $\hat{VaR} = \hat{Y}_{\xi_i}^B(k_i)$. We will consider that the best

remedy is to account for inner scenarios that will have the maximum amount of changes to the actual loss estimated conditioned to a given outer scenario. In other terms, we will try to maximize the chances to jump over $\hat{y}_\alpha \pm \hat{\sigma}_i$ depending on the initial position of the loss-estimate under consideration and the current estimator of the VaR. The previous treatment will maximize the chances that an order change will occur, and this jump is more likely to be considered relative to the real value of the VaR. In other terms, we are hoping to identify an allocation of the additional inner scenario that may clear the fog and position the loss outside the uncertainty domain $([\hat{y}_\alpha - \hat{\sigma}_i, \hat{y}_\alpha + \hat{\sigma}_i])$.

If we were to perform the additional sample in a given scenario ξ_i , this would result in a new loss estimate given by:

$$\hat{Y}_{\xi_i}(k_i + 1) = \frac{1}{k_i + 1} \sum_{j=1}^{k_i+1} \hat{Z}_{i,j} = \frac{1}{k_i + 1} \hat{Z}_{i,k+1} + \frac{k_i}{k_i + 1} \hat{Y}_{\xi_i}(k_i) \quad (2.15)$$

$\hat{Z}_{i,j}$ is the value of the simulated inner scenario j corresponding to the outer scenario i .

This additional sample will have maximum impact if it changes the order of the set of loss estimators. The event of sample order change, according to the previous condition when considering uncertainty, will be called event A . At this level, it is easy to see that A is a union of two disjointed events: A_1 and A_2 . Where A_1 is the settings where the uncertainty domains of the scenario and the VaR estimator determined by the corresponding standard deviation are overlapping. A_2 is the case where both the estimator of the VaR and the loss conditioned of a scenario are different enough to ignore the effect of uncertainty.

Noting that $d = \hat{\sigma}_{[\alpha L]}$ and $m = \hat{y}_\alpha$ Observe that:

$$\mathbb{P}(A_1) = \mathbb{P}\left(|\hat{Y}_{\xi_i}(k_i + 1) - m| \geq d \mid |\hat{Y}_{\xi_i}(k_i) - m| \leq d\right) \quad (2.16)$$

$$\mathbb{P}(A_2) = \mathbb{P}\left(\left(\hat{Y}_{\xi_k}^B(k_i) - m\right) \left(\hat{Y}_{\xi_k}^{B+1}(k_k) - m\right) \leq 0 \mid |\hat{Y}_{\xi_i}(k_i) - m| \geq d\right) \quad (2.17)$$

Using the approximation $k_i(m+d-\hat{Y}_{\xi_i}(k_i))+(m+d-\mu) \approx k_i(m+d-\hat{Y}_{\xi_i}(k_i))$ where $\mathbb{E}[\hat{Z}_{i,k}] = \mu$, the one side Chebyshev inequality, and by denoting $\sigma_i = \mathbb{V}ar[\hat{Z}_{i,k}]$ we can establish that:

$$\begin{aligned}\mathbb{P}(A_1 | \hat{Y}_{\xi_i}(k_i) \leq m) &\leq \left(1 + \frac{k_i^2}{\sigma_i^2} (m+d-\hat{Y}_{\xi_i}(k_i))^2\right)^{-1} \\ \mathbb{P}(A_1 | \hat{Y}_{\xi_i}(k_i) \geq m) &\leq \left(1 + \frac{k_i^2}{\sigma_i^2} (m-d-\hat{Y}_{\xi_i}(k_i))^2\right)^{-1}\end{aligned}\quad (2.18)$$

And finally, using equivalent development and notation we can establish that:

$$\mathbb{P}(A_2) \leq \left(1 + \frac{k_i^2}{\sigma_i^2} (m-\hat{Y}_{\xi_i}(k_i))^2\right)^{-1} \quad (2.19)$$

Before detailing the idea of the optimization algorithm, we will need to divide the set of outer scenarios into three subsets I_1 , I_2 , and I_3 where:

$$I_1 = \left\{i \in 1..L \mid m-d \leq \hat{Y}_{\xi_i}(k_i) < m\right\} \quad (2.20)$$

$$I_2 = \left\{i \in 1..L \mid m \leq \hat{Y}_{\xi_i}(k_i) < m+d\right\} \quad (2.21)$$

$$I_3 = \left\{i \in 1..L \mid |\hat{Y}_{\xi_i}(k_i) - m| \geq d\right\} \quad (2.22)$$

From this set of equations, we can observe that we have an optimal solution that maximizes the probability of order change (event A). Thus let n_1^* , n_2^* and n_3 be a triplets of integer that satisfies :

$$\begin{aligned}n_1^* &= \underset{i \in I_1}{\operatorname{argmin}} \left\{ \frac{k_i^2}{\sigma_i^2} (m+d-\hat{Y}_{\xi_i}(k_i))^2 \right\} \\ n_2^* &= \underset{i \in I_2}{\operatorname{argmin}} \left\{ \frac{k_i^2}{\sigma_i^2} (m-\hat{Y}_{\xi_i}(k_i))^2 \right\} \\ n_3^* &= \underset{i \in I_3}{\operatorname{argmin}} \left\{ \frac{k_i^2}{\sigma_i^2} (m-\hat{Y}_{\xi_i}(k_i))^2 \right\}\end{aligned}\quad (2.23)$$

Recall, that n_i^* ; $i = 1..3$ is the minimum over 3 disjointed outer scenario spaces I_i ; $i = 1..3$. Them each minimum is computed over one of the three subspaces. This could have implications for the memory allocation and CPU time in the imple-

mentation of the algorithm. Moreover, this distinction between outer scenarios is the cornerstone of the procedure dealing with the uncertainty of the intermediate estimator of the quantile. More specifically, only when the scenario ξ_k is in the subspace I_3 , are we sufficiently confident that the corresponding loss is different from a possible VaR scenario and therefore we can only hope for a scenario that may affect the order of the losses. However, in both cases when the scenario ξ_k is either in I_1 or I_2 , the scenario is a possible candidate to be the VaR scenario. Hence, we hope that the additional inner scenario will not only affect the order of the losses but also eliminate this foggy situation. The confusion will be eliminated if the additional inner scenario makes the loss jump outside the uncertainty domain.

Finally, in the estimation algorithm, the additional scenario will be attributed to $i^* = \inf(n_1^*, n_2^*, n_3^*)$ as this is the scenario that is most likely to impact the current estimator of the VaR.

To summarize the first sequential algorithm then becomes:

Algorithm 2 VaR Sequential estimator sampling

```

procedure SEQUENTIAL ( $\Gamma, L_0, N_0$ )
2:   for  $l \leftarrow 1, L_0$  do
      Generate scenario  $\xi_l$ 
4:   Evaluate the accrued value at  $H$  of the interim cash flows
      Estimate the closed form price for position  $p_0$ 
6:   Conditioned on the scenario  $\xi_l$  generate i.i.d inner samples
       $\hat{Z}_{l,1}, \dots, \hat{Z}_{l,N_0}$  of portfolio losses
      Compute an estimate of portfolio loss in scenario  $l$   $\hat{Y}_l = \frac{1}{N} \sum_{i=1}^{N_0} \hat{Z}_{l,i}$ 
8:   end for
      Compute the remaining Budget  $\Gamma_r = \Gamma - L_0(N_0\gamma_1 + \gamma_0)$ 
10:  while  $\Gamma > \gamma_1$  do
      Compute  $n_1^* = \operatorname{argmin}_i \left\{ \frac{k_i^2}{\sigma_i^2} \left( m + d - \hat{Y}_{\xi_i}(k_i) \right)^2 \right\} \forall i \in I_1$ 
12:  Compute  $n_2^* = \operatorname{argmin}_i \left\{ \frac{k_i^2}{\sigma_i^2} \left( m - d - \hat{Y}_{\xi_i}(k_i) \right)^2 \right\} \forall i \in I_2$ 
      Compute  $n_3^* = \operatorname{argmin}_i \left\{ \frac{k_i^2}{\sigma_i^2} \left( m - \hat{Y}_{\xi_i}(k_i) \right)^2 \right\} \forall i \in I_3$ 
14:   $i^* = \inf(n_1^*, n_2^*, n_3^*)$ 
      Generate an additional sample for the scenario  $\xi_{i^*}$  then compute the new
      loss estimate  $\hat{Y}_{i^*} = \frac{1}{N} \sum_{i=1}^{N_{i^*}+1} \hat{Z}_{l,i^*}$ 
16:   $\Gamma \leftarrow \Gamma - \gamma_1$ 
      end while
18:  Compute an estimate of the VaR,  $\hat{y}_\alpha = \hat{Y}_{[\alpha L]}$ 
end procedure

```

Perhaps, the closest nested simulation algorithm to ours is the one developed by Broadie et al., 2011 and Lee and Glynn, 2003. However, some differences should be noted between the two class of algorithms

1. First, their work is based on the evaluation of probabilities and distribution with special attention on the tail of distribution as it encloses the least expected scenarios. Ours, on the contrary, is for the evaluation of quantile-based risk measures such as Value-at-Risk which reflects extreme losses under normal market conditions. This is a measure commonly used among practitioners and must be reported according to regulation such as Basel II, Basel III and Solvency II for the internal model approach.
2. Second, our procedure overcomes a great difficulty in computing the Value-at-Risk is that the desired loss is the unknown to be estimated. In fact, classical sequential optimization techniques will seek to create a rule of thumb to allocate budget around a fixed threshold. For example, the threshold defining a big loss is used for estimating the probability of large losses. In our setting, the uncertainty of the intermediate estimator of the threshold has to be accounted for. Moreover, the algorithm is allocating budget sequentially and is given a loss estimate at every step. In practice, it is easy to keep track of the evolution of the estimator, and it is possible to stop the computation at any point if the time constraint is tightened during the simulation phase.
3. Finally, this algorithm is taking the bias due to the inner simulation into consideration, again by distinguishing between outer scenarios yielding a loss, whether within the uncertainty domain or not, in the intermediate simulation steps and trying to minimize its effect on budget allocation. Such uncertainty could lead to increasing drift on the estimator bias if ignored by myopic budget allocation. This issue may be encountered by any sequential simulation for quantile measuring.

Furthermore, the sequential algorithm requires a larger computational budget for execution compared to the crude uniform algorithm. Within each step of the sequential algorithm, an array of upper bounds has to be computed according to equations (2.18) and (2.19) which determine the upper bounds for the probabili-

ties of order changes. However, it is important to note that only the value of the loss estimate within the updated scenario has to be computed in the case of no changes in the VaR estimate. Moreover, the relative additional computing time will only represent a small fraction of the budget needed to price a large portfolio, which is usually required in financial institutions. This additional burden could be lightened by using the proper data structure (see for example, SSJ library in Java with data structure adopted for Monte Carlo simulations). Additionally, the sequential algorithm will require more memory than the uniform algorithm as it is necessary to keep track of the σ_i of every scenario as well as the upper bounds for the probabilities of changing order. In fact, the uniform estimator could be implemented in such a way that it will consider each scenario separately. In this particular case, we will only need to store the $\lceil \alpha * L \rceil$ greater losses and update them sequentially. As the intermediate simulations are discarded, then the memory consumed will become constant during the simulation and will only depend on the number of outer simulations and the loss probability. However, the sequential algorithm needs to keep track of all the loss in every scenario along with sum and the sum of squares of each inner scenario that will be compulsory to compute standard deviation. However, this flaw will not prevent the use of this algorithm for practitioners as the development of hardware technologies allows for the allocation of the needed amount of data. It may be important to note, that our procedure should be faster than the sequential algorithm proposed by Broadie et al., 2011 for estimating the probability of large losses. This gain in efficiency is derived from the fact that we are dividing the outer scenario into three subsets. Both algorithms, in some steps, will need to identify the minimum of an array of L numbers. However, ours will divide the set into three subsets I_i , $i = 1..3$ represented in three arrays with joint length $= L$. The division of the problem is known to reduce the computational burden in most cases. This classic setting is known as divide and conquer in algorithmic.

Budget allocation of the Sequential algorithm

In this section, we will provide a theoretical justification that the sequential algorithm cleverly distributes budget. For that purpose, we need to compute the probability of the event that we will call $\mathbb{A}_i$. $\mathbb{A}_i$ is that the budget for a new inner scenario

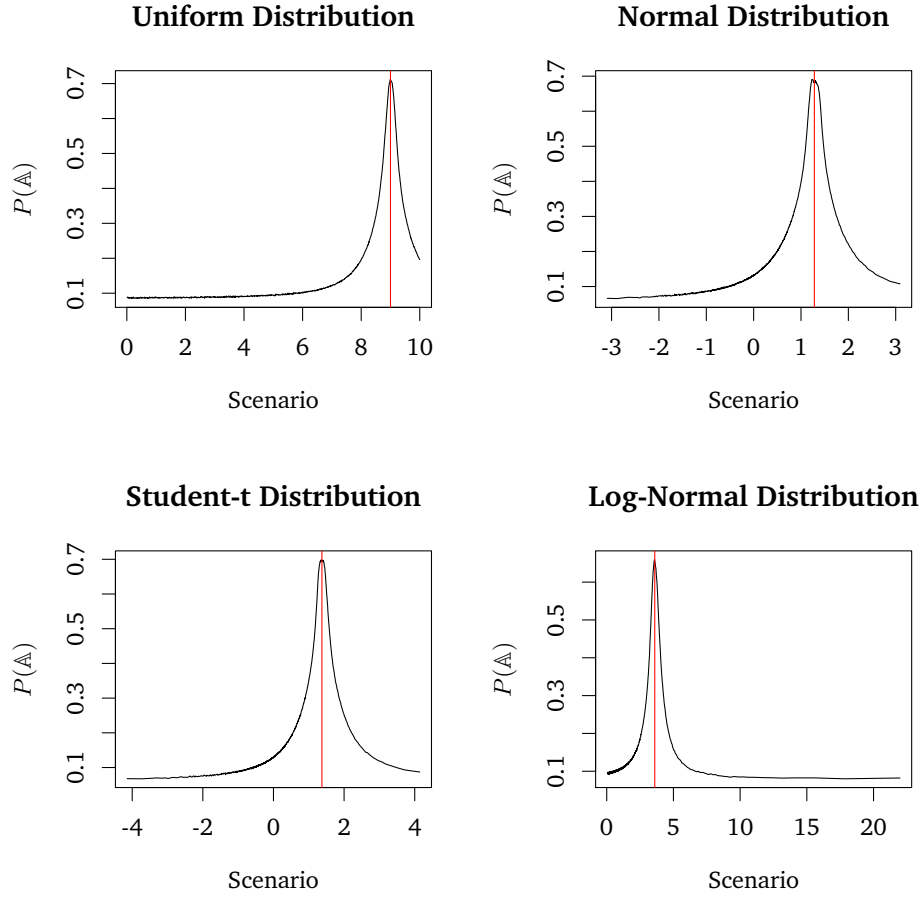


Fig. 2.4.: This figure shows the value of the upper bound of $P(\mathbb{A})$. The expected values of outer scenarios μ_i are drawn from a uniform, normal, student-t and log-normal distribution. k_i are from a Poisson distribution and $\sigma_i = 1 + \frac{a}{k_i}$. The figures are averaged over 1000 simulations. The vertical red line represent the theoretical value of m in each simulation.

is allocated to the inner simulation with the outer scenario i in the sequential algorithm. The value of $P(\mathbb{A}_i)$ can be seen as a proxy of the number of inner simulation in the outer simulation i

Theorem 1 Assuming that $\hat{Y}_i, \forall i$ is a normal distribution with mean μ_i and standard deviation σ_i and that m is the expected value of the estimator of the VaR we have that:

$$P(\mathbb{A}_i) \leq \prod_{\substack{j=0 \\ j \neq i}}^N \left(\frac{k_j^2}{\sigma_j^2} \frac{1+b_j}{N+1} \right)^{\frac{1}{N}} \alpha \left(\frac{\pi}{N\beta} \right)^{\frac{1}{2N}} \quad (2.24)$$

Figure (2.4) shows the values of an upper bound of $P(\mathbb{A})$ as evaluated by 2.24. The expected values of outer scenarios μ_i , $i \in 1..N$ are drawn from a uniform, normal, Student-t and log-normal distributions. k_i representing the number of inner scenarios in the outer simulation i is represented by a Poisson distribution. Finally, $\sigma_i = 1 + \frac{a}{k_i}$.⁵ This representation of σ_i should take into consideration the increase of precision with each additional inner scenario. The vertical red line represents the theoretical value of m in each simulation. The most important feature in all sub-figures of figure (2.4) is that the distribution of $P(\mathbb{A})$ seems to peak around the theoretical value of m . Remember that m is the value to estimate by the proposed algorithms. The sequential algorithm, as anticipated, gives more importance to scenarios that are around the theoretical value of the quantile. We do not claim that this argument represents a theoretical justification of the supremacy of the sequential algorithm over its uniform version. Nevertheless, we showed that the sequential algorithm provides what seems to be a smarter allocation of budget as it gives more importance to scenarios around the theoretical value. Taking into consideration that the bias is a function of the number of inner simulations, we expect the proposed algorithm to reduce it.

2.3.4 Stratified Sequential Simulation

The main objective of the sequential algorithm is to reduce the bias due to the inner level uncertainty by cleverly allocating budget between outer scenarios. A structural flaw in this algorithm is that the number of outer scenarios is exogenous and is not the object of an optimization process. Therefore, this important parameter must be chosen according to some empirical rule or benchmark.

Moreover, it is well known that for a fixed budget, the variance of the estimator grows when the number of inner simulations grows as the number of outer simulations drops for the uniform estimator. However, thanks to the allocation rule the sequential algorithm will keep the value of the bias at constant levels and may even cause it drop when increasing the number of outer simulations. This feature

⁵The parameters of the uniform, normal, Student-t and log-normal distribution were omitted because different experiments suggests that the results seem to be insensitive to the choice of those values

will allow us to increase the number of real-world scenarios with a limited trade-off between variance and bias.

To overcome this problem, a stratified procedure is proposed. Roughly, the idea is to divide the simulation into different steps or strata and spend the computational budget accordingly. We will begin by spending a fraction of the total budget on a uniform estimator. Then, we will execute the sequential estimator on a set of existing scenarios. This is repeated until we exploit all the potential in the generated scenarios. In other words, we will generate additional scenarios, if the computational budget remains, whenever the marginal change of the VaR estimator is inferior to a certain threshold. The intuition behind this idea is that the estimation error generated by the simulation will decay exponentially with the number of inner simulations. Hence, the marginal change of the estimator will decrease as inner simulations are added, and when a certain level is reached, it will become more interesting to add a new outer scenarios. Another reason behind the idea of the stratified sequential algorithm is that in a real situations financial institutions need to evaluate the level of risk undertaken by their managers using the VaR criteria. The institutions will be running the simulation just a single time and at best twice. As the outer scenarios are randomly generated, an unfortunate sampling can contain outer scenarios that are different from the scenario corresponding to the unbiased value of VaR. In that setting, no matter how significant is the budget allocated to the inner simulations, a certain level of inaccuracy will remain because of the unlucky drawing of the outer scenarios. In the situation described, an additional number of outer scenarios added to the sample is welcomed. The second algorithm that we propose is a remedy to that kind of situation as it will detect, according to a threshold, the level at which additional inner scenarios are no longer acquired and when a bigger sample of outer scenarios is required to improve the quality of the estimator.

The idea's main contribution is that the number of scenarios will be decided using the optimization process. The threshold is the only parameter which should be fixed according to the desired number of outer scenarios and could be fixed based on exploratory simulations.

To introduce the notion of strata, we need to add some key variables to our previous algorithm:

- γ : is the minimum relative variation of the estimator between two strata.
- p : is the number of scenarios that will be added if the relative variation of the estimator is below the threshold.
- S : is the number of inner simulations within each strata.

The proposed relative variation at strata i , δ_i is then:

$$\delta_i = \frac{V\hat{a}R(B) - V\hat{a}R(B - B_S)}{V\hat{a}R(B - B_S)} - \lambda * \delta_{i-1} \quad (2.25)$$

Where B_S , is the computational budget consumed at each strata. λ is a weighting coefficient that will make sure the history of the evolution of the estimator is monitored. In fact, depending on the length of the strata S , it is possible that the estimator might not be affected between two strata and consequently its relative evolution will be null. However, this will probably mean that we simply need more simulations to affect the estimator directly. λ should be a coefficient < 1 to guarantee that the weight of previous strata will decrease when advancing in the simulation.

Algorithm 3 VaR Sequential estimator sampling: marginal evolution

```
procedure SEQUENTIAL THRESHOLD ( $\Gamma, L_0, N_0, \gamma$ )  
   $I = \{i \in \mathbb{N}, 1 \leq i \leq L\}$   
3:  for  $l \in I$  do  
    Generate scenario  $\xi_l$   
    Compute an estimate of portfolio loss in scenario  $\xi_l$   $\hat{Y}_l = \frac{1}{N} \sum_{i=1}^{N_0} \hat{Z}_{l,i}$   
6:  end for  
    Compute the VaR estimate  $\hat{VaR}$  and  $V \leftarrow \hat{VaR}$   
    repeat  
9:    for  $i \leftarrow 0, S$  do  
      Find  $i^*$  that maximize  $\mathbb{P}(A)$   
      Compute an estimate of portfolio loss in scenario  $\xi_{i^*}$   
12:    end for  
      Compute the VaR estimate  $\hat{VaR}$  and  $V \leftarrow \hat{VaR}$   
      if  $\delta_i < \gamma$  then  
15:        Generate p additional real world scenarios  
      end if  
      Compute remaining Budget  $\Gamma$   
18:    until  $\Gamma < \Gamma_{Min}$   
  end procedure
```

The major flaw of these algorithms is the dependence between the estimators of loss in each scenario. This dependence is the consequence of the allocation rule, and it may lead to difficulties estimating the efficiency of the algorithm.

2.4 Numerical Results

In this section, we present the numerical results that shows the gains of using a non-uniform nested simulation.

2.4.1 Experimental Settings

Case 1 : Gaussian Portfolio

The experimental setting is set in the context of evaluating the MSE generated by each algorithm. Therefore the theoretical VaRs should be computable in order to enable the evaluation of the bias and a comparison of the accuracy of each algorithm. We will also use the same context of testing as Broadie et al., 2011 for benchmarking purposes.

The first set of tests consists of a Gaussian portfolio where both the inner and outer scenarios are generated from a normal distribution. More precisely we will consider a portfolio with initial value $X_0 = 0$ at time 0 and the future value $X_\tau(\omega) = \omega$ at risk horizon τ . The first assumption is that the real-valued risk factor ω is normally distributed with zero mean and a unit standard deviation ($\sigma^2 = 1$). Hence, the portfolio loss will be $Y(\omega) = X_0 - X_\tau(\omega) = -\omega$ is a standard normal variable. Conditioned to the scenario ω_i , each inner loss sample will take the following shape $\hat{Z}_{i,j} = -\omega_i + \sigma_{inner} W_{i,j}$ where $W_{i,j}$ is a standard normal random variable and where $\sigma_{inner} = 5$ is the standard deviation of the inner stage samples.

In this context of study and given a loss probability α , the Value at Risk VaR is given by: $VaR = \phi^{-1}(1 - \alpha)$ given that the pay-offs are normally distributed and that the mean of the inner distribution is 0. If we choose solvency probabilities equal to 10%, 1% and 0.1% the corresponding Value-at-Risk is 1.282, 2.326 and 3.090. It is important to note that the previous values are computed analytically and are accurate. This motivates our choice of examples. We need the exact value of the VaR to be able to compute the Bias and therefore assess the quality of each algorithm.

Case 2: European style put option

The second example consists of a portfolio of a single long position on a European put option. The difficulty of the example arises in the non-linearity of the portfolio cash flows along with their skewness that varies substantially with outer scenarios. The asset follows a geometric Brownian motion having an initial price $S_0 = 100$. Other parameters governing the portfolio cash-flows are as follows:

Drift under the real world distribution	$\mu = 8\%$
The annualized volatility	$\sigma = 20\%$
The risk-free rate	$r = 3\%$
The strike of the put option	$K = 95$
The maturity of the put option	$T = 0.25 \text{ years}$
The risk horizon	$\tau = 1/52 \text{ years}$

Using these settings the initial value of the portfolio could easily be given by the Black and Scholes formula : $X_0 = 1.669$ Concerning the simulation settings, let us consider $S_\tau(\omega)$ the underlying asset price at the risk horizon τ . This random variable is generated according to $S_\tau(\omega) \triangleq S_0 e^{(\mu - \sigma^2/2)\tau + \sigma\sqrt{\tau}\omega}$ and considers that ω is a standard normal variable.

Conditioned to the value of the real world scenario ω a second random variable relative to the inner simulation has to be generated to compute the value of cash-flows beyond the risk horizon. This random variable generated is:

$$S_T(\omega, W) \triangleq S_\tau(\omega) e^{(r - \sigma^2/2)(T - \tau) + \sigma\sqrt{T - \tau}W}$$

Where W is an independently distributed random variable. The portfolio loss will be defined by:

$$L(\omega) = X_0 - \mathbb{E}[e^{-r(T - \tau)} \max(K - S_T(\omega, W), 0) | \omega]$$

At the end, given an outer scenario ω_i the inner loss sample will have the shape:

$$\hat{Z}_{l,i} = X_0 - e^{-r(T-\tau)} \max(K - S_T(\omega, W_{l,i}), 0)$$

given that $W_{l,i}$ is an independent standard random variable. It is worth noting that the outer stage scenario is generated using the real-world distribution governed by the drift μ while the inner scenario used to evaluate the future price of the option conditioned to the outer scenario ω is the risk-neutral distribution subject to a drift r .

One of the reasons behind the choice of such examples is the possibility of computing the theoretical value of the VaR in order to assess the quality of each estimator by computing its MSE. In the particular case of the put option, it is possible to note that the loss is strictly increasing on the risk factor ω . The VaR at a probability level α could be computed regarding that $VaR = \inf\{y : \mathbb{P}(Y \leq y) \geq 1 - \alpha\}$, then we can easily establish that:

$$VaR = L(\omega^*) \mid \omega^* = \Phi^{-1}(1 - \alpha) \quad (2.26)$$

For more detail refer to Appendix B. For example, if we consider the loss probability 10%, 1% and 0.1% then the theoretical corresponding VaR will be 0.859 , 1.221 and 1.390 respectively.

Case 3: Multivariate Gaussian Portfolio

In the third example, we want to test the case of a the multivariate Gaussian portfolio. More precisely, the dimension of the scenarios governing the portfolio value is chosen to be equal to 25. $\Omega_i = (\omega_k)_{1 \leq k \leq 25}$ is the vector of outer scenario. The first assumption is to consider the portfolio initial value $X_0 = 0$ and that value of

the portfolio is the sum of the value of all risk factors at risk horizon τ . In other words:

$$X_\tau(\Omega_i) = \sum_{k=1}^{25} \omega_k \quad (2.27)$$

Where, ω_i is normally distributed with a unit standard deviation $\sigma_i^2 = 1$ for all $i = 1..25$.

Therefore conditioned on the scenario Ω_i , the portfolio loss will be $Y(0) = X_0 - X_\tau(\Omega_i)$. Finally, each inner loss will have the following shape:

$$\hat{Z}_{ij} = -X_\tau(\Omega_i) + \sigma_{inner} \sum_{k=1}^{25} W_{i,j,k} \quad (2.28)$$

Where $W_{i,j} = (w_{i,j,k})_{1 \leq k \leq 25}$ is a vector of 25 normal random distribution and σ_{inner} is the standard deviation of each component of the vector $W_{i,j}$.

Again, it is possible to compute the analytical VaR of this setting as the sum of uncorrelated normal random variables follows a normal distribution. The importance of this test is to evaluate the performance of the sequential algorithm and stratified sequential algorithm for multidimensional risk factors with uncorrelated components. If we choose the solvency probabilities to be equal to 10 %, 1 % and 0.1 % the corresponding VaR would be 6.407, 11.631 and 15.451.

Case 4: Basket option

The last test example will be a portfolio with a single long position on a European style basket option of two highly correlated assets. This problem combines the difficulties of non-linearities of pay-offs plus the multidimensionality of risk factors. Both assets follow a geometric Brownian motion and have the initial price $S_0 = 100$. To simplify the problem, the two assets will have the same parameters of evolution.

The other parameters governing the portfolio value evolution for the two assets are :

Drift under the real world distribution	$\mu = 0\%$
The annualized volatility	$\sigma = 30\%$
The risk-free rate	$r = 4\%$
The strike of the put option	$K = 200$
The maturity of the put option	$T = 0.25years$
The risk horizon	$\tau = 1/52years$
Corrolation between assets	$\tau = 90\%$

Taking into consideration those settings, the initial value of the portfolio is $X_0 = 19.375$. Concerning the nested simulation to evaluate the VaR, let us consider $S_\tau^i(\omega_i)$ the underlying price of the asset i at the risk horizon τ . This random variable is generated according to the rule $S_\tau^i(\omega_i) \triangleq S_0 e^{(\mu - \sigma^2/2)\tau + \sigma\sqrt{\tau}\omega_i}$ and that by taking ω_i also a standard normal random variable. It also important to stress that the two variables ω_1 and ω_2 are two correlated random variables and that the correlation between the two variable is 90%. Again, conditioned to the value of the real world scenario ω_i , the value of the asset i at the horizon T after drawing the inner scenario W_i from a normal distribution would be:

$$S_T^i(\omega_i, W_i) \triangleq S_\tau^i(\omega_i) e^{(r - \sigma^2/2)(T - \tau) + \sigma\sqrt{T - \tau}W_i}$$

At the end, the portfolio loss will be :

$$L(\omega) = X_0 - \mathbb{E}[e^{-r(T - \tau)} \max(K - (S_T^1(\omega_1, W_1) + S_T^2(\omega_2, W_2)), 0), 0) | \omega]$$

Again, it is possible to compute the VaR of the portfolio analytically. In order to do so, we will try to find the equivalence between the case of a classic put option and a European-style basket option. A moment matching technique will be used to obtain an equivalent distribution of the sum of two log normal distributions.

I.e., the idea is to obtain a distribution $S(\omega)$ that has the same first and second order moments, at the given horizon t , as the sum of two distributions $S^1(\omega)$ and $S^2(\omega)$. By keeping in mind that, $\mathbb{E}[S^1(\omega) + S^2(\omega)] = \mathbb{E}[S^1(\omega)] + \mathbb{E}[S^2(\omega)]$ and that $\mathbb{V}[S^1(\omega) + S^2(\omega)] = \mathbb{V}[S^1(\omega)] + \mathbb{V}[S^2(\omega)] + 2\text{Cov}[S^1(\omega), S^2(\omega)]$, we will try to solve:

$$\begin{cases} \mathbb{E}[S(\omega)] = \mathbb{E}[S^1(\omega) + S^2(\omega)] \\ \mathbb{V}[S(\omega)] = \mathbb{V}[S^1(\omega) + S^2(\omega)] \end{cases} \quad (2.29)$$

Finally the computed moment will be⁶:

$$\begin{cases} \mathbb{E}[S(\omega)] = \mu - \frac{1}{2}\sigma_S^2 \\ \sigma_S^2 = \frac{1}{t} \ln \left(\frac{S_1(0)^2 e^{\sigma_1^2 t} + S_2(0)^2 e^{\sigma_2^2 t} + 2S_1(0)S_2(0)e^{\rho\sigma_1\sigma_2 t}}{(S_1(0) + S_2(0))^2} \right) \end{cases} \quad (2.30)$$

It is important to note that, the equivalent log-normal distribution described by equations 2.30 is not a risk-neutral process in the sense that the moment matching is valid only at time t . For the sake of computing the VaR of the portfolio, the same matching should be performed in the horizon τ and T in order to reprice the option at each time horizon. This technique appeared to yield excellent results for positivity correlated assets and those with very close volatilities which are exactly the setting of our study case.

In the end, after matching the two processes into one equivalent process, we have transformed the basket option into a classic European put option. Therefore, we can apply the same methodology to compute the VaR as described in the case of a put option. If we consider the solvency level to be equal to 10 %, 1 % and 0.1 % the corresponding VaR would be 3.942, 6.693 and 8.483.

Variance Estimator

Both the stratified sequential algorithm and the sequential algorithm require the estimation of the value of σ_i the standard deviation of the inner samples $\hat{Z}_{i,1}, \hat{Z}_{2,1}, \dots$

⁶The proof is given in Henriksen, 2008

However, in the case of put options and basket options, in the right tail of the loss distribution where the options has little chance to be exercised and by consequence the loss will be the initial value of the option, it is hard to estimate the standard deviation with a small number of inner scenarios. In fact, when most of the inner scenarios lead to non-exercise of the option the standard deviation will be almost null. In order to increase the reliability of the estimator of the standard deviation, we will introduce a second estimator which tries to find the balance between an ensemble estimate and a local estimate. This variation will be:

$$\tilde{\sigma} = \frac{N_i}{N_i + b} \hat{\sigma}_i + \frac{b}{N_i + b} \hat{\sigma} \quad (2.31)$$

And we define:

$$\hat{\sigma}_i = \left[\frac{1}{N_i - 1} \sum_{j=1}^{N_i} (Z_{i,j} - \hat{L}_i)^2 \right]^{\frac{1}{2}}$$

and $\sigma = \sum_{i=1}^L \sigma_i$.

For $b = 0$, no weight is given to the estimate which corresponds to the usual standard deviation estimator. However, in our experimental setting, we used $b = 5$.

2.4.2 Simulation Results

For the purpose of bias comparison simulation, we generated 5000 outer stage scenarios. Then, for the uniform estimator, we varied the inner stage simulation numbers from 20 to 200. Whereas, for both the sequential and marginal algorithms the initial simulation had a fixed budget of 100000. Then we numerically computed the bias. The results are averaged over 1000 trials to reduce the effect of randomness.

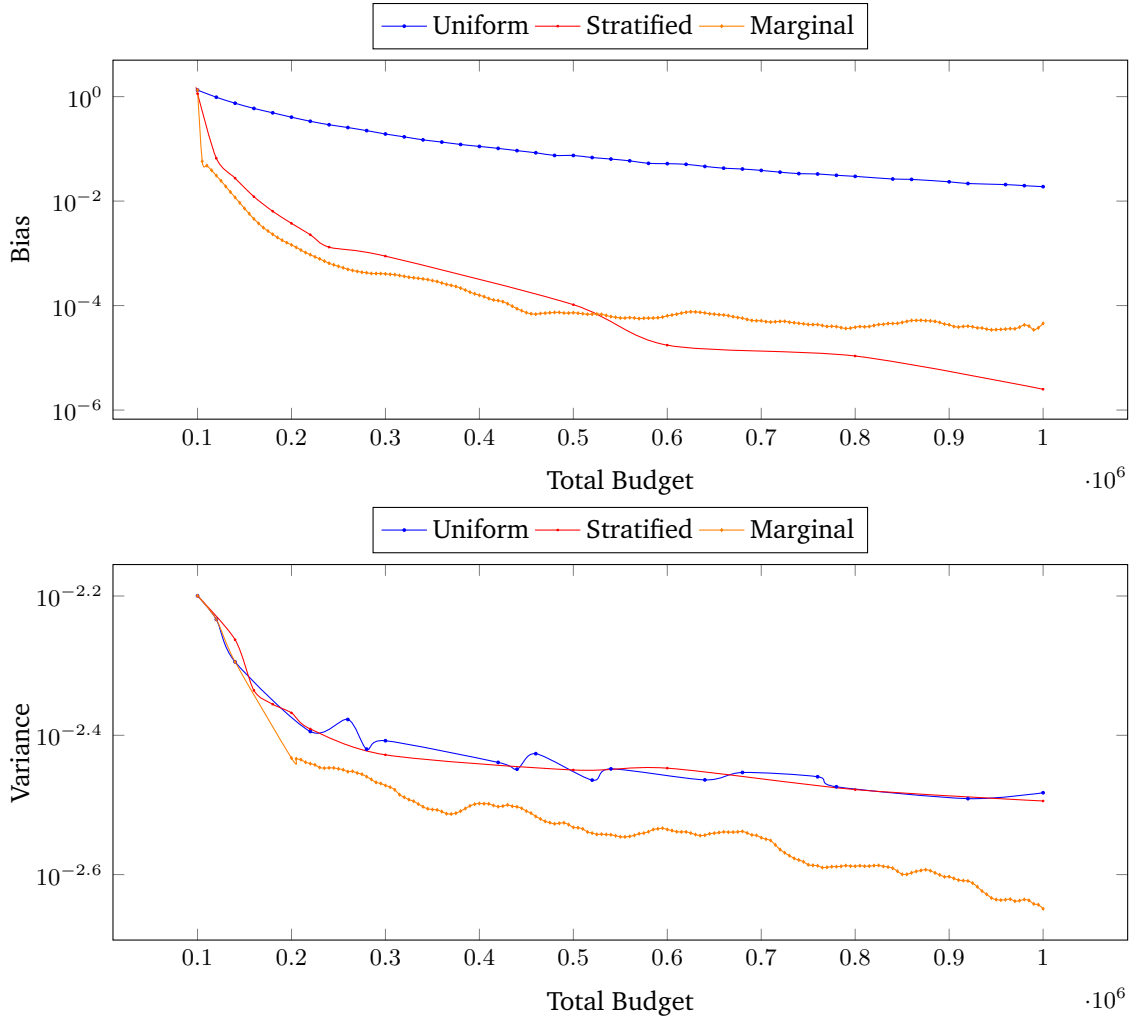


Fig. 2.5.: Bias and variance of the uniform, the sequential and the marginal estimator as a function of the number of inner scenarios in the example of one dimensional Gaussian portfolio. The test is conducted for 5000 outer scenarios and for an initial budget equal 100,000 for both sequential and marginal simulation. The bias is averaged over 1000 trails

The benefits of the marginal and stratified algorithms are visible in Figure 2.5. In fact, the bias generated by the uniform algorithm is higher over the whole testing range. A second important finding is that the slope of the non-uniform algorithm is greater than the slope of the uniform algorithm. Despite the fact that both slopes

seem to get closer when the budget increases, the initial difference will generate an advantage for the non-uniform algorithm that will be cumulated. Second, this difference of slope will be of great importance as we can achieve a much higher level of accuracy by using a small fraction of additional budget

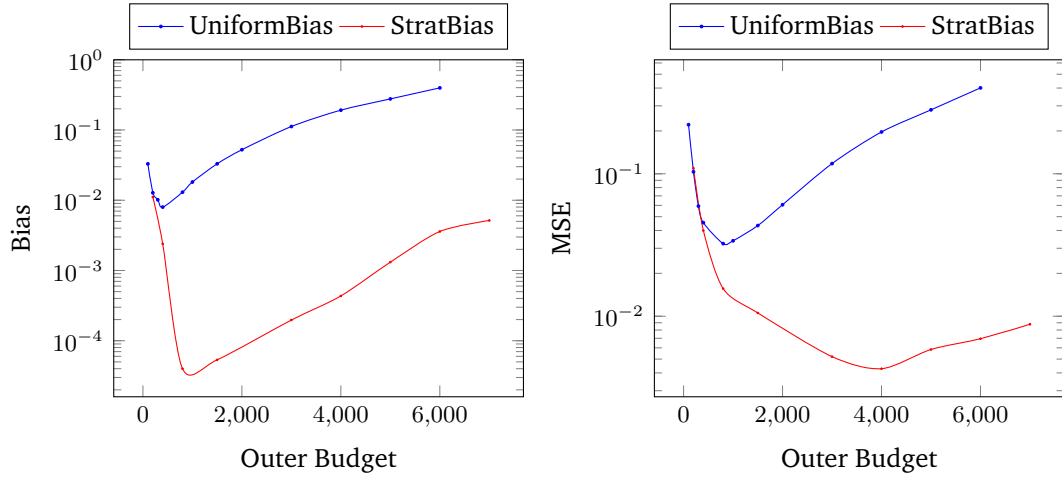


Fig. 2.6.: Bias and MSE of the uniform, the stratified estimator as a function of the number of outer scenarios. The test is conducted for a range of outer scenarios for a total budget of 250 000 and an initial budget equal 100,000 for the sequential. The variance is not represented as both algorithm yield the same variance and have the same evolution of variance. The MSE is only driven by the variation of the Bias.aps The bias and MSE is averaged over 1000 trails

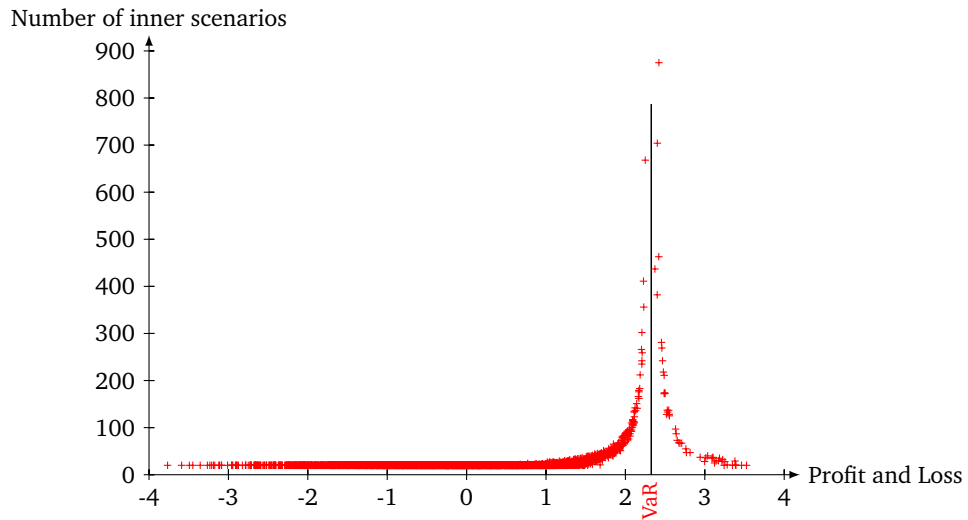


Fig. 2.7.: Distribution of the number of inner stage samples over the different levels of risk for the sequential algorithm. The figure illustrates how the budget is allocated around the actual value of the VaR. The overall budget used is 1000000. The uniform estimator is using a constant number of 200 inner scenario.

Figure (2.7) gives a view of how the stratified sampling algorithm allocates the resources between scenarios. Here the results are always averaged over 1000 experiments. We note that the closer we are to the VaR scenario the greater the

amount of the budget is allocated to the inner scenarios. We also note that the minimum budget was allocated to the left tail of the distribution where extreme losses are highly improbable. This figure features the same patterns that were shown in figure 2.4 which shows that the budget distribution in the simulations matches the expected allocation evaluated theoretically.

Tab. 2.1.: This table shows the results of Bias, Variance and MSE for 6 test cases: one dimensional Gaussian, Single put option for different level of probabilities. Test where performed for an overall budget of 1000000. The column Normalized illustrates the l_{n2} of the uniform budget required to reach the same level of bias and variance divided by the budget used.

		<i>variance</i>	<i>Bias</i> ²	<i>MSE</i>	Normalized
Gaussian	$\alpha = 10\%$	Uniform	$1,51 \cdot 10^{-2}$	$2,89 \cdot 10^{-2}$	$4,40 \cdot 10^{-2}$
		Sequential	$2,34 \cdot 10^{-2}$	$2,22 \cdot 10^{-4}$	$2,36 \cdot 10^{-2}$
		Sequential Sequential	$4,76 \cdot 10^{-4}$	$1,40 \cdot 10^{-4}$	$6,16 \cdot 10^{-4}$
Gaussian	$\alpha = 1\%$	Uniform	$3,29 \cdot 10^{-3}$	$1,89 \cdot 10^{-2}$	$2,22 \cdot 10^{-2}$
		Sequential	$3,20 \cdot 10^{-3}$	$2,50 \cdot 10^{-6}$	$3,21 \cdot 10^{-3}$
		Sequential Sequential	$2,24 \cdot 10^{-3}$	$4,56 \cdot 10^{-5}$	$2,29 \cdot 10^{-3}$
Gaussian	$\alpha = 0,1\%$	Uniform	$1,45 \cdot 10^{-2}$	$2,03 \cdot 10^{-3}$	$1,66 \cdot 10^{-2}$
		Sequential	$1,73 \cdot 10^{-2}$	$4,79 \cdot 10^{-4}$	$1,78 \cdot 10^{-2}$
		Stratified	$5,38 \cdot 10^{-3}$	$1,68 \cdot 10^{-3}$	$7,06 \cdot 10^{-3}$
Put Option	$\alpha = 10\%$	Uniform	$2,58 \cdot 10^{-4}$	$2,03 \cdot 10^{-3}$	$2,29 \cdot 10^{-3}$
		Sequential	$1,47 \cdot 10^{-4}$	$1,51 \cdot 10^{-6}$	$1,49 \cdot 10^{-4}$
		Stratified	$1,24 \cdot 10^{-4}$	$1,00 \cdot 10^{-3}$	$1,12 \cdot 10^{-3}$
Put Option	$\alpha = 1\%$	Uniform	$2,11 \cdot 10^{-4}$	$4,05 \cdot 10^{-3}$	$4,27 \cdot 10^{-3}$
		Sequential	$2,63 \cdot 10^{-4}$	$2,74 \cdot 10^{-6}$	$2,66 \cdot 10^{-4}$
		Stratified	$1,32 \cdot 10^{-4}$	$4,41 \cdot 10^{-7}$	$1,32 \cdot 10^{-4}$
Put Option	$\alpha = 0,1\%$	Uniform	$5,37 \cdot 10^{-4}$	$4,02 \cdot 10^{-3}$	$4,56 \cdot 10^{-3}$
		Sequential	$5,48 \cdot 10^{-4}$	$5,10 \cdot 10^{-5}$	$5,99 \cdot 10^{-4}$
		Stratified	$1,61 \cdot 10^{-4}$	$6,61 \cdot 10^{-6}$	$1,67 \cdot 10^{-4}$

Tab. 2.2.: This table shows the results of Bias, Variance and MSE for 6 test cases: Multi dimensional Gaussian and Basket option for different level of probabilities. Test where performed for an overall budget of 1000000. The column Normalized illustrates the ln_2 of the uniform budget required to reach the same level of bias and variance divided by the budget used.

		<i>variance</i>	<i>Bias</i> ²	<i>MSE</i>	Normalized
Multivariate Gaussian	$\alpha = 10\%$	Uniform	$1,38 \cdot 10^{-2}$	$1,45 \cdot 10^{-1}$	$1,59 \cdot 10^{-1}$
		Sequential	$1,83 \cdot 10^{-2}$	$2,90 \cdot 10^{-5}$	$1,83 \cdot 10^{-2}$
		Stratified	$1,47 \cdot 10^{-2}$	$1,55 \cdot 10^{-3}$	$1,62 \cdot 10^{-2}$
Multivariate Gaussian	$\alpha = 1\%$	Uniform	$7,49 \cdot 10^{-2}$	$4,45 \cdot 10^{-1}$	$5,20 \cdot 10^{-1}$
		Sequential	$6,64 \cdot 10^{-2}$	$4,47 \cdot 10^{-6}$	$6,64 \cdot 10^{-2}$
		Stratified	$5,15 \cdot 10^{-2}$	$1,61 \cdot 10^{-5}$	$5,16 \cdot 10^{-2}$
Multivariate Gaussian	$\alpha = 0,1\%$	Uniform	$4,61 \cdot 10^{-1}$	$6,62 \cdot 10^{-1}$	$1,12$
		Sequential	$4,41 \cdot 10^{-1}$	$6,00 \cdot 10^{-3}$	$4,47 \cdot 10^{-1}$
		Stratified	$1,26 \cdot 10^{-1}$	$6,83 \cdot 10^{-5}$	$6,83 \cdot 10^{-5}$
Basket Option	$\alpha = 10\%$	Uniform	$2,72 \cdot 10^{-2}$	$1,07$	$1,10$
		Sequential	$1,15 \cdot 10^{-2}$	$1,78 \cdot 10^{-3}$	$1,33 \cdot 10^{-2}$
		Stratified	$1,23 \cdot 10^{-2}$	$6,96 \cdot 10^{-3}$	$6,96 \cdot 10^{-3}$
Basket Option	$\alpha = 1\%$	Uniform	$2,28 \cdot 10^{-2}$	$1,05$	$1,08$
		Sequential	$2,18 \cdot 10^{-2}$	$4,11 \cdot 10^{-4}$	$2,22 \cdot 10^{-2}$
		Stratified	$1,68 \cdot 10^{-2}$	$3,50 \cdot 10^{-4}$	$1,72 \cdot 10^{-2}$
Basket Option	$\alpha = 0,1\%$	Uniform	$9,77 \cdot 10^{-2}$	$1,48$	$1,58$
		Sequential	$1,50 \cdot 10^{-2}$	$1,33 \cdot 10^{-3}$	$1,63 \cdot 10^{-2}$
		Stratified	$8,77 \cdot 10^{-3}$	$6,18 \cdot 10^{-3}$	$1,50 \cdot 10^{-2}$

Tests for the stratified sequential algorithm were performed using $\gamma = 0.005$, $S = 150$ and $\lambda = 0.5$.

2.4.3 Algorithm Comparison

Uniform algorithm The uniform sampling is the simplest algorithm and the easiest to implement. In addition, it has the advantage of consuming a low relative amount of hardware memory compared to the other algorithm. Gordy and Juneja, 2010 demonstrated that is possible to have an optimal choice of inner and outer scenarios for the uniform estimator. However, it is not clear how to compute the optimum value in practice. We note that a small shift away from the optimal values will have significant impact on the performance of the estimator. Also, as the budget is divided equally between the outer scenarios, the computational efforts needed to obtain an acceptable level of bias is significant.

Sequential estimator The sequential estimator varies from the uniform estimator in the rule of inner scenario's allocation. In fact, the number of inner scenarios is not equal between outer scenarios, and much importance is attributed to outer scenarios that may have a bigger impact on the VaR estimator. The main comparative advantage of this estimator is the substantial reduction of bias relative to the uniform estimator. The variance that essentially depends on the number of outer scenarios will not be changed. However, we can also see that an improvement of bias is notable even for a small amount of additional budget allocated to the sequential step, and the same performance could be reached using a considerable budget for the uniform estimator. Moreover, we have noted that, for the sequential estimator, there seems to be an optimal number of outer scenarios to be used and at the same time the cost of a shift from optimality is not as significant as the uniform estimator which may give greater flexibility for a practitioner and lower the risk of wasting the budget. A more practical advantage of the sequential estimator is that it is possible to have intermediate results at every step of the simulation. Hence, a practitioner could have a good approximation of the risk being undertaken while waiting for more accurate results, and this could be of great importance for risk management purposes. Nevertheless, this algorithm presents some drawbacks compared with the uniform one. More precisely, the memory consumed and the additional compu-

tation burden compared to the uniform algorithm need to be considered but not to the extent that it prevents the implementation of this approach.

Marginal estimator The marginal estimator is based on the sequential estimator. The sequential estimator does not have much impact on variance and is only designed to have an impact on bias. The marginal estimator tries to reduce the variance by holding the same level of bias. Therefore, the number of outer scenarios is decided using the criteria of the marginal evolution of the estimator from different steps. The main advantage of this idea is that it is possible to reduce the variance without having to go through the bias-variance trade-off. As this estimator will have more outer scenarios by construction, the memory needed to execute it will grow whenever new scenarios are added.

Optimality of both the sequential and the marginal estimator The numerical experiments show that the sequential and the marginal estimator are superior to the uniform algorithm in those simple settings. The theoretical justifications in theorem 1 also highlighted that the computational budget is distributed in a way we think is more efficient. However, it is important to note that we do not claim the absolute optimality of the algorithms proposed in this chapter. Gordy and Juneja, 2010 computed the optimal allocation of budget but their results remain inapplicable in practice. It remains important for the literature on nested simulations to establish the optimal allocation without the need of explicitly knowing the probability distribution of losses. To obtain such results we need to overcome number of difficulties mainly the interdependence between scenarios in any allocation rule. Meanwhile, this chapter proposes applicable heuristics to compute VaR in the same spirit as Broadie et al., 2011.

Conclusion

In this paper, we investigated the possibilities of improving the performance of the crude Monte Carlo estimator for measuring quantile-based risk measures such as Value-at-Risk. We developed a sequential algorithm based on the idea that an additional inner scenario will not have the same level of impact on the final estimator

conditioned on the risk factor. We created a rule of thumb based on the evaluation of an upper bound of the probability that adding a single scenario will change the order of losses estimated. Moreover, our setting takes into consideration the uncertainty overlaid in the inner simulation in the computation of the probability of order change and the selection of the candidate outer scenario to receive an additional budget. Experimental settings set-out in the last section confirm the comparative advantage of the sequential simulation over the crude uniform estimator. This paper is unique because the algorithms proposed are tested in the delicate cases of multidimensionality and highly correlated risk factors.

However, this research does not cover the following gaps:

1. Theoretical developments need to be established to demonstrate the supremacy of the marginal estimator over the uniform estimator. Theoretical bias and variance characterization similar to that of Gordy and Juneja, 2010 computed for the uniform estimator has to be carried out for the case of the sequential estimator.
2. This work should be retested in the context of heavy-tailed risk estimators as some work suggests that sequential estimators may lose their competitiveness when dealing with such loss distributions

A. Proof of Equations 2.18 and 2.19

It is known that : $|\tilde{Y}_{\xi_i}(k_i+1)-m| \geq d$ is equivalent to $\tilde{Y}_{\xi_i}(k_i+1)-m \geq d$ Or $\tilde{Y}_{\xi_i}(k_i+1)-m \leq -d$

If we were to perform the additional sample in a given scenario ξ_i , this would result in a new loss estimate given by :

$$\tilde{Y}_{\xi_i}(k_i+1) = \frac{1}{k_i+1} \sum_{j=1}^{k_i+1} \hat{Z}_{i,j} = \frac{1}{k_i+1} \hat{Z}_{i,k+1} + \frac{k_i}{k_i+1} \tilde{Y}_{\xi_i}(k_i)$$

Recall that $\hat{Z}_{i,j}$ is the value of the simulated inner scenario j corresponding to the outer scenario i . By consequence : $|\tilde{Y}_{\xi_i}(k_i + 1) - m| \geq d$ is equivalent to

$$\begin{cases} \frac{1}{k_i+1} \hat{Z}_{i,k+1} + \frac{k_i}{k_i+1} \tilde{Y}_{\xi_i}(k_i) - m \geq d \\ \frac{1}{k_i+1} \hat{Z}_{i,k+1} + \frac{k_i}{k_i+1} \tilde{Y}_{\xi_i}(k_i) - m \leq -d \end{cases}$$

$\Rightarrow$

$$\begin{cases} \hat{Z}_{i,k+1} - \mu \geq k_i(m + d - \tilde{Y}_{\xi_i}(k_i)) + (m + d - \mu) \\ -\hat{Z}_{i,k+1} + \mu \leq k_i(m - d - \tilde{Y}_{\xi_i}(k_i)) + (m - d + \mu) \end{cases}$$

Using the approximation :

$$\begin{cases} k_i(m + d - \tilde{Y}_{\xi_i}(k_i)) + (m + d - \mu) \approx k_i(m + d - \tilde{Y}_{\xi_i}(k_i)) \\ k_i(m - d - \tilde{Y}_{\xi_i}(k_i)) + (m - d + \mu) \approx k_i(m - d - \tilde{Y}_{\xi_i}(k_i)) \end{cases}$$

This approximation is acceptable as for the purpose of simulations usually $k_i \gg 1$

Finally we can establish using the one sided Chebyshev inequality, and by denoting $\sigma_i = \text{Var}[\hat{Z}_{i,k}]$, that :

$$\begin{aligned} \mathbb{P}(A_1 \mid \tilde{Y}_{\xi_i}(k_i) \leq m) &\leq \left(1 + \frac{k_i^2}{\sigma_i^2} (m + d - \tilde{Y}_{\xi_i}(k_i))^2\right)^{-1} \\ \text{Or} \\ \mathbb{P}(A_1 \mid \tilde{Y}_{\xi_i}(k_i) \geq m) &\leq \left(1 + \frac{k_i^2}{\sigma_i^2} (m - d - \tilde{Y}_{\xi_i}(k_i))^2\right)^{-1} \end{aligned}$$

We can observe that the upper bound depends on the relative position of $\tilde{Y}_{\xi_i}(k_i)$ and m on the real axis.

Regarding equation (2.19) we should notice that :

$$(\tilde{Y}^B(\xi_k) - m) (\tilde{Y}^{B+1}(\xi_k) - m) \leq 0$$

Could be rewritten as

$$\begin{cases} \text{if } \tilde{Y}^B(\xi_k) \geq m \Rightarrow \tilde{Y}^{B+1}(\xi_k) \leq m \\ \text{Or} \\ \text{if } \tilde{Y}^B(\xi_k) \leq m \Rightarrow \tilde{Y}^{B+1}(\xi_k) \geq m \end{cases}$$

The using similar development and the approximation $k_i(m - \tilde{Y}_{\xi_i}(k_i)) + (m + d - \mu) \approx k_i(m - \tilde{Y}_{\xi_i}(k_i))$ we can establish equation (2.19)

B. Proof of equation (2.26)

$$y_\alpha = VaR_\alpha[Y] = \inf\{y : P(Y \leq y) \geq 1 - \alpha\}$$

Considering that the loss function is strictly increasing on the risk factor then, $\forall y$ within the loss domain $\exists \omega^*$ satisfying that $P(Y \leq y) = P(\omega \leq \omega^*) = \Phi(\omega^*)$ where Φ is the Normal cumulative distribution function

$$\Rightarrow VaR = \inf\{y : P(\omega \leq \omega^*) = \Phi(\omega^*) \geq 1 - \alpha\}$$

$$\Rightarrow VaR = \inf\{y = L(\omega^*) : \Phi(\omega^*) \geq 1 - \alpha\}$$

Where $L(\omega)$ is the loss of the portfolio for the risk factor ω .

$$\Rightarrow VaR = L(\omega^*) \text{ where } \omega^* = \Phi^{-1}(1 - \alpha)$$

C. Sensitivity based optimization algorithm

Hong (2009) and Hong and Liu (2008) established methods to compute the sensibility of quantiles related to a given risk factor α_i . We will use their estimator to imagine an optimization algorithm. The idea is that we will use the sensitivities to divide the set of real world scenarios into different subsets. Within the same subset, we will suppose that scenario will have roughly the same loss and hence it is more efficient to tribute equivalent budget between these subsets. Within each subset, we will use one of the non-uniform allocation algorithms specified in the previous section.

Let us define the set I_ω by :

$$I_\omega = \{\omega_s \in \Omega : \sum_{i=1}^{dim(\Omega)} (q'_{\alpha_i})^2 |\omega_s e_i - \omega e_i|^2 \leq r\} \quad (2.32)$$

Since the set of outer scenario is countable, It is possible to find a subset of scenarios I_g that we will call the set of generators that verify: $\forall$ scenarios ω , $\exists w_g \in I_g$ where $\omega \in I_{w_g}$.

The algorithm proposed is then :

Algorithm 4 VaR Sensitivity based estimator sampling

procedure SENSITIVITY(L , Total budget Γ)
 Generate L outer scenarios $\omega_i, i \in [1..L]$
 Create a set I of generating family of all outer scenarios
 Allocate budget $B = \frac{\Gamma}{\text{card}[I]}$ for each set $I_\omega, \omega \in I$
5: Use the optimal sampling algorithm for each set of scenarios I_ω
 Compute the estimate of the VaR
end procedure

This algorithm is proposed as an extension to this work. The result were implemented and presented in the ICMCM in 2013.

A. Proof of Equations 2.24

The aim of this section is to compute the probability of the event that we will call $\mathbb{A}$. $\mathbb{A}$ is that the budget for a new inner scenario is allocated to the inner simulation with the outer scenario i in the sequential algorithm. In other words, according to the specification of the sequential algorithm the realization of $\mathbb{A}$ is equivalent to:

$$i = \underset{j=1..N}{\operatorname{argmin}} \frac{k_j}{\sigma_j} |\mu - \hat{Y}_j| \quad (2.33)$$

We will assume that $\hat{Y}_i \forall i$ is a normal distribution with mean μ_i and standard deviation σ_i . Although we are aware that μ is the value to estimate but for the purpose of this proof, we will also assume that μ is a known constant. Then we have that $(m - \hat{Y}_i)^2$ is a non central χ^2 distribution $\chi'^2(\lambda_i, 1)$. Where $\lambda_i = \left(\frac{\mu_i - m}{\sigma_i}\right)^2$.

Using those notations we have that:

$$P(\mathbb{A}) = \int_0^{+\infty} f_i(x) \left(\prod_{\substack{j=0 \\ j \neq i}}^N (1 - F_j(x)) \right) dx \quad (2.34)$$

Where f_i is the PDF of the i^{th} order statistics and F_j is the CDF of the j^{th} order statistics.

To evaluate $P(\mathbb{A})$, we use the first approximation:

$$P(\chi'^2|\nu, \lambda) \approx P\left(\frac{\chi^2}{1+b}|\nu^*\right) \quad (2.35)$$

Where $\nu^* = \frac{a}{1+b}$ given that $a = \nu + \lambda$ and $b = \frac{\lambda}{\nu + \lambda}$

$$\nu^* = \frac{a}{1+b} = \frac{1 + \left(\frac{\mu_i - m}{\sigma_i}\right)^2}{1 + \frac{\left(\frac{\mu_i - m}{\sigma_i}\right)^2}{1 + \left(\frac{\mu_i - m}{\sigma_i}\right)^2}} = \frac{\left(1 + \left(\frac{\mu_i - m}{\sigma_i}\right)^2\right)^2}{1 + 2\left(\frac{\mu_i - m}{\sigma_i}\right)^2} \quad (2.36)$$

Using a first order Taylor approximation of the numerator we have that:

$$\nu^* \approx \frac{1 + 2\left(\frac{\mu_i - m}{\sigma_i}\right)^2}{1 + 2\left(\frac{\mu_i - m}{\sigma_i}\right)^2} = 1 \quad (2.37)$$

With $\nu^* = 1$ we can have that:

$$P(\mathbb{A}) = \int_0^{+\infty} \frac{1}{\sqrt{2\pi}} \left(\frac{X}{1+b_i} \frac{\sigma_i^2}{k_i^2}\right)^{-1/2} e^{-\frac{X}{1+b_i} \frac{\sigma_i^2}{2k_i^2}} \prod_{\substack{j=0 \\ j \neq i}}^N \operatorname{erfc}\left(\frac{\sqrt{2}}{2} \frac{\sigma_j}{k_j} \sqrt{\left(\frac{x}{1+b_j}\right)}\right) dx \quad (2.38)$$

We rewrite $P(\mathbb{A})$ in the form:

$$\begin{aligned} P(\mathbb{A}) = & \int_0^{+\infty} \frac{\sqrt{(2\pi)^{\frac{N-1}{2N}}}}{\sqrt{2\pi}} X^{\frac{N-1}{2N}} X^{-\frac{1}{2}} \left(\frac{1}{1+b_i} \frac{\sigma_i^2}{k_i^2}\right)^{-\frac{1}{2}} \\ & \left[\prod_{\substack{j=0 \\ j \neq i}}^N \left(\frac{1}{1+b_j} \frac{\sigma_j^2}{k_j^2}\right)^{\frac{1}{2N}} \right] e^{-x \left[\frac{\sigma_i^2}{2k_i^2} \frac{1}{1+b_i} - \frac{1}{N} \sum_{\substack{j=0 \\ j \neq i}}^N \frac{-\sigma_j^2}{2k_j^2} \frac{1}{1+b_j} \right]} \\ & \prod_{\substack{j=0 \\ j \neq i}}^N \left[\frac{1}{\sqrt{2\pi}} \left(\frac{X}{1+b_j} \frac{\sigma_j^2}{k_j^2}\right)^{-1/2} e^{-\frac{X}{1+b_j} \frac{\sigma_j^2}{2k_j^2}} \prod_{\substack{j=0 \\ j \neq i}}^N \operatorname{erfc}\left(\frac{\sqrt{2}}{2} \frac{\sigma_j}{k_j} \sqrt{\left(\frac{x}{1+b_j}\right)}\right) \right] dx \quad (2.39) \end{aligned}$$

We will use Holder inequality to compute an upper bound of $P(\mathbb{A})$

$$P(\mathbb{A}) \leq \frac{\alpha}{\sqrt{\pi}} \left[\int_0^{+\infty} x^{-\frac{1}{2}} e^{-x \frac{\beta}{N}} dx \right]^{1/N}$$

$$\prod_{\substack{j=0 \\ j \neq i}}^N \int_0^{+\infty} \left[\frac{1}{\sqrt{2\pi}} \left(\frac{X}{1+b_j} \frac{\sigma_j^2}{k_j^2} \right)^{-1/2} e^{-\frac{x}{1+b_j} \frac{\sigma_j^2}{2k_j^2}} \prod_{\substack{j=0 \\ j \neq i}}^N \operatorname{erfc}^N \left(\frac{\sqrt{2}}{2} \frac{\sigma_j}{k_j} \sqrt{\frac{x}{1+b_j}} \right) dx \right]^{1/N}$$
(2.40)

With

$$\alpha = \frac{1}{\sqrt{(2\pi)}} \left(\frac{1}{1+b_i} \frac{\sigma_i^2}{k_i^2} \right)^{-\frac{1}{2}} \left[\prod_{\substack{j=0 \\ j \neq i}}^N \left(\frac{1}{1+b_j} \frac{\sigma_j^2}{k_j^2} \right)^{\frac{1}{2N}} \right]$$
(2.41)

and

$$\beta = \frac{\sigma_i^2}{2k_i^2} \frac{1}{1+b_i} - \frac{1}{N} \sum_{\substack{j=0 \\ j \neq i}}^N \frac{-\sigma_j^2}{2k_j^2} \frac{1}{1+b_j}$$
(2.42)

$$\int_0^{+\infty} \left[\frac{1}{\sqrt{2\pi}} \left(\frac{1}{1+b_j} \frac{\sigma_j^2}{2k_j^2} x \right)^{-\frac{1}{2}} e^{-\left(\frac{1}{1+b_j} \frac{\sigma_j^2}{2k_j^2} x \right)} \operatorname{erfc}^n \left(\frac{\sqrt{2}}{2} \frac{\sigma_j}{2k_j} \sqrt{\frac{x}{1+b_j}} \right) dx \right] = \left[\frac{k_j^2}{\sigma_j^2} \frac{1+b_j}{N+1} \right]$$
(2.43)

$$\alpha \int_0^{+\infty} \left(x^{-\frac{1}{2}} e^{-N\beta x} \right) dx = \alpha \left(\sqrt{\frac{\pi}{N\beta}} \right)$$
(2.44)

At the end we have that :

$$P(\mathbb{A}) \leq \prod_{\substack{j=0 \\ j \neq i}}^N \left(\frac{k_j^2}{\sigma_j^2} \frac{1+b_j}{N+1} \right)^{\frac{1}{N}} \alpha \left(\frac{\pi}{N\beta} \right)^{\frac{1}{2N}}$$
(2.45)

Financial Institutions Externalities and Systemic Risk: Tales of Tails Symmetry

MEasuring negative externalities of banks is a major challenge for financial regulators. We propose a new risk management approach to enhance the financial stability and to increase the fairness of financial transactions. The basic idea is that a bank should assume as much risk as it creates. Any imbalance in the tails of the distribution of profit and losses is a sign of the banks failure to internalize its externalities or the social costs associated with its activities. We link those asymmetries to contribution of banks to the systemic risk. In this paper, we develop a theoretical model to show the importance of tail symmetry on the sustainability of the financial system. We also propose a mathematical definition and a measure of tail imbalance based on Extreme Value Theory(EVT). This measure could help regulators and policymakers to have an insight into the contribution of each bank to the overall risk hidden in the financial system and waiting to burst in the context of a crisis to serve as an early warning indicator. We also propose econometric techniques to overcome the data availability issue and create an early warning system to monitor the state of the financial system.

Introduction

Banks and more generally financial institutions play a major role in the economy. The ability of the financial system to intermediate between those who are willing to lend and others in need of borrowing is a key determinant of growth and economic welfare. In the absence of this system of intermediation, it would be difficult for most companies to fulfill their need for investment and for individuals to invest in durable goods and consume non-durable goods. Unfortunately, it was a global breakdown of the financial system that was required to witness enough attention from researchers to the issue of systemic stability.

Andrew Crockett,¹ pointed out that some rational or even desirable decisions at the individual level may have an unwelcome collateral effect on the macro level (Crockett (2000)). This observation was his motivation to urge regulators years before the crisis of 2008 toward the need of balancing micro and macro-management approaches when regulating the financial system. In fact, he highlighted the importance of reinforcing the traditional Basel accords by marrying micro-risk management with macro-risk management.

The objective of this paper is to propose a theoretical and practical approach for measuring the negative externalities or social costs that are generated by banks' activities that may contribute to the embedded stress in the financial system and consequently increase its fragility. Several approaches have been published to fill up the gap on individual risk measures. A survey by Bisias et al. (2012) listed up to 30 systemic risk measures in the literature. The originality of our approach is that it accounts for all the internal decisions of the bank and then considers their net effect on the system. The first apprehension is that decisions that may have the greatest impact on the financial system in the time of distress are probably those seeking gains and more precisely large gains. Notably, both sides of the profit and loss distribution will be examined in our analysis with a focus on the tails. The key concept is that outstanding positive returns should have their importance in assessing systemic risk as should extreme losses.

¹General manager of the Bank of International Settlements from January 1994 until March 2003

In this paper, we suggest a new rule regarding the management of systemic risk by regulators: a financial activity with no negative externalities on the financial system should not alter the symmetry of the tails of the profit and losses distribution of both signatories. Even though externalities connote a powerful image of economic and financial plagues, it is hard to define it scientifically, by consequence the measurement of externalities is troublesome for a regulator. Nevertheless, we consider that potential losses that an economic agent A can suffer after contracting a transaction with an agent B as being externality generated by the latter if at the moment of the signature A is not totally aware of the unfortunate event leading to losses while B is expecting that gain. For the sake of parsimony, we always designate negative externalities by simply externalities. A prototypical example of externalities would be the implicit to big too fail guarantee. In fact, banks will be keeping the gains from the very risky position but potentially very profitable while the government will intervene to prevent the losses. As the implicit guarantee is potentially paid by tax payer, not the bank we consider it as an externality according to our definition.

Several arguments suggest that banks should have tails symmetry properties.

The main intuition behind the idea of looking into gains is that most derivatives are zero-sum games: for every winner, there is a loser. Gains of one financial institution should be reflected in the losses part of one or several banks holding the opposite position on the same derivatives. In the presence of perfect symmetry of information ², market actors should have the same assessment of risk and more importantly the same expected extreme payoffs. First, it is crucial to draw the attention to the fact that the derivatives market has known a substantial growth in the last few decades. In fact, the total value of the outstanding notional amounts on derivatives is larger than the world's GDP by a scale of magnitude ³.

Based on this rule, the left and right tails of the profit and loss distribution of a bank should have the same tail fatness. Notice that we consider that the notion of

²The symmetry of information also implies that both financial institutions have equivalent computational resources and human capital to process the available information and draw conclusions. Arora and Barak, 2009 pointed out that computational complexity can increase the information asymmetry.

³According to the World Bank the total GDP for 2010 is around 63 trillion of US\$. Also according to the Bank of International Settlements, the total outstanding notional amount for the derivatives market is 601 trillion of US \$ as of December 2010.

symmetry, due to the zero-sum positions, is something that is only relevant to the tails. In fact, it is possible and even healthy for the financial market to have actors with different opinions and anticipations based on the same set of information. By contrast, the extreme events are by definition unpredictable and most important of all unseen in history. Hence, the anticipation of extreme payoffs based on rare market movements should be identical among all financial institutions.

The second argument supporting the idea of tail symmetry is that often the pre-crisis unusual gains are the results of hazardous financial innovation that will have negative effects in the future. In measuring systemic risk, it is also important to pay special attention to the run-up phase, in which systemic risk and bubbles are built up in the background and waiting to burst during a financial crisis. In this phase, it is important to identify actors that are exposing the system to hidden risks to generate excessive positive returns. In this context, it is crucial not only to look at the correlations that may create a contagion of losses between financial institutions, it is also crucial to detect imbalances that are built up during the pre-crisis phase. Brunnermeier and Oehmke (In Press) pointed out the difficulties in accessing the intensity of the risk hidden in the system and waiting for a trigger to materialize in the financial system in the shape of a financial crisis. Our idea is that looking into gains should identify institutions which are creating invisible risks in the system. It is the extreme asymmetry between potential gains and possible losses that should trigger the concerns of regulators about the bank's level of externalities. For example, Beltratti and Stulz (2012) found evidence supporting that large gain during the pre-crisis period is negatively correlated with their performance during the crisis. At first glance, this approach seems more comprehensive as it aggregates a wider set of information compared to loss based measures. In times of financial stability and growth, most investment decisions are observed on the profit side of the P&L distribution rather than on the loss side. In addition, losses are highly monitored and regulated by the financial authorities with different incentives to limit exposures to downside risk. Meanwhile, little attention is allocated to gains. The race to boost gains may create important pervert effects and the origin of potential systemic risk lies in the failure to internalize them.

So far, systemic risk is associated with the correlation between the left tails of both the bank and the financial market. However, it is possible that companies behind the financial system's fragility are not the ones that suffer the most from the possible damages of the crisis. It is more likely that the abnormal profits that those companies yield in normal times shield them from a critical decrease in capital value during the turmoil. A prototypical example is the one presented by Bernard et al. (2013) about *Ambac*, a US company providing financial guarantee. If *Ambac* fails, many guarantees will become riskier, and by consequence, its counterparts will see their risk increase dramatically leading to an overall increase in risk. However, it is obvious that the activity of the company should increase the overall stability of the system. On the other hand, a company like Goldman Sachs made profit helping its clients take a position in the housing sector via the creation of Collateralized Debt Obligation(CDO). These activities of were unseen on the loss side. However, it is obvious that they contributed to systemic risk.

From a purely conceptual point of view, every financial institution should pay for the risk generated by its investments. Expressed differently, no financial institution should benefit from the exposure to positive shocks without bearing the risk of losing an important amount of money in the occurrence of some negative shocks. We develop this concept through the idea of tail symmetry. The symmetry perception should reflect both the notion of information symmetry and the internalization of the possible damaging impacts on the financial system. Said differently, the regulator should make sure that the bank internalizes all the negative effects on the system, i.e. its externalities. Moreover, our methodology has an ethical dimension. It seems unfitted that a bank *A* hides its exposure to extreme positive gains from its counterpart *B*. In this situation, it is bank *B* that ultimately pays the price of the risk taken by *A*. Moreover, in the spirit of Allen and Gale (1997) intermediaries such as banks ensure an intertemporal smoothing of no diversifiable risks. The problem arises when banks fail in the risk transfer function and hide the unsustainable risk in the near future. The concept of tail symmetry will help regulator to detect when banks are borrowing future gains at the expense of the future system stability.

The last argument in favor of tail symmetry is related to internal risk management within banks. In fact, banks should know that excessive gains are the results of

a favorable outcome of an exposure to some risk factors. However, a negative downturn in the evolution of those risk factors could have a disastrous impact on the value of the bank's assets. For that reason, banks should also monitor their profits to identify exposures to risk factors that are only visible in the gains part.

The content of this paper will be organized as follows. First, we will present the fast evolving literature dealing with systemic risk, and we will try to point out where our approach stands out from the crowd. The second section is dedicated to the presentation of a theoretical framework to justify the idea of tail symmetry. Then we will introduce a toy model to test some theoretical claims via Monte Carlo simulation. In the third section, we will present how to benefit from the large literature of extreme value theory to compute risk measures based on tail symmetry. Finally, using publically available market data, we will present the results of the newly designed measure and study the evolution of the externalities measure before, within and after the financial crisis.

3.1 State of the Art

We think that this work is at the crossroad between the literature on measuring systemic risk and the one about detecting bubbles and banks externalities.

While we all agree that strategies to mitigate the effect of a systemic crisis should be developed, no consensus has emerged on the definition of systemic risk. This divergence is even present among regulators across countries. The Bank of England considers a very broad definition where any possible threat to the financial system is a systemic risk event. The FED governor Daniel Tarelli defined the systemic risk with the spotlight on the financial system stability. In fact, he considered institutions to be systemically risky when any situation of distress of the bank may endanger the overall financial system. By contrast, the focus of the ECB is the impact on the real economy as it relates the systemic risk to the possible economic impact. According to it, the systemic risk is any event in the financial system that may affect the consumption, welfare, and growth of the real economy (Hartmann et al. (2009)). The ECB also characterizes its perspective of the systemic risk as being a "vertical" approach. Others developed a more specific definition and focused on

threats that affect the public confidence on the financial system (Caballero (2010)). Despite the difference, those definitions share a retrospective vision of systemic risk by looking into its effects. We propose a prospect definition of systemic risk related to the pre-crisis situation as we identify extreme asymmetries as signs of the system's fragility.

For measuring systemic risk, one can distinguish two major approaches. The first is related to network analysis and focuses on identifying relationships between financial institutions to forecast contagious chains. The second is to access the interaction between a single institution and the financial market to detect capital shortages during crises. This strand of literature based on networks borrows from the large one on epidemiology. The similarities between a crisis and the spread of an epidemiology is almost straightforward.

For a complete overview on networks for accessing systemic risk please refer to Cont et al. (2013), Elsinger et al. (2013) and Cont et al. (2013). It is also of interest the work of Barigozzi and Brownlees (2013) and Dungey et al. (2012) who designed econometric approaches to identify network interaction using publicly available data. Those approaches are best suited to examine spillover effects and identify clusters of banks that may fall together during crises. It is important to cite the pioneer work of Eisenberg and Noe (2001) who designed a mechanism of clearing in case of default of one or several financial institutions.

The second mainstream on measuring systemic risk is probably the answer of researchers to the question of the FED Governor on how to identify systemically important banks within the system and their contribution to the overall risk. The first work is by Adrian and Brunnermeier (2016) who proposed to evaluate the systemic risk as the total loss encountered by the system whenever a bank is in distress. They called their measure $\Delta CoVaR$. The idea is to evaluate the possible increase of the VaR of the system when a bank is under stress. A similar approach by Acharya et al. (2017) defined the "Systemic Expected Shortfall (SES)" as the average return of the firm when the overall system is stressed. Econometric methods to evaluate the SES were developed by Brownlees and Engle (2011). For a complete assessment of the statistical properties of both measures and a discussion about their differences,

readers can refer to Bernard et al. (2013). Both techniques have the advantage of being applicable using market data. Those data have no confidential aspect and are updated on a high-frequency basis.

While both measures introduced by Adrian and Brunnermeier (2016) and Acharya et al. (2017) capture aspect of systemic failure, both methods are focusing on the time of distress. Second building policies based on such measures can be hard to accept by bankers. It can be implied that both $\Delta CoVaR$ and SES penalize banks for being successful. In fact, part of the systemic importance that a bank may suffer is because they were able to innovate and create a financial product that had some success among other banks. Our approach, however, discriminates between transactions that create asymmetry and others that enable the bank to internalize completely its externalities. The main issue is that a systemic crisis has the annoying characteristic of being unique. In fact, the continuous innovation of financial engineering made the structure of each new crisis differs from the previous ones. Therefore, it is more important to pay attention to the build up phase to detect fragilities.

In this context, it is also interesting to review the literature on bubbles, financial crises and the study of financial channels. In fact, what we try to propose could also be interpreted as a bubble detection technique based on extreme heterogeneous beliefs. For a complete historical overview of crises and the contagion channels during this crises, please refer to Brunnermeier (2008) and Xiong (2013).

A great inspiration to our work is the article Engle (2011). While long-term risk is rather ignored in the discipline of risk management, it turned to be an important factor in the recent financial crisis. In Engle (2011), skewness and asymmetry were associated to term long risk. Our work is a proposition to include the gain side of the Profit and Losses distribution in the design of a risk measure. Authors such as Valderrama et al. (2012) and Jondeau (2010) are the first to consider both gains and losses asymmetrically in the context of systemic risk. However, the main purpose of their works was to identify the response of financial institutions toward positive and negative shocks in order to detect the system instabilities. Ours, by contrast, is introducing a new vision of externalities based on tail imbalances gener-

ated by information asymmetry. The difference in tail coefficient is from our point of view a failure for one bank to assume all the risk created by the bank's position.

We believe that the idea of tail symmetry shares a common ground with the concept of fragility and anti-fragility introduced in Taleb (2012) in a philosophical essay about randomness and developed in a more technical fashion in Taleb and Douady (2012) and Taleb et al. (2012). In short, fragility is the important nonlinear exposure to negative shocks that could be viewed as a concave loss function. By contrast, anti-fragility could express the opposite behavior where rare positive events are tremendously beneficial without suffering from fragility problems. The similarity between our approach and Taleb's idea could be highlighted in the chief ethical rule expressed in Taleb (2012): *Thou shalt not have anti-fragility at the expense of the fragility of others*

3.2 Relation between Tails' symmetry and extreme losses

3.2.1 A model of the banking system

Our model of the banking system is based on the theoretical model designed by Eisenberg and Noe (2001) using the formulation and notation of Rogers and Veraart (2012). This theoretical framework was tested by Elsinger et al. (2006) on the Austrian banking system and concluded that contagion should be the major concern of regulators. They also find out that the costs to prevent contagions are surprisingly small. We will extend the definitions by adding the notion of a crisis to the model and use its framework to establish the theoretical importance of tails symmetry on the measuring of banks externalities.

Let us consider a set $\mathcal{N} = \{1..N\}$ of financial institutions. Notice that in the paper, without making any specific distinction a financial institution is also called a bank. Each bank $i \in \mathcal{N}$, has liabilities to other banks in the system. Those liabilities are defined by an $N \times N$ matrix.

Definition 1 *Liabilities matrix* : the liabilities matrix is given by $\mathbf{L} \in \mathbb{R}^{N \times N}$, where the entry L_{ij} is the total liabilities of bank i toward bank j . Here we assume that $L_{ij} \geq 0 \forall i, j$. Also, $L_{ii} = 0$ because the bank could not have liabilities towards itself.

The model assumes the priority of debt claims and that all debts have the same seniority.

Definition 2 *Total liabilities* : The total nominal liabilities of one bank i toward the financial system is denoted by $\bar{L}_i$. $\bar{L}_i$ is given by $\bar{L}_i = \sum_{j=1}^N L_{ij}$

Definition 3 *Net asset* : Let us denote by e_i the net asset of the bank i from sources outside the financial system. The corresponding vector of net asset is $\mathbf{e}$

The net asset e_i is always positive. In other words, we exclude costs paid by the bank outside the financial system such as operational costs. Notice that such assumption is not restrictive because it is always possible to create a fictive institution inside the financial system that will have no obligation toward other banks and it will adsorb all the operational losses that are subjected by banks.

$\mathbf{e}$ represents the link between banks and the real economy. The role of the financial system is to provide liquidity and mitigate risks via the inter-banking market represented by the liability matrix.

Definition 4 *Value of equity v_i* : is given by the total incomes less the value of liabilities paid to creditor

$$v_i = \sum_{j=1}^N L_{ji} + e_i - \bar{L}_i$$

Definition 5 *Financial System* : given a liabilities matrix $\mathbf{L}$ and a vector of net assets $\mathbf{e}$. The couple $(\mathbf{L}, \mathbf{e})$ is a financial system.

This definition of the financial system supposes the knowledge of all the liabilities of banks toward other members of the system. In practice, such knowledge is hardly available even for the regulator and central banks. Nevertheless, this framework is an excellent starting point to study the interaction between banks and possible defaults. The famous Diamond and Dybvig (1983) model is well suited to study liquidity contagion, and bank runs in a multi-period framework where bilateral clearing issues are trivial. Nevertheless, in the context of a multilateral network with cyclical liabilities, the solution of this problem is less obvious and the Eisenberg and Noe (2001) model is more adapted. In fact, Eisenberg and Noe (2001) proved that, under some mild condition, it is possible to find a vector of obligation L^* which respects the limited liabilities of banks and the proportional sharing in case of defaults. In other terms, it is possible to find the expected payment to each bank in case of default of one or several financial institutions and of course the banks defaulting from the first series of collapse.

3.2.2 Definition of tails symmetry

Most of the literature studying the tails of distribution focus on one side of any event distribution. Usually, more focus is attributed to the disastrous side of any given distribution because extreme positive events are usually welcomed. Therefore, practitioners feel little need to study both tails simultaneously only the probability of negative events is measured to mitigate their effects. This explains the absence of the idea of tail symmetry in the literature about heavy tailed distributions. Incorporating both tails in the analysis is to the best of our knowledge unique in the discipline of risk management, and it is the main originality of our work.

Definition 6 A probability distribution $\mathbb{P}$, having a zero mean, is said to be long tail symmetric if $\exists \kappa > 0$ where $\forall \mu > \kappa$ we have :

$$\mathbb{P}(X > \mu) = \mathbb{P}(X < -\mu) \quad (3.1)$$

κ is called the tail symmetry threshold.

The idea of tail symmetry is different from the idea of skewness. While skewness is a measure of the symmetry of the entire distribution, the tail symmetry is only considering the extreme components. It is trivial that symmetrical distributions with zero skewness are also tail symmetric. Nevertheless, it is not possible to infer the value of skewness starting from the symmetry of tails. To illustrate the difference between both skewness and tail symmetry we will present the following example of distribution. Given one positive real number κ :

$$\begin{cases} f_X(x) = \alpha_1 x + \beta_1 & \text{for } b \leq x < 0 \\ f_X(x) = \alpha_2 x^2 + \beta_2 & \text{for } 0 \leq x < \kappa \\ f_X(x) = e^{-|x|} & \text{otherwise} \end{cases}$$

α_1 , α_2 , β_1 and β_3 are chosen such as f_X is a probability distribution. It can be shown that the choice of α_1 , α_2 , β_1 and β_3 is always possible and that it is unique under the condition that f_X is continuous at all points.

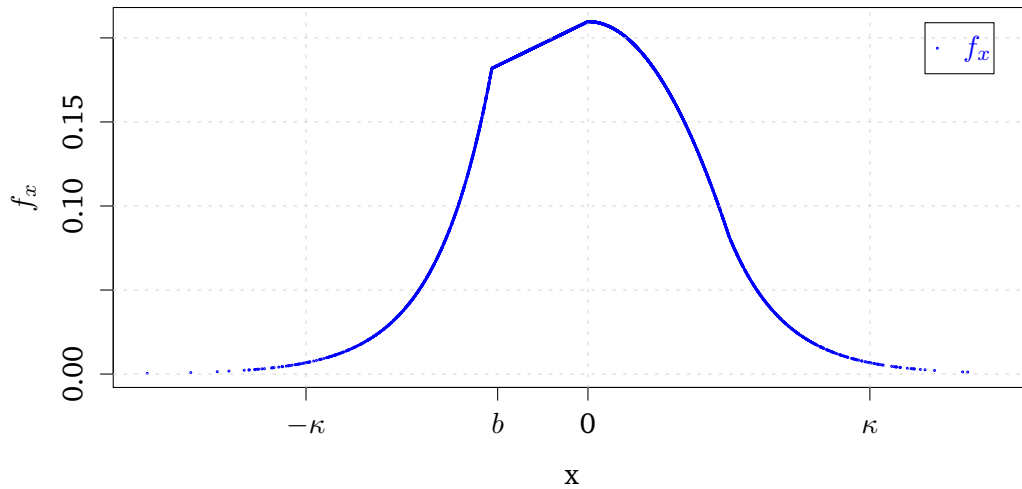


Fig. 3.1.: The figure illustrates the probability distribution f_x . It is visible that the distribution is skewed however tail symmetric

The probability distribution f_x is skewed. However, the distribution is tail symmetric. In fact, $\forall \mu > \kappa$, it is easy to verify that $P(X > \mu) = P(X < -\mu)$. The idea of tail symmetry is only specific to the behavior of the distribution far on the tails without paying attention to the symmetry of the distribution around the mean value.

Therefore, the value of the skewness, usually dominated by asymmetry around the mean, is not a real indicator of the tails balance. The importance of introducing such definition of tail symmetry is to stress on the fact that we are only interested in tails as by definition systemic risk events are extreme events. Finally, it is important to notice that zero-skewed distributions can also have unbalanced tails. The skewness of the distribution was lately associated with long-term risk in the paper Engle (2011). In his paper, he argues that negative skewness is an indicator of long-term risk which was a major component in the 2008 financial crisis. Our paper tries to push this rationale to the limits and only focuses on extreme imbalances. While skewness is associated with the long-term component of risk that could be mitigated, we think that tail imbalances are risks that are beyond mitigation and that the regulator should be well armed to face them in case they trigger a systemic crisis. A simple example that we can give about externalities is that of a chemical company that invests in high-risk facilities that can increase dramatically the incomes of the company. Meanwhile, insurance companies are not pricing this risk and continue to impose low premiums for the chemical company. In other words, the later company will cash in gains and is not willing to pay for the potential losses.

3.2.3 Implication of tail symmetry in distress probability :

To incorporate the idea of a crisis in our model we will begin by presenting the implication of a crisis on the banking system.

Definition 7 *The condition of a systemic crisis is defined by an important negative shock to the net assets e_i of all banks in the financial systems.*

Rogers and Veraart (2012) argued that given an initially solvent financial system, only a substantial negative shock on the bank's net assets from outside the financial system could lead to the situation of default of one or several financial institutions. This substantial negative shock is what defines a financial system crisis in the context of our paper. A change in the pay-off structure inside the financial system could not engender the crash of the system as a whole since value is always inside the system. The impact resulting from the decrease of the assets e_i can be expressed in the

following way: the bank i will fail to pay its obligations toward other banks based on the expected revenue of assets outside the financial system. A straightforward simplification of the system that we will adapt in the context of crisis is to neglect the amount of the net assets e_i relative to the liabilities of the banks toward the overall financial system. Notice that the model implicitly suppose that shocks that lead to banking failure are originated from the real economy and not from inside the financial system. The inter-banking activities cannot create value. Banks raise funds to finance the lending and investing operation outside the financial system. Ex ante, they expect to have a positive equity balance after reimbursing debts. Ex post when uncertainty is resolved and the state of nature turns out to be a situation of crisis, some banks will have a shortage of liquidity and all the money they collected will be provided to the clearing system. In this situation, we also witness a destruction of value inside the financial system. Of course, we should assume that we have a perfect claim-enforcement technology.

An important feature of the definition of systemic crisis that we adopt is that the origin of the shock is assets. This excludes all type of crises originating from funding shocks like bank runs in the famous paper of Diamond and Dybvig (1983). This choice can be explained by two reasons essentially. The first is that the model that we propose of the financial system is not rich enough to describe funding of banks through deposits. The system is a closed system and banks mainly finance their operations through the interbank liquidity markets. The second and most important reason is that we designed the model to later justify the importance of tail symmetry for banks. The stability of funding is not reflected in the distribution of profit and losses and will not have any impact on the asymmetry of the distribution. It can also be seen that the implicit assumption that we have in this model is that banks have a constant amount of stable deposits that are not subject to runs and that the only additional source of liquidity is the inter-banking liquidity market. The impact of confidence in banks on the stability of the system is outside the scope of this chapter. Finally the magnitude of the shock on net assets should be interpreted in relative terms compared to the total liability of the bank. Consequently, we consider that in the situation of a crisis the net assets fall to a level where it becomes negligible compared to the total liabilities of the bank itself. But from a general perspective big banks will remain to have important net assets compared to smaller banks.

Theorem 2 *Given a financial system (L, e) with at least two participants. Under the condition of systemic crisis E as defined in definition (7) and tail symmetry of all financial institutions $i, i \in \mathcal{N} = \{1..n\}$ acting in a financial system (L, e) , and the assumption that the bank can only pay a fraction of its liabilities in case of default, we have :*

$$\begin{aligned} \exists \psi > 0 \text{ where} \\ \forall i, j \in \mathcal{N} \text{ and } \forall \delta > \psi \quad \mathbb{P}(v_i \leq -\delta | E) = \mathbb{P}(v_j \leq -\delta | E) \end{aligned} \tag{3.2}$$

The results of theorem (2) are of course sensitive to its underlying assumptions. The first one is that all banks in the system have a distribution of profit and loss which is perfect tail symmetric. It is rather a strong assumption to be satisfied by all banks in the financial system and we show later that this is definitely not verified empirically. However, one should bear in mind that the objective of this theorem is to show what would be the implication on financial stability if we live in this perfect financial system in which externalities are non-existent from all actors. We do not claim that this is a realistic assumption, but we show that under those conditions all banks will have the same probability of failure resulting from unpredictable risks. In those settings, if some banks are more prone than others to fail in systemic crisis means that the condition of tail asymmetry is not satisfied. The second assumption is the definition of the crisis itself. Besides ignoring funding costs, if you obtain equal probability of failure in systemic crisis is also because we assume that the net asset of banks are very small to cover its liabilities toward other banks. It is important to note that we do not assume those net assets to be zero across all banks but only a milder version. The assumption remains realistic if we take the liquidity of those assets into consideration. In fact, the model is a two period model. In the second period, banks need to liquidate their assets to cover their engagement. The failure to provide short term funding due to the liquidity of the assets can also result on the failure of the bank. This assumption can still be representative in a financial system in which banks continue to have important maturity mismatch between their assets and liabilities.

The importance of tail symmetry for the stability of the financial system is a direct implication of theorem 2. In fact, this theorem stipulates that given the tail symmetry of the financial system, all banks will have similar probabilities of realizing extreme losses. To give more intuition about the policy implication of this theorem, one should bear in mind that the situation of systemic crisis represents unpredictable events by both regulators and a set of financial institutions in the system. The banks failing to identify the crisis are innovation followers who invest in derivatives designed by the competition and not fully understood by their risk managers. It is the leading banks who make a profit by exposing their portfolios to positive events while hiding part of the potentially disastrous effects from their counterparts. The role of regulators is to make sure that banks can have homogeneous extreme beliefs. Said differently, the homogenization of extreme risk assessment across all banks will lead to tail symmetry. The threshold defining extreme losses(ψ) in this context is also the maximum of tails symmetry thresholds κ_i of all the financial system. In other words, if regulators can impose on banks a limit of asymmetry of profit and loss distribution the risk of experiencing events of extreme losses will be equivalent between all banks. It also implies that all banks had been successful at internalizing their externalities due to their financial activity in the financial system. Put differently, whenever a bank is changing its strategy to be exposed to some profitable event that same decision should make the bank vulnerable to other negative event that will balance both tails of the distribution. This idea can have a close link to the concept of no-arbitrage. In fact, the bank should not be potentially profiting from some unpredictable events in the market while other financial institutions are bearing the risk of extreme losses due to the same event.

The tails symmetry thresholds ψ choice can have important implications for the financial system, and that choice should reflect the regulator policy on the management of the financial system under his jurisdiction. In fact, we believe that setting ψ to be very high raises two important issues. First, the regulator could be missing threatening financial products that have an effect below the symmetry threshold. In the long run those products could result in a major systemic failure. Moreover, the regulator can have a hard time monitoring the tail symmetry of bank if the thresh-

old is high simply because it will be hard to measure this effect due to the curse of black swans in the extremes.

By contrast, if the regulator chooses to set the tail symmetry threshold to be at a low level, the banks should adjust their portfolios to meet the symmetry condition starting from relatively frequent and predictable outcomes. In fact, transactions in the financial markets are driven by the dispersion of beliefs between the market participants. Setting the threshold limit to a low level can lead to the homogenization of beliefs and by consequence limit the evolution of the financial market with all the economic implications that may arise from such restriction. A direct implication of the homogenisation of beliefs is to slow down financial innovation which is an important factor to sustain growth and support the fast-changing business world. Moreover, such restriction can have a negative impact on competition between banks and by consequence increase the costs of financial services. In the end, this could lead to further income inequalities as argued by Beck et al. (2010).

In conclusion, the regulator needs to strike a balance and choose the right level of tails symmetry threshold according to their policy about financial innovation and growth.

Presently, we believe that regulators are implicitly proposing an extreme threshold in the Basel III regulations. In fact, proposing that banks should predict 99% of the losses means that they consider that 1% are unpredictable. By consequence, regulator tolerates that banks are unable to predict 2% of the P&L distribution. According to theorem 2, central banks should also make sure that banks are tail symmetric starting from the 99% quantile.

3.2.4 System stability illustrated by simulations

To illustrate the relationship between tail symmetry and the stability of the financial system we will borrow the model of the banking system proposed by Ichiba and Fouque (2013). A similar model was also presented in Fouque and Sun, 2013, the authors were able to illustrate the limitation of diversifications in the case of systemic events. In our case, we will use the model to show via Monte-Carlo simu-

lations that the financial system can become unstable in the presence of tail asymmetry.

Let us consider the following financial system $Y := (Y_t := (Y_t^{(1)}, \dots, Y_t^{(n)}), 0 \leq t < \infty)$ of $(N \geq 2)$ banks. $Y_t^{(i)}$ is the log-monetary reserve of bank $i, i \in \llbracket 1, N \rrbracket$ at time t .

It is important to highlight that we do not claim that the model is a full representation of the interactions between banks. In fact, the model is ill suited to study spillovers in the real financial system where each bank and each connection is unique in the system. Nevertheless, network analysis has shown that individuals tend to have similar behaviors in the situation of crisis and panics. For this reason, we believe it is safe to implicitly assume that banks have similar behaviors in our model and use this model to extract conclusions about extreme random behaviors of the financial system.

In the absence of interaction between banks, where no lending and borrowing is possible between market actors, $Y_t^{(i)}, i = 1 \dots N$ are independent. Given these settings we assume that the banks in the system are only driven by a Brownian motion:

$$dY_t^{(i)} = \sigma dW_t^{(i)} \quad (3.3)$$

Where $(W_t^{(i)}, i \in \llbracket 1, N \rrbracket)$ are independent standard Brownian motions that start at time $t = 0$ from $Y_0^{(i)} = y_0^{(i)}, i \in \llbracket 1, N \rrbracket$. Also, for the purpose of this study we choose to use a fixed and identical diffusion coefficient σ .

To model interaction between banks, it is important to distinguish between two channels of shocks transmission that are identified in the contagion literature. The first is through direct interbanking claims as defined by Allen and Gale (2000). By contrast, Diamond and Rajan (2005) argue that contagion is possible even in the absence of direct links. Because they refill their liquidity supplies throughout the interbanking liquidity market, the shrinkage of the later due to a situation of a stressed bank can have negative effects in other banks sharing the same market. Empirical evidence suggests that both types of contagion exist simultaneously, hence in this section we show the externalities generated by tails asymmetry of both topologies

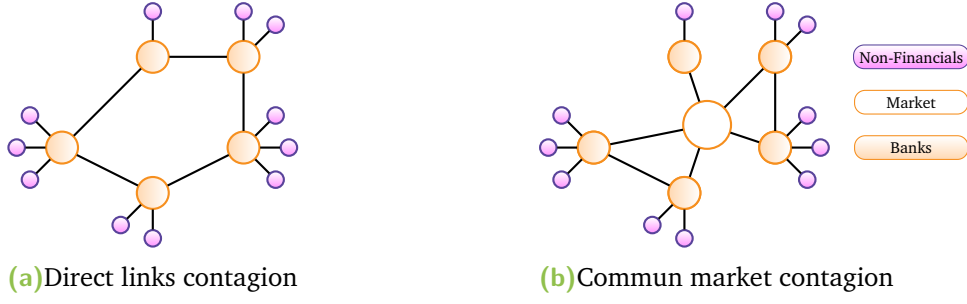


Fig. 3.2.: Network representation of the contagion channels in financial markets. Figure (a) represents market dominated by direct links channels. Figure (b) is an illustration of contagion based on a common liquidity market

of the financial system. To focus our study on the effect of tails asymmetry, both structures of the financial system will be studied separately. This treatment will also help to disentangle between fragility related to a specific type of structure. Figure (3.2) illustrates the two types of topology of the financial market that we will consider in this section.

First, we describe the modeling fashion when we only consider direct links between banks as the leading contagion channel as in the model of Allen and Gale (2000). In the latter, this topology of networks will be called direct networks. For that purpose, the interaction between banks is introduced throughout a drift term in the diffusion process. The drift $(Y_t^{(i)} - Y_t^{(j)})$ is proportional to the rate at which the bank i borrows from bank j . The dynamic of interactions is in line with the previous modeling of the banking system. In fact, the link between banks in the inter-banking market is the driver of contagions and interaction between actors in the system in both situations. This model is suited to study systemic risk events considering that the failure of banks within a crisis context is usually coupled with a dramatic fall in its monetary reserve with a failure to get liquidities from the inter-banking money market.

Here the dependent model is then :

$$dY_t^{(i)} = \frac{1}{N} \sum_{j=1}^N \alpha_{ij} (Y_t^{(i)} - Y_t^{(j)}) dt + \sigma dW_t^{(i)}, \quad i \in \llbracket 1, N \rrbracket \quad (3.4)$$

Where α_{ij} is a Bernoulli random variable with a parameter p . In fact, $p = 0$ represents the independent system where all banks are independent of each other. By

contrast, $p = 1$ is a complete network in the sense of Allen and Gale (2000). In the latter configuration, each bank in the system has links with all other banks throughout the lending/borrowing mechanism.

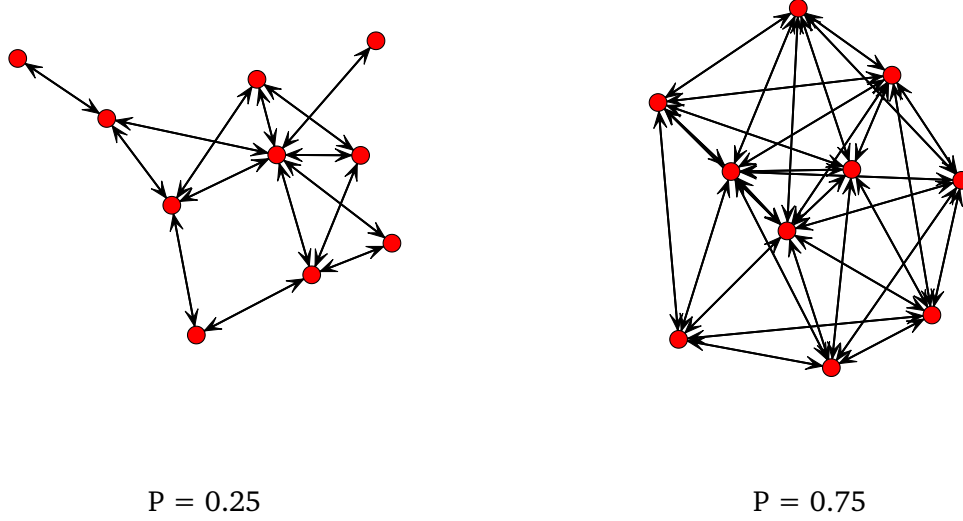


Fig. 3.3.: Graphical representation of the links between banks in a direct network for $p = 0.25$ and $p = 0.75$

Because the diffusion process of all banks is identical in the model introduced in equation (3.4), the banking system presents the characteristics of tail-symmetric banks. To study the effect of tail asymmetry on the fragility of the system, we will compare the symmetric system generated with equation 3.4 to a modified version where we introduce tail asymmetry to one of the financial institutions.

The model will be then :

$$\begin{cases} dY_t^{(i)} = \frac{1}{\sum_{j=1}^N \alpha_{ij}} \sum_{j=1}^N \alpha_{ij} (Y_t^{(i)} - Y_t^{(j)}) dt + \sigma dW_t^{(i)}, i \in \llbracket 2, N \rrbracket \\ dY_t^{(1)} = \frac{1}{\sum_{j=1}^N \alpha_{1j}} \sum_{j=1}^N \alpha_{1j} (Y_t^{(1)} - Y_t^{(j)}) dt + \sigma dW_t^{(1)} + J(t) dq(t) \end{cases} \quad (3.5)$$

Where $J(t) > 0$, is the jump intensity. $dq(t)$ is a Poisson counter process variable with intensity λ such as $\mathbb{P}[dq(t) = 1] = \lambda dt$. The jump process is identical to the jump component introduced by Merton (1976) in the price dynamic.

The realization of $q(t) = 1$ represents a rare event that leads to excessive gains for bank 1 in the system. Of course, the jump introduced in the diffusion process of bank 1 will induce a distortion of the right tail of the gains distribution.

In both financial systems defined by the process (3.4) and (3.5), we consider that the bank is defaulting if the value of its log monetary reserve falls below a certain threshold η .

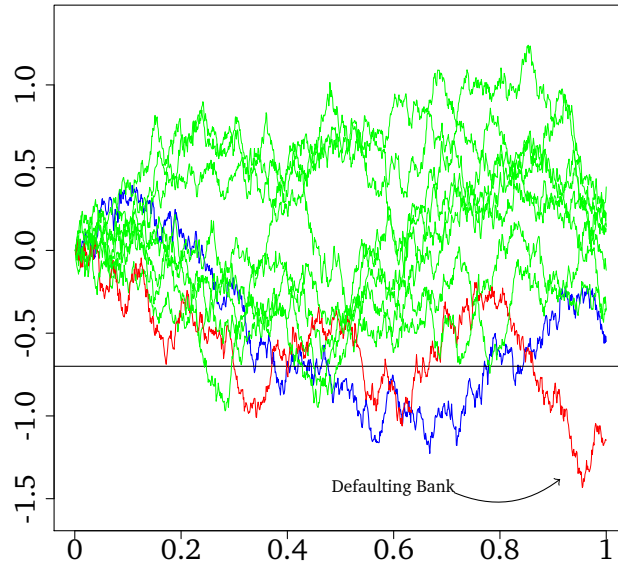


Fig. 3.4.: One realization of the trajectories of the coupled diffusion (3.5) with $\alpha = 1$. The solid line represents the default level $\eta = -0.7$

Figure (3.5) illustrates the impact of introducing a jump component to the diffusion process of bank 1. The left figure clearly shows that the quantiles of the right tail are higher than those of the normal distribution. However, such difference is less relevant in the case of the right plot where no jump component is introduced in the diffusion. Of course, the positive jump has very little impact on the left tail.

We will use Monte-Carlo simulations to compare the coupled diffusion (3.4) with the diffusion presented by the system (3.5). The aim of the simulation is to study the effect of introducing possible rare but extreme gains to one bank on the overall stability of the system. The latter will be proxied by the distribution of the number

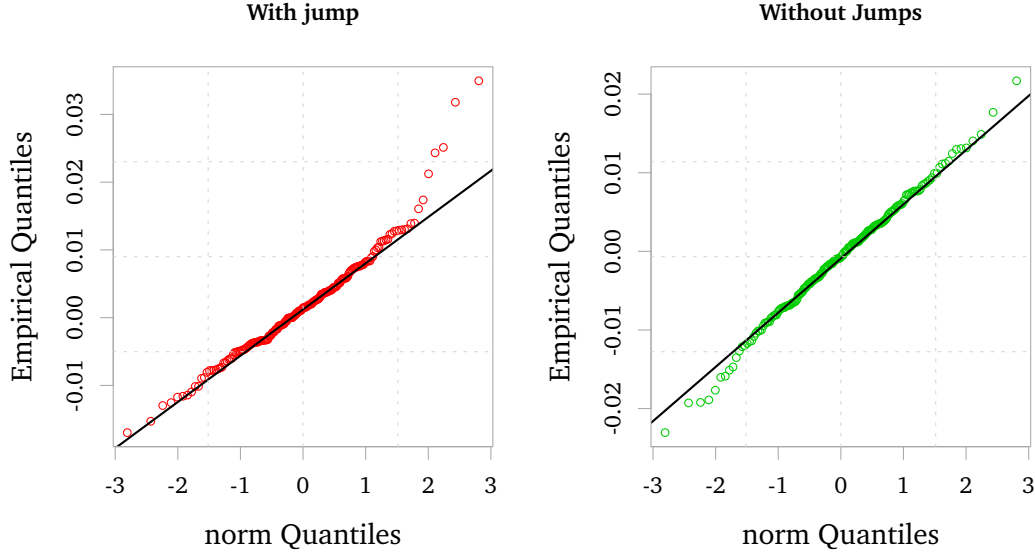


Fig. 3.5.: Q-Q plot for the coupled diffusion introduced in equation (3.5). The right plot is for the bank with a jump component $J(t) = 0.02 Y_{t-1}^{(1)}$. The left plot represents the quantiles of a bank with no jump component on the diffusion.

of failing banks in the system according to our default criteria. This distribution will be called the default distribution for the sake of parsimony

For the simplicity of our simulation, we assume the following parameters: a common $\sigma = 1$, $N = 20$, and we used the Euler scheme with time-step $\Delta = 10^{-3}$, up to time T . Finally, we assume $y_0^i = 10, i = 1..N$ and that η

We can see from the illustration of default distribution for both diffusions with and without positive jump probability, that the asymmetry of the tail of one bank seems to weaken the system. In fact, the loss distribution corresponding to diffusion (3.5) has a higher right skewness compared to the system where all banks have symmetric tails. Tail asymmetry have also an impact on the expected value of failing banks which is equal to 6.946 in the presence of jumps and is only 5.343 in the other configuration ⁴.

Figure 3.7 illustrates the variation of the expected number of the failing bank E_F in each configuration modeled by the diffusion process 3.4 and 3.5. Two important conclusions could be drawn from this figure. First, the financial system with a toxic bank in its premises is riskier compared to a system described by equation (3.4)

⁴The difference is $< 1\%$ significant following a t-test and a Wilcoxon test

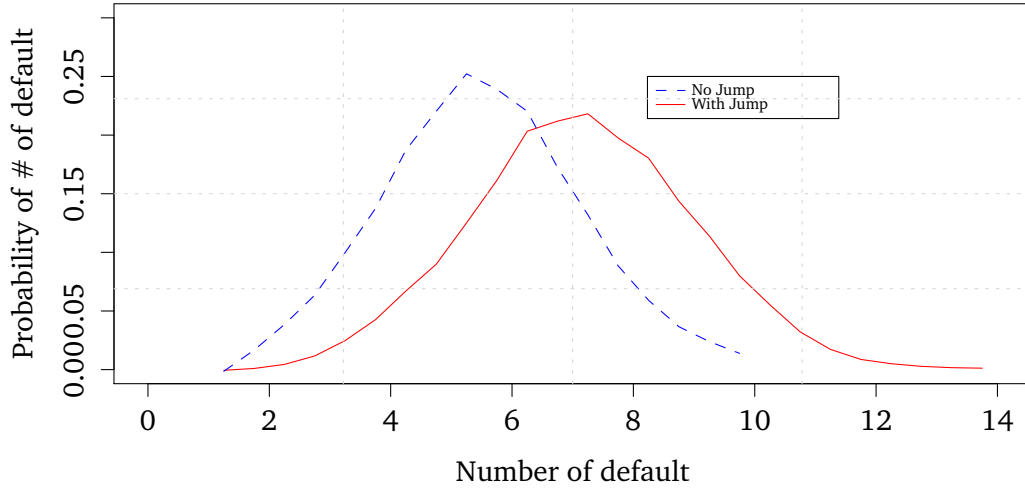


Fig. 3.6.: The figures shows the Probability distributions of the number of defaulting banks in the system with $p = 0.8$ and jump component $J(t) = 0.02 Y_{t-1}^{(1)}$.

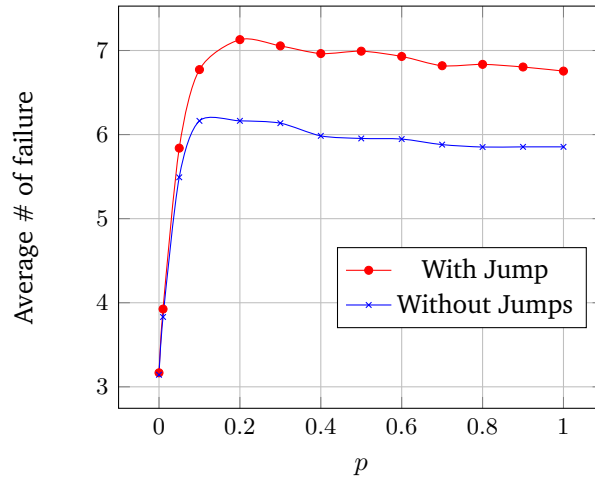


Fig. 3.7.: The expected number of failure as a function of the sparsity parameter p . 5000 simulations for $N = 20$ banks were executed to compute this figure. All other settings are identical to those of figure 3.8

in the whole range of p . Second, and most importantly we can notice that the evolution of E_F is characterized by two different regimes. E_F starting from a low value of corresponding to the independent system $p = 0$ will continue to increase until reaching its maximum at p around 20%. Then, E_F will decrease however with a slower rate until $p = 1$ for the complete network. The change of monotonicity is a due to the trade-off between the beneficial effect of links in the financial network as highlighted by Allen and Gale (2000) and the role of links as a channel of financial contagion.

To check the robustness of these results, we will consider the case of a banking system where the major source of contagion is the inter-banking liquidity market following the model of Diamond and Rajan (2005). The proposed dynamic of liquidity reserves Y^i , $i \in \llbracket 1, N \rrbracket$

$$\begin{cases} dY_t^{(i)} = \frac{\alpha}{1+Y_t^{(i)}} \left(Y_t^{(i)} - \frac{1}{N} \sum_{j=1}^N Y_t^{(j)} \right) dt + \sigma dW_t^{(i)}, & i \in \llbracket 2, N \rrbracket \\ dY_t^{(1)} = \frac{\alpha}{1+Y_t^{(1)}} \left(Y_t^{(1)} - \frac{1}{N} \sum_{j=1}^N Y_t^{(j)} \right) dt + \sigma dW_t^{(1)} + J(t)dq(t) \end{cases} \quad (3.6)$$

Where we assume that the mean reversion rate $\alpha > 0$. α is a parameter that expresses the level of dependence between banks. In fact, $\alpha = 0$ represents the independent system, reader interested in discussions about the impact of parameter α can refer to Fouque and Sun (2013). Notice that the larger the parameter α the more stability is observed on the system, but the impact of a systemic event will be more destructive to the system in that case.

Two important features are considered in the design of the interaction between banks in equation (3.6). $Y_t^{(i)}$ is driven by the rate of lending and borrowing between the banks and the average available liquidity in the market $\sum_{j=1}^N Y_t^{(j)}$. The bank's liquidity target is again the average available liquidity in the inter-banking market. Notice that, the banks will perceive positive interests whenever its liquidity is higher than the average. It will also borrow to reach its target and pays interest whenever it has low liquidity provisions. Moreover, average interest paid by banks rates is decreasing whenever the banks have higher available liquidity $Y_t^{(i)}$. The rational behind this choice is that the liquidity market is organized in successive rounds. Each round will see the exchange of a fixed amount of liquidity and rates. The interest rates are decreasing between rounds. Banks with higher cash available (needs) should participate in more rounds to satisfy their needs. Therefore, the average interest rates will be lower for banks engaging in several rounds of liquidity exchanges.

The crucial fact of those simulations is that regardless of the nature of the banking interaction channels, introducing an asymmetry of the tail's distributions results on an increase of the overall instability of the system and leads to greater chances of

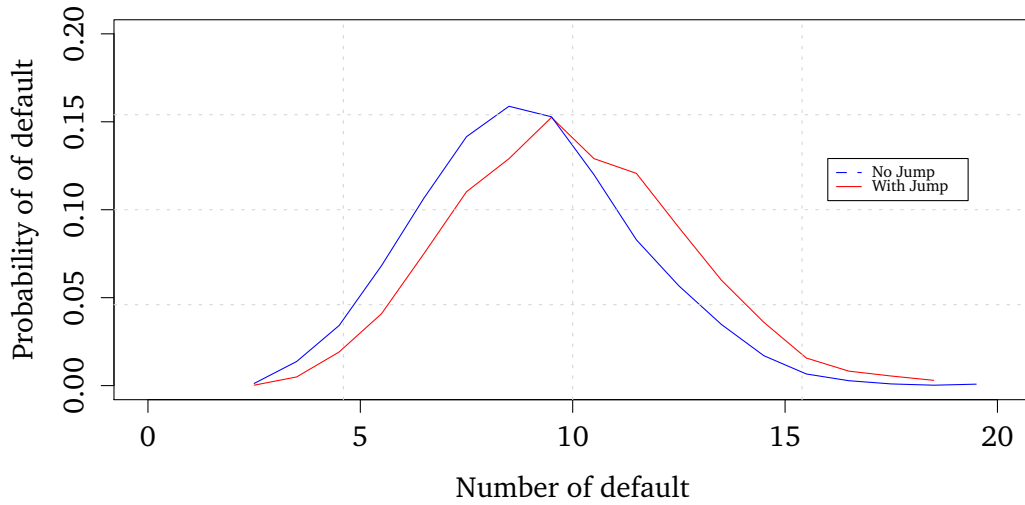


Fig. 3.8.: The figure shows the Probability distributions of the number of defaulting banks in the system with $p = 0.8$ and jump component $J(t) = 0.02 Y_{t-1}^{(1)}$.

systemic crisis. We do not claim that those simulations are some sort of proof that asymmetries are the origin of fragility in the system. Nevertheless, we can conclude that the asymmetry amplifies the negative effects of regular contagion channels.

3.3 Measure of tail asymmetry and capital provisions

3.3.1 Tail index

The empirical study of this research will focus on identifying evidence of externalities based on the study of a financial institutions' stock price. At this level, it is convenient to call for the large literature about tails and extreme events developed in extreme value theory(EVT).

In the approach of EVT, we focus on the tails regardless of their behavior around the mean value. This approach has the great advantage of the possibility to characterized the tails with the need of only one parameter for a wide range of distributions. The main idea is that a very large class of probability distributions could be approached in the tails by a probability distribution called Generalized Extreme Value

distribution (GEV). The cumulative distribution function of the GEV distribution could be expressed as :

$$F_{\xi}(x) = \begin{cases} \exp\{-(1 + \xi x)^{-\frac{1}{\xi}}\} & \text{for } \xi \neq 0 \\ \exp\{-e^{-x}\} & \text{for } \xi = 0 \end{cases} \quad (3.7)$$

The particular case of $\xi > 0$ characterize the heavy-tailed distribution to which we acquire a special interest. In fact, authors such as Guillaume and Dacorogna (1997), Longin (1996) and Loretan and Phillips (1992) found empirical evidence that series of return in the stock market or the foreign exchange market usually present heavy tails. It is safe then to assume that series of returns in finance are heavy tailed and we will focus our study on the particular case of heavy tails.

Thanks to Gnedenko (1970), we can write a simpler formulation of the heavy tailed distribution F .

$$\overline{F}(x) = 1 - F(x) = x^{-\frac{1}{\xi}} L(x) \quad (3.8)$$

Where L is a slow varying function.⁵

The importance of the formulation in equation (3.8), is that it is easier to understand the signification of the parameters ξ . Distributions with a large value of ξ have fatter tails and by consequence the occurrence of extreme events is more frequent in that case. A series of distribution such as the *Student-t*, *Pareto* are in the class of heavy-tailed distribution. Negative values of ξ indicate there is a short tail distribution, which means that the maximum values are capped. $\xi = 0$ indicates distributions with exponentially decaying tails such as the Normal distribution. Finally, $\xi > 0$ is for the class of heavy-tailed distribution. The greater the value of ξ the slower the tail decay and by consequence the greater the probability of extremes occurring.

A widely used technique to estimate the tail factor is a semi-parametric approach that uses a Hill type estimator. The strategy of the estimator is defined in Embrechts

⁵ The function L is said to be slow varying if

$$\lim_{x \rightarrow \infty} \frac{L(tx)}{L(t)} = 1, \forall t > 0$$

et al. (1997). The idea is to choose a threshold point x_0 such that all observations exceeding that point are considered to be from a *Pareto* distribution. The set of observations exceeding the threshold will be used to form a maximum likelihood estimator for the tail parameter $\alpha = \xi^{-1}$ that we need to estimate.

We can set x_0 to be the p quantile of gains and losses observations $(X_j)_{1 \leq j \leq n}$. We choose to denote it by VaR_p to be in line with the risk management notations. Then assuming that all observations exceeding the VaR_p belong to the tail that can be approached by a *Pareto* like distribution, we can estimate the tail factor $\hat{\alpha}_{p,n}^H$ as being :

$$\hat{\alpha}_{p,n}^H = \left[\frac{1}{\sum_{j \in Exceeds} 1} \left(\sum_{j \in Exceeds} \ln(X_{j,n}) - \ln(VaR_p) \right) \right]^{-1} \quad (3.9)$$

Where *Exceeds* is the set of observations that exceed the VaR_p . It is common to choose the risk horizon for losses to be 10 days and the probability level to be 99% due to the regulatory requirement of the BIS. However, in the context of the choice of the optimal threshold to compute the tail factor, it is recommended to use a Hill-plot to visualize the region where the risk factor is robust for the choice of threshold. The convergence properties of the Hill estimator are very thoroughly studied in the literature about heavy-tailed distributions. The consistency of the estimator is established under some technical conditions. For further discussion about the Hill estimator please refer to Resnick and Stărică (1995), Resnick and Stărică (1998) and Embrechts et al. (1997).

3.3.2 Tail imbalance factor

In heavy tail theory, the modeling of extremes is usually focused on one side of the distribution. However, the theoretical foundation allows for the evaluation of the tail index for both tails. Our proposed heuristic is based on the balance between

the tail index in both tails of the distribution. Assuming that both tails of the profit and losses distribution have fat tails, following equation (3.8) we have that :

$\exists \xi_g, \xi_l > 0$, and two slowly varying function L_1, L_2 such that:

$$\begin{cases} \overline{F}(x) = x^{-\frac{1}{\xi_g}} L_1(x) \\ F(-x) = (-x)^{-\frac{1}{\xi_l}} L_2(-x) \end{cases} \quad (3.10)$$

ξ_g and ξ_l are called respectively the gains tail factor and the losses tail factor.

In this context we will define the tail imbalance factor λ to be :

$$\lambda = \frac{(1 + \xi_g)}{(1 + \xi_l)} - 1 \quad (3.11)$$

Where $\alpha_g = \frac{1}{\xi_g}$ and $\alpha_l = \frac{1}{\xi_l}$

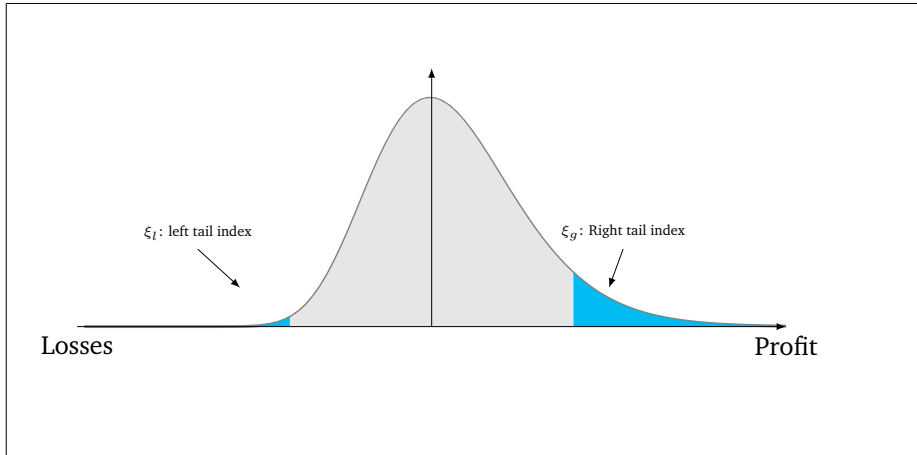


Fig. 3.9.: Tails imbalance in the distribution of gains and losses for a financial institution

The tail imbalance factor is a measure of the asymmetry of the tails as it expresses the ratio of both tails factors. Different intuitions are behind the choice of the formulation of the tail imbalance factor. It is easy to see that the first order Taylor expansion of λ is actually the difference between ξ_g and ξ_l . Second using Karamata's theorem ⁶, this formulation leads to equal conditional expectations for values ex-

⁶Embrechts et al. (1997)

ceeding thresholds. In other words, the difference between conditional gains and conditional losses will decrease when the imbalance measured by λ is fading. Distributions with symmetric tails will yield approximately the same tail factors ξ_g and ξ_l for both tails. In those conditions, the tail factor will be close to zero. Distributions with tail index ($\xi = 0$) have exponentially decaying tails such as the normal distribution. Hence we consider the zero point as the reference for no detectable externalities.

3.4 Measuring the bank externalities based on asymmetry

Risk measures such as the VaR or Expected Shortfall (ES) are very popular in the financial industry because they are easy to grasp and then derive policy implication based on them. The reason is that they are measured in monetary units and represent real potential losses. While the tail imbalance factor can detect externalities based on its sign, the absolute value is hard to interpret. Moreover, this measure does not take into consideration the size of the financial institution and by consequence can only compare banks of equivalent sizes. To remediate to such shortages, we will introduce a new measure of banks externalities that we call Value of Externalities (VoE). The proposed measure is inspired from the estimator of the VaR in the case of extreme value theory and based on the tail factor (see for example Embrechts et al. (1997)). Again in the context of extreme value theory, it is possible to have an estimate of the Value at Risk based on the empirical distribution of losses and the Hill estimate of the risk factor. First let $X^{(1)} \geq X^{(2)} \geq \dots \geq X^{(n)}$ be the order statistics of a historical sample of losses of size n . Assuming $u = X^{(k)}$ a very high threshold and $\frac{k}{n}$ the probability associated with u (from the empirical distribution) A proposed estimate of the VaR is :

$$\widehat{VaR}_q(X) = X^{(k+1)} \left(\frac{n}{k} (1 - q) \right)^{-\hat{\xi}} \quad (3.12)$$

This estimator assumes a Pareto type shape of the tail. In equation 3.12 the historical quantile $X^{(k+1)}$ is corrected to take into account the tail fatnesses. The

correction factor is greater than one for positive value of ξ . This formula will be our inspiration for the design of the *VoE*. Consequently, the estimator of the Value of Externalities will be defined as follows :

$$\widehat{VoE}_q(X) = \left| X^{(k+1)} - X^{(n-k+1)} \right| \left(\frac{n}{k} (1 - q) \right)^{-\hat{\lambda}} \quad (3.13)$$

Where $\hat{\lambda}$ is :

$$\hat{\lambda} = \frac{\max(1 + \xi_g, 1 + \xi_l)}{\min(1 + \xi_g, 1 + \xi_l)} - 1 \quad (3.14)$$

Reason for the choice of this form for *VoE* The most straight-forward measure of externalities that we can use is simply the difference in the α -quantile and $(1 - \alpha)$ -quantile corrected by the tail fatness each time⁷. There are two reasons behind the choice of this form of equation to measure externalities. The first is that we wanted to translate the tail imbalance factor $\hat{\lambda}$ directly into a measure that is easier to grasp. We think that $\hat{\lambda}$ carries important information about the business model of banks and its risk profile and we wanted to include it directly into the equation and we will simply loose this important parameter if we adopt other form of externalities measure. The second and most important argument of the choice of this equation is that it actually gives more weight to the difference in tail fatness compared to the difference in the quantiles itself. The measure that we propose actually corrects the difference first to the fatness of the left tail and later to the fatness of the right tale separately. If we correct the right and left quantiles each by the corresponding tail fatness and then take the difference, we risk that the *VoE* will be dominated by the difference in quantiles and not by the tail asymmetry. In such case, we are back to the traditional drawback of *VaR* which that it does not detect risk hiding in the tail.

Complementarity of *VoE* and *VaR* We must keep in mind that the *VoE* is a not an alternative standard downside risk measure, such as the *VaR*, but is rather complementary. *VoE* is not designed to monitor losses but should always be considered in

⁷ $\widehat{VaR}_q(X) - \widehat{VaR}_{1-q}(X)$

the presence of a standard risk measure based on losses like VaR or ES. In fact, VoE cannot detect tail fatness in both sides. It is the difference that has an impact on the value of externalities. For example, banks taking excessive risks concentrated in one type of assets should have a very high VaR . But simultaneously, if those risks can also yield important gains in favorable conditions the potential losses will be completely offset by the potential gains and the VoE will be close to zero. However, it is obvious that the bank in this situation is endangering his viability and maybe the viability of the financial system itself. However, measures like VaR can easily detect those risky behaviors of banks. The VoE has advantages in detecting negative externalities only when combined with downside risk measurement. Under banking regulation, financial institutions should already monitor their losses. Hence the loss tail should be thinned to avoid costly capital provision. The value of the VoE should be zero for tail symmetric distribution which means limited externalities as well. One straightforward implication of the VoE , is that it may prevent from the risk hiding in the tails that may be an effect of regulations and the traders compensations system. In fact, the compensation system for trades usually provides asymmetric incentives because traders will receive high bonuses in case of highly risky strategies that will pay-off. In the same time, an equivalent amount cannot be drawn back from them simply because their salaries cannot go below zero. Such behavior can create highly asymmetric distribution encouraged by the regulation that tends to ignore risk far in the losses tail which is the case of regulation based on measures such as the VaR . Eventually, an important tail asymmetry could be a strong signal for regulators and policy makers that the banks portfolio is hiding a large amount of risk that may endanger the system once revealed.

Notice that in equation (3.13) we used $\hat{\lambda}$ to characterize asymmetries instead of using λ . First of all, $\hat{\lambda}$ a slightly modified version of λ that measures the absolute tail asymmetry without any specific attention to the direction of imbalance (whether the right tail is heavier or the opposite). The rational behind this choice is that as opposed to individual risk measures we think that systemic risk indicators should consider on its equation the system adaptation and mutation due to the regulation based on the indicator itself. The best way to satisfy such features by the VoE is that a bank cannot decrease its own externalities without decreasing the total externalities hidden in the system. This measure should not tolerate risk transfer

between banks. In other words, the increase of a single institution's welfare should be coupled with an increase in the welfare of the system as a whole.

3.4.1 Data description

We construct a sample of publically listed banks and financial institutions in general included in the NYSE financial index. The sample includes the major US financial institutions traded on the NYSE. About 180 financial institutions are included in the sample. Prices and market capitalizations are downloaded from yahoo finance for the period from 2000 to 10-2014.

Tab. 3.1.: Summary statistics of the data of 10-days log returns for the period beginning in 2000. *1%Diff* is the difference between the 99% quantile and the 1% quantile: a first indicator of tail symmetry. *Annulized Vol* is the annualized volatility. *Skweness* is the skewness of the distribution of the 10 days log returns

2000-2003				2004-2006		
	1% Diff	Vol	Skweness	1% Diff	Vol	Skweness
10% Quantile	-3.559	26.914	-0.567	-4.675	15.116	-0.827
Median	1.361	35.045	0.0361	0.840	21.103	-0.0464
90% Quantile	7.389	54.726	0.584	3.625	32.829	0.472
2007-2009				2010-2014		
	1% Diff	Vol	Skweness	1% Diff	Vol	Skweness
10% Quantile	-19.491	37.967	-1.312	-4.514	18.961	-0.826
Median	-5.442	62.111	-0.435	-1.583	29.111	-0.411
90% Quantile	1.224	108.186	0.134	2.418	43.266	0.170

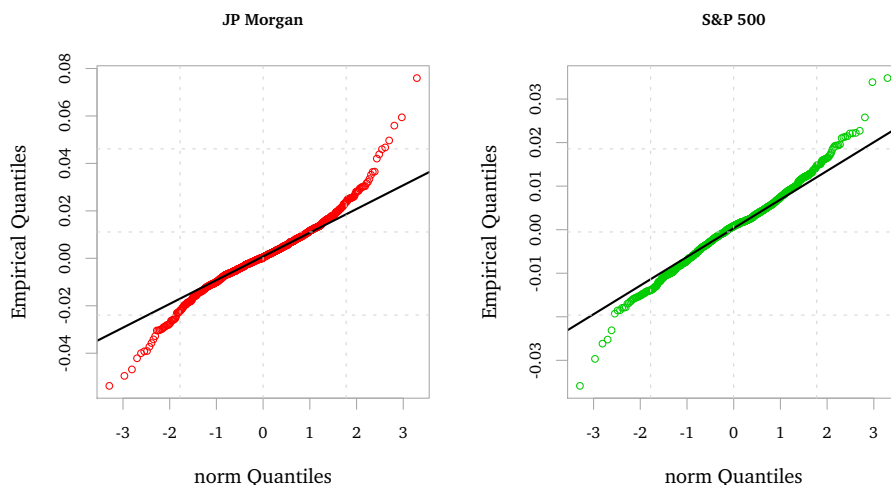


Fig. 3.10.: In the left, we show a Q-Q plot for the returns of *JPMorgan* for the period 2002 – 2006. The right plot is the returns of the *S&P500* index for the same period.

Table (3.1) gives a quick overview on the symmetry structure of the returns distributions for different periods from 2000 to 2013. The 1 % *Diff* variable which is the difference between the 99% quantile and the 1%*quantile* indicates a sign a tail asymmetry in the data. Nevertheless, the median value is close to *zero*, which suggests that the market has some aspect of tail symmetry on average. This idea is also visible in the $Q - Q$ plot for the return of the *S&P* 500 index returns displayed in figure 3.10. The table also exhibits the changes that occurred to the symmetry structure between different periods. The indicative measure of asymmetry presented suggests that the embedded stress in the system due to an imbalance between losses and gains is reduced in the time of high volatility.

Although table (3.1) and figure (3.10) suggests that the overall system may have some characteristics of balanced distribution with symmetric tails, it also strengthens the feeling that some banks do have strong tail asymmetry that can be a source of fragility and embedded stress in the financial system. This asymmetry is visible in the QQ plot of the returns of *JP Morgan* for example.

3.4.2 Results

Results on the value of the *VoE* are summarized in table (3.2). More specifically, we reported the banks that made it to the top 10 of the ranking for contribution to the system fragility throughout externalities according to the metric of *VoE*. We only report the top 10 of each period. For instance, institutions like *AIG* are not included in 2006 but are part of the reported banks in 2007. The results are reported in term of relative contribution to the overall embedded stress in terms of percentage of the total *VoE* of the institutions in our sample. Because we only show the top 10 the reported values does not sum up to 100%

Obviously, any analysis based on returns cannot discriminate effect based on the size of the form, we corrected to *VoE* to the size by simply multiplying by the relative size of each bank in the *NYSE* financial index.

At this stage, it is worth making some observations about figures in table (3.2). The most important finding is that the top 10 banks are the of origin more of than 50%

Tab. 3.2.: Table represents the top 10 US financial institutions for the periods of 2006 to 2013 according to the VoE metric. SVoE is the percentage of the VoE compared to the total VoE of the financial institutions in our sample. Note that *VoE* is corrected by the relative size of each institution in the system

2006		2007		2008		2009	
Ticker	SVoE	Ticker	SVoE	Ticker	SVoE	Ticker	SVoE
MS	3.30	WFC	3.09	BBT	1.42	STI	2.16
MET	3.32	AIG	3.36	BEN	1.74	GS	2.37
PRU	3.48	SCHW	3.50	GS	2.63	JPM	2.57
BEN	3.56	MET	3.94	COF	3.23	PRU	2.91
WFC	4.38	AON	4.71	AXP	4.70	BLK	2.93
GS	4.63	PRU	5.05	MS	4.89	AXP	4.13
AXP	4.77	GS	5.05	JPM	6.95	MS	5.74
C	4.89	TRV	5.19	WFC	8.83	WFC	6.16
BAC	5.76	JPM	5.75	AIG	9.91	BAC	10.58
JPM	5.98	AXP	6.03	BAC	11.24	AIG	12.45
BLK	7.72	BLK	6.87	C	21.76	C	16.81
2010		2011		2012		2013	
Ticker	SVoE	Ticker	SVoE	Ticker	SVoE	Ticker	SVoE
COF	2.17	BK	2.16	AIG	2.64	MS	2.30
USB	2.56	USB	2.20	OCN	2.70	PNC	2.49
WFC	2.59	ACE	2.58	COF	2.95	AIG	2.92
RF	2.70	MS	2.67	GS	3.71	BLK	2.95
BK	3.31	AXP	3.06	BLK	4.38	OCN	3.07
STI	3.70	PNC	3.09	MET	4.53	COF	3.99
AXP	4.29	GS	4.14	PRU	6.04	USB	5.60
MS	6.77	COF	5.10	BAC	6.39	C	6.38
AIG	7.94	JPM	12.04	WFC	6.51	JPM	9.51
C	15.29	BAC	13.08	USB	8.95	BAC	13.07
BAC	18.27	WFC	19.40	JPM	10.22	WFC	18.98

of the externalities existing in the system and that contributes to the overall system fragility. What can be even more striking is that the top 5 institutions capture almost the third of the total *VoE*. It is also important to notice that financial institutions that are listed in our ranking are also the most important banks in terms of market capitalization and activities. One possible explanation of the externalities' concentration is the implicit too big to fail guarantee (TBTF). Recall that, we consider that what give the bank the opportunity to make gains and not fully paid by it is an externality. This concept also applies to the TBTF that is a guarantee fully supported by the tax payers and not directly by the financial institution. We also notice that this concentration is more relevant starting from 2008 as it was apparent to the market after the fall of Lehman Brothers that the government would not allow for a second failure and that intensified the implicit guarantee for the TBTF banks.

Moreover, it is interesting to follow the evolution of the ranking of some of the biggest commercial banks in the US, for instance, Bank of America(BoA) and Citi bank. Before 2007, BoA was considered to be a conservative institution relative to another big bank in the US. However, at the end of 2008, the bank was ranked top 2 relative to externalities which correspond to the acquisition of Merrill Lynch and

Wachovia. Citi bank and AIG were both ranked in the top 3 at the end of 2008. Both financial institutions were the first to feel the heat of the financial crisis because of their deep implication in the mortgage business and are also one of the first to benefit from the bailout program. Again, according to our definition, those bailout could be seen as externalities which in turn can explain those important ranking according to our measure.

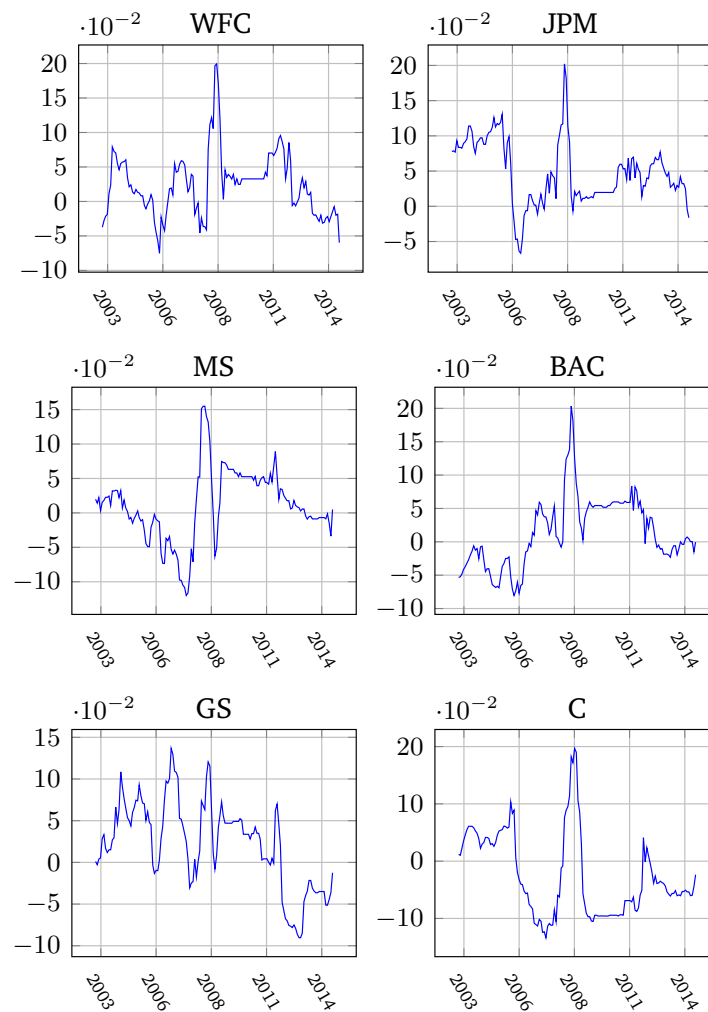


Fig. 3.11.: Those figures show the evolution of the tail imbalance factor λ for the period of 2003-2014 for six major US banks. The value is updated monthly with a fixed timespan of 4 years. The value for the year 2006 means that we include all the returns up to 12-2006.

Figures (3.11) and (3.12) show the time series evolution of both $\hat{\lambda}$ and VoE between 2003 and 2014. The important conclusion is that all banks feature a spike in their externalities around mid-2008. Precisely in 2007, we see that the value of

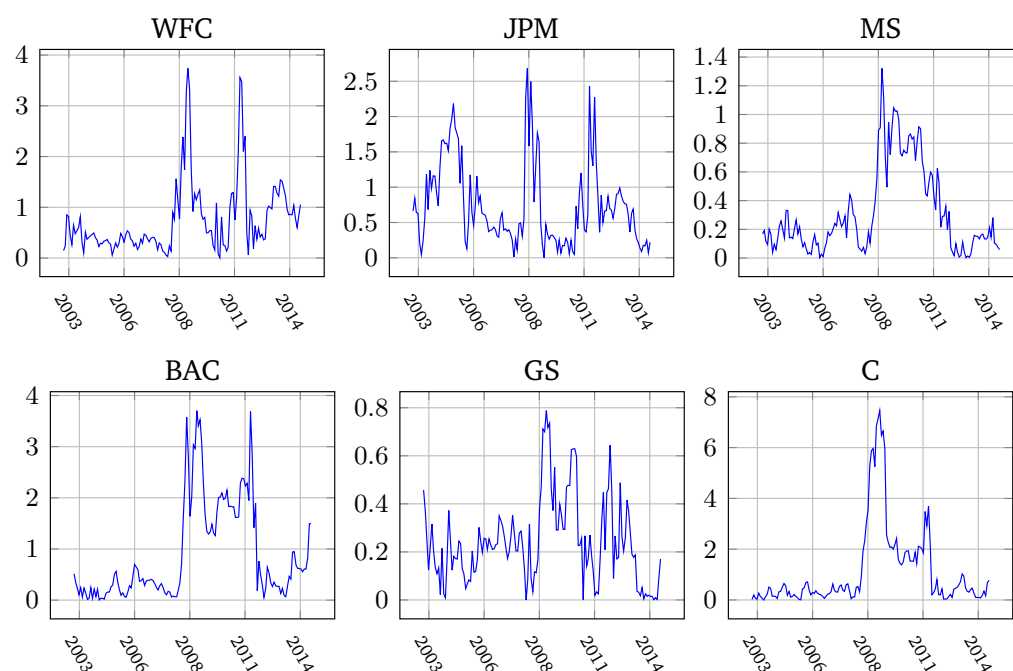


Fig. 3.12.: Those figures show the evolution of the VoE for the period of 2003-2014 for six major US banks. The value is updated monthly with a fixed timespan of 3 years. The VoE is computed for $q = 99\%$. The value for the year 2006 means that we include all the returns up to 12-2006.

externalities of all banks started to increase sharply which could be translated by an important externalities and systemic fragility. These spikes fall right after 2008 which mean that the hidden stress has materialized into a full-scale systemic crisis. According to these figures, in 2007 the regulator should have been alarmed by the declining health of the financial system. Of course, retrospectively we know that the market already had indication about the situation of the financial system starting from 2007. But at that time, opinions were mitigated about the scale of the hidden stress in the system. Using an indicator such as the VoE could have shown that banks were reaching unforeseen level of externalities and that the system would soon or later collapse to this stress.

At this stage, it is important to draw the attention to the fact that the methodology that we propose should not be interpreted as an alternative to the measures of systemic risk proposed by Acharya et al. (2017) and Adrian and Brunnermeier (2016). In fact, the last two approaches could be seen as measures of simultaneous failure of the system and a financial institution and by consequence focus on the

system in time of a crisis. The main insight that both measures propose to regulators is the identification of financial institutions that require special attention in time of distress either because they can be identified as SIFI or that their survival is at stake.

Our, by contrast, propose a measure of externalities of banks based on the principle of tail asymmetry that can be the source of system fragilities. The constant gathering of those fragilities may lead to a systemic crisis. It is the evolution of the measures that we propose that can enlighten regulators on financial institutions that adventure at activities with high externalities and are pushing the system toward its collapse. Regulators should react when they observe the increase of the level of asymmetry in time of economic stability as this can be seen as a premonition of future turmoil. Moreover, this indicator of symmetry can encourage banks to re-examine their relationships with their counter-parties that are highly exposed to positive shocks at the expense of the system fragility. This peer evaluation by banks to other banks' activities will undoubtedly contribute to the overall stability of the system.

Nevertheless, the approaches of Acharya et al. (2017) and Adrian and Brunnermeier (2016) do have some common grounds. We also rely on publicly available market data to be able to compute and evaluate the system stability based on our measures. While we share all the advantages of using market data, we also suffer from the curse of all the critics that could be addressed to methodologies using stock prices.

However, it is important to notice that the key concept of tail symmetry could also be applied to bottom-up approaches such as the Basel methodology to compute VaR. In theory, banks using their internal risk assessment models to evaluate their risk measures can also rely on the same model and simulations to deliver information about their right tails. In addition, the perfect conditions would be regulators who have access to all transactions of the financial institutions and then able to reconstruct their portfolios. Then using the available information and unified pricing model compute metrics about tail imbalances and react accordingly. With the use of unified pricing models, tail imbalance could be easily unjustified and even rep-

rehensible by regulators. This ideal situation may seem out of reach in the present conditions. Nevertheless with the development strategy of the OFR to construct a complete data warehouse recording all the transactions in the banking system, and the power given by the Dodd-Frank (2010) to this office, the application of a monitoring system of tails imbalance fairness seems possible. Meanwhile, relying on market data appear to be the best alternative.

3.5 The post-financial crisis fines

Two years after the peak of the financial crisis, the US market has experienced a major shift itself a result of a change in Obama's administrations' policy. In fact, after the biggest bailout procedure in the history of the banking industry, Wall Street banks and the major foreign banks operating in the US have paid out more than \$ 100 billion in fines. To grasp the magnitude of those fines, it is important to highlight that the Supervisory Capital Assessment Program (SCAP) conducted stress tests in 2009 on the 19 largest US banks. They concluded that the banks needed to raise \$ 74.6 billion if the economy were to get worse. In addition, these fines to the banking industry set a record in the history of legal settlements in the US and are only exceeded by the historical litigation of the tobacco industry however on a period of 25 years ending in 2025. The fines reflected that the willingness of the Obama administrations to persuade the general public that the bankers would not get off lightly for their role in the ignition of the financial crisis. Besides the political motivation of those fines, they are much in the spirit of our analysis of externalities in this paper. The fines cover the banking practices in different business areas ranging from lending to fraudulently issuing mortgage-backed securities and market manipulation.

Of course measuring *ex-ante* the externalities that resulted from the banks' activities was generally considered to be a difficult task, the fines that resulted from a thorough analysis of the collateral damage of those externalities could be seen as an *ex-post* credible measure. The data on fines and penalties billed to the financial industries are collected and updated regularly by the Financial Times on their web-

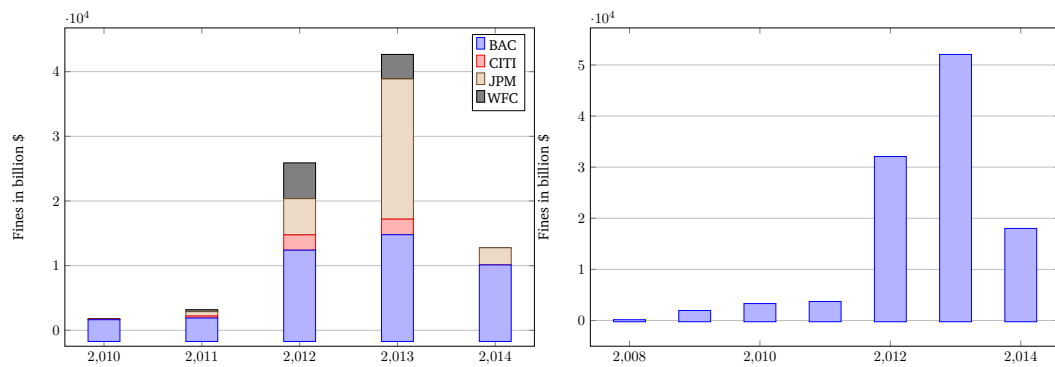


Fig. 3.13.: Fines paid by 4 US banks from **Fig. 3.14.:** Total fines claimed from the banking industry 2010-2014

site ⁸. The data is collected between 05-2007 and 05-2014. The 4 biggest US banks (JP Morgan & Chase, Wells Fargo, Citigroup, Bank of America) total as much as 57.1 billion \$ of fines. Foreign-based banks with activities in the US such as HSBC and Deutsche Bank have a bill of over 15.5 billion \$.

Using fines to proxy externalities in the financial industry is also one of the originality of this work. Nevertheless, the idea of linking fines to social welfare and externalities in industries such as coal or tobacco industries goes back to 1920 in Pigou (2006). Similar proxies are unused in other disciplines like the transportation industry where fines are considered as the cost paid by consumers and the industry to cover the social costs of environmental externalities (see for example Hultkrantz et al. (2012)).

The question in this section is to what extent our newly designed measure captures the level of externalities measured *ex-post* via the fines imposed on banks. More precisely, we will also compare the *SVoE* to other measures of systemic risk such as the *MES* of Acharya et al. (2017) and *SRISK* by Brownlees and Engle (2011). For that purpose, table 3.3 provides an OLS regression analysis that explains the fines with regressors *Voe*, *MES* and *SRISK*. The *Voe* is based on the *Voe* computed following the methodology explained in the previous section and scaled by the share of the domestic activities of the US banks. The domestic activity is evaluated from the ratio of foreign assets divided by the total assets of the banks. The data for US banks are extracted from the FR Y-9C report ⁹. For non US banks, we used the

⁸see <http://blogs.ft.com/ftdata/2014/03/28/bank-fines-data/>

⁹reports are available at <http://www.ffiec.gov/>

geographical distribution of revenue to proxy the activities in the US. Data were extracted from the annual report of each bank available on their websites. *MES* and *SRISK* were extracted from the V-Lab website ¹⁰ where those measures are regularly maintained and published by a large panel of banks. The extraction was made as of 31-12-2007 and return for the years of 2004-2007 were used to compute the value of *VoE*. The choice of time frame is due to the fact that we want to access the ability of each of those measures to forecast *ex-ante* the observed externalities of the banks after the crisis.

Tab. 3.3.: The dependant variable is the amount of fines. Model (1) to (3) are OLS regression with a single variable in the model. Model (4) is an OLS regression with *MES*, *SRISK* and *VoE* as regressors. *MES*, *SRISK* are computed as of Dec-2007 and *VoE* is evaluated as of dec-2007 with a 4 years window.

	(1)	(2)	(3)	(4)
MES	8.176			−6.624
SRISK		0.9881*		5.601
VoE			6.035***	5.429***
Adjusted R^2	−1.952%	12.91%	44.29%	43.06%

The striking result from Table 3.3 is that *VoE* is highly significant in regression (3) and (4). For example in regression (3) where only *VoE* is included as an explanatory variable the *t* – value is 5.929 with an *adjusted* – R^2 equal to 44.29%. It is also important to notice that the variable *SRISK* had little significance in the regression (2) with a *t* – value of 2.925 and an *adjusted* – R^2 equal to 12.91%. Nevertheless, this variable lost all its significance when regressed with the *VoE*. Also, note that the *adjusted* – R^2 had a small drop between regression (3) and (4) which indicated that *MES* and *SRISK* had a small marginal added-value to predict externalities compared to *VoE*. The important point is that the *VoE* seems to better capture the externalities measured via fines and penalties than *MES* and *SRISK*. Of course, this is not a fair horse-race. The *MES* and *SRISK* were specially designed to capture the shortage of capital in the financial institutions during the periods of turmoil. They do not take into consideration the externalities or the social costs associated with the banker’s activities on the system. Nevertheless, *MES* does a good job forecasting the capital shortage and seems to predict for example the results of the stress test that were performed on banks in 2009 by the SCAP. The bottom line

¹⁰<http://vlab.stern.nyu.edu/>

of this analysis is that the VoE designed to capture externalities should not be interpreted as an alternative to systemic risk measures such as the MES and $SRISK$ and the policy implication of both approaches are different. While high $SRISK$ indicates that the banks have a capital shortage that could require a government intervention during a crisis situation, a high value of the VoE suggests that the banks have important externalities and have an important marginal contribution to the system's fragility.

Conclusions

The systemic crisis of 2008 initiated a series of research to identify important financial institution that the system cannot survive their failure. Nevertheless, this research had little focus on the pre-crisis era where risk is building up in the background to burst in the form of a crisis. This paper introduces the idea that banks should have symmetric tail in order to limit the systemic risk created by their activities. While we consider that skewness is acceptable in financial markets because it is the result of different expectations, tail asymmetry is not tolerated because if extreme losses are unpredictable so should extreme gains. The failure to pay attention to the imbalance between potential gains and potential losses in prosperous times allowed banks to increase their externalities. In fact, we translate the difference between tail asymmetry into a measure of banks externalities that could result in a systemic crisis. It was shown using both a theoretical model of banks interaction and Monte Carlo simulations that the asymmetry of tails of banks can lead to a riskier financial system. In addition, this paper proposes an estimation procedure to overcome the data availability problem. In fact, we were able to propose a measure of externalities based on publicly available price data. The measure proposed is building on extreme value theory and the hill estimator of tail factor. We also performed an *ex-post* test of the $SVoE$ measure considering the fines and penalties paid by the banking industry after the crisis of 2008. We show that the $SVoE$ can have greater explanatory power than measures of systemic risk such as $SRISK$.

The results of this paper can have direct policy implications. As tail asymmetry is proven to be potentially harmful to the long run survivorship of the financial system

and the sustainability of the financial services, it is suggested that regulators should regularly monitor both right and left tails of the profit and losses distribution of banks. It was suggested that long period of financial prosperity and growth can give premonition of future crisis. This paper provides the theoretical and empirical evidence that support such a claim.

A. Proof of theorem 2

Proof of theorem 2 The proof will be done via induction on the size N on the financial system. Without any loss of generality we will suppose that the Profit and Losses distribution have zero means for all financial institutions.

Part 1 : Case of size 2 In the context of two banks financial system the value of equity v_i , $i = 1, 2$ of each bank is :

$$v_1 = L_{21} - L_{12} + e_1 \text{ and } v_2 = L_{12} - L_{21} + e_2$$

Starting from $\mathbb{P}(v_1 \leq -\delta|E) = \mathbb{P}(L_{21} - L_{12} + e_1 \leq -\delta|E)$ and given the assumption of a systemic financial crisis, we have that $v_1 = L_{21} - L_{12} + e_1 \approx L_{21} - L_{12}$ and $v_2 = L_{12} - L_{21} + e_2 \approx L_{12} - L_{21}$. Then we can conclude that $\mathbb{P}(v_1 \leq -\delta|E) = \mathbb{P}(L_{21} - L_{12} \leq -\delta|E) = \mathbb{P}(L_{12} - L_{21} \geq \delta|E)$. Finally this leads to :

$$\mathbb{P}(v_1 \leq -\delta|E) = \mathbb{P}(v_2 \geq \delta|E)$$

On the other hand, we have that the Profit and losses distribution of bank 2 is tail symmetric. In other words, we can find $\psi > 0$ where $\forall \alpha > \psi$, we have: $\mathbb{P}(v_2 \leq -\alpha|E) = \mathbb{P}(v_2 \geq \alpha|E)$

With the combination of the results on crisis conditions and tail symmetry we have that :

$$\delta > \psi \quad \mathbb{P}(v_1 \leq -\alpha|E) = \mathbb{P}(v_2 \leq -\alpha|E)$$

Part 2 : Case of size 3

In the context of three banks in system the value of equity v_i , $i = 1, 2, 3$ and of each bank is :

$$v_i = \left(\sum_{j=1}^3 (1 - \delta_{ij}) L_{ji} - \sum_{j=1}^3 (1 - \delta_{ij}) L_{ij} \right) + e_i$$

Where δ_{ij} is the kronecker factor. We can start from :

$$\mathbb{P}(v_1 \leq -\delta|E) = \mathbb{P}((L_{12} + L_{32}) - (L_{21} + L_{22}) - (L_{32} - L_{23}) + e_1 \geq \delta|E)$$

Given the assumption of a systemic financial crisis, we have that

$$v_1 = \left(\sum_{j=1}^3 (1 - \delta_{ij}) L_{ji} - \sum_{j=1}^3 (1 - \delta_{ij}) L_{ij} \right) + e_1 \approx ((L_{12} + L_{32}) - (L_{21} + L_{22}) - (L_{32} - L_{23}))$$

Then we can write that $\mathbb{P}(v_1 \leq -\delta|E) = \mathbb{P}(v_2 \geq \delta + (L_{32} - L_{23})|E)$. Regarding that δ is rather on the tail and is a significant of a big loss then we can reasonably assume that $\delta + (L_{32} - L_{23}) \approx \delta$. Again, with the combination of the results on crisis conditions and tail symmetry for bank (2) we have that as in part (1) we can conclude that

$$\delta > \psi \quad \mathbb{P}(v_1 \leq -\alpha|E) = \mathbb{P}(v_2 \leq -\alpha|E)$$

Part 3 : Induction

Next, we claim by mathematical induction that this Theorem 2 is true for all financial system with size $\leq k - 1$, and let us prove that this equality is valid for a banking system with size k . First, let us consider a financial system $S = (L, e)$ with k banks acting in the system. Let us imagine that bank C is the fusion of bank 1 and 2. Without any loss of generality, the bank should be chosen as a solvent bank. The case of the merger of two solvent banks will have no impact on the solvency and cash-flows on other banks in the system because every bank will continue to pay its obligation toward the bank. However, if the bank 2 is insolvent the merger of the two banks 1 and 2 will only be a bank C where expected payments from the bank C are the expected payments of bank 2 and the obligation of any financial institution with regards to C are the obligation toward the second bank plus the net obligation toward the failing bank. Because the bank, can only pay a fraction of its obligation in case of default. In other word, a default will always have its costs on the system. Such merger is always possible, and bank will have the incentive to perform it as presented in Rogers and Veraart (2012).

After the fusion of banks 1 and 2 into a single bank C , we have a new financial system $S' = (L', e)$. With the merger of two banks, the new financial system is of size $k - 1$

According to the induction hypothesis, we have that :

$$\exists \psi_1 > 0 \text{ where} \tag{3.15}$$

$$\forall i, j \text{ a bank } \in S \text{ and } \forall \delta > \psi_1 \quad \mathbb{P}(v_i \leq -\delta|E) = \mathbb{P}(v_j \leq -\delta|E)$$

So far, we have established the desired equality for $k - 2$ bank in the initial system S . Choosing a different financial system S'' , where this time we will merge the bank 1 with a different bank will make the relationship applicable for the bank 2. This is true because the system S and S'' will have banks in common. Finally, to prove the extreme losses property for the financial system including the non-defaulting bank 1. We will distinguish

the case where we have a second non-defaulting bank from the case where all banks beside bank 1 are defaulting. In the first case, to demonstrate that equation 3.15 holds even if we are including the bank 1 we can simply merge the second non-defaulting bank with any of the banks except bank1. Then the equality becomes evident. Finally, the case were all banks besides bank 1 will default is also simple. Notice that the equation 3.15 holds for the merger of bank 1 and 2. But the bank 2 is defaulting without inducing any costs on the system. Therefore, the value of the bank C is equal to the value of the bank 1. Thus the equation 3.15 holds for the bank 1 as well.

Finally, via mathematical induction we can establish Theorem 2

A. Illustration of λ as a measure of externalities

While the main purpose of this paper is to measure externalities in the financial sector, the measurement procedure could be applied to any series of returns. It is possible to apply this technique to other sectors of the economy. The idea is that the concept of externalities in the financial sector is very similar in principle to pollutions and greenhouse gas emissions in industry. The analogy is based on the idea that both pollution and bank's externalities are side effects that are not fully assumed by the emitting entity. The rational behind this empirical test is that we can observe the behavior of λ for sector that their externalities are identified and measured.

We evaluated the value of λ for different sectors of the economy. In fact, our analysis is based on the return of index funds that tries to track some sector indices globally. The funds are managed by BlackRock, and their returns are published on a daily basis.

Tab. 3.4.: Tail imbalance factor computed for a list of global sector indices based on ETFs managed by Ishares BlackRosck

λ	Sector
Energy	
0.01	Global Energy
-0.03	Global Clean Energy
0.04	Global Nuclear Energy
-0.04	MSCI Global Energy Producers
0.02	MSCI Emerging Markets Energy Capped
Financials	
0.02	MSCI Europe Financials
0.09	Global Financials
-0.01	MSCI Emerging Markets Financials
-0.09	MSCI Far East Financials
Materials	
0.07	Global Timber & Forestry
0.13	MSCI Global Metals & Mining Producers
-0.06	MSCI Global Agriculture Producers

Table (3.4) reports the value of λ for some indices in the energy, financial and raw material sectors. Positive values indicate sectors with important externalities, which means

that these industries failed to limit the impact of their activity on the environment and the society in general. Looking closer at the results for the energy sector we can notice that the clean energy index have negative λ as opposed to the positive value of the nuclear sector. This result indicates that the value of λ is coherent with our expectations regarding both nuclear and clean energy. A quick overview of recent history should indicate that the nuclear energy sector can have important negative externalities. For the raw material sector, mining and metals producers have positive λ which is in line with the highly polluting dimension of those industries. Regarding the financial sector, Far east and emerging market banks seem to have limited externalities compared to global banks. To conclude, this table seems to confirm that λ can be used to detect externalities. In fact, we were able to point out sectors in which their polluting nature is identified and are easily measured. This confirms the intuition that we can use this approach to detect externalities for the banking sector.

A game theory approach for systemic risk and international regulatory coordination

THE aim of this research is to find a theoretical justification toward the mutual benefits for members of a banking union in the context of a strategic interaction model. We use a unique contagion dynamic that marries the rich literature of game theory, contagion in pandemic crisis and the study of collaboration between regulators. The model is focused toward regulating asset classes, not individual banks. This special design address moral hazard issues that could result from government intervention in the case of crisis. The framework that we propose is flexible enough to be used in different settings. In a country such as the USA, the financial system is regulated by several regulatory bodies with different mandates and asset classes to supervise. While the Dodd-Frank Act developed the legal framework for collaboration between Federal agencies, this paper proposes to provide the academic justification.

Introduction

Unity makes strength is an old saying used by many countries like Belgium as their National motto. However, it is meaningless when costs are coupled with collaboration. The fortunate will balance benefits with expenditure to continuously help the misfortunate. The question of union becomes less evident when it concerns the financial system and the coordination in the situation of crises. The last ones dramatically highlighted that financial distress could quickly transform from a regional problem into a global phenomenon. It also underlined that sometimes remote region from the hypo-center of the crisis bear higher and mostly of all unexpected costs. In this paper, we ask about the importance of international coordination to mitigate financial crises in the presence of costs. We also question to what extent *Unity makes strength* or should rich countries restrict their policy regarding financial stability into their national scope.

Financial and economic reforms are key determinants of the stability of the financial system. Most of those reforms are dictated or suggested by supra-national regulators, group of experts like the Basel Committee on Banking Supervision (BCBS) in the case of banking regulation and IFIs. However, those reforms require important skills and resources to be implemented effectively. Imperfections in reforms implementation due to economic or political reason paired with more interdependent global financial system can create incentives for international collaboration.

The examples of contagion of crises are numerous in the recent history. The last financial crisis is probably a prototypical example where shocks on the housing market in the US has rapidly spread to foreign financial systems like Europe.

A second incentive for collaboration is that sharing the loss burden of an international bank *ex-post* requires challenging negotiation between the involved governments¹. The failure of the French-Belgian bank Dexia and the Dutch-Belgian Fortis Bank clearly pointed out that negotiation without a per-established cost sharing plan can be highly inefficient.

¹In this paper government and regulator are used interchangeably. We make no conceptual distinction between the two. Both terms refer to any authority capable of implementing a regulation than can have an impact on the risk of some financial assets

The later problem was the subject of huge discussion in the last years to establish *ex-ante* the resolution rules specially in the framework of the banking union in Europe with the introduction of bail-ins as a new resolution tool to use by authorities. The result was the creation of a Single Resolution Mechanism (SRM) under the umbrella of the ECB. Nevertheless, a handful of studies refers to the international collaboration prior to the crisis and the trade-off between the *ex-ante* costs of collaboration and the *ex-post* benefits from fewer contagion effects.

This paper tries to cover the gap in this strand of the literature. We propose a strategic theoretical model based on contagion in the natural environment to justify collaboration between similar actors in the financial networks such as central banks. Although the model is flexible to cover several settings, we focus on describing strategic interactions between regulators to sponsor financial reforms in other countries given the cross borders exposures and interlinkages. The literature on banking has studied extensively the subject of a banking unions in the resolution of crises. This particular strand extends the pioneering model of Diamond and Dybvig (1983) and later by Allen and Gale (2000) on multi-regional liquidity crises and conclude that a banking union leads to higher social welfare and lower crisis costs. Nevertheless, all those model fail to address an important issue in the financial regulation i.e. moral hazard. In fact, moral hazard arises because banks uses funds i.e. deposits that are guaranteed by the government. Systemic banks also that enjoy the very comfortable position of a too big to fail and the implicit guarantee that follows create moral hazard. With the introduction of bail-in, relying on bailouts for crisis resolution is outdated. Although the conclusion of prior models remains mostly valid their policy implications are to be revised. While some models try to account for the lack of monitoring of assets due to moral hazard, they all consider costly bailouts as the main resolution tool available to central banks. In our paper, we explicitly exclude any direct program that a central bank can initiate to protect a single bank or prevent its failure. Thus we rule out the bailout options from the regulator's toolbox. Moreover, this paper is inspired by the very rich literature on contagion in the natural world. We are not the first to do that. However our approach is unique and should be able to address both moral hazard and most of the critiques addressed by biologist to economist.

Economist and journalist always borrowed terminologies from the natural world to describe contagion in the financial system. However, the word contagion is new to the economic literature. In fact, before 1990, only 17 published works in economics had the word contagion in their text (Edwards (2000)). The interest of economist in contagion culminated after 1990 with the spread of crises from emerging countries affecting others with solid fundamentals. The July 1997 crisis that started in Thailand and then spread across the world was described as the "Asian Flu". Later, the Russian outbreak of 1998 was also known as the "Russian virus".

The definition of contagion in the biological world is straightforward referring to the transmission of a disease by direct or indirect contact. Economists did not reach any consensus on the definition of contagion in the financial context. Kaminsky and Reinhart (2001) consider contagion as the case which the knowledge of the existence of a crisis elsewhere increases the probability of a crisis at home. This definition could refer to a different type of contagion according to the understanding of the "elsewhere" and home. For others such as Edwards (2000), contagion refers to situations where the magnitude of the shock transmitted internationally exceeds what was expected *ex-ante* focusing only on spillovers. In this paper, we will stick with the first definition of contagion as it covers a wider spectrum of crisis transmission mechanisms and is commonly used by policy makers and international regulatory bodies.

Several examples of pre-crisis collaboration agreements exist between governments. For example, the major currency central banks i.e. the FED, the ECB, the Bank of England and the National Swiss Bank signed currency swaps. This arrangement allows those central banks to provide emergency funds to their banks not only in their home currency but also in other currencies like the Euro or the GBP for the case of the FED. A second example of collaboration is the IFIs sponsoring and supervision financial reforms in developing countries. While the main objective of those programs is not to prevent future contagion events but to help economic development, they can also contribute to mitigating the effect of financial turmoils originating in the beneficiary regions. Although the economic literature stress the importance of collaboration and the optimality of the central planner solution, nu-

merous examples of crisis handling situation can be used as an argument against collaboration.

In the famous case of AIG bailout, only 44 billion of dollars out of a total of a 100 billion dollars went to counter-parties establishing their headquarters in the US. Banks in countries such as France and Germany got the lion-share of the international payments. Although this may seem as American taxpayers were bailing out rich countries, it is even possible that this operation prevented a second round of failure that could have a higher cost on the US soil after accounting for international contagion. This type of revelation was the object of important discussion in the Congress that will probably have an impact on the future of resolution plans during the failure of internationally-integrated financial companies.

In this paper consider that a monetary unit of the risky asset is the atomic unit susceptible to contagion. Our approach contrasts with the previous works in the literature on financial intermediation that consider banks as the atomic unit in those models. This design has the advantage of eliminating the important moral hazard problem out of the regulator decision. This will allow for the possibility of analyzing general protection measures that are not oriented to help or protect a single financial institution deemed to be too big to fail. This modeling approach is coherent for example with the growing belief after the 2008 crisis that central banks should restrain themselves from targeting few financial institutions with their emergency measures. For example two Senators Warren and Vitter are trying to pass a bill unpopular in the financial industry in addition to Dodd-Frank act that goes in that direction. From a technical perspective, we propose to base this paper on the mathematical framework enriched by the biological literature to study contagion of deceases. Readers familiar with the later literature will find similarities between the dynamic of contagion that we propose and the famous SIR contagion model (Susceptible/Infected/Recovered or Removed) famous in biostatistics.

The remaining of this paper is as following. After a review of the literature about contagion and networks in the finance, we give a short introduction to the history of contagion in the financial world with an overview of the principal channels of transmission that were identified by the literature. In the next section, we present

the model that we propose while discussing the important hypothesis and rational behind each choice. Later, we present the solution of the strategic interaction problem.

4.1 State of the Art

Beginning from 1990, literature on financial contagion started to grow both on empirical and theoretical sides. The interest of researchers spiked after each episode of financial contagion as it reveals the importance of linkages in the transmission of crises. Most of those works try to uncover the role of a certain fragility in the financial system topology or dynamics in the spread of shocks. This theoretical strand is based on the pioneering paper by Diamond and Dybvig (1983) that was extended in the work of Allen and Gale (2000) and Diamond and Rajan (2005). In those papers, authors argue that some imperfection of the financial system such as liquidity shortage or poor monitoring in time of crisis can lead to imperfect coordination. As a result, it is possible to witness the failure of perfectly solvent banks. It is also worth mentioning the work by Eisenberg and Noe (2001) on cascade failure of banks. They study the negative effect of the failure of a single bank on the financial system. The paper also provides a clearing mechanism in that case. Papers like Elsinger et al. (2006) tried to use this methodology to studies vulnerabilities in existing national banking networks.

Empirical works can be divided into two important categories. First, papers like Degryse and Nguyen (2007), Artzner et al. (1999) and Gropp et al. (2009) used detailed banking data on the national level to identify structural deficits in the financial system. They also established the importance of contagion on the spread of shocks. Nevertheless, those studies seem to lack generality to draw policy recommendations outside the scope of the country under study.

The second stream of empirical work focuses on the analysis of granular data using the network toolbox and methodology. In fact, those papers try to identify network structural properties that relate banks via different channels. Their objective is to infer some policy implication in a time of crisis or detect vital nodes in the network.

Those papers typically try to highlight the important connection and clusters and the central nodes that play the role of a hub in the network. For details, see for example the work of Boss et al. (2004) about the Austrian inter-banking market and

Cont et al. (2013) who explored a very detailed database of the Brazilian banking system. Those works usually apply simulation exercises based on the extracted network properties to study the impact and spread patterns of financial shocks.

The main weakness of those models is that the network structure tends to change dramatically and usually unpredictable in the time of crises which makes the simulation and conclusion based on pre-crisis data inefficient. For example, financial network tends to be highly complex with banks forming several connections during normal time, but those connections usually shrink when banks are under stress. Central nodes become more important to the system under those circumstances. Authors usually rely on a set of strong approximation to reach policy conclusion that can be hardly justified to the financial industry if ever adopted by regulators. It is not obvious for example the choice of the distribution of banking connections which makes some simulation exercise based on few networks parameters heavily dependent on the model's assumption.

In the same spirit, Agent-Based Models (ABM) were developed in order to infer resilient financial structures and the impact of interaction rules on the system. This type of modeling considers the system as a collection of autonomous interacting agents. The model defines to each agent communication rules and the way they react to information. The main features of those models is that agent equipped with rules that dictate their behavior independently from other agents (Howitt, 2008). To model, the financial system ABM models consider agents of type banks, households, investors and central banks. ABM modeling independence property allows for the implementation for very flexible rules but limits the possibility of using strategic interaction which is at the heart of this paper.

All the previous works help to identify contagion channels and the transmission mechanisms both empirically and theoretically. They also helped to gave insight about the links between those fragilities and systemic risk. However, little interest was given to collaboration of central banks. The policy recommendation was generally to encourage coordination due to the negative effects of contagion without considering the costs of this collaboration.

This paper is closely related to the literature about the financial system network and propagation of crisis, The originality of our paper ,besides the strategic dimension, reside in the methodology used and mainly in the conception of the financial network itself.

In fact, the mathematical framework is borrowed from the work of Mamani et al. (2013). Basically, they were able to demonstrate using a game theory approach and a model of contagion of deceases that governments can benefit from the coordination of their expenditure and vaccination policy. Their results also suggest that wealthy government sponsoring the vaccination of individuals in foreign countries could have positive results on the national level. The benefit of central coordination is substantial in the configuration where countries are highly connected. Other related work on epidemiology, such as the work of Sun et al. (2009) and Wang et al. (2009) tackled the issue of decentralized selfish vaccination policies and its implications. They established conditions when satisfied the general interest coincide with those of governments considering only the in border interests. Moreover, their work also contributes to the general discussion regarding the benefits of sharing information between governments. In fact, imperfect information symmetry could have important pervert effects on the general welfare.

Despite the obvious link between financial contagions and pandemic crisis, researchers in the biology field advice financial ones to be very cautious regarding their use of biology models. As pointed out by Peckham (2013), epidemiological models suffer from the lack of generality. As concepts have a precise meaning in the biological model and sometimes are very hard to translate into the financial context. In addition, King (2013) argues that a major difference between viral infection and financial contagion is that the latter is intentionally transferred between individuals. Individuals infected by a virus usually have very little gain from the transfer process while in the financial context contagion can lighten the burden on the infectious individuals and creates incentives that encourage contagions. Finally, outside the conventional confines of epidemiological concept (Haldane and May (2011), Toivanen (2013)), agents are usually translated into banks in the financial context. However, banks carry a very annoying property of being fundamentally different, and most of all the effect of a single bank is far from being negligible on the system especially for

the Systemically Important Financial Institutions (SIFI). This later assumption is in the heart of most epidemiological models which limit their application in the financial context. The advantage of the model that we propose in this paper is that it is not subjected to the previous criticism addressed by bio-statisticians to economists. The main reason is that our model does not violate the assumption that individuals are small and negligible.

Papers like Toivanen (2013) and Garas et al. (2010) used the SIR model in the financial context to study and model the propagation of international crisis. Those authors replaced the individuals in the biological models by banks while being aware of the main difference between the two. Their modeling approach reduced the applicability of the SIR model to simulation exercises. In fact, those papers violated the key assumptions used in the natural world and by consequence limited the possible benefits that the economist can have from the large literature on epidemic contagion.

4.2 History and channels of contagion

History of contagion in the financial crisis goes back to the 19th century. The crisis of 1825 is probably known as the first modern example of contagion of financial turmoil (Bordo and Murshid (2000)). The evaporation of the gold reserve of the Bank of England led to a series of bankruptcy in England which spread to the Latin America. As a consequence, several sovereigns in the region defaulted despite important gold and silver natural resources. Ever since examples of contagion in the economic context have multiplied. Later several global financial crises erupted after the failure of railroad companies especially in England and the US, but the scope of events was large enough to cover the Latin America and other European countries. For a complete survey about crisis transmission in the 19th and early 20th refer to Bordo and Murshid (2000). The recent year's global financial integration led to spectacular contagion patterns. One example of contagion was illustrated by the Mexican secretary of the treasury during the Russian crisis: "Ninety percent of Mexicans have never heard of the Duma, and yet the exchange rate and interest rates that they live with every day were being driven by people with names like

Kiriyenko and Chernomydrin and Primakov." GURRÍA (1999) The most famous and severe ones are the great depression of 1930 and the great recession. Both events were reflected in stock market crashes around the globe.

The most recent one is the sovereign crisis in Europe as fragilities of the sovereigns in the peripheral countries raised concerns about the creditworthiness of the core countries and required a series of interventions of the ECB. The common feature of this series of contagion is that it induced costs and most of the times unbearable costs on countries with solid fundamentals.

The common feature between those unfortunate events is that they highlighted the importance of international collaboration to resolve a crisis. This coordination is more efficient when prepared, and rules are set prior to the burst of the crisis. Situations of stress will usually increase selfishness among individual governments and will narrow the scope of interventions.

The main difference is that several channels were in action during each episode of contagion that is presented above.

Economist has classified contagion channels into two distinct categories: fundamental and others related to investor's behavior. Contagion can occur due to different types of fundamental causes. The first type is a global shock such as a big movement in commodity prices or a sudden change of interest rates for one of the major currencies. This event can trigger capital outflows and increase co-movement of asset prices. Masson (1998) call it the "monsoonal effect". The second and most straightforward type of fundamentals channels of contagions is trade linkages that regroups the effects of direct links and competitive devaluation. In fact, a crisis in one country can cause a big reduction in demand for imports of good produced in other countries and by consequence affecting the export and the trade balance of unfortunate trade partners. In addition, when a crisis in one country causes a high devaluation of its currency it creates a competitive advantage in international trades over other countries. As a result, this can lead to a reduction of export labeled in the non-affected currency. This effect triggers a cascade of devaluation of several currencies.

When dealing with financial stability, the most important fundamental cause is probably due to financial linkages. Globalization and financial integration makes those links more and more important and are the subject of attention of regulators. For example, a crisis in one country specially developed ones can reduce dramatically the amount of capital supply available in the market and then reduce the financial system resiliency and lending capacity in a second country. This channel is illustrated by Diamond and Rajan (2005) who argues that the failure of a major financial institution can trigger contagion to others even in the absence of direct link due to the shrinkage of the inter-banking capital market. Moreover, a crisis in one country usually makes banks refocus their business on their more profitable units and simultaneously withdrawing from other markets. Thus exporting the turmoil to the later markets. In addition, the outflow of capital can be coupled with an increase in borrowing costs and a sharp depreciation of a country's currency which creates the competitive advantage discussed above. Those links are easily identifiable however harder to mitigate their effects. Hence, those links should be considered as input for regulators dealing with financial stability because of their exogenous characteristic.

The second major cause of contagion is those related to investors behavior. The literature classified those events into five subclasses: market coordination problems, informational asymmetries, incentive problems, investor reassessment and liquidity problems. Although, the name investors behaviors suggests that those are issues related to behavioral aspects all of them are bound to be *ex-ante* individually rational. It is the herding behavior that creates the excess leading to the crisis. An example of crisis due to investors behaviors is that financial breakdown in a region of the world can reveal fragilities in others countries. This sudden revelation will lead investors to withdraw from the later market triggering a full-scale crisis. For a complete survey about contagion related to investor's behavior refer to Claessens and Forbes (2004). Often it is hard to classify the contagion channel due to the overlap that may exist between the results of different categories. The general rule that is established by the literature is that if investor's behaviors are collectively and individually *ex-ante* rational the contagion is usually classified to be of fundamental origins.

4.3 Description of the model

In this section, we will describe the proposed model of a multinational financial system. The model aims at describing ex-ante the strategic decisions of governments of different countries in term of the financial stability policy. The model is flexible enough to cover other configurations in the financial system where different agents are faced with the decision of costly collaboration. Examples can be banks pooling together to save a major bank against failure and prevent contagion effects. Other example of application of the model is when different agencies that are responsible for supervising different segments of the market need to cooperate to prevent transmission of shocks. In the remainder of this paper, we only focus on the case of governments' collaboration but most of the conclusion remain valid for the other cases.

We firstly assume that a source country (labeled country 0) is identified. We define the source country as the one where the outbreak is anticipated to start and then spread to other financial systems. The source country is identified prior to the burst of the crisis. The country 0 could designate an economic region or a financial segment. This model is only relevant when contagion is possible between countries because then country 0 cannot be completely isolated from other countries. The source country designate typically, trade partners, a country of origin of foreign investors or a destination for international capital flows. Each of the $M+1$ countries has a population of N_i value units that we will call $m_{ij}(s)$ $i \in \llbracket 0..M \rrbracket$ and $j \in \llbracket 0..N_i \rrbracket$. We attribute to each asset unit a risk profile that we call P . The behavior of the value unit will differ with the risk profile. Both concepts will be explained in details in the next section.

As mentioned earlier, we borrow a mathematical model from the epidemiological literature. Two main reasons motivate this choice. First, we think that there are enough similarities between the two settings which justify this choice. The functioning of some channels of contagion is very similar in both worlds. The increase of international traffic between two countries strengthens the chances of contagion. The same concept is very close to foreign investment which can be a source of economic growth in good times and an accelerator of financial contagion between two

countries in bad ones. Nevertheless, contagion for non-fundamental reasons is a feature special to the financial system. It is hard to find the equivalent channel of contagion to those driven by loss of confidence in a foreign country in the natural world. In principle, this should not prevent the use of biological models in finance. But from a model perspective then, it is important to take into consideration the non-fundamental channel of contagion in the calibration parameters. This is a layer of difficulty that researchers in biology should not specifically address. Second, researchers using this type of models developed a rich mathematical toolbox that can be used in different type of settings. One of the main contributions of this chapter is in the model itself. Combining the advances made in the study of transmission of diseases with this model can have a multitude of application in the regulation of financial markets. We focus only on the question of international coordination but other interesting research question can also be tackled with this model. An example, of other application can be the implementation of regulation and surveillance in the presence of budgetary pressure under which supervisors are constrained to prioritize between each regulation and surveillance scope.

4.3.1 Elementary asset value units

Elementary asset value units or asset value units (we use the terms value units and asset unit sometimes in this text for the sake of parsimony) are defined as a monetary unit of a risky asset. The value of each of those units is equal to 1\$ before the start of the outbreak. The financial system is the set of all asset units and its value is the sum of those of asset units. The model excludes the effects of inflation on the value of the system. We also assume that countries share the same base currency. The effect of currency devaluation during crisis is not explicitly modeled however its effect can be expressed in the form of a failure of a asset unit. We also exclude returns that can increase the value of an elementary asset unit. The main reason is that the focus of the model is the situation of crisis where the objective is to weather the storm and rarely to make profit.

Ceteris paribus, the value of the financial system remains constant over time. Otherwise, during a crisis, we will witness a destruction of wealth and a reduction of the value of financial assets. Regardless of the debate about the existence of risk-free

assets, this concept does not exclude them from the model as they are considered as elementary asset units with the particular risk profile that are not susceptible to contagion or failure.

The value unit can have four different states over time:

- State S_0 : is called the normal state. It is characterized by the fact that the value of the asset remains at 1\$. Asset units at the state S_0 are exposed to contagion from other units. All the asset units are remaining at state S_0 in a country i means that the contagion could not reach that country. Such unit at the state S_0 can only fail through contagion in this model. In this state, units bare all the risk intrinsic to their profile without any type of regulatory intervention to reduce this risk.
- State S_1 : is called the stressed state. The value of the elementary asset unit has fallen below 1. There is also a chance of the asset destruction or total loss realization where the value of the value unit will become 0 after the outbreak. A value unit at this state is also a source of contagion as it could spread the stress to other units. The number of simultaneous value units in state S_1 indicates the amplitude of the crisis and the speed of recovery of the financial system. In our model this is always a temporary state as value units will either become loss or recover from failure (State S_2 and S_3).
- State S_2 : is called the failure state. In this case the value of the elementary asset unit is estimated to be 0. In this case, it will be removed from the population. S_2 is an absorbing state. The recovery is not possible. The units of this state are not susceptible of contagion and cannot contaminate other units. Only those at state S_1 can spread the stress to other units in their home country or a foreign one.
- State S_3 : is called the recovery or liquidation state. In this state the asset value unit has recovered from the stress. The unit may not recover to its initial value however it becomes non contagious. This state is similar to S_2 in the sense that they do not represent any future threat to the financial system. Risk unit in the state S_3 are also immune to contagion. In other words, any

asset unit can only be contaminated once and then it develops some sort of immunity. This state particularly describes assets that the market feared their failure during a crisis and then those fears disappeared in the future after the discovery of new information or a public announcement. It also the case of asset that were liquidated or reimbursed. In other words, their value is transferred to a highly liquid and risk-free vehicle such as cash. S_3 is also an absorbing state.

One possible critic that can be addressed to using asset values as the atom on which we build the model is that similar financial assets seems to behave similarly in bad times. Government bonds will be hit simultaneously in sovereign crisis notwithstanding amplification mechanism like fire sales and herding behavior that can increase this type of simultaneous failure. Therefore, it can be seen as unrealistic the assumption that assets will fall individually. However, the model does not exclude the possibility of having similar assets falling together. Assets can be divided into clusters or risk profiles that we describe in detail in the coming section. If we set the probability of contagion within cluster fairly high (or 1 if necessary) we can reproduce the herding behavior described earlier. Using asset units adds some flexibility as we can distinguish between asset classes and set different contagion probabilities within those clusters and also attach different recovery and failure probability to each one of them.

4.3.2 Description of the risk profile

To each elementary asset value units, we attach a risk profile P that can be seen as a categorization of assets according to their behavior during crisis. More particular, is the characterization of going into stress as well as the contagion potential of the asset unit sharing the same risk profile. Asset value that share the same risk profile also have the same probability of recovering from the stress and go into failure. In this concept, we encapsulate assets that have similar features in the financial system. We can classify, for example, risk free assets immune to failure and contagion in the same risk profile. Debt of financial institutions with different seniorities will have different risk profiles reflecting their position in the liquidation chain.

Before the start of the crisis, regulators decide to intervene on some risk profiles and apply a form of protection on those categories of assets. The objective of the protection is to increase market confidence in the quality of the asset under protection. By consequence, we consider that this will reduce the probability that an asset unit goes into the stressed state. It is important to highlight that the decision of protecting risk units is costly and regulators can only choose to protect a limited number of risk units in their country due to resource scarcity. We also use the terms hedged or vaccinated risk profiles to indicate those that were the subject of regulatory interventions. We use vaccination mainly because it is easy to draw some analogies between government actions to *ex – ante* protect financial assets and vaccination in the epidemiology models. In our design, we intend risk profiles to group similar asset and but not an ownership structure. For example, a mutual fund or assets of the same bank should not be interpreted as having the same risk profile.

4.3.3 Hedging and recovery mechanism

As opposed to the natural world where vaccination and recovery require no further explanation, those concept translated by hedging and recovery respectively in the economic world requires precisions and also clarification of their mechanisms.

Any form of protection that law makers can establish to reduce risks on a class of financial assets is considered to be a form of hedging of a risk profile and the cost of those reforms are perceived in the broadest context. Not only direct costs but also negative externalities are added to the general bill of intervention. Moreover, those interventions are rules that are established *ex – ante* and should have the same effect on all assets sharing the same risk profile. Several governments or central bankers' interventions can decrease the risk of failure and contagion of a financial asset. The obvious example is probably the explicit guarantee that governments offer on deposits. Almost in every country in the world a special fund is created to provide insurance to deposits in case of the failure of their custodian. This program offers a form of protection to a class of assets i.e. deposits that reduces dramatically their failure probability and limit their spillovers to other assets. Nevertheless, those programs are costly even when we account only for operating and administrative

expenses required to run the deposit insurance funds like the FDIC in the US. Other types of asset classes benefit from this explicit guarantee. The government can decide for example to subsidize lending to households by offering a guarantee to agencies that promote such activities.

The government also offers an implicit guarantee to financial assets. Those are insurances that governments do not explicitly protect, but market actors are expecting a form intervention in the case of failure of those assets. An example is the implicit guarantee that is usually attributed to the too-big-to-fail banks. Other form of asset protection are regulations. In fact the role of regulation is to increase the resiliency of some assets by increasing their monitoring. Not only the human and capital resources required to such activities are costly but other indirect costs are also coupled with those regulations. For example restraining lending to certain type of activities deemed to be too risky can have negative impact on growth and by consequence induce costs on governments budgets. Rules of disclosure on certain types of activities are also a form of protection of an asset class. Those rules are usually costly as they require the use of human and IT capital to produce the information to publish and the same type of resources for the analysis and the verification of the disclosed information.

Short term recovery and the liquidation of financial asset after distress without complete asset destruction can happen with or without external intervention of governments by injecting additional funds into the system. Spontaneous recovery can occur for example when new information about asset deemed to be under stress reveal that the risk is unreal. This is usually the case when a major member of an industry fails and uncertainty is raised against its competitors.

Moreover, we intentionally exclude any recovery mechanism that includes *ex – post* funds injection like bail-outs. All form of recovery rules should be established before the failure with predetermined cost sharing agreements between governments. The main reason behind this design is to eliminate the possibility of bail-outs and favor another mechanism of recovery like bail-ins. The bail-ins allows for investors holding the bank's debt to utilize those funds to participate in the losses both on a voluntary or mandatory basis. The concept of risk failure can cover the bail-in

cases. In this model this can be easily transposed as failed bank equity spreading contagion to debts eligible for bail ins and inducing them into a failure status. In fact, the bank debts can be distributed among several risk profile each with different susceptibility to the bank's equity stress and with different costs in case of contamination. Moreover, creditors have further incentives to monitor the bank's managers which in term should increase the quality of its asset and reduce the probability of the bank's failure. The possible haircut applied to those debt applied after the absorption of losses are then considered to be failure costs. Again, regulation is decided before any outbreak starts which also mean that governments are unable to alter the risk profile during the crisis. This rule is particularly reinforced with the intention that every financial institution is supposed to prepare a resolution plan that detail how it will go through a potential bankruptcy and how to unwind the position taken by the bank without inducing additional costs to tax payers. Thus, the model of financial system that we propose is perfectly suited to study the new design of crisis management embraced by regulators including resolution plans of big banks and bail-in.

Moreover, in this model, we consider that the writing down of impaired asset is a form of recovery although this case is technically a failure. To account for the losses related to such event we consider that any losses in case of default are included in the cost of failure. Later, we consider that the asset has recovered while bearing losses due to its transition to the default state. Finally, we consider that all assets that have recovered from the stress are shortly liquidated after that and transformed into a liquid and non-risky asset such cash are are immune to future contagions.

4.3.4 State transition dynamic

Figure(4.1) shows the different states that an asset unit can have from the starting states (before the financial crisis) to the final states (after the end of the financial crisis). The system reaches the final states when no evolution of any asset unit is possible. As asset unit in the stressed state is the source of contagion in the model, the final state is reached when the number of stressed units hits 0. This is achieved throughout two possible scenarios: all the stressed asset units have recovered, and further contagions are impossible, or all units experienced a stressed period and are

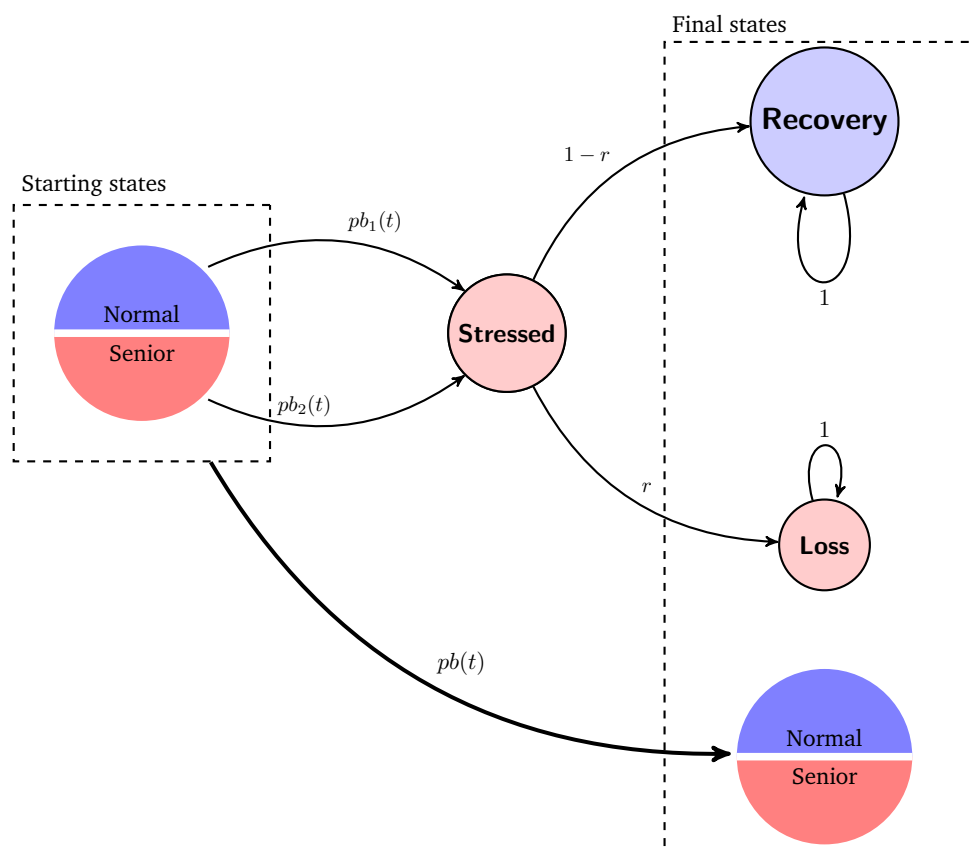


Fig. 4.1.: Different possible states of a asset unit and the transition possibilities. Senior and Normal indicates assets units with two different risk profiles. Senior risk profile are hedged by regulators reducing its susceptibility of failure ($pb_2 < pb_1$) and potential of contagion.

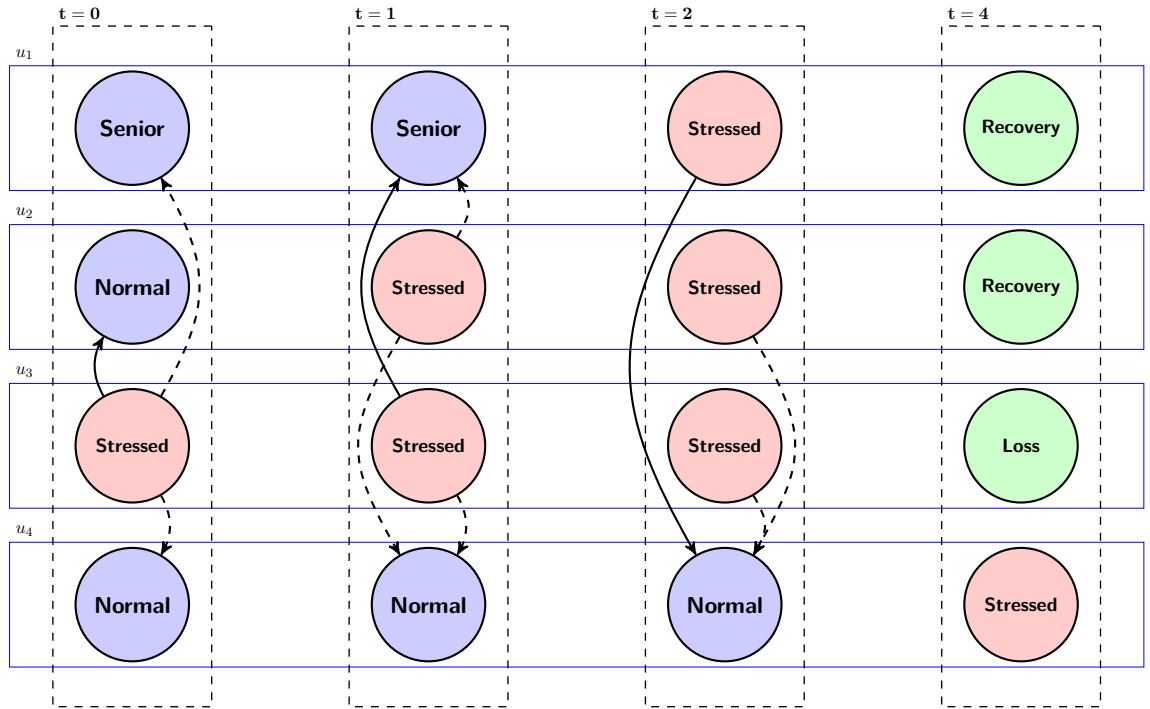


Fig. 4.2.: Possible evolution of different asset units (u_1, u_2, u_3, u_4) between time intervals (from time $t = 0$ to $t = 4$). The Solid lines indicate actual contagion while the dashed one refers to possible but not realized contagions in this scenario. Senior and Normal indicates assets units with two different risk profiles. Senior risk profile are hedged by regulators reducing its susceptibility of failure and potential of contagion.

now either recovered and liquidated or in a loss state and are immune to further contagion. The transition through the stressed state is not compulsory: any asset units in the normal or senior risk profiles can go through a crisis without any transition in their states. This is usually the case for high-quality asset units or for others outside the scope of contagion of a local crisis. Finally, the stressed state is always a transition state and failed risk units should either recover or be recognized as a loss and have a 0 value. The model assumes that the risk profile remains constant during the crisis. Although, usually assets tend to behave differently during financial stress relative to normal trading days as market conditions change dramatically. In fact, co-movements of asset prices will increase, and liquidity evaporates with an increase price impact. Those conditions can increase contagion probability and susceptibility to becoming under stress. Because, we only focus on modeling the system during a crisis, all those factors need to be taken into account and design risk profiles to be in-line with the stressed behavior of asset. It is important to measure transition probabilities during the crisis episodes. In practice, this poses an empirical challenges to overcome problems like tail estimations and the limited number of observations.

Figure (4.2) illustrates an evolution scenario for a population of 4 elementary asset units and starting from a unit in the state S_1 (asset unit u_3). Each horizontal box regroups states of the same asset unit while dashed vertical box delimits a set of asset units at a given point in time. We also distinguish between asset units that have different risk profile. Senior indicates an asset unit with a risk profile monitored by regulators which decrease its failure probability and potential of contagion compared with the asset unit Normal.

Those scenarios are the different contagion possibilities that a financial system with a population of 4 asset units can witness in the presence of one contagious element. In each time step dashed lines are for potential contagion. Solid arrows represent realized contagion in the represented scenario.

Three important facts are to be noted. After stress two outcomes are possible: loss or recovery. Each state has a different economic meaning, but they share important features from a modeling perspective. First, both states is absorbing states. A risk

unit in the state S_3 or S_4 is immune to failure again. Second, those units can be removed from the population without any impact on the number of future failures.

Second, transitions between state are not Markovian. More precisely, the history of each unit is relevant for the future transition probabilities. In fact, the asset units u_1 and u_3 at time $t = 2$ can have a different impact on the population because the senior asset unit that is under stress (u_1) has less chance of contaminating other units compared to other stressed asset units (u_3). Moreover, the risk of failure is not affected by the length of the period where the unit remained in its starting state. However, the unit u_4 had higher chance of fail at period 4 simply because at period 3 the system contained more failed units compared to its past states.

4.3.5 Crisis and government interventions

We recall that government intervention is oriented toward altering a specific risk profile and cannot target individual asset units. In the following, we refer to the possibility of being under stress as susceptibility and the potential to transfer the stress to other units as infectiousness. To quantify the effect of hedging (we sometimes use vaccination due to the obvious analogy) let $(1 - \theta)$ be the impact of vaccination on susceptibility in a risk profile P . In other terms, the probability of going under stress of an asset unit with profile P is multiplied by θ ($\theta < 1$). Next, let $(1 - \phi)$ be the effect of government interventions (vaccination of asset units) on infectiousness. It is the impact on the probability that an asset unit at the stressed state transfers the stress to other risk units. Thus, the combined effect of vaccination on asset units is $\psi = 1 - \theta\phi$.

The effectiveness of the government policy to prevent the spreading of a financial crisis also depends on the underlying dynamics of the crisis. In this model, this concept is transposed into the parameter R_0 . It is defined as the expected number of secondary stress situations transmitted from a single stressed asset unit in a complete population of non-vaccinated asset units. R_0 determines what fraction of the global assets will be affected by the crisis in the complete absence of government's interventions and coordination. The larger R_0 is, the higher the fraction of losses are. Thus, the mitigation effects of *ex-ante* government prevention policies are lim-

ited for large values of R_0 . Here the model acknowledges that some future crisis will be severe enough to be mitigated by governments. In this extreme situation, collaboration will lose its meaning because all financial system are uncapped and the possibilities of spare budget to guarantee assets outside the national borders are very limited.

R_0 combines the effect of three different phenomena that characterize the transmission of crises: the transmission probability of the initial shock, the degree of linkages between asset units and the duration of the crisis. First, R_0 is partially determined by the type of crisis that hits the original country and more precisely the severity and widespread of the initial shock.

Second, the degree of linkage between risk units in one country can increase with the correlation between asset classes. It may vary depending on different factor related to the topology of the country's financial market. For example, the number of trading venues, the degree of usage of securitization and derivatives, quality of the legal system and the financial services' infrastructure. Finally, the duration of the crisis is also a determinant factor for the dynamic of the crisis. In fact, some asset units can only be affected after some period from the beginning of the turmoil. A prototypical example would be high-quality assets that the bank tried to keep in its portfolio because of their high and guaranteed future yield. Nevertheless, the bank's manager was obliged to fire-sale those assets to secure liquidity if the crisis continues after a certain amount of time. Longer crisis can also cover the case where new asset units are discovered risky after the failure of similar asset. The perfect example is the reassessment of the risk profile of sovereigns after the near bankruptcy of some government bonds.

If R_0 is the average number of contagions in a population of asset unit without any government intervention, then $(1 - \psi)R_0$ is the average number of contagion from a single asset unit if all units are in a risk profile that was protected by regulators.

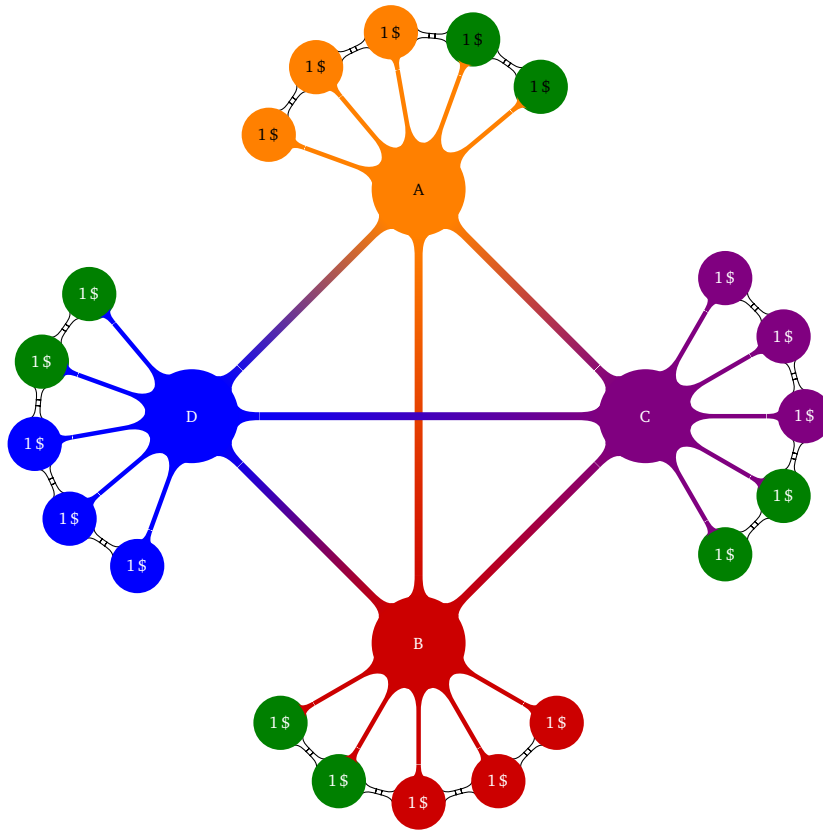


Fig. 4.3.: Connections Graph of national and international contagion within a risk population

4.4 General transmission dynamics

In this section, we will present the dynamics of transmission of shocks in our model. Models of contagion in the financial literature usually rely on the network theory. Those models use banks in the nodes and rely on market price or exposure networks to estimate the network structure.

We choose to model transmission differently in this paper. First, our modeling of the financial system is not based on banks as the atomic unit of the model. We rather consider directly the portfolio of asset with banks or governments at the higher level. We choose this approach mainly to avoid moral hazard in the regulator's decisions and to be able to benefit from the intensive literature on epidemiology and risk transmission. We also decide on this approach for practical reason. Network model using banks as nodes rely on historical data to calibrate their models and infer lesson about future crisis and the best mitigation policy to be applied by

regulators. It is hardly believable that banks characteristics remain unchanged over a long period of time. Those institutions tend to change their source of financing, exposures and business models which create structural breaks on the data. However, asset classes should have a more stable risk profiles which make the inference exercises more significant

The starting point of the network model of financial contagion is the famous model of epidemic spread called the standard Susceptible(S), Infected/ Infectious(I) and Removed (R) or SIR model. Here we assume that the intervention policy is specific to each country. Each government can decide individually on its crisis mitigation policy. Although individual countries may endeavor into reforms that are imposed by international regulators, we consider that they do not bare the costs of those reforms and should not be taken into account in the decision-making process. Those general reforms will reduce the probability of contagion between countries and within the same country. We start by describing the model without the effects of government interventions or vaccination.

The cross border contagion dynamics is modeled by the factor $R_{ij}(P_m, P_n)$. R_{ij} is the number of secondary infected asset units with the risk profile P_m in country j directly affected from a randomly selected infectious asset unit having the risk profile P_n in country i for every $i, j \in \{0 \dots M + 1\}$ and $P_m, P_n \in \{P_1, \dots, P_p\}$ the set of possible risk profiles assuming that all units belong to a risk profile with no intervention. R_{ij} describes the potential of infection of population i and the susceptibility of the population j . Without loss of generality, we will regroup the state of recovery and failure. Elementary asset units in those states have no impact in the population because they lost all their infectiousness potential and are immune to future change of state.

The matrix $R = [R_{ij}(P_m, P_n)]$ is a generalization of R_0 to the case of multiple population (or countries in this context) and several risk profiles. In the epidemic literature, it is shown that the stability of the crisis dynamic is directly related to the value of the largest eigenvalue of $R \forall n, m$. In fact, if $R_0 < 1$ the crisis is early contained as the number of infected individuals will decrease over time. However,

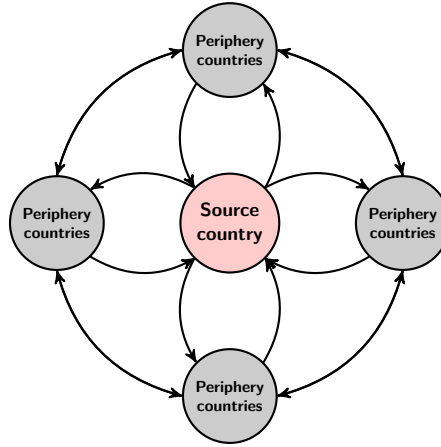


Fig. 4.4.: The complete model with all the possible contagion channels

if $R_0 > 1$ the first infectious asset units will spread their risk to more units and this could result in a larger scale financial outbreak.

At the end of the crisis, let T_i be the total number of infected elementary asset units. This represents both units that were at the origin of the crisis and other that were affected by the different phases of contagion. Let us suppose $S_i(0)$ be the fraction of the susceptible asset units at the start of the crisis. In other words, it is the fraction of risk units that can go into the failure state. Let $I_i(0)$ be the fraction of infectious at time 0 or the fraction of the risk unit that have contagion effect on other risk units.

Without any loss of generality, we will consider that any stressed risk unit has an equal probability to recover or switch to the loss state. For the sake of parsimony, we will call R_{C_i} the subset of risk units that recovered or failed in the country i . The units of R_c have no influence on the evolution of the crisis.

We propose a dynamic of a complete model of contagion. In the complete model, contagion can occur between the source country and the peripheral countries and can also happen between two countries both different from the source country. The model is called complete because it accounts for all the possible transmissions between countries i and j as long as the $R_{ij} > 0$

In this illustrative model we consider that all asset units share the same risk profile P such as $R_{ij}(P_m, P_n) = R_{ij}$. γ_i and κ_i are respectively the recovery rate and the failure rate for the profile P .

The dynamic of the complete model can be described via a set of differential equations:

$$\frac{dS_i}{dt} = -\frac{1}{n_i} S_i(t) \sum_j R_{ij} n_j I_j(t) \quad (4.1)$$

The model is a continuous time model. In order for this assumption to hold, the number of individuals should be sufficiently large to be represented via a continuous variable. This is in line with the assumption that individuals in our model are asset value units. Nevertheless, it also means that considering banks as the individuals violates this assumption. The contagion is spread between asset units which implicitly suppose a sufficient linkages between those assets. R_{ij} also defines the rate to which a risk unit under stress from the population i spread the risk to a risk unit in the country j .

In equations (4.1) and (4.2), $S_i(t) R_{ij} n_j I_j(t)$ is the rate of appearance of stressed asset unit at time t in country j from stressed asset unit in country i . The increase in stressed units is compensated by a decrease in the susceptible population. Moreover, individuals in country j are infected from several other countries. Then the total rate of appearance of new stressed risk unit is $S_i(t) \sum_j R_{ij} n_j I_j(t)$.

$$\frac{dI_i}{dt} = \frac{1}{n_i} S_i(t) \sum_j R_{ij} n_j I_j(t) - (\gamma_i + \kappa_i) I_i(t) \quad (4.2)$$

Simultaneously, individuals are constantly exiting the stressed state either to recover or become permanent loss according to respectively to the recovery and failure rates. The exiting from the stressed states is dictated by the rates of recovery or failure.

$$\frac{dRc_i}{dt} = (\gamma_i + \kappa_i) I_i(t) \quad (4.3)$$

The size of the recovery and failed population continue to increase because it is a final state. In the equation (4.3), it is only possible to enter this state without any

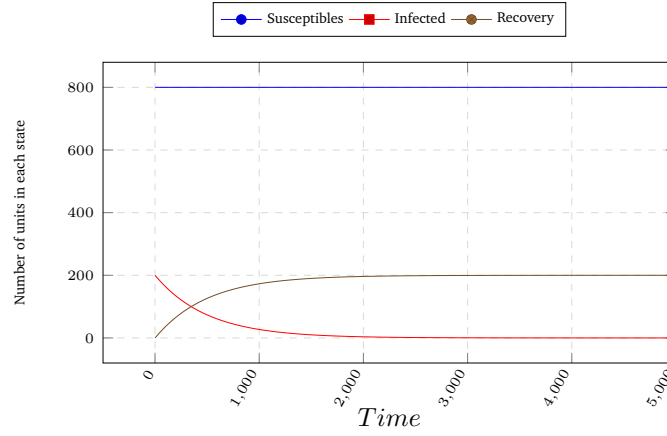


Fig. 4.5.: Evolution of the number of Susceptible, Infectious and Recovered/Removed as function of time. The simulation is done on a population of a 1000 asset units in the same risk profile. The initial number of infectious is 200. All other variables are randomly assigned. $R_{ij} = 0$: no contagion is possible. The initially infected units will recover without additional spread of the crisis.

existing mechanism. The system reaches a steady state when $I_i(t) = 0$. In the end of the considered phase of outbreak, the asset value units either remains in the susceptible states or transit threw the stressed state to end-up in the recovery or the failure state. In other words, $R_{ci}(t_f) + S_i(t_f) = m_i$ with t_f is the minimum time where no more evolution of the system is possible.

Figures (4.5),(4.6),(4.7) and (4.8) illustrates different realization of the contagion dynamic. All figures share the same model parametrization except for the susceptibility and crisis intensity parameter $R_{ij} \in \{0.04, 0.02, 0.01, 0\}$. Depending on the value of R_{ij} the stress spread to the all asset units in the financial system ($R_{ij} = 0.04$) (4.8) or to a fraction of the total asset units ($R_{ij} \in \{0.02, 0.01\}$). Figure (4.5) shows the evolution of the crisis when no contagion is possible ($R_{ij} = 0$). Obviously, stressed risk units will continue to recover or fail according to the corresponding rates without seeing any increase in their number. This is an extreme case where crisis hits a complete isolated cluster of the global financial system but is of no interest to this analysis simply because there is no incentive for collaboration. The simulation in the figure (4.7) shows an intermediate case. Contagion is possible but still very mild. We notice that the total number of stressed units remains around its initial value (200) and that the total number of recovery continues to increase. More precisely, we witness two phases in the evolution of the crisis: before and after time period 1000. In the first part, the size of the stressed asset units is

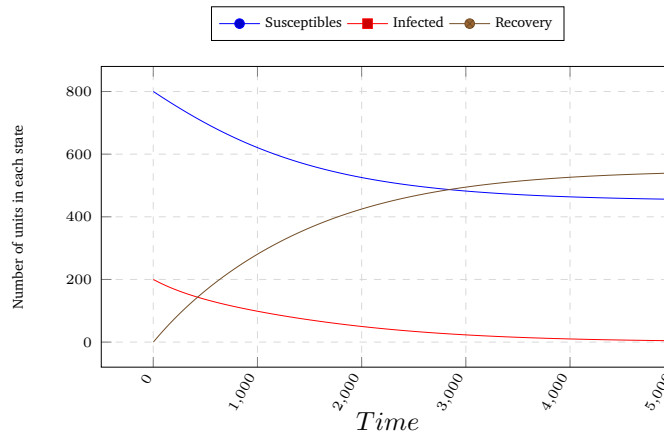


Fig. 4.6.: Evolution of the number of Susceptible, Infectious and Recovered/Removed as function of time. The simulation is done on a population of a 1000 asset units in the same risk profile. The initial number of infectious is 200. All other variables are randomly assigned. $R_{ij} = 0.01$: mild contagion is possible. The crisis did not affect most of the asset units.

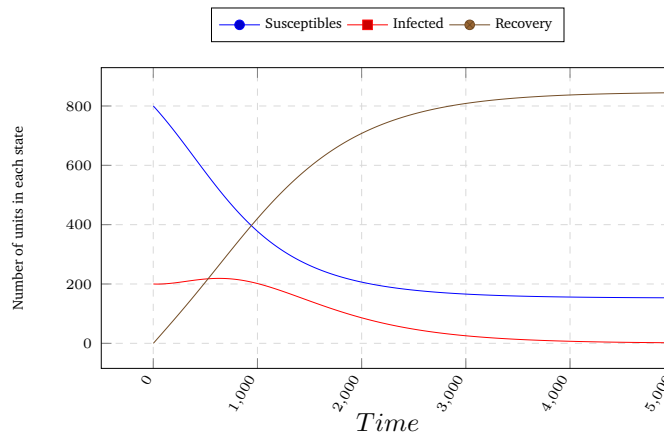


Fig. 4.7.: Evolution of the number of Susceptible, Infectious and Recovered/Removed as function of time. The simulation is done on a population of a 1000 asset units in the same risk profile. The initial number of infectious is 200. All other variables are randomly assigned. $R_{ij} = 0.02$. A severe crisis but not all asset units were affected

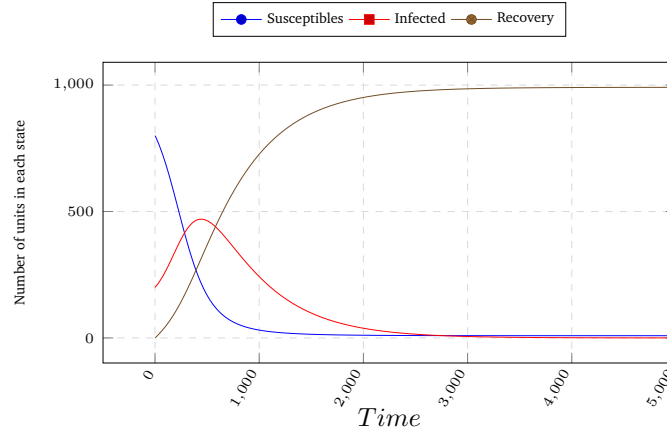


Fig. 4.8.: Evolution of the number of Susceptible, Infectious and Recovered/Removed as function of time. The simulation is done on a population of a 1000 asset units in the same risk profile. The initial number of infectious is 200. All other variables are randomly assigned. $R_{ij} = 0.04$. A full-scale financial crisis where all risk units are affected by the initially stressed assets.

stable around 200 which leads to conclude that the number of recovery is close to the number of new infected units. It should not be implied that the crisis is mild and contained in this case. In fact, the total number of stressed units remains high at the end of the outbreak (around % 80). This is typically the scenario of mild but long-lasting crisis that can be as damaging as a short but highly severe one.

Finally, figure (4.8) shows the case of the most contagious crisis in the presented simulations. The size of the set of stressed units continues to increase due to higher infection rate compared to the recovery rate. After the burst period, contagion slowed because the population of susceptible asset units shrunk dramatically in size. At the end of the crisis all the financial system went through the stressed state. This is the situation of a very severe financial crisis where the lost of confidence reached all financial assets.

p_i is defined as the fraction of failed asset units at the end of a crisis. p_i is also called the attack rate for the country i . p_i can be derived from the solution of the system of PDE described in (4.1), (4.2) and (4.3). The detailed proof is given in Longini et al. (1978).

$$p_i = S_i(0) \left(1 + \frac{I_i(0)}{S_i(0)} - e^{-\sum_{j=0}^M R_{ij} p_j} \right) \quad (4.4)$$

The total number of infections T_i is related to p_i by the following relationship: $T_i = N_i p_i$. The equation (4.4) describes a system of $M + 1$ equation with $M + 1$ unknown: p_i . No explicit solution to this system can be obtained however it is easy to find a numerical solution via an iterative procedure.

Next, we let f be the vector representing the government vaccination policy. f_i represents the fraction of elementary asset units that a government i decides to protect or monitor. After the regulator's decision, the financial system will consist of asset units distributed among two profiles. Each profile reflects the crisis behavior of the asset unit taking into account the protection or monitoring of the regulators. T is a function of f as the foreign regulatory policy have an impact on the number of infected in the local population due to international contagion dynamics. Asset units with the risk profile that regulator decides to protect are called senior asset units. The remaining asset units are refereed to as normal units.

To include the effect of vaccination on the dynamic of interaction, we consider the senior asset units and the normal asset units separately in the model dynamic each specific to the corresponding risk profile. In each population i let S_i^f , I_i^f and R_i^f represents the number of susceptible, infected and recovered in the sub-population of the protected risk profile. S_i^{nf} , I_i^{nf} and R_i^{nf} are the corresponding susceptible, infected and recovered for the sub-population i of the remaining or normal asset units.

The corresponding dynamic can be written then:

$$\frac{dS_i^f}{dt} = -\frac{1}{n_i f_i} S_i(t) \theta_i \sum_j R_{ij} n_j f_j \phi_j I_j^f(t) - \frac{1}{n_i f_i} S_i(t) \theta_i \sum_j R_{ij} n_j (1 - f_j) I_j^{nf}(t) \quad (4.5)$$

The system of equations (4.5), (4.6), (4.7), (4.8), (4.9) is similar to the system of equations (4.1), (4.2) and (4.3) in many dimensions.

The main difference is that the population of susceptible and stressed asset is divided in two risk profiles. The population of stressed units in the regulated profile S_i^f can decrease either because of infection from the contagious units in the nor-

mal profile $-\frac{1}{n_i f_i} S_i(t) \theta_i \sum_j R_{ij} n_j (1 - f_j) I_j^{nf}(t)$ or being contaminated by a unit in the risk profile $-\frac{1}{n_i f_i} S_i(t) \theta_i \sum_j R_{ij} n_j f_j \phi_j I_j^f(t)$. The latter being less important because the regulations hamper the contagious potential of asset unit and the impact is represented by ϕ_j .

$$\frac{dS_i^{nf}}{dt} = -\frac{1}{(1-f_i)n_i} S_i(t) \sum_j R_{ij} n_j f_j \phi_j I_j^f(t) - \frac{1}{(1-f_i)n_i} S_i(t) \sum_j R_{ij} n_j (1-f_j) I_j^{nf}(t) \quad (4.6)$$

The same dichotomy applies for susceptibles in the normal risk profile S_i^{nf} .

$$\begin{aligned} \frac{dI_i^f}{dt} = & -\frac{1}{n_i f_i} S_i(t) \theta_i \sum_j R_{ij} n_j f_j \phi_j I_j^f(t) - \frac{1}{n_i f_i} S_i(t) \theta_i \sum_j R_{ij} n_j (1-f_j) I_j^{nf}(t) \\ & -(\gamma_i + \kappa_i) I_i^f \end{aligned} \quad (4.7)$$

The population of stressed units with the protected risk profile (I_i^f) can increase from two sources of contagion: contagion from other risk units with the same profile $\left(\frac{1}{n_i f_i} S_i(t) \theta_i \sum_j R_{ij} n_j f_j \phi_j I_j^f(t)\right)$ and from stressed risk units having the non protected risk profile $\left(\frac{1}{n_i f_i} S_i(t) \theta_i \sum_j R_{ij} n_j (1-f_j) I_j^{nf}(t)\right)$. The parameter ϕ_j accounts for the attenuation of the potential of contagion of asset unit in the protected risk profile. The population of non-protected stressed units also increase from both sources of contagion as before with similar increase rates.

$$\begin{aligned} \frac{dI_i^{nf}}{dt} = & \frac{1}{(1-f_i)n_i} S_i(t) \sum_j R_{ij} n_j f_j \phi_j I_j^f(t) + \frac{1}{(1-f_i)n_i} S_i(t) \sum_j R_{ij} n_j (1-f_j) I_j^{nf}(t) \\ & -(\gamma_i + \kappa_i) I_i^{nf} \end{aligned} \quad (4.8)$$

$$\frac{dRc_i}{dt} = (\gamma_i + \kappa_i)(I_i^f + I_i^{nf}) \quad (4.9)$$

Figure 4.9 illustrates the impact of the regulatory policy on the number of infectious. The figures show that governments intervention as captured by f in our model have a positive impact on the total number of infected units.

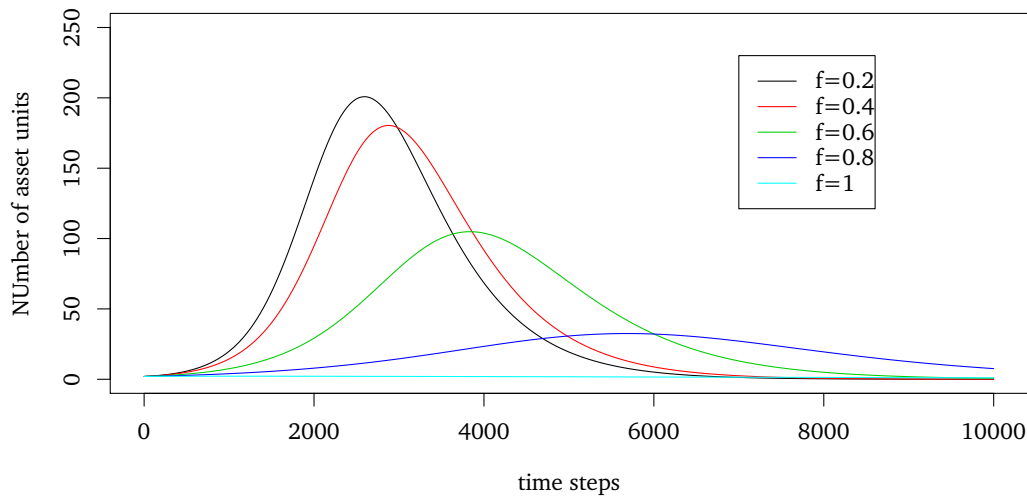


Fig. 4.9.: This figure shows the effect of government intervention f on the number of infectious and their evolution over time. The simulation takes into consideration 1 population of size N that contains both units in the normal risk profile and $f * N$ units in the senior risk profile. The different lines indicate results when varying f and keeping all the remaining parameters as constant.

Although we choose to focus our attention in this paper on collaboration between countries or central banks, this model is well suited to consider the strategic interaction between banks. The countries in the model can be replaced by banks. In the later context, the model can be used to describe contagion dynamics between banks and more importantly to design strategic contracts between them in the case of default or situations of a stress of one bank. An example of application would be to design a contract where a pool of banks can intervene to re-capitalize a bank under stress to avoid spillover effect to the financial system. This type of contract can emerge instantaneously between agents in the financial system or be imposed by a central planner. This paper should provide the theoretical justification to the regulator to impose such contracts on the agents of the financial system. The model can be easily extended to consider the case of different simultaneous failure of banks like the last financial crisis.

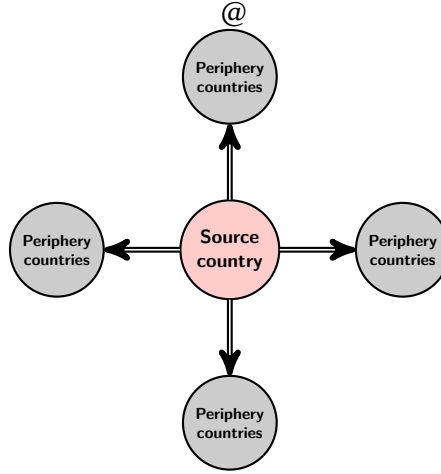


Fig. 4.10.: Star model when $R_{ij}R_{jk} \approx 0$. $i \neq j$ and $j \neq k$

4.5 Strategic decision of individual countries

In order to determine the strategic response of each country in term of regulation expenditure in response to other countries choices, we need to apply some simplification to the model that was described in the previous section.

Here, to keep the model realistic enough to represent real world transmission scenario and at the same time resolvable. We seek a good approximation model to the transmission dynamic. We assume then that contagion is conducted in the form of two distinct waves. The first wave of contagion is cross-border between the source country and peripheral countries in the first place and then between peripheral countries between each other. In the second wave, we assume that the within border transmission is the dominant form of contagion. Furthermore, we assume that in the first wave transmission is bounded to follow the star model illustrated in the figure. 4.10. This assumption is technically satisfied if we consider that $R_{ij}R_{jk} \approx 0$. Assuming the star model does not exclude transmissions between two peripheral countries. It states that the probability that two consecutive direct infection crossing borders between countries i and j and then j and k is small. For example, assuming that $R_{ij}R_{jk} \approx 0$ and if i is the United States, j is the UK and k is Belgium implies that the number of asset unit in Belgium directly contaminated from an asset unit in the UK directly contaminated from those stressed in the US is rather negligible.

In the second wave of transmission, it becomes possible to approximate the fraction of total asset units that went into a stress. This can be achieved via (4.4) given that we assign the proper initial conditions for the fraction of susceptible, stressed and recovers/lost asset units.

$$p_i = 1 - \psi f_i - (1 - \chi_i)(1 - \psi f_i)e^{-R_{ii}p_i} \quad (4.10)$$

Where χ_i is the fraction of asset units in country i that become stressed after the first wave of contagion. We also assume that no further contagion from the source country is possible after the first wave of contagion. For mathematical proof of 4.10 and how to compute χ refer to A.

Decision of individual countries: the game problem

In order to solve the decision problem we need to establish the cost of the crises endured by the individual countries. The regulation efforts are costly and let v_i represents the average unit cost of vaccinating a single asset unit. v_i should cover all the cost related to the government intervention: the potential cost of a bailing procedure, the labor costs of monitoring the financial market the possible negative effects on growth from restricting the bank's activity. Without any loss of generality, we consider that there are no economies of scale related to hedging and that an increase of f_i can only improve the stability of the financial system. Any negative side effects of regulation are out of scope of our analysis. Moreover, let b_i be the average direct and indirect cost from the stress and potential failure of a single elementary asset unit. The asset units are defined in monetary term, so the maximum natural value of b_i is 1. We allow for a higher value of b_i because governments may want to account not only for the financial losses but for the spillover effects to the real economy.

To account for the fact that financial contagion is different from the one in the biological world, we will introduce c_i as the benefit to the source country that results from contagion. In fact, banks operating in the source country have incentives to transfer part of the financial stress to their subsidiaries in other countries in

order to relieve their core activities. Countries like the US have a long history of exporting crises to other continents like Europe. The benefit of contagion should be more pronounced when banks of the source country have important foreign direct investments. This also should be limited when the home bias factor is very important.

In what follows we consider the simple case where in each country asset divided between two type of risk profile. First, the government made some efforts to protect and monitor a fraction of the wealth of the system and those asset units then belong to the senior profile. The second risk profile is what we call normal where regulators are not doing any sort of effort to decrease the risks of the asset within this profile. The strategic decision is then what is the optimal distribution of asset unit between those two profiles to reduce the impact of a financial crisis. The total cost a financial crisis is:

$$\left\{ \begin{array}{ll} b_0 T_0(f) + v_0 f_0 N_0 - c_0 \sum_{j=1}^M y_j(1, f) & \text{for country } 0 \\ b_j T_j(f) + v_j f_j N_j & \text{for country } j \in 1, \dots, M \end{array} \right\} \quad (4.11)$$

$y_j(1, f)$ is the first generation of infected asset units. In other terms, they are units directly infected from others in the source country. The costs modeled in equation (4.11) exhibit the dependence of each country in T_j to the policy decision of other countries. Nevertheless, governments and central banks want to focus their strategy on decision variables that are within their scope of action.

We want to examine the regulation expenditures of each individual country in the absence of a central planner. We define therefore the game problem to be a one-shot game between the different governments that use the total number of infectious asset units $T_j(f)$ to decide on their policy. The objective of each country is to minimize the total perceived costs of the financial outbreak GF_i . GF_i includes

the financial costs, economic costs, the costs of the hedging policy and finally the possible benefits of contagion.

$$\left\{ \begin{array}{l} \min_{0 \leq f_0 \leq 1} GF_0 = b_0 T_0(f_0) + v_0 f_0 N_0 - c_0 \sum_{j=1}^M y_j(1, f) \\ \min_{0 \leq f_j \leq 1} GF_j = b_j T_j(f_0, f_j) + v_j f_j N_j \end{array} \right\} \quad (4.12)$$

To be able to characterize the solution of this problem we need to assume some regularity condition for the cost function GF_i . In addition to the technical reason to impose those restrictions, we can also attach an economic foundation to each of them.

Assumption 1 • $\frac{\partial GF_0}{\partial f_0} \Big|_{f_0=0} < 0$

• $\frac{\partial GF_i}{\partial f_i} \Big|_{f_i=0} < 0$ for all $i \in 1..M$

Assumption 1 is to make sure that the initial cost of regulation does not exceed the benefits on GF_i $i \in 0..N$. It supposes that every individual country should have the incentive to engage in regulatory efforts. In case that assumption (1) is violated regulators should engage in important monitoring efforts before it can be beneficial to the economy. This can serve as a deterrence to considering regulating the financial system.

Assumption 2 if $T_i() = N_i p_i$ then:

- p_i is strictly decreasing in f_i for all $i \in 1..M$
- p_i is strictly decreasing in f_0 for all $i \in 1..M$

Assumption 2 guarantee that regulation can only have a positive impact on the stability of the financial system, We exclude from this paper the discussion about the unwanted negative effect of vaccination

Assumption 3 • For all $i \in 1..M$, there exists a value $\bar{f}$ where $\frac{\partial p_i^2}{\partial f_0 \partial f_i} \geq 0$ for $f_i \geq \bar{f}_i$ and $\frac{\partial p_i^2}{\partial f_0 \partial f_i} \leq 0$ for $f_i \leq \bar{f}_i$

- p_i is strictly convex then strictly concave in $f_i \forall i$
- all first and second derivatives exists and are continuous.

Assumptions 3 are more technical and are required to ensure the existence of a solution of the system 4.12 but also have an economic interpretation. Here we assume that the marginal stability improvement due to more regulation expenditure is increasing then decreasing after reaching the level of regulatory efforts $\bar{f}$. In other words, the first regulatory efforts have the maximum impact on stability, the pace of increase of those benefits slowly decrease until reaching 0 then become negative after the value $\bar{f}$. $\bar{f}$ should not be interpreted as the optimal level of regulatory expenditure because costs of regulation are not taken into account in those assumptions. Regulatory efforts beyond $\bar{f}$ will only yield less benefit compared to those before reaching $\bar{f}$.

The game problem solution in (4.12) satisfies:

$$f_0^G = \sup \left\{ f \in [0, 1] : b_0 \frac{\partial}{\partial f_0} T_0(f_0) \Big|_{f_0=f} + v_0 N_0 - c_0 \frac{\partial}{\partial f_0} \sum_{j=1}^M y_j(1, f_0, f_j^G) \Big|_{f_0=f} < 0 \right\} \quad (4.13)$$

$$f_j^G = \sup \left\{ f \in [0, 1] : b_0 \frac{\partial}{\partial f_j} T_0(f_j, f_0^G) \Big|_{f_j=f} + v_j N_j < 0 \right\} \quad (4.14)$$

In equation (4.14), it is important to note that if $b_0 \frac{\partial}{\partial f_0} T_0(f_j, f_0^G) + v_j N_j < 0$, then the benefits of regulating an monitoring assets in the peripheral countries (countries that are not the source country) is so much higher than the government intervention costs to stabilize the financial system. In this context, the government should aim at protecting all the asset units in the financial system (f_j^G). If this condition is not met, the optimal fraction f_j^G is less than 1 and is the unique solution of $b_0 \frac{\partial}{\partial f_0} T_0(f_j, f_0^G) + v_j N_j = 0$. This interpretation is slightly different for country 0. In fact, the benefit of vaccination is attenuated by the possible positive effect of first order contagion. But again, if $b_0 \frac{\partial}{\partial f_0} T_0(f_0) + v_0 N_0 - c_0 \frac{\partial}{\partial f_0} \sum_{j=1}^M y_j(1, f_0, f_j^G) < 0$ the

optimal vaccination level will be reached at 1. Otherwise, the unique solution is when $b_0 \frac{\partial}{\partial f_0} T_0(f_0) + v_0 N_0 - c_0 \frac{\partial}{\partial f_0} \sum_{j=1}^M y_j(1, f_0, f_j^G) = 0$.

In this paper, the relevant discussion is in the second case when the optimal hedging policy f_j^G is less than 1 mainly because the policy implication when regulation is highly beneficial compared to costs are straightforward (government should aim at protecting all asset units) and also the comparison between the central planner solution and the game problem is not relevant as both will suggest the same outcome. We rather focus on the other solution because governments are required to make trade-offs between costs and benefits of regulation where this model can become handy.

The optimal vaccination response in the absence of contagion benefit should be higher compared to the solution with contagion benefits for the source country. In fact, the source country has the motive to reduce its efforts to stabilize the global financial system when it has the possibility to export part of the stress outside its borders. We implicitly assume that there is no limit to the capital outflows that can be imposed by regulators in case of crisis to prevent contagion.

Needless to say that we do not impose resources restriction in this model where governments are unable to implement their optimal solution due to lack of available resources.

Coordinated decision

In this section, we will study the decision of a central planner with a global stability perspective. The planner is interested in minimizing the overall financial and economic costs of the system as a whole. This type of central planner can be a supranational regulator such as the European Central Bank (ECB) in Europe who coordinates between different national regulators. The FED can also be seen as a central planner after the Dodd-Frank act as it is considered to be the coordinator between different federal agencies with different mandates and asset classes to supervise. The costs for the central planner is defined by summing the total cost as perceived by each government. It is important to highlight that the central planner

should not consider the benefits of contagion to the source country for obvious reason but in the total cost, we will continue to account for it. The reason behind this choice is that the source country does not bare all the cost of the failed risk units. In fact, not subtracting the benefits of contagion means that we will consider that cost twice. Another way of modeling it would be to consider that the average cost of the county 0 in the system problem is different from the average cost in the game problem. However, for consistency reason, we decide to maintain the benefits of contagion in the system problem without changing the value of b_0 .

The problem can be formulated as following:

$$\min_{0 \leq f_0 \leq 1} SF(f) = b_0 T_0(f_0) + v_0 f_0 N_0 - c_0 \sum_{j=1}^M y_j(1, f) + \sum_{i=1}^M b_i T_i(f_i, f_0) + v_i f_i N_i \quad (4.15)$$

To be able to characterize the solution of this problem we need to assume that the expected costs of vaccination do not exceed the eventual gains. The system problem solution satisfies:

$$f_0^S = \sup \left\{ f \in [0, 1] : b_0 \frac{\partial}{\partial f_0} T_0(f_0) \Big|_{f_0=f} + v_0 N_0 - c_0 \sum_{j=1}^M \frac{\partial}{\partial f_0} y_j(1, f_0, f_j^S) \Big|_{f_0=f} + \sum_{i=1}^M b_i \frac{\partial}{\partial f_0} T_i(f_i^S, f_0) \Big|_{f_0=f} < 0 \right\} \quad (4.16)$$

$$f_j^S = \sup \left\{ f \in [0, 1] : b_0 \frac{\partial}{\partial f_0} T_0(f_j, f_0^S) \Big|_{f_0=f} + v_j N_j < 0 \right\} \quad (4.17)$$

Figure 4.11 and 4.12 show the optimal fraction of asset units to hedge compared to the size of the financial system both for the game problem and the central planner problem. The first figure (figure 4.11) focus on the source country while the second figure (figure 4.12) focus on one of the peripheral country. The figure shows that the central planner always decides on more intense regulation in the source country and

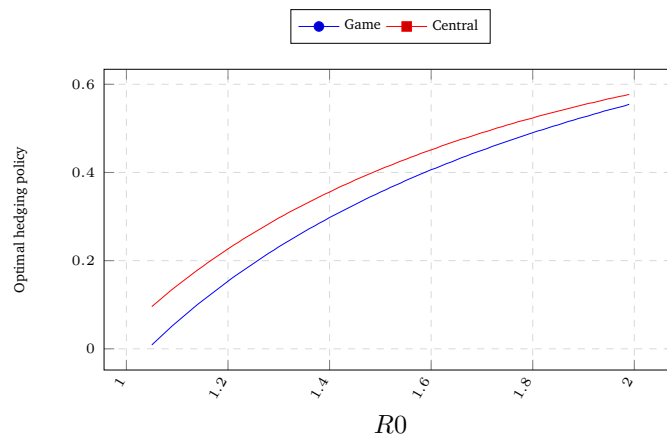


Fig. 4.11.: The optimal Hedging strategy for the source country in the Game problem and the Central planner solution.

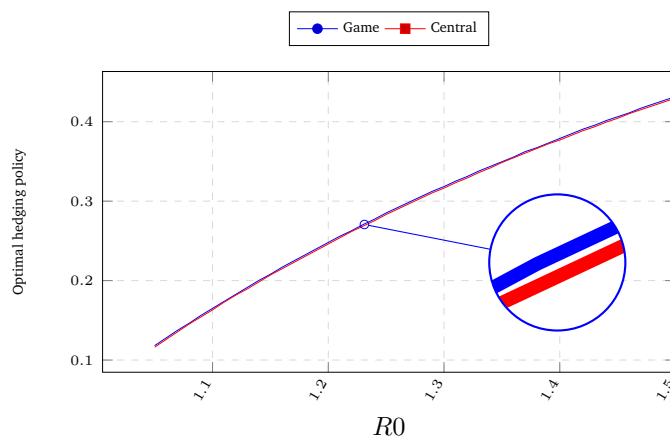


Fig. 4.12.: The optimal Hedging strategy for a peripheral country in the Game problem and the Central planner solution.

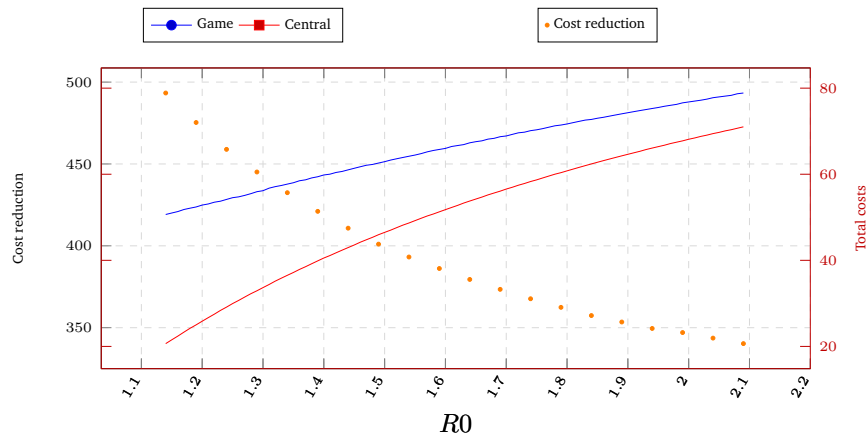


Fig. 4.13.: The evolution of the total costs including the losses due to a financial crisis and contagion on the system both in the case of the game problem or the central planner solution represented per 1000 of asset units as a function of the severity of the crisis. The right axis represents the gains from the central planner solution per 1000 of asset unit. Other simulation parameters are chosen randomly.

protecting less asset unit on the peripheral country. The difference is noticeable for the source country and shows the importance of the regulation in the country where the crisis starts. The difference is less visible for peripheral countries mainly because they have smaller impact on the overall costs and also because the effect of the additional effort made by the source country is divided among all other countries. Figure 4.13 shows the cost reduction of the central planner solution compared to the game problem. The benefits decrease when the contagion channels are more important between countries or the increasing severity of the crisis.

Coordination contract between individual governments

Given that the central planner and individual supervisors have different game solution, it is important to design a collaboration contract. The aim of the contract is to resolve the issue of the difference of incentives that exists between the central planner and individual governments and push the individual decision toward the system-optimal. The idea of the contract is very intuitive. The non-source country should subsidize the source country regulatory action to offer a form of guarantee to more risk units. The additional funds received by the source country should be used to reduce the failure probability and infectiousness potential. The non-source country expects benefits from subsidizing the source country. In return, the addi-

tional efforts conducted by the source country should have benefits on the number of failed risk units inside their borders. In addition, we also expect that governments interacting within the financial system will accept to split the costs saving that resulted from the contract or the shift toward the system's optimal solution.

The individual subsidy paid by the non-source country to the source country is :

$$G_i(f_0) = (\alpha_i - 1) GF_i(f_i^s, f_0) + \alpha_i \sum_{\substack{j=0 \\ j \neq i}}^M GF_j(f_j^s, f_0) \quad (4.18)$$

Where α_i $i \in 1..M$ are chosen such that: $\alpha_i, i \in (0, 1) \forall i$ and $\sum_{i=1}^M \alpha_i < 1$. The basic idea is that any country will pay a fraction of the cost reduction of the crisis seen from the system perspective. The subsidy by country i increase with the total benefit from the extra regulation in country 0, $\sum_{j=0}^M GF_j(f_j^s, f_0)$ and decrease with the country i own costs. α_i is the fraction of contribution of each country that can be fixed according to different criterion. The exposure of country i to the source country or any other form of indicators. It can also be fixed to improve the individual costs of the crisis.

Theorem 1 *If country 0 chooses a vaccination policy of f_0 and if each non source country pays a subsidy of $G_i(f_0)$ for $i \in 1, 2, \dots, M$ then*

- *The total costs to country i is equal to $\alpha_i SF(f^s)$ for $i \in 0..M$ and where $\alpha_0 = 1 - \sum_{i=1}^M \alpha_i$*
- *Additionally if $\alpha_i = \frac{GF_i(g_i^G, f_0^G)}{SF(f^G)}$ then the total cost of crisis of each country is improved compared to the solution of the individual vaccination policy of the game problem.*

The total subsidy received by the source country is:

$$G_0(f_0) = \sum_{i=1}^M G_i(f_0) = (1 - \alpha_0) GF_0(f_0) - \alpha_0 \sum_{i=1}^M GF_i(f_i^s, f_0). \quad (4.19)$$

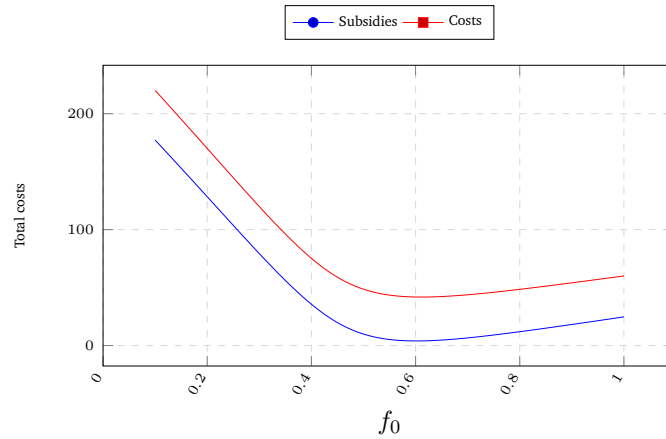


Fig. 4.14.: The line costs represents the evolution of the total costs to the source country (country 0) of a financial crisis a function of the regulatory effort f_0 . Subsidies represents the total subsidies of the non-source country to the hedging effort to the country 0

The total subsidy is a weighted average between the cost of the crisis to the source country and the others impacted via contagion by the original failure. It can also be seen that the subsidy contract tries to push the solution toward the system solution. In fact, if each country will adopt the system solution for its policy of financial regulations the subsidy to the source country will become 0. However, the more regulators deviate from the optimal central planner solution, and the higher is the subsidy to country 0 to reduce the costs of selfish behaviors.

Figure (155) shows the total subsidies received by the source country to enhance the stability of its financial system that was given by the non-source country susceptible to contagion. The figure shows a decreasing and an increasing pattern. In the first region, subsidies are high but decreasing this is mainly because the source country already benefits from increasing the regulatory efforts because this should lower the costs of the crisis which is relatively high for low values of f_0 . In the second region, subsidies are low but increasing which translates that the source country has no longer the incentives to increase its efforts because this would benefit more the peripheral countries that increase their subsidies to encourage the source country to increase its efforts of hedging. Overall, the cost to the source country taking the subsidies into consideration is decreasing while f_0 increases. The conclusion to draw from this figure and the policy implications are that a source country should not be incentivized to have a minimum form of financial monitoring and regulations. However, this country will reach its optimal level of regulation before the system

level and will stop investing in the stability of the financial system despite the positive impact on its economic partners. For that reason, the peripheral countries should share the costs of this additional regulations. The subsidy is more important for higher value of f_0 because the marginal benefit of regulation decrease for the source country in the second region of the the curve. In other words, country 0 will require more financial incentives to increase its financial safety because it will have little additional benefits itself. In a nutshell, the conclusion is that the source country need no incentive to invest in its financial stability but it should be helped when the marginal benefits of regulation to country 0 does not exceed the costs but the the marginal benefits to the system remains interesting.

In a nutshell, the policy message is that the source country need no incentive to invest in its financial stability but it should be helped when the marginal benefits of regulation to country 0 does not exceed the costs but the marginal benefits to the system remains interesting.

4.6 Policy implication and discussion of the model

The policy implication of the model are strongly related to one of the main assumption in this chapter which stipulates that the source country is identified ex-ante and the international coordination is then derived based on that information. It is true that in history some crisis started from non-anticipated countries with what seemed to be a strong and growing economy. However, we think that this framework can also be applied in practice without such a prior knowledge of the mechanism of the crisis. For the purpose of financial stability national regulators need to identify all the possible threats resulting from international contagion. It is possible to apply this framework based on exposures. In fact, if the national financial system is highly exposed to some country it is interesting to consider the benefits of collaboration between both regulators in the both countries. For example, core countries in Europe can consider the question of collaboration with peripheral european countries and even if deemed necessary subsidizing regulation in those countries even in the absence of sign of upcoming crises originating in the periphery. The exposure of the financial system to those countries can justify considering the benefits of contagion.

It is better to be prepared to those bad events because cleaning the mess ex-post can be costly. Moreover, the same type of simulation can be conducted considering the source country to be a region or a class of assets and then national regulator can decide on their collaboration strategy based on the benefits which can be yielded from considering each time a relevant source country individually. We admit that a better setting would be a model in which a crisis starts at random in one of the countries and then try to consider the benefits of regulatory expenditures in each country based on such a more comprehensive model. Results of the first part of the model can be extended to cover this dynamic but it is beyond the scope of the game theory results that we presented in this chapter.

The model also has other type of limitation and potentially caveats. We assume that the potential of contagion and susceptibility are parameters that a supervisor can correctly evaluate before a financial crisis. In practice, this is rather difficult. It may be possible to quantify the effect of rational contagion based on financial exposures and financial linkages. It may be possible to design a model that translates all the channels of fundamental contagion into potential losses due to contagion, but the applying the same is difficult for other channels that have a more behavioral aspect. In the later transmission mechanism, behaviors such as herding and loss of confidence are difficult to pin down to correctly estimate their impact. Underestimating those channels is always a risk difficult to mitigate and also overestimation to have a good security margin can yield unnecessary expenditures. We have to admit in this case that some expert judgement is important to complement the conclusion of the strategic interaction model that we propose in this chapter. A second caveat that the model does not consider explicitly is that there is a risk that subsidizing the source country, although rational if we consider costs, can create a problem of moral hazard for both the peripheral and source country. The source country knowing that other countries are inclined to provide financial help will wait for subsidies to improve the stability of its financial system. And peripheral countries can count on their neighboring countries to stabilize their system and neglecting theirs. Only the problem of moral hazard for the source country is somehow addressed in this chapter. Because the initial marginal benefit of regulation for the source country is so high that this probably mitigate any moral hazard issues. Only later when additional expenditure yields small benefits that subsidies are required.

Conclusion

Financial regulation faces new challenges after the last financial crisis that highlighted fragility in the increasingly integrated global financial system. Many regulators had then endeavored in a series of reforms to avoid the costs of similar crisis episodes. One of the main questions in those reforms is the framework of collaboration with entities outside the national scope. The main contribution of this paper is the theoretical framework that we propose to model interaction between asset classes in the financial system and modification to the risk profile for each of those classes that could result from regulation.

This paper provided the theoretical justification via the adaptation of the famous SIR model extensively used in the biological world. We also provide a contract to subsidize the source country of the outbreak to prevent contagion and reduce risks in other countries. The results should encourage regulators to consider the international dimension in their regulatory efforts expenditure. Depending on the level of interconnectedness, peripheral countries should help the source country in its regulatory effort beyond its selfish optimal level in the game problem. In fact, the country where the crisis begins has no incentive to stabilize more its financial system, but this can have important external effects on other countries. We also proposed a contract for collaboration that should reduce the costs of a financial crisis to both the source country and the peripheral country. The latter should be the best argument for collaboration and even subsidizing other countries.

Beyond the results that are presented in this paper, the framework of the financial system offers possibilities for other application related to regulation contract between several actors in the financial system in the presence of costs. It could be used to justify for example the financing of a central clearing system by several banks to manage third party risk and avoid contagion to other entities in the system. Fees and interest paid to the central clearing houses can be incorporated in this model as the form a subsidy that the source of the stress received to increase its safety. Despite the strong theoretical results, this approach can face a major challenge which is calibration. It is in practice difficult to estimate the linkages between financial systems and predetermine the costs of regulation and the expected posi-

tive effect of increasing that regulation. Although statistician working on the SIR model for the biological world had developed a very rich literature on the subject of estimation the outbreak calibration parameters, economists need to study to what extent those statistical advances could be transposed into the financial context.

List of Symbols

N_i The size of the population of value units in country i .

P_i The risk profile i .

R_0 The expected number of secondary stress situations transmitted from a single stressed asset unit in a complete population of non-vaccinated asset units.

R_{ij} The number of secondary infected asset units with the risk profile P_m in country j directly affected from a randomly selected infectious asset unit having the risk profile P_n in country i .

Rc_i The subset of risk units that recovered or failed in the country i .

S_0 The state S_0 is the normal state.

S_1 The state S_1 is the stressed state.

S_2 The state S_2 is the failure state.

S_3 The state S_3 is the recovery state.

T_i The total number of failure.

ϕ The effect of government interventions (vaccination of asset units) on infectiousness in a risk profile P .

ψ The combined effect of vaccination on value units..

θ The impact of vaccination on susceptibility in a risk profile P .

b_i The average direct and indirect cost from the stress and potential failure of a single elementary asset unit.

c_i The benefit to the source country that results from contagion.

f The vector representing the government vaccination policy.

f_i The fraction of elementary asset units that a government i decides to protect or monitor.

m_{ij} The value unit j in country i .

p_i The fraction of failed asset units at the end of a crisis.

v_i The average unit cost of vaccinating a single asset unit.

y_j The first generation of infected asset units..

Country 0 The country where the outbreak is anticipated to start.

M The number of countries excluding the source country.

A. Proof of Equations (4.10)

In this section, we present how to compute χ_i . Where χ_i is the fraction of asset units in country i that become stressed after the first wave of contagion.

First, we define χ as :

$$\chi = \frac{R_{i0} \mathcal{J}_i^{(g_0)}(f_i, f_0)}{N_i} \quad (4.20)$$

Let's start by assuming that the vector $\mathbf{y}$ represents the value of expected number of infections at generation g . The time between generation is the mean average for infection. An asset unit infected at generation g was in the susceptible state at generation $g - 1$. We

assume, without any loss of generality, that we have only a protected and a non-protected risk profiles in each country

$$\mathbf{y}(g) = [y_{00}(g), y_{10}(g), \dots, y_{0M}(g), y_{1M}(g)]^T \quad (4.21)$$

Where $y_{ij}(g)$ is the number of infections at generation g in country i in the risk profile j .

Following the next generation model(it account for contagion within countries and inter-countries) we can write that :

$$\mathbf{y}(g+1) = \mathbf{N}^g \mathbf{y}(0) \quad (4.22)$$

Where $\mathbf{N}^g$ is the next generation matrix.

$\mathbf{N}^g$ can be written as :

$$\mathbf{N}^g = \begin{bmatrix} R_{00}(1-f_0) & R_{00}\phi(1-f_0) & \cdots & R_{0M}(1-f_0) & R_{0M}\phi(1-f_0) \\ R_{00}\theta f_0 & R_{00}\phi\theta f_0 & \cdots & R_{0M}\theta f_0 & R_{0M}\phi\theta f_0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ R_{M0}(1-f_M) & R_{M0}\phi(1-f_M) & \cdots & R_{0M}(1-f_M) & R_{0M}\phi(1-f_M) \\ R_{M0}\theta f_M & R_{M0}\phi\theta f_M & \cdots & R_{MM}\theta f_M & R_{MM}\phi\theta f_M \end{bmatrix} \quad (4.23)$$

Given the assumption that the crisis starts only in country 0. $\mathbf{y}(0)$ can be written as:

$$\mathbf{y}(0) = [y_{00}(0), y_{10}(0), 0, 0, \dots, 0, 0]^T \quad (4.24)$$

Then we can show via mathematical induction that for any generation g of infected asset units can be written as:

$$y_{0i}(g) + y_{1i}(g) = R_{i0}\mathcal{J}_i^{(g)}(f_i, f_0) + \mathcal{G}_i^{(g)}(\mathbf{f}) \quad (4.25)$$

Where $\mathcal{J}^{(g)}(f_i, f_0)$ is a function of f_i and f_0 but none of the other f_j $j \neq i$. All other terms involving all other regulation fraction f_j for any $j \notin \{0, i\}$. Those terms has a coefferient $R_{ab}R_{bc}$ for some $a \neq b \neq c$, where $a, b, c \in \{0, 1, \dots, M\}$.

Starting from equation (4.22), we regroup all the terms with f_i and f_0 we can write $\mathcal{J}_i^{(g)}(f_i, f_0)$ as:

$$\begin{aligned}\mathcal{J}_i^{(g)}(f_i, f_0) = & (y_{00}(0) + \phi y_{00}(0)) (1 - \psi f_0) \\ & \left(R_{ii} \left(\mathcal{J}_{ii}^{g-1}(\mathbf{f}) + \mathcal{K}_{ii}^{g-1}(\mathbf{f}) \right) \right) \\ & + \left(R_{00} \left(\mathcal{J}_{00}^{g-1}(\mathbf{f}) + \mathcal{K}_{00}^{g-1}(\mathbf{f}) \right) \right)\end{aligned}\quad (4.26)$$

Where we define $\mathcal{J}_{ii}^g(\mathbf{f})$ and $\mathcal{K}_{ii}^g(\mathbf{f})$ by

$$R_{ij} \mathcal{J}_{ij}^g(\mathbf{f}) = (1 - \psi f_j) \sum_{k=0}^M R_{ik} R_{kj} \mathcal{J}_{ij}^{g-1}(\mathbf{f}) \quad (4.27)$$

$$R_{ij} \mathcal{K}_{ij}^g(\mathbf{f}) = (1 - \psi f_j) \sum_{k=0}^M R_{ik} R_{kj} \mathcal{K}_{ij}^{g-1}(\mathbf{f}) \quad (4.28)$$

for all $i, j \in \{0, 1, \dots, M\}$ and $g > 1$. And we define the initial conditions to be $\mathcal{J}_{ij}^1(\mathbf{f}) = 1 - f_j$ and $\mathcal{K}_{ij}^1(\mathbf{f}) = \theta f_j$.

Those are all the equation that can be required to compute Ξ_i .

4.2 Parameter used for figures (4.5), (4.6), (4.7) and (4.8)

The impact of vaccination on susceptibility in a risk profile	$\theta = 0.9$
The effect of government interventions on infectiousness in a risk profile	$\phi = 0.2$
The time period for the simulation	$t = 5$
The descretization unit of time	$h = \frac{1}{1000}$
The size of the population	$N = 1000$
The recovery rate	$\gamma = 2$

Conclusions

The financial regulations is in perpetual change shaped by series of financial crises. The Basel agreement had known several revisions since the first 1988 accords. Since then, risk measurement becomes the cornerstone of banks health. It appears necessary to access the risk inherent within each asset class in general and individual assets in particular. Risk measures like Value-at-Risk and Expected Shortfall made famous in the banking industry bring to the surface the question of computational complexity. In the first chapter of this thesis, we tried to contribute to the literature on the subject and more specifically propose some enhancement to the direct nested-simulations technique.

We first started by introducing a sequential algorithm based on the intuition that the marginal improvement of extending inner-scenarios decrease exponentially and that a budget saving can result from better allocation laws. We started the algorithm from a rule of thumb that adding a single additional inner scenario should increase the probability of impacting the final value of the estimator. Then we showed via relatively complex sets of simulations (multidimensional and nonlinear payoffs functions)that the idea could be verified in practice. In addition, we provided theoretical proof that the technique that we propose effectively yield more efficient allocation of the computational effort.

Nevertheless, we have to point out few important caveat to bear in mind. First, the algorithms that we propose improve only the efficiency of the computational part and should not have any effect on the quality of the estimator itself from a risk management perspective. In fact, the implementation of what we propose suppose that the scenario-generating technology and the pricing technique are unchanged. In other words, using the output of this paper can increase the number of real-world scenarios but cannot expand the scope of coverage of the measurement in the first place. Moreover, this paper is designed for large and complex derivatives. In fact,

an implicit assumption in the simulation analysis is that the burden of repricing the portfolio under each scenario is significant. As a result, the extra operation that we need to find the best allocation is negligible compared to the repricing burden which explains why risk managers engage in simulations optimizations in the first place.

Regulators can also draw some lessons from this first chapter. First, they should be aware that computing risk measures is highly cumbersome from a computational perspective and that the cost of increasing the number of real-world scenarios is important. This should be kept in mind when they require banks to supplement additional results for stress- tests for example. In fact, if the number of those scenarios increase, banks are obliged to apply some simplification to the pricing algorithms to accommodate the extra-regulatory requirements. At the end, it could seriously impact the quality of the risk assessment. Second, regulators should also be aware that literature on simulation also provided some solutions to the issue of nested simulations and that the trade-off between precision and practicality is not always fated.

The second chapter, introduces the idea that banks should directly pay for any risk they take and especially extreme risks. In fact, we transpose the difference between tail asymmetry into a measure of banks' externalities that could result in a systemic crisis. While we consider that skewness is acceptable in financial markets because it is the result of different expectations, tails asymmetry is not tolerated. In fact, because if extreme losses are unpredictable so should extreme gains. We show via different techniques that tails asymmetry can contribute to the overall fragility of the financial system. We argue that this feature should be limited and also monitored by regulators. We do not limit the scope of this chapter to the theoretical dimension, we propose also a measure of externalities based on publicly available price data. The measure proposed is building on extreme value theory (EVT). We finally show that the externalities measure that we propose better explains proxies of negative-externalities than classical measures of systemic risk such as the total fines paid *ex post*.

Nevertheless, some extra-caution is also advised if this approach is ever to be implemented. Gains and more precisely financial innovation should not be stigmatized as they are also a indicator of the good health of the financial intuition. Monitoring gains, should not incentivize banks to curtail their risk behavior to the point that harms economic growth. Moreover, banks can also apply a sort of regulatory arbitrage where they hide their gains or divert them into other jurisdiction where profits are not monitored. Such behavior, can harm investors confidence in banks and be a source of financial instability itself.

The last chapter of this thesis covers a major challenge that emerged to supervisors: the international collaboration. The Basel committee is the living proof that the question of collaboration was raised few decades ago but the recent globalization of the financial system and the emergence of global banks made the question resurface again as a global priority. The major contribution of this paper is the theoretical framework that we propose to model the interaction between asset classes in the financial system coupled with regulatory efforts to stabilize this system. This model is then explored to justify theoretically the importance of international collaboration to weather the effects of crises.

We also describe a contract to subsidize the source country of the outbreak to reduce contagion risks to other peripheral countries. The outcome of the model encourages supervisors to consider the international dimension in their regulatory efforts expenditure. Depending on the level of interconnectedness, peripheral countries can have financial benefits to help the source country in its regulatory effort beyond the selfish optimal level of the later in the game problem. In fact, the source country reaches its optimal level of regulation before the optimal point from a central planner view. Thus the first has no incentive to go beyond even if it decrease dramatically the global costs of a crisis. It is the role of the peripheral country then to incentivize the source country to enhance the stability of the global financial system.

Beyond the applications of the model to the financial system that are presented in this paper, the framework offers possibilities for other application related to collab-

orations contract between several actors in the financial system in the presence of costs.

It could be used to justify for example the financing of a central clearing system by several banks to manage third party risk and avoid contagion to other entities in the system. Fees and interest paid to the central clearing houses can be seen as a form of subsidy to the central counterpart. Despite the strong theoretical results, this approach can face a challenge which is calibration. It is in practice difficult to estimate the linkages between financial systems and predetermine the costs of regulation and the expected positive effect of increasing that regulation. Although statistician working on the SIR model for the biological world had developed a very rich literature on the subject of estimation the outbreak calibration parameters, economists need to study to what extent those statistical advances could be transposed into the financial context. Data availability and confidentiality of interlinkages data can also represent a major challenge.

In this report we tried to cover several aspects of financial stability. Each chapter, tried to consider the issue with different hats: individual financial institution, national regulators and international regulatory bodies. However, we recon that other aspects of the financial regulation remain beyond the somehow large scope of the document. The electronic security of the banking industry is more important than ever with the increase development of the computers use ranging from everyday banking, international transaction settlement to risk management. The failure of the cyber-network can result on a catastrophic scenario only imagined so far in blockbuster Hollywood movies.

Moreover, new actors on the financial system are tacking an increasing importance in the network of financial intermediation. More precisely, finetechs and shadow banking provides alternatives to the traditional banking topologies. They increase the accessibility and the efficiency of the financial industry and sometimes with huge cost reduction. However, regulation must follow to prevent a major financial breakdown originating in those nodes of the network.

Bibliography

- Acharya, Viral V., Lasse Heje Pedersen, Thomas Philippon, and Matthew P. Richardson (2017). „Measuring Systemic Risk“. In: *The Review of Financial Studies* May (cit. on pp. 71, 72, 100, 101, 103).
- Adrian, Tobias and Brunnermeier (2016). „CoVaR“. In: *American Economic Review* (cit. on pp. 71, 72, 100, 101).
- Allen, F and D Gale (1997). „Financial markets, intermediaries, and intertemporal smoothing“. In: *Journal of Political Economy* (cit. on p. 69).
- (2000). „Financial contagion“. In: *Journal of political economy* (cit. on pp. 82–84, 87, 113, 117).
- Arora, Sanjeev and Boaz Barak (2009). „Computational complexity and information asymmetry in financial products“. In: ... *Center for Computational* ... (Cit. on p. 67).
- Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber, and David Heath (1999). „Coherent Measures of Risk“. In: *Mathematical Finance* 9.3, pp. 203–228 (cit. on pp. 18, 117).
- Bali, Turan G., Nusret Cakici, and Fousseni Chabi-Yo (2011). „A Generalized Measure of Riskiness“. In: *Management Science* 57.8, pp. 1406–1423. eprint: <http://mansci.journal.informs.org/content/57/8/1406.full.pdf+html> (cit. on p. 18).
- Barigozzi, Matteo and Christian T Brownlees (2013). „NETS: Network Estimation for Time Series“. In: pp. 1–33 (cit. on p. 71).
- Bauer, Daniel, A. Reuß, and D. Singer (2012). „On the calculation of the solvency capital requirement based on nested simulations“. In: *ASTIN Bulletin* 42.2, pp. 453–499 (cit. on pp. 18, 26).
- Beck, Thorsten, Ross Levine, and Alexey Levkov (2010). „Big bad banks? The winners and losers from bank deregulation in the United States“. In: *The Journal of Finance* 65.5, pp. 1637–1667 (cit. on p. 81).
- Beltratti, Andrea and René M. Stulz (2012). „The credit crisis around the globe: Why did some banks perform better?“ In: *Journal of Financial Economics* 105.1, pp. 1–17 (cit. on p. 68).
- Bernard, Carole, Eike Christian Brechmann, and Czado Claudia (2013). „Statistical Assessments of Systemic Risk Measures“. In: *Handbook on systemic risk*. Ed. by Jean-Pierre Fouque and Joseph A. Langsam. Cambridge: Cambridge University Press, pp. 165–179 (cit. on pp. 69, 72).

- Bisias, Dimitrios, Mark Flood, A Lo, and Stavros Valavanis (2012). „A survey of systemic risk analytics“. In: *US Department of Treasury*, ... (Cit. on p. 66).
- Blanco, C. and S Blomstrom (1999). „VaR applications : Setting VaR Based limits“. In: *Financial engineering associate* (cit. on p. 15).
- Bordo, Md and Ap Murshid (2000). „Are Financial Crises Becoming Increasingly More Contagious?“. In: *International Financial Contagion* 3 (cit. on p. 120).
- Boss, Michael, Helmut Elsinger, Martin Summer, and Stefan Thurner (2004). „Network topology of the interbank market“. In: *Quantitative Finance* 4.6, pp. 677–684. arXiv: 0309582 [cond-mat] (cit. on p. 117).
- Bouchard, Bruno and Xavier Warin (2012). „Monte-Carlo Valuation of American Options: Facts and New Algorithms to Improve Existing Methods“. English. In: *Springer Proceedings in Mathematics* 12. Ed. by René A. Carmona, Pierre Del Moral, Peng Hu, and Nadia Oudjane, pp. 215–255 (cit. on p. 19).
- Broadie, M., Y. Du, and C. C. Moallemi (2011). „Efficient Risk Estimation via Nested Sequential Simulation“. In: *Management Science* 57.6, pp. 1172–1194 (cit. on pp. 7, 16, 20, 31, 33, 38, 39, 45, 58).
- Brownlees, C and R Engle (2011). „Volatility, correlation and tails for systemic risk measurement“. In: *Available at SSRN 1611229* October (cit. on pp. 71, 103).
- Brunnermeier, Markus K. (2008). „bubbles“. In: *The New Palgrave Dictionary of Economics* (cit. on p. 72).
- Brunnermeier, Markus K. and Martin Oehmke (In Press). „Bubbles, Financial Crises, and Systemic Risk“. In: *Handbook of the Economics of Finance*. Amsterdam: Elsevier (cit. on p. 68).
- Caballero, RJ (2010). „The" other" imbalance and the financial crisis“. In: (cit. on p. 71).
- Claessens, Stijn and Kristin Forbes (2004). „International Financial Contagion: The Theory, Evidence, and Policy Implications“. In: *The IMF's role in emerging market economies: Reassessing the adequacy of its resources*, pp. 18–19 (cit. on p. 122).
- Cochran, James J., Louis A. Cox, Pinar Keskinocak, Jeffrey P. Kharoufeh, and J. Cole Smith (2010). „Robust External Risk Measures“. In: (cit. on p. 18).
- Cont, Rama, Amal Moussa, and Edson B. Santos (2013). „Network Structure and Systemic Risk in Banking Systems“. In: *Handbook on systemic risk*. Ed. by Jean-Pierre Fouque and Joseph A. Langsam. Cambridge: Cambridge University Press, pp. 327–365 (cit. on pp. 71, 118).
- Crockett, Andrew (2000). „Marrying the micro-and macro-prudential dimensions of financial stability“. In: *BIS speeches* September (cit. on p. 66).
- Crouhy, M D.Galai R. Mark (2000). *Risk management*. McGraw-Hill (cit. on p. 18).
- Cuoco, Domenico, Hua He, and Sergei Isaenko (2008). „Optimal Dynamic Trading Strategies with Risk Limits“. In: *Operations Research* 56.2, pp. 358–368 (cit. on p. 15).
- Degryse, Hans and Grégory Nguyen (2007). „Interbank Exposures: an Empirical Examination of Systemic Risk in The Belgian Banking System“. In: *International Journal of Central Banking* 3.2, pp. 123–171 (cit. on p. 117).

- Devineau, Laurent and S Loisel (2009). „Risk aggregation in Solvency II: How to converge the approaches of the internal models and those of the standard formula?“ In: *Bulletin Français d'Actuariat*, pp. 1–40 (cit. on pp. 18, 26).
- Diamond, DW and PH Dybvig (1983). „Bank runs, deposit insurance, and liquidity“. In: *The journal of political economy* 91.3, pp. 401–419 (cit. on pp. 75, 78, 113, 117).
- Diamond, DW and RG Rajan (2005). „Liquidity shortages and banking crises“. In: *The Journal of finance* (cit. on pp. 82, 88, 117, 122).
- Dodd-Frank (2010). *Dodd-Frank Wall Street Reform and Consumer Protection Act*. 111th Congress Public Law 203 (cit. on p. 102).
- Dungey, Mardi, Matteo Luciani, and David Veredas (2012). „Ranking Systemically Important Financial Institutions“. In: *SSRN Electronic Journal*, pp. 1–33 (cit. on p. 71).
- Edwards, Sebastian (2000). „Contagion“. In: February (cit. on p. 114).
- Eisenberg, Larry and TH Noe (2001). „Systemic risk in financial systems“. In: *Management Science* 603 (cit. on pp. 71, 73, 75, 117).
- Elsinger, Helmut, Alfred Lehar, and Martin Summer (2006). „Risk Assessment for Banking Systems“. In: *Management Science* 52.9, pp. 1301–1314 (cit. on pp. 73, 117).
- (2013). „Network Models and Systemic Risk Assessment“. In: *Handbook on systemic risk*. Ed. by Jean-Pierre Fouque and Joseph A. Langsam. Cambridge: Cambridge University Press, pp. 287–303 (cit. on p. 71).
- Embrechts, Paul, Claudia Klupperlberg, and Thomas Mikosch (1997). *Modelling Extremal Events*. Vol. 33. Springer (cit. on pp. 90–93).
- Engle, R. F. (2011). „Long-Term Skewness and Systemic Risk“. In: *Journal of Financial Econometrics* 9.3, pp. 437–468 (cit. on pp. 72, 77).
- Fouque, Jean-Pierre and Li-Hsien Sun (2013). „Systemic Risk Illustrated“. In: *Handbook on systemic risk*. Ed. by Jean-Pierre Fouque and Joseph A. Langsam. Cambridge: Cambridge University Press, pp. 444–542 (cit. on pp. 81, 88).
- Garas, Antonios, Panos Argyrakis, Céline Rozenblat, Marco Tomassini, and Shlomo Havlin (2010). „Worldwide spreading of economic crisis“. In: *New Journal of Physics* 12.11, p. 113043. arXiv: 1008.3893 (cit. on p. 120).
- Gnedenko, BV (1970). „Limit theorems for sums of a random number of positive independent random variables“. In: *Proceedings 6th Berkeley Symposium on Math. Stat. ...* Pp. 537–549 (cit. on p. 90).
- Gordy, M. B. and S. Juneja (2010). „Nested Simulation in Portfolio Risk Measurement“. In: *Management Science* 56.10, pp. 1833–1848 (cit. on pp. 7, 16, 19–21, 28, 30, 31, 57–59).
- Gropp, Reint, Marco Lo Duca, and Jukka Vesala (2009). „Cross-Border Bank Contagion in Europe“. In: *International Journal of Central Banking* 5.1, pp. 97–139 (cit. on p. 117).
- Guillaume, DM and MM Dacorogna (1997). „From the bird's eye to the microscope: A survey of new stylized facts of the intra-daily foreign exchange markets“. In: *Finance and Stochastics* 129. September 1996, pp. 95–129 (cit. on p. 90).
- Guojun Gana, X. Sheldon Lin (2015). „Valuation of large variable annuity portfolios under nested simulation: A functional data approach“. In: *Insurance: Mathematics and Economics* (cit. on p. 19).

- GURRÍA, J A (1999). *Identifying the Pending Agenda for Stability*, *Global Emerging Markets* (cit. on p. 121).
- Haldane, Andrew G and Robert M May (2011). „Systemic risk in banking ecosystems.“ In: *Nature* 469.7330, pp. 351–5 (cit. on p. 119).
- Hartmann, P, O De Bandt, P Hartmann, P Molyneux, and J Wilson (2009). „The concept of systemic risk“. In: *FINANCIAL STABILITY REVIEW* December 2006, pp. 134–143 (cit. on p. 70).
- Henriksen, Pal Nicolai (2008). „Lognormal Moment Matching and Pricing of Basket Options“. In: *journal* (cit. on p. 50).
- Hong, L. J. (2009). „Estimating Quantile Sensitivities“. In: *Operations Research* 57.1, pp. 118–130 (cit. on pp. 19, 61).
- Hong, L. J. and G. Liu (2008). „Simulating Sensitivities of Conditional Value at Risk“. In: *Management Science* 55.2, pp. 281–293 (cit. on p. 61).
- Howitt, Peter (2008). „Macroeconomics with intelligent autonomous agents“. In: *Macroeconomics in the Small and the Large: Essays ...* (Cit. on p. 118).
- Hultkrantz, Lars, Jan-Eric Nilsson, and Sara Arvidsson (2012). „Voluntary internalization of speeding externalities with vehicle insurance“. In: *Transportation Research Part A: Policy and Practice* 46.6. “Transportation Economics”, pp. 926 –937 (cit. on p. 103).
- Ichiba, T and JP Fouque (2013). „Stability in a model of inter-bank lending“. In: *preprint*, pp. 1–17 (cit. on p. 81).
- Jondeau, Eric (2010). „Asymmetry in tail dependence of equity portfolios“. In: *NCCR-FINRISK Working Pa-per Series. Swiss: The ...* August, pp. 1–50 (cit. on p. 72).
- Kaminsky, Graciela L. and Carmen M. Reinhart (2001). „Bank lending and contagion : Evidence from the Asian crisis“. In: *East Asia* 153.4, pp. 1–44 (cit. on p. 114).
- King, Nicholas B (2013). „Contagion: a misnomer for financial crisis.“ In: *Journal of public health (Oxford, England)*, pp. 1–2 (cit. on p. 119).
- Lee, S.-H (1998). „Monte Carlo computation of conditional expectation quantiles“. PhD thesis. Stanford University (cit. on p. 18).
- Lee, Shing-Hoi and Peter W. Glynn (2003). „Computing the distribution function of a conditional expectation via monte carlo: Discrete conditioning spaces“. In: *ACM Trans. Model. Comput. Simul.* 13.3, pp. 238–258 (cit. on pp. 18, 31, 33, 38).
- Liu, Ming and Jeremy Staum (2010). „Stochastic kriging for efficient nested“. In: *Spring* 12.3, pp. 3–27 (cit. on p. 19).
- Longin, FM (1996). „The asymptotic distribution of extreme stock market returns“. In: *Journal of business* 69.3, pp. 383–408 (cit. on p. 90).
- Longini, Ira M., Eugene Ackerman, and Lila R. Elveback (1978). „An optimization model for influenza A epidemics“. In: *Mathematical Biosciences* 38.1, pp. 141 –157 (cit. on p. 141).
- Longstaff, F.A. and E.S. Schwartz (2001). „Valuing American options by simulation: A simple least-squares approach“. In: *Review of Financial Studies* 14.1, pp. 113–147 (cit. on p. 19).

- Loretan, M. and P.C.B. Phillips (1992). *Testing the Covariance Stationarity of Heavy-Tailed Time Series: An Overview of the Theory with Applications to Several Financial Datasets*. Working papers 9208. Wisconsin Madison - Social Systems (cit. on p. 90).
- Mamani, H., S. E. Chick, and D. Simchi-Levi (2013). „A Game-Theoretic Model of International Influenza Vaccination Coordination“. In: *Management Science* 59.7, pp. 1650–1670 (cit. on p. 119).
- Masson, Paul R. (1998). *Contagion: Monsoonal Effects, Spillovers, and Jumps Between Multiple Equilibria*. Vol. d, pp. 265 – (cit. on p. 121).
- Merton, RC (1976). „Option pricing when underlying stock returns are discontinuous“. In: *Journal of financial economics* (cit. on p. 84).
- Morrison, Steven (2009). *Nested Simulation for Economic Capital*. Tech. rep. (cit. on p. 26).
- Peckham, Robert (2013). „Contagion: epidemiological models and financial crises.“ In: *Journal of public health (Oxford, England)*, pp. 1–5 (cit. on p. 119).
- Pigou, A.C. (2006). *The Economics of Welfare*. Cosimo classics economics vol. 1. Lightning Source Incorporated (cit. on p. 103).
- Resnick, S and C Stărică (1995). „Consistency of Hill’s estimator for dependent data“. In: *Journal of Applied Probability* (cit. on p. 91).
- (1998). „Tail index estimation for dependent data“. In: *The Annals of Applied Probability* 8.4, pp. 1156–1183 (cit. on p. 91).
- Rogers, L. C. G. and L. a. M. Veraart (2012). „Failure and Rescue in an Interbank Network“. In: *Management Science* 1909, pp. 1–17 (cit. on pp. 73, 77, 107).
- Sun, Peng, Liu Yang, and Francis de Véricourt (2009). „Selfish Drug Allocation for Containing an International Influenza Pandemic at the Onset“. In: *Operations Research* 57.6, pp. 1320–1332 (cit. on p. 119).
- Taleb, Nassim Nicholas (2012). *Antifragile: Things That Gain from Disorder*. Random House of Canada, Limited (cit. on p. 73).
- Taleb, Nassim Nicholas, Elie R.D. Canetti, Tidiane Kinda, Elena Loukoianova, and Christian Schmieder (2012). „A New Heuristic Measure of Fragility and Tail Risks: Application to Stress Testing“. In: *SSRN Electronic Journal* (cit. on p. 73).
- Taleb, NN and R Douady (2012). „Mathematical Definition, Mapping, and Detection of (Anti) Fragility“. In: *arXiv preprint arXiv:1208.1189*, pp. 1–19 (cit. on p. 73).
- Toivanen, Mervi (2013). „Contagion in the interbank network: An epidemiological approach“. In: (cit. on pp. 119, 120).
- Valderrama, Laura, Antonio Moreno, and Antonio Rubia (2012). „Systemic Risk and Asymmetric Responses in the Financial Industry“. In: (cit. on p. 72).
- Wang, Shouqiang, Francis de Véricourt, and Peng Sun (2009). „Decentralized resource allocation to control an epidemic: a game theoretic approach.“ In: *Mathematical biosciences* 222.1, pp. 1–12 (cit. on p. 119).
- Xiong, Wei (2013). „Bubbles, Crises, and Heterogeneous Beliefs“. In: *Handbook on systemic risk*. Ed. by Jean-Pierre Fouque and Joseph A. Langsam. Cambridge: Cambridge University Press, pp. 663–706 (cit. on p. 72).

List of Figures

2.1	Illustration of the Nested Simulation sampling	23
2.2	Illustration of the Stratified Nested Simulation	34
2.3	Illustration of the Stratified Nested Simulation	34
2.4	The value of the upper bound of $P(\mathbb{A})$	40
2.5	Bias and variance of the uniform, the sequential and the marginal estimator	52
2.6	Bias and MSE of the uniform, the sequential and the marginal estimator	53
2.7	Distribution of the number of inner stage sample	53
3.1	The figure illustrates the probability distribution f_x	76
3.2	Network representation of the contagion channels in financial markets	83
3.3	Graphical representation of the links between banks in a direct network for $p = 0.25$ and $p = 0.75$	84
3.4	Trajectories of the coupled diffusion	85
3.5	Q-Q plot for the coupled diffusion introduced in equation (3.5)	86
3.6	Probability distributions of the number of defaulting banks	87
3.7	The expected number of failure as a function of the sparsity parameter	87
3.8	Probability distributions # banks in the system with $p = 0.8$	89
3.9	Tails imbalance in the distribution of gains and losses for a financial institution	92
3.10	Q-Q plot for the returns of <i>JPMorgan</i> and <i>S&P500</i>	96
3.11	λ for the period of 2003-2014 for six major US banks	99
3.12	VoE for the period of 2003-2014 for six major US banks	100
3.13	Fines paid by 4 US banks from 2010-2014	103
3.14	Total fines claimed from the banking industry	103
4.1	Different possible states of a asset unit and the transition possibilities .	130
4.2	Possible evolution of asset units (u_1, u_2, u_3, u_4) in time	131

4.3	Connections Graph of national and international contagion within a risk population	135
4.4	The complete model with all the possible contagion channels	137
4.5	Simulation of the number of Infectious and Recovered/Removed as function of time, $R_{ij} = 0$	139
4.6	Simulation of the number of Infectious and Recovered/Removed as function of time $R_{ij} = 0.01$	140
4.7	Simulation of the number of Infectious and Recovered/Removed as function of time $R_{ij} = 0.02$	140
4.8	Simulation of the number of Infectious and Recovered/Removed as function of time $R_{ij} = 0.04$	141
4.9	The effect of government intervention f on the number of infectious .	144
4.10	Star model when $R_{ij}R_{jk} \approx 0$, $i \neq j$ and $j \neq k$	145
4.11	The optimal Hedging strategy for the source country in the Game problem and the Central planner solution.	152
4.12	The optimal Hedging strategy for a peripheral country in the Game problem and the Central planner solution.	152
4.13	The evolution of the total costs including	153
4.14	The evolution of costs and subsidies as a function of regulation	155

List of Tables

2.1	The results of Bias, Variance and MSE for one dimensional Gaussian, Single put option cases	55
2.2	The results of Bias, Variance and MSE for Multi dimensional Gaussian and Basket option cases	56
3.1	Summary statistics of the data	96
3.2	The top 10 US according to the VoE metric	98
3.3	Regression of fines with <i>MES</i> , <i>SRISK</i> and <i>VoE</i>	104
3.4	Tail imbalance factor computed for a list of global sector indices based on ETFs managed by Ishares BlackRosck	108

