

**Extreme Value Theory with an Application to Bank Failures
through Contagion**

By

Rashid Nikzad

Supervisor

David McDonald

A Thesis
submitted to the Faculty of Graduate and Postdoctoral Studies
in partial fulfilment of the requirements
for the degree of
Master of Science in Mathematics and Statistics

©Rashid Nikzad, Ottawa, Canada, 2011

Abstract

This study attempts to quantify the shocks to a banking network and analyze the transfer of shocks through the network. We consider two sources of shocks: external shocks due to market and macroeconomic factors which impact the entire banking system, and idiosyncratic shocks due to failure of a single bank. The external shocks will be estimated by using two methods: (i) non-parametric simulation of the time series of shocks that occurred to the banking system in the past, and (ii) using the extreme value theory (EVT) to model the tail part of the shocks. The external shocks we considered in this study are due to exchange rate and treasury bill rate volatility. Also, an ARMA/GARCH model is used to extract iid residuals for this purpose. In the next step, the probability of the failure of banks in the system is studied by using Monte Carlo simulation. We calibrate the model such that the network resembles the Canadian banking system.

Acknowledgements

I would like to thank my thesis supervisor David McDonald for his invaluable insight, motivation and support that guided me towards my Master's in Mathematics and Statistics.

I am also grateful to the committee members, Professor Mahmoud Zarepour and Professor Jason Nielsen, for their understanding and comments.

My greatest appreciation also goes to Jesse Collingwood for his invaluable comments during the past two years.

Dedication

This thesis is dedicated to my wife, Golnaz, without her support I could not have made this effort, and to our son, Nick, who brought joy in our life.

Contents

Abstract	i
1- Introduction	1
2- Literature review of failures in the banking system	3
2.1- External shocks and bank failure	3
2.2- Contagion in banking system	6
3- Theory and application of time series techniques on exchange rates and treasury bill rates	10
3.1- An introduction to AR, ARMA, and GARCH modeling	10
3.2- Application of GARCH/ARMA to exchange rates	12
3.3- Bivariate analysis	17
4- Simulation results	23
5- Theory and application of the extreme value theory	29
5.1- Block maxima approach	30
5.2- Exceedances over a threshold	35
5.3- Extensions and special cases	38
5.4- Random variable generation and importance sampling	41

5.5- Application of EVT to exchange rate and treasury bill rate changes	45
6- Monte Carlo simulation using EVT	50
7- Conclusion	55
References	57
Appendix 1- Measures for shocks to the Canadian banking system	60
Appendix 2- List of the Canadian banks	63
Appendix 3- A sample of importance sampling program in R	69

Tables and graphs

Graph 1 - Schematic view of a banking network	1
Graph 2 - Financial structure of a bank	6
Graph 3 – US exchange rate versus Canadian dollars	13
Graph 4 – Histogram, qq-plot, ACF, and PACF of r_t (nlogUSXch.ts)	14
Table 5- GARCH(1,1) with ARMA(1,1) filtering of r_t	15
Graph 6- Residuals (ϵ_t) and standardized residuals (z_t) of the model	16
Graph 7- Properties of the standardized residuals (z_t)	17
Graph 8- Bank of Canada 3-month treasury bill rates (%) and its negative changes (weekly data)	18
Graph 9- Some properties of the treasury bill negative returns	19
Table 10- AR(2) estimation of treasury bill negative returns	20
Graph 11- Properties of the residuals of the AR(2) model of treasury bill negative returns	20
Graph 12- Comparison of negative returns on US exchange rates and treasury bill rates	21
Graph 13- Scatter plot of the residuals of the two models	22

Graph 14- Probability of failure and contagion by reserve value	26
Graph 15- Histogram of failure and contagion by bank	27
Graph 16 – Block maxima (left panel) and exceedances over a threshold (right panel) approaches	31
Graph 17 - Plot of z_p against $\log y_p$ for different ξ	34
Graph 18 - Original distribution function and estimation of the tail part	35
Graph 19 - qq-Plot of z_t against the exponential distribution	46
Graph 20 - Exceedances over a threshold estimation; $u = 2$	47
Graph 21 - Exceedances over a threshold estimation; $u = 1.5$	48
Graph 22 - The shape parameter for $u = 1.5$ when removing the outliers	49
Graph 23 - Exceedances over a threshold estimation of treasury bill negative returns; $u = 0.000317$ (weekly data)	50
Table 24- Simulation results of different λ' for the reserve ratio equal to 5%	53
Table 25- Simulation results for the reserve value equal to 5% and $B = 2.59$ when variance is high	54

Chapter 1

Introduction

This study attempts to quantify the shocks to a banking network and analyze the transfer of shocks through the network. We consider two sources of shocks: external shocks due to market and macroeconomic factors which impact the entire banking system, and idiosyncratic shocks due to failure of a single bank. The external shocks will be estimated by using two methods: (i) non-parametric simulation of the time series of shocks that occurred to the banking system in the past, and (ii) using the extreme value theory (EVT) to model the tail part of the shocks. In the next step, the probability of the failure of banks in the system will be studied.

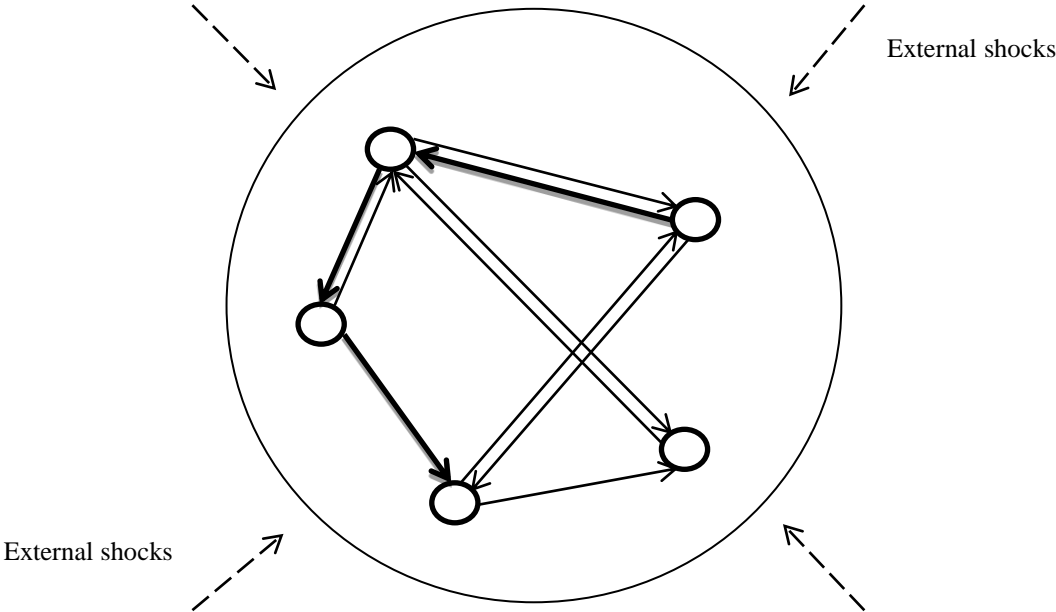
Graph 1 presents a schematic view of a banking network. In this graph, each node represents a bank, and the links between the banks represent interbank loans and their directions. The banking network is subject to external shocks. This banking network could represent a network of domestic banks or a network of international banks. In either case, the network receives shocks that could cause the failure of a bank, and potentially through contagion, the failure of the entire network. In this study, we calibrate the model such that the network resembles the Canadian banking system¹. Then, we analyze the contagion in this context.

The structure of this paper is as follows. Chapter 2 presents a short introduction to the failure of the banking system and related literature. Chapter 3 presents the theory and application of time series modeling on US-Canada exchange rates and Bank of Canada treasury bill rates. The objective of this Chapter is to use time series models to transform the residuals of these

¹ The main objective of this study is to show different methods that can be used to simulate the failure of a banking system. Therefore, more emphasis was given to the procedure and methods, and less to the accuracy of the numbers used to resemble the Canadian banking system. Paying attention to this issue is important when interpreting the results.

series to iid sequences. Chapter 4 presents the results of the Monte Carlo simulation analysis. Chapter 5 presents the theory and application of the extreme value theory. The extreme value theory will be used to model the tail of the series. Chapter 6 presents the simulation results using the extreme value theory. Chapter 7 concludes.

Graph 1- Schematic view of a banking network



Chapter 2

Literature review of failures in the banking system

This Chapter reviews the literature on the banking system failure and the way the probability of failure is estimated. The Chapter consists of two parts. In the first part, we explain the shocks to the banking system and the way failure happens. Second part discusses the process of failure and contagion, and the way their probabilities are estimated.

2.1. External shocks and bank failure

As mentioned earlier, bank failure may happen due to losses from market shocks or due to contagion as a consequence of other banks' failure. We follow the terminology of Elsinger, Lehar, and Summer (2006) to call the former type of failure a fundamental default and the latter a contagious default. Identifying and capturing these two sources of failure are the main modeling challenges. Contagion defaults are not independent of fundamental ones since contagion is more likely to happen in situations where the banking system has already been weakened by external shocks. In what follows, we first identify the external shocks the market may impose to the system, and then, we will show how these shocks may be spread out in the system. Section 2.1 takes a closer look at the mechanism through which contagion happens.

In terms of market shocks, Illing and Liu (2003) have identified highly stressful financial events in the Canadian banking system. They define financial stress as “the force exerted on economic agents by uncertainty and changing expectations of loss in financial markets and institutions”. In their study, financial stress is measured with an index called the Financial Stress Index (FSI). The extreme values of FSI are called financial crises. Illing and Liu (2003) claim

that the FSI better explains financial crises in developed countries than the other measure, early-warning indicators (EWIs), which is usually used to model financial crises in developing countries. According to them, the main four sources of stress in the Canadian banking system are as follows:

- Banking crises/stress
- Foreign exchange crises/stress
- Debt crises/stress
- Equity crises/stress

This study formulates the US-Canada exchange rate shocks and Bank of Canada treasury bill shocks by using time series and EVT techniques. US-Canada exchange rate constitutes 86% of the Canadian exchange rate index. Therefore, it can be a proxy measure for the foreign exchange crises. Moreover, Bank of Canada treasury bill shocks can be a measure for the debt crises (Illing and Liu, 2003).

If the shocks are strong enough, they may cause a bank, and eventually, the entire banking system to fail. Appendix 1 presents some of the measures Illing and Liu (2003) suggest to estimate these four shocks. Chapters 3 to 6 use the US-Canada exchange rates and treasury bill rates to show how non-parametric simulation and extreme value theory could be used to simulate failure in the banking system.

Considerable effort has been made by the academia and central banks to model banking crises and predict financial contagion. The aim of these studies is to better understand the nature of banking crises and mitigate their impacts at national and international levels. Examples of this kind of research include Allen and Gale (2000, 2007), Eisenberg and Noe (2001), Santor (2003), Li (2009), May (2010), and Gai and Kapadia (2010). Moreover, Gauthier, Lehar, and Souissi

(2010), Elsinger, Lehar, and Summer (2006), and Boss et al. (2004) are examples of empirically applying these theories to national banking systems.

Financial contagion means the transmission of financial shocks from one financial entity to other interdependent entities. The transfer of shocks among banks could normally occur through financial linkages. However, banking contagion is possible even when the banks are independent (Santor, 2003). This study assumes contagion happens through interbank linkages, i.e. through interbank loans and borrowings.

Most studies rely on “counterfactual simulations” to estimate the likelihood of contagion arising from a default in repaying interbank loans. One reason is that troubled banks are often bailed out by central banks rather than letting them fail. This limits the use of other methods of study, including event analysis to estimate the probability of failure. The down side of counterfactual simulation is the implication of strong assumptions we have to make to define different scenarios. Upper (2007) has done a survey on these studies, their methodologies, and the results. He concludes that though contagions in banking systems are unlikely, their possibility cannot be fully ruled out. This needs authorities’ attention since the cost of contagion to the society could be very high.

Moreover, Upper (2007) mentions that most studies focus only on the contagion that results from the failure of individual banks (for example due to a fraud). However, this represents only a small fraction of all bank failures. Most failures happen when several banks are hit by an external shock at the same time and become insolvent since their net value becomes negative. This is in contrast to the former models that consider only one source of risk, i.e. interbank linkages, and ignore other sources, e.g. macroeconomic factors. Elsinger, Lehar, and Summer (2006) and Gauthier, Lehar, and Souissi (2010) are among the few studies that integrate

both idiosyncratic (e.g. a fraud) and aggregate shocks (i.e. economy wide shocks) to analyze contagion in the banking system.

This study will follow a similar approach as Gai and Kapadia (2010) and May (2010). However, contrary to these models, we also consider both idiosyncratic and aggregate shocks to estimate the probability of failure. In this sense, our approach will be similar to Elsinger, Lehar, and Summer (2006).

2.2. Contagion in banking system

This Chapter establishes a simple system of a banking network and the way shocks are transferred in the system. This representation is based on Gai and Kapadia (2010) and May (2010). As before, we assume each node in Graph 1 represents a bank and arrows show the direction and magnitude of interbank loans. Graph 2 presents the structure of assets and liabilities of each bank.

Graph 2- Financial structure of a bank

Liabilities	Assets
Deposits d_i	External assets e_i
	Interbank loans l_i
Interbank borrowing b_i	Reserves y_i

Assume interbank borrowing b_i comes from j other banks and interbank loans l_i goes to k other banks. If a bank fails (i.e. it goes bankrupt), it will impact the entire network by being unable to repay its debts. A bank is insolvent if its net value becomes negative, i.e., $e_i + l_i + y_i < d_i + b_i$. We assume each bank keeps a proportion of its assets, y_i , as a reserve to protect itself

against shocks. The external assets e_i could consist of q subclasses of assets with different interest rates and risk degrees. Later, we divide external assets into foreign assets, loans, and other assets to analyze the impacts of different shocks.

To analyze the impact of external shocks, we assume when an external asset is hit by the shock, its value will drop to e_i' . Then, the bank survives if $e_i' + l_i + y_i \geq d_i + b_i$. However, if the bank fails due to the shock to its external assets, a second (internal) shock will be generated through the interbank linkages: since the bank is bankrupt, the creditor banks lose an amount f of their interbank loans. The value of f depends on our recovery assumption (zero recovery or the resale price of the bankrupt bank's assets) and the amount of the creditor bank's loan to the failed bank. In this case, the creditor bank j survives if $e_j' + l_j + y_j - f_j \geq d_j + b_j$. If the creditor bank does not survive, the shock will spread out to the network through the same interbank linkage mechanism to generate a second round of shocks. This process continues until the network gets to a steady-state position, i.e. all banks fail or no more banks fail.

We may make two assumptions when a bank fails. A simplifying assumption is zero recovery, meaning that the value of the insolvent bank will become zero and creditors will lose all of their loans to the bank. The other assumption is that the insolvent bank sells its assets, probably in a lower market price, to repay its creditor. In this study, we follow the second approach and assume the depositors are cleared first, and then, the rest of the asset is distributed proportionally to the creditor banks.

We should note that when a bank sells its assets, prices drop, which causes a depreciation of other banks' assets and net value. This feedback effect increases the probability of the bank's default as its net value decreases. Gauthier, Lehar, and Souissi, (2010) assumed that illiquid assets at each bank can at most lose 2% in value even when banks sell all their holdings. Cyclical

interdependence among the banks is another factor that could increase the probability of contagion. We ignore these two effects in this study.

Matrices are a convenient way to show interbank loans when it comes to simulation of contagion (e.g. Upper, 2007; Boss et al., 2004). If there are N bank in the system that may lend to each other, interbank loans can be represented by an $N \times N$ matrix as follows:

$$X = \begin{bmatrix} 0 & \cdots & x_{1i} & \cdots & x_{1N} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{i1} & \cdots & 0 & \cdots & x_{iN} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{N1} & \cdots & x_{Nj} & \cdots & 0 \end{bmatrix}$$

where x_{ij} represent the loan of bank i to bank j . The sum of rows $l_i = \sum_j x_{ij}$ represents bank i 's total loans to other banks and $b_j = \sum_i x_{ij}$ is bank j 's total liabilities to other banks.

The construction of matrix X is straightforward if the data on interbank loans are available. Otherwise, we need to make some assumptions on banks' balance sheets to fill in this matrix. Banks' balance sheets are usually accessible, though they show only total interbank borrowing and lending. A simple assumption in this case is that interbank lending and borrowing is equally distributed among other banks (Upper, 2007). Another assumption is that interbank lending and borrowing are proportional to other banks' assets (for example, Gauthier, Lehar, and Souissi, 2010). We have used the second approach in constructing interbank loans in this study. Appendix 2 introduces major banks in the Canadian banking system and their estimated balance sheets and interbank loans.

When the matrix of interbank loans has been constructed, we need to specify the shock to the system and its impacts. Furtine (2003) has developed a simple approach to simulate the shocks and the contagion. This approach has been widely used in subsequent studies. His steps are as follows:

- i) Bank i is (randomly) chosen to fail.

- ii) Any bank j that has interbank loans to the bank i may fail. The usual assumption is that bank j fails if its interbank loan x_{ji} multiply by a given parameter for “loss-given-default” (LGD) exceeds its net assets (the amount f as discussed before). Assuming LGD equal to 35% may be a good start (Upper, 2007; Elsinger, Lehar, and Summer, 2006), but other values should also be checked for the robustness of the results.
- iii) A second round of contagion may happen if bank j with interbank liabilities to other banks fails. This process continues until no more banks fail.

Based on these shocks and their impacts, the probability of contagion is estimated. We assume that the portfolio of bank holdings do not change during this process. At the end, the probability of contagion is usually depicted against the value of LGD.

We will follow a similar approach in this study. However, we will extend it in two directions. First, instead of assuming a random/exogenous idiosyncratic shock to the banking system, we define shocks to the banking system as extremes of the stresses to the Canadian banking system suggested by Illing and Liu (2003) either non-parametrically or by applying the extreme value theory to their measures. Then, the impacts of these shocks on each bank’s asset will be calculated. Banks that default in the first round of shocks are fundamentally insolvent. We will show that using the extreme value theory gives us a better estimate of the shocks to the system than the empirical distributions used in Elsinger, Lehar, and Summer (2006) and Gauthier, Lehar, and Souissi (2010). Second, we will combine importance sampling with this process to improve the estimate of the probability of failure even further. Since contagion is a rare event in the system, the usual Monte Carlo approach used in previous studies is not an efficient way to estimate its likelihood in the banking system. Next Chapter introduces time series modeling and its application in financial time series.

Chapter 3

Theory and application of time series techniques on exchange rates and treasury bill rates

This Chapter explains how times series techniques can be applied to exchange rate and treasury bill rate shocks to obtain iid residuals. The objective is to use these residuals for the Monte Carlo simulation in Chapter 4 and EVT modeling in Chapters 5 and 6. The Chapter starts with a short introduction to some of time series techniques we will use to obtain iid residuals. Then, we use these techniques to model the US-Canada exchange rate shocks (univariate case) and Bank of Canada treasury bill shocks (bivariate case).

3.1. An introduction to AR, ARMA, and GARCH modeling

An autoregressive (AR) model is a type of random process which is used to model time series processes. The AR(p) representation of a time series is as follows:

$$y_t = c + \sum_{i=1}^p \varphi_i y_{t-i} + \varepsilon_t, \quad (1)$$

where $\varepsilon_t \sim \text{iid}(0, \sigma^2)$. An autoregressive model can be extended to an autoregressive-moving average (ARMA) model by including the lagged error terms. An ARMA(p, q) representation of a time series is as follows:

$$y_t = c + \sum_{i=1}^p \varphi_i y_{t-i} + \sum_{j=1}^q \psi_j \varepsilon_{t-j} + \varepsilon_t. \quad (2)$$

A property many financial time series possess is that their variance conditional on their past history may change over time. In other words, they may show time varying conditional heteroskedasticity. The “generalized autoregressive conditional heteroskedasticity” (GARCH) modeling is a method to capture this volatile behaviour of financial time series. The serial

correlation in squared returns, or conditional heteroskedasticity, can be modeled using a simple autoregressive process for squared residuals. A formulation of an “autoregressive conditional heteroskedasticity” (ARCH) model is as follows:

$$y_t = c + \varepsilon_t, \quad (3)$$

$$\varepsilon_t = z_t \sigma_t,$$

$$\sigma_t^2 = a_0 + a_1 \varepsilon_{t-1}^2 + \dots + a_p \varepsilon_{t-p}^2,$$

where c is the mean of y_t , and $z_t \sim iid$ with mean zero is the standardized residual of the series.

The conditional variance can be generalized to include past variances as follows:

$$\sigma_t^2 = a_0 + \sum_{i=1}^p a_i \varepsilon_{t-i}^2 + \sum_{j=1}^q b_j \sigma_{t-j}^2, \quad (4)$$

where the coefficients $a_i, i = 0, \dots, p$, and $b_j, j = 1, \dots, q$, are all assumed to be positive to ensure that the conditional variance σ_t^2 is always positive. This formulation is called a GARCH(p, q) model. A GARCH(p, q) model may be combined with an ARMA(r, s) model as follows:

$$y_t = c + \sum_{i=1}^r \varphi_i y_{t-i} + \sum_{j=1}^s \psi_j y_{t-j} + \varepsilon_t, \quad (5)$$

$$\varepsilon_t = z_t \sigma_t,$$

$$\sigma_t^2 = a_0 + \sum_{i=1}^p a_i \varepsilon_{t-i}^2 + \sum_{j=1}^q b_j \sigma_{t-j}^2.$$

The parameters of the model can be estimated using the quasi-maximum likelihood estimator. Let $\boldsymbol{\theta} = (\mathbf{a}', \mathbf{b}')'$ be the vector of the parameters of the model with the initial value $\boldsymbol{\theta}_0 = (\mathbf{a}'_0, \mathbf{b}'_0)'$, where $\mathbf{a} = (c, \varphi_1, \dots, \varphi_r, \psi_1, \dots, \psi_s)'$ and $\mathbf{b} = (a_0, a_1, \dots, a_p, b_1, \dots, b_q)'$. Assume $z_t \sim \text{Normal}(0,1)$. Given that ε_t follows Gaussian distribution conditional on past history, the quasi-likelihood of $\varepsilon_1, \dots, \varepsilon_T$ is,

$$L_T(\boldsymbol{\theta}) = \prod_{t=1}^T \frac{1}{\sqrt{2\pi\sigma_t^2(\boldsymbol{\theta})}} \exp\left(-\frac{\varepsilon_t^2(\boldsymbol{\theta})}{2\sigma_t^2(\boldsymbol{\theta})}\right). \quad (6)$$

Then,

$$\log L_T(\boldsymbol{\theta}) = \sum_{t=1}^T l_T(\boldsymbol{\theta}) = -\frac{T}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^T \log \sigma_t^2(\boldsymbol{\theta}) - \frac{1}{2} \sum_{t=1}^T \frac{\varepsilon_t^2(\boldsymbol{\theta})}{\sigma_t^2(\boldsymbol{\theta})}. \quad (7)$$

It can be shown that under some regularity conditions, $\boldsymbol{\theta}$ is asymptotically normal (Andersen et al., 2009; Francq and Zakoïan, 2010),

$$\sqrt{T}(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}) \xrightarrow{d} N(0, (EZ^4 - 1)J^{-1}), \quad (8)$$

where $J = E \frac{1}{\sigma^4} \frac{\partial^2 \sigma^2}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}$, $z_t \sim \text{Normal}(0,1)$, and $\hat{\boldsymbol{\theta}} = \text{argmax} L_T(\boldsymbol{\theta})$

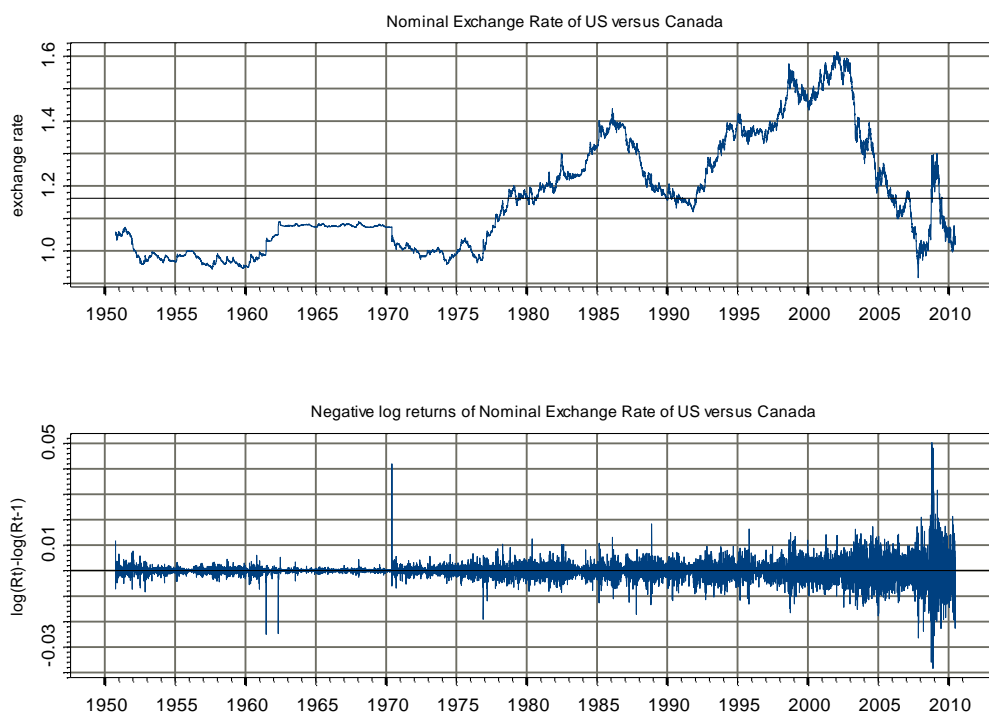
In the next Chapter, we filter the US exchange rate with a GARCH(1,1)/ARMA(1,1) to obtain standardized iid residuals.

3.2. Application of GARCH/ARMA to exchange rates

This Chapter is an analysis of the nominal US exchange rate versus Canadian dollars by using EVT. Since the US dollar constitutes 86% of the value of the “Canadian effective exchange rate”, we concentrate on modeling this variable to analyze the impact of foreign exchange rate shocks on the banking system in this study. The link of a change in exchange rates to the change in bank assets is as follows. Suppose A is the bank’s foreign assets in foreign currency. Let R_t be the exchange rate of the foreign currency at time t . Then, the value of the foreign assets in Canadian dollars will be $S_t = A.R_t$. Consequently, the change in the value of foreign assets due to a change in exchange rate will be $\Delta S_t = A. \Delta R_t$, where $\Delta R_t = R_t - R_{t-1}$. If the value of the foreign exchange rate depreciates, i.e. $\Delta R_t < 0$, the bank will lose part of its value, ΔS_t , which may lead to its bankruptcy. We assume all other variables are constant to simplify the analysis. Since we are interested in the maximum of losses, we model the negative changes in the exchange rate.

The data covers the daily noon spot US exchange rate from November 2, 1950 to June 24, 2010. The data set contains 15031 observations. In what follows, we analyze the data first, and then, use GARCH/ARMA to obtain iid residuals. Let R_t be the nominal daily exchange rate of US versus Canada (denoted by USXch in the graphs). Consistent with the literature, we define the negative returns of exchange rate as $r_t = -\log(R_t / R_{t-1}) = -[\log(R_t) - \log(R_{t-1})]$.

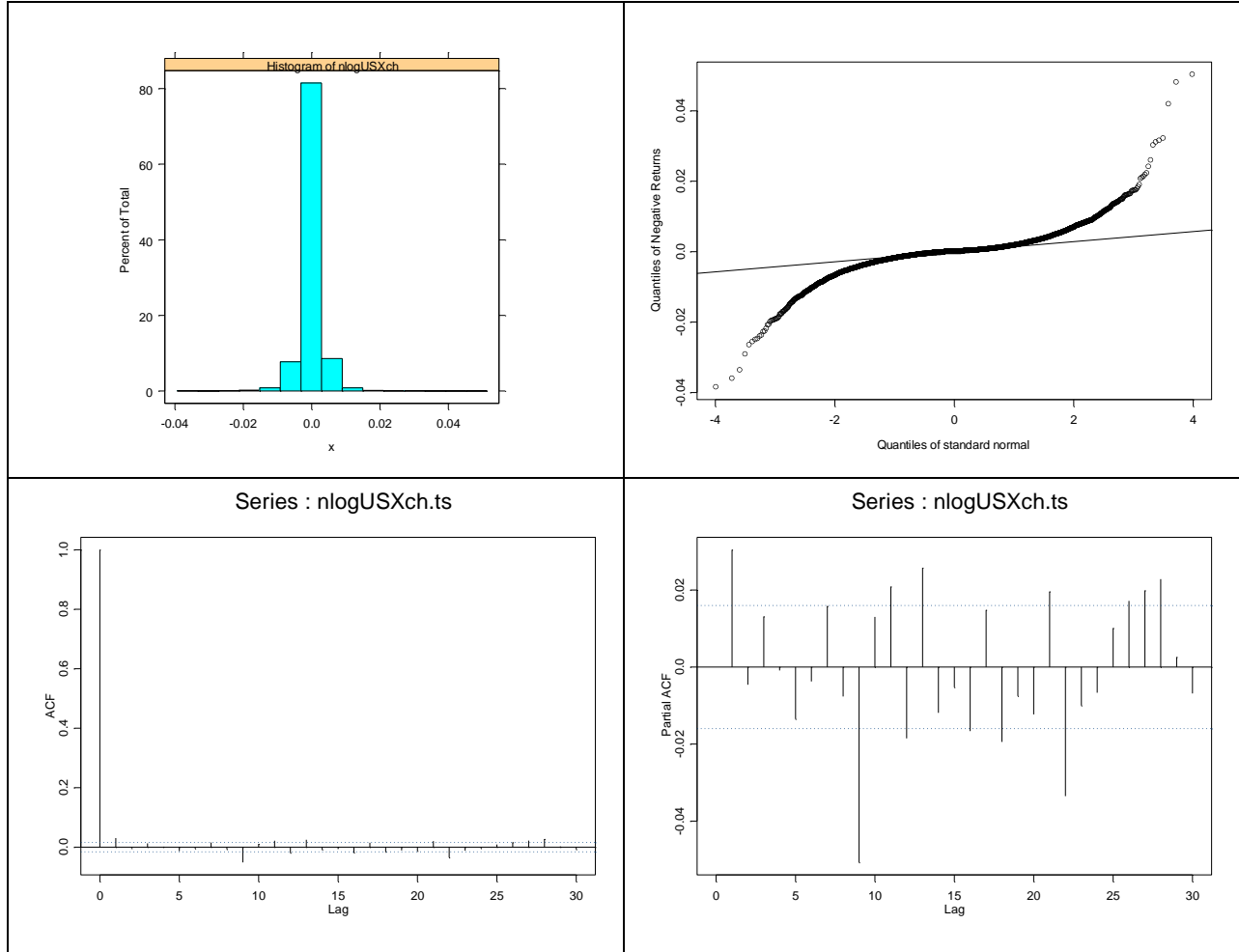
Graph 3 – US exchange rate versus Canadian dollars



Graph 3 presents the plots of the exchange rate R_t and the negative returns r_t . Unit root tests suggest that the exchange rate R_t is not mean stationary, but the negative return r_t is. However, the ARCH test and the autocorrelation function (ACF) and partial autocorrelation functions (PACF) in Graph 4 suggest that there are long- and short-term dependencies in r_t . As will be explained in Chapter 4, the EVT modeling cannot be applied on this series. Graph 4 also presents the histogram and qq-plot of r_t against the normal distribution. The mean, standard

deviation, and skewness of r_t are zero, 0.0032, and 0.303. The kurtosis of r_t is 26, which suggests the negative returns of the exchange rate are heavy-tailed.

Graph 4 – Histogram, qq-plot, ACF, and PACF of r_t (nlogUSXch.ts)



We use a GARCH(1,1) with ARMA(1,1) model to remove the long- and short-term dependencies in r_t . Table 5 presents the estimated results. As shown in Table 5, all coefficients are highly significant. The model can be represented as follows:

$$r_t = c + \alpha r_{t-1} + \beta \varepsilon_{t-1} + \varepsilon_t, \quad (9)$$

where

$$\varepsilon_t = \sigma_t z_t, \quad (10)$$

$$\sigma_t^2 = a + \lambda \varepsilon_{t-1}^2 + \mu \sigma_{t-1}^2, \quad (11)$$

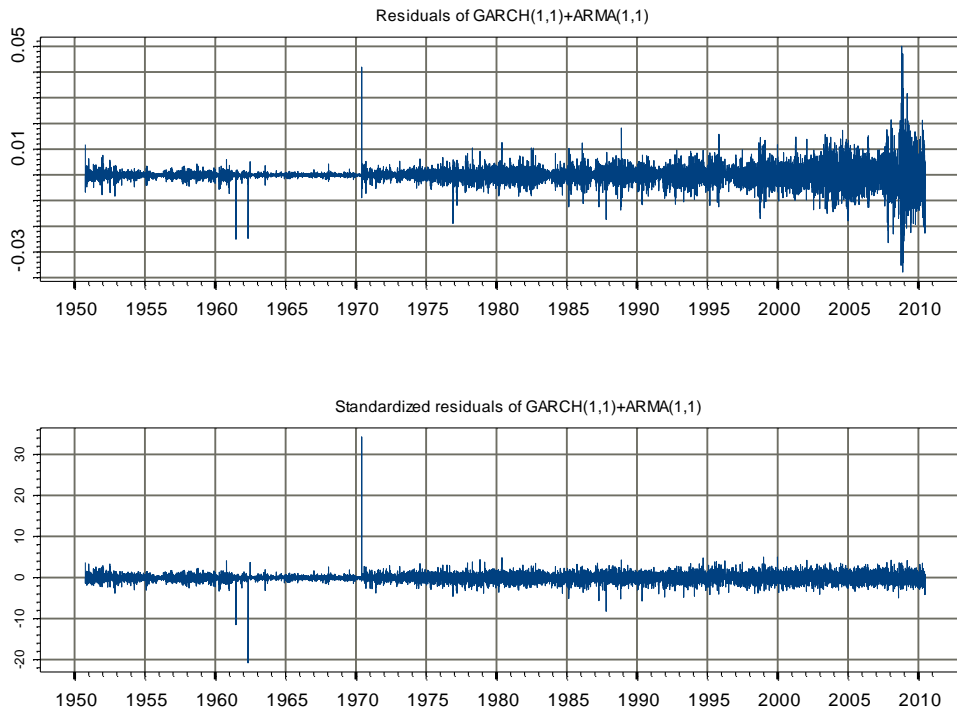
$$z_t \sim iid(0, 1). \quad (12)$$

Table 5- GARCH(1,1) with ARMA(1,1) filtering of r_t

Mean Equation: nlogUSXch.ts ~ arma(1, 1)				
Conditional Variance Equation: ~ garch(1, 1)				
Conditional Distribution: Gaussian				
Estimated Coefficients:				
		Value	Std.Error	t value Pr(> t)
C	9.091e-005	2.067e-005	4.399	5.487e-006
AR(1)	-3.526e-001	1.057e-001	-3.337	4.246e-004
MA(1)	4.205e-001	1.028e-001	4.089	2.178e-005
A	1.936e-007	5.132e-009	37.730	0.000e+000
ARCH(1)	1.207e-001	3.421e-003	35.286	0.000e+000
GARCH(1)	8.592e-001	3.616e-003	237.583	0.000e+000

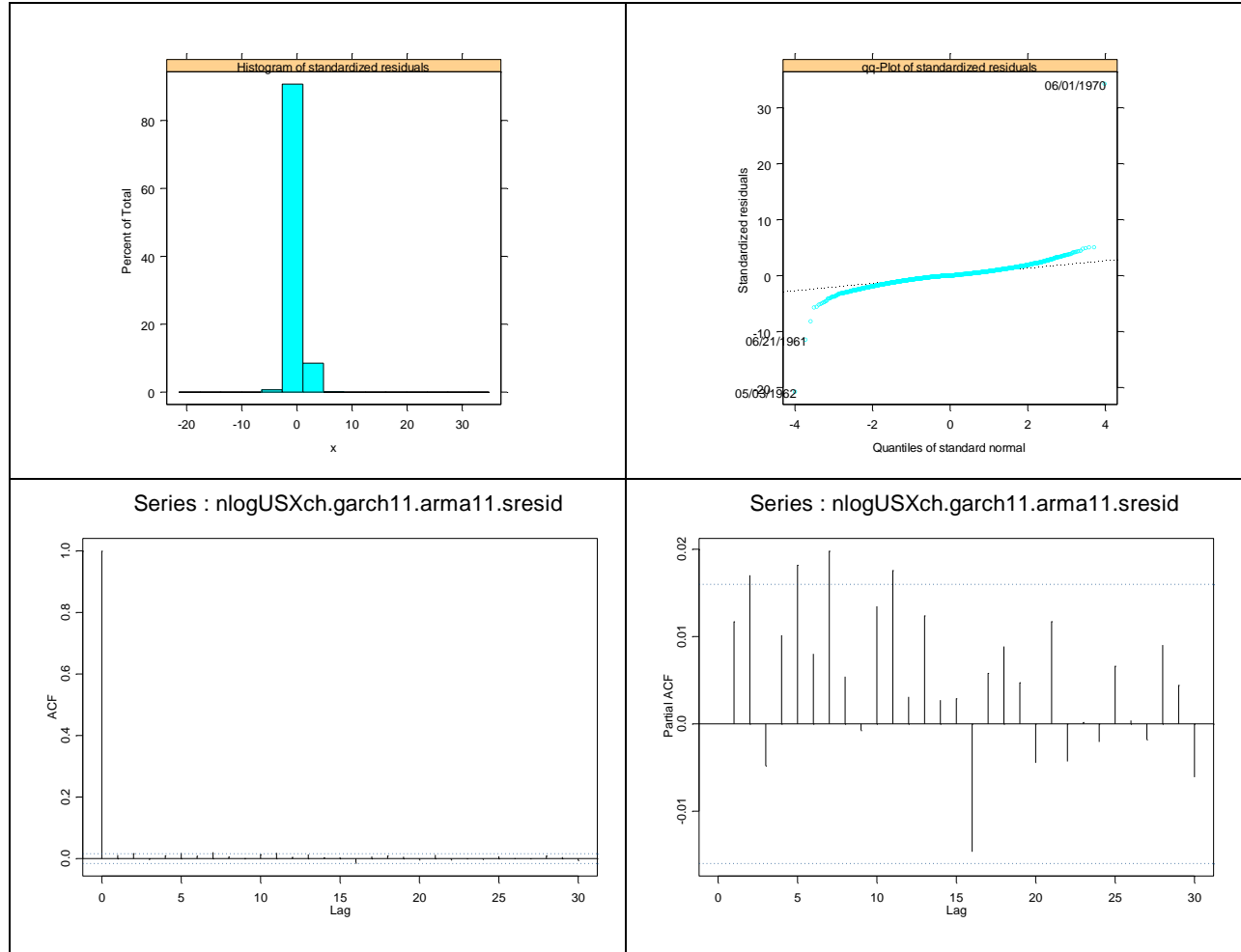
The term “ $c + \alpha r_{t-1} + \beta \varepsilon_{t-1}$ ” is the conditional mean, and “ $\sigma_t^2 = a + \lambda \varepsilon_{t-1}^2 + \mu \sigma_{t-1}^2$ ” is the conditional variance. The estimated parameters are as follows: $c = 0.0001$, $\alpha = -0.3526$, $\beta = 0.4205$, $a = 0.0000002$, $\lambda = 0.1207$, and $\mu = 0.8592$. These estimates are based on the assumption that the residuals are normal. Assuming heavy-tailed distributions such as Student-t and log-normal for the residuals does not change the results much, but since the package suggested the estimations did not converge, we present the results with the normal assumption. We denote the residuals and standardized residuals of the model by ε_t and z_t . Graph 6 compares these two series.

Graph 6- Residuals (ε_t) and standardized residuals (z_t) of the model



Graph 7 presents the histogram, qq-plot against the normal distribution, ACF, and PACF of z_t . As the graphs suggest, there is no autocorrelation left in the data and the moving average effect has been significantly reduced. Also, the ARCH test suggests that no ARCH effect is left in z_t . Therefore, we can conclude that the GARCH(1,1)/ARMA(1,1) model filtered out all long- and short-term dependencies in the data. The mean, standard deviation, skewness, and kurtosis of z_t are -0.035, 0.95, 2.2, and 133. We will use these iid residuals in the following sections for the Monte Carlo and EVT analysis.

Graph 7- Properties of the standardized residuals (z_t)

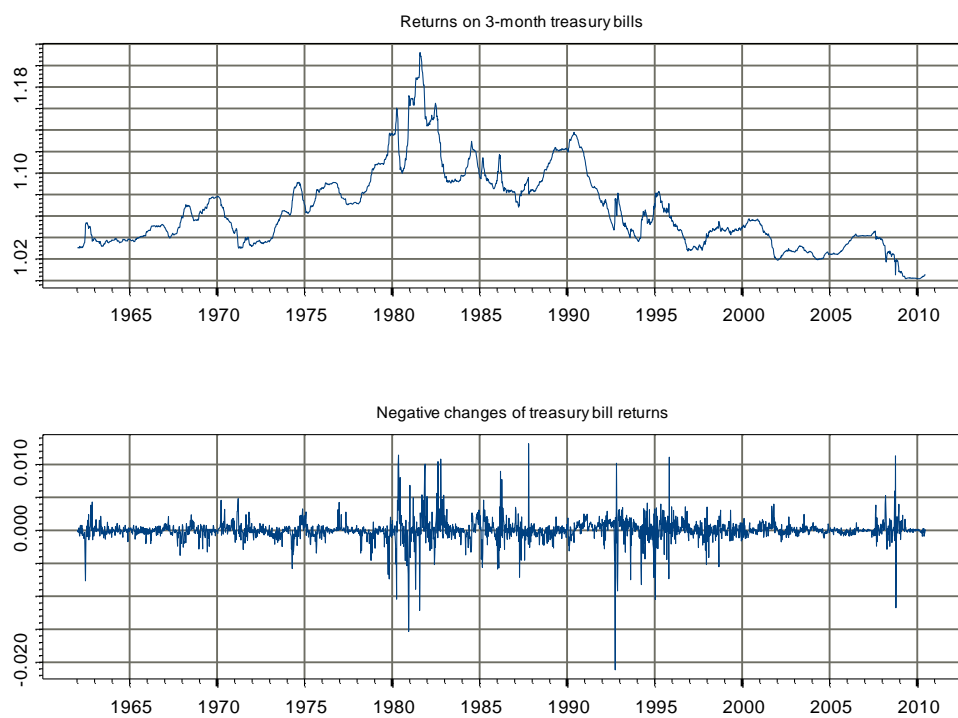


3.3. Bivariate analysis

This Chapter assumes the banking system has to deal with more than one type of shock. This could increase or decrease the probability of failure depending on whether the shocks are positively or negatively correlated. We extend the analysis of the previous section by considering two possible shocks to the system: the exchange rate shock and the debt market shock. These two shocks are among the four important shocks to the Canadian banking system identified by Illing and Liu (2003).

A measure for the shocks to the debt market is the shocks due to the changes in the Bank of Canada treasury bill rates. Let TR_t be the Bank of Canada 3-month treasury bill rate at noon on day t . We define debt market shocks to the banking system as extreme drops in treasury bill rates, i.e. extremes of $tr_t = -(TR_t - TR_{t-1})/TR_{t-1}$.

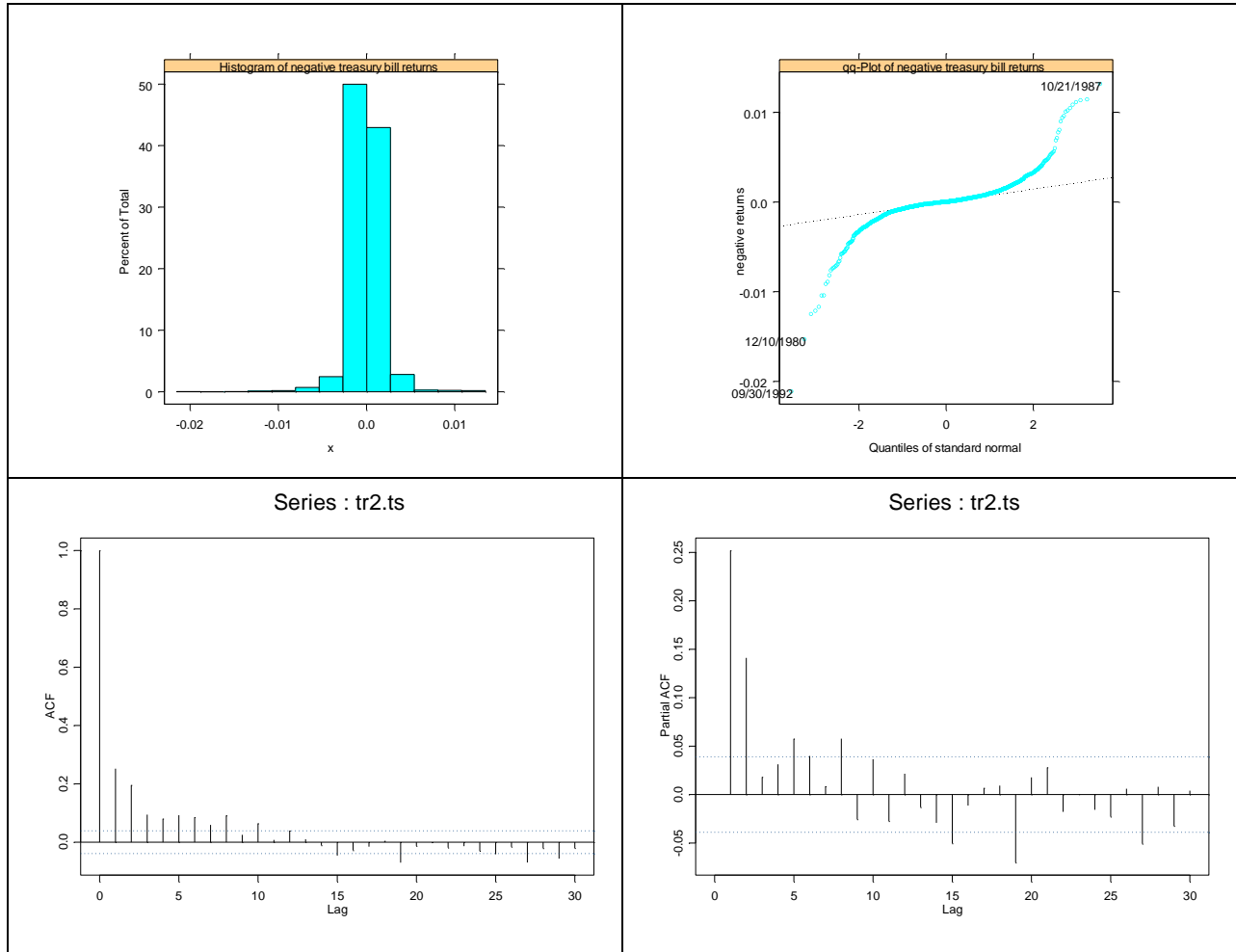
**Graph 8- Bank of Canada 3-month treasury bill rates (%) and its negative changes
(weekly data)**



Since Bank of Canada treasury bill rates are available only on a weekly basis, we have to limit ourselves to a weekly analysis in this Chapter. To be consistent with the treasury bill data, we pick the exchange rates of the same day the treasury bill rates have been collected (i.e. Wednesdays of each week). The weekly data covers 2350 observations from January 3, 1962 to

June 23, 2010. Graph 8 presents the Bank of Canada 3-month treasury bill rates (in percent) and its weekly negative changes.

Graph 9- Some properties of the treasury bill negative returns

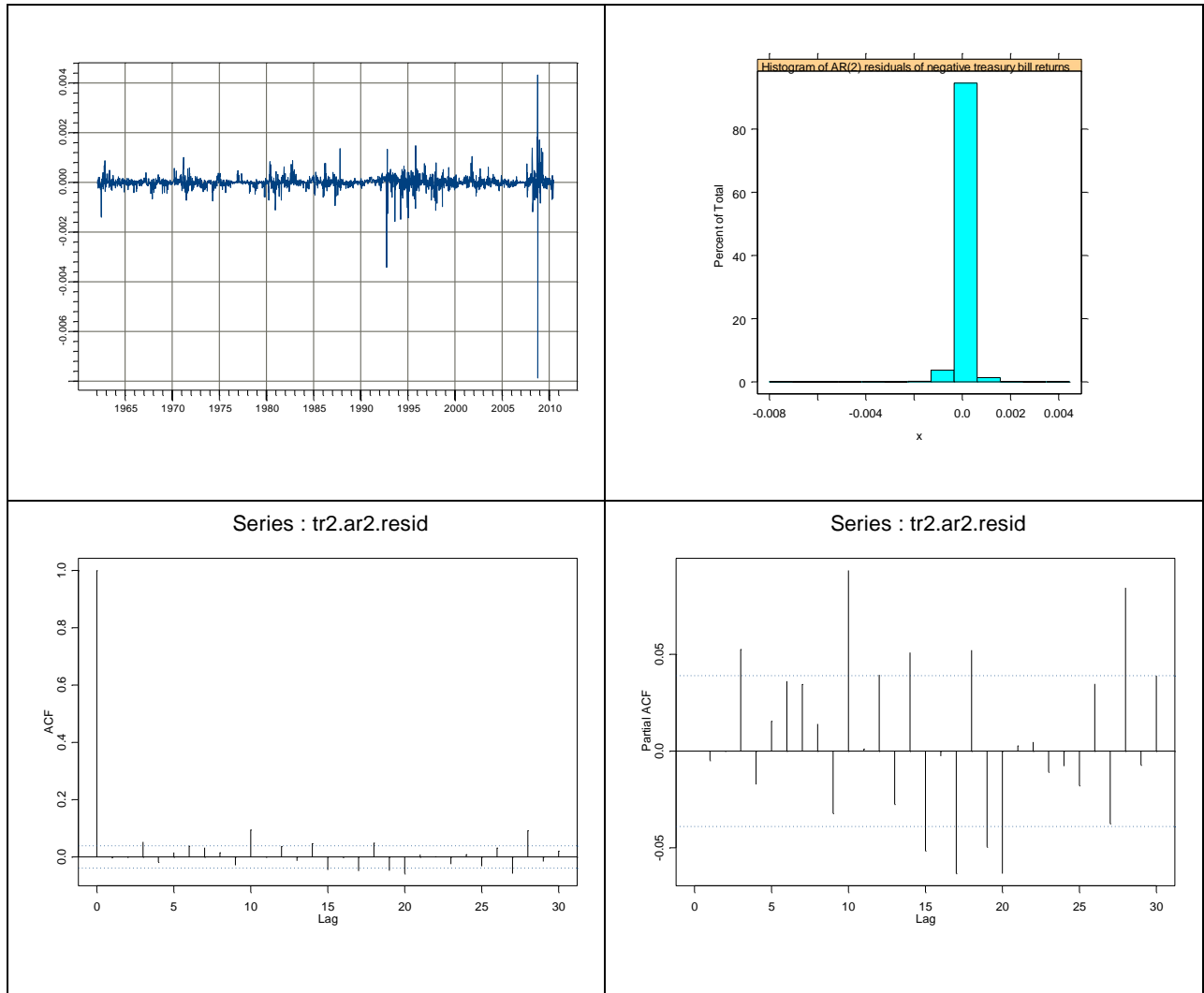


Graph 9 presents the histogram, qq-plot against the standard normal, and ACF and PACF of the treasury bill negative returns. The negative returns are fat-tailed with autocorrelation and moving average components. The mean, standard deviation, skewness, and kurtosis of this series are zero, 0.001704, -0.9199, and 27.82.

Table 10- AR(2) estimation of treasury bill negative returns

Estimated Coefficients:					
	Value	Std. Error	t value	Pr(> t)	
(Intercept)	0.0000	0.0000	0.1644	0.8695	
lag1	0.2164	0.0197	10.9807	0.0000	
lag2	0.1411	0.0197	7.1620	0.0000	

Graph 11- Properties of the residuals of the AR(2) model of treasury bill negative returns

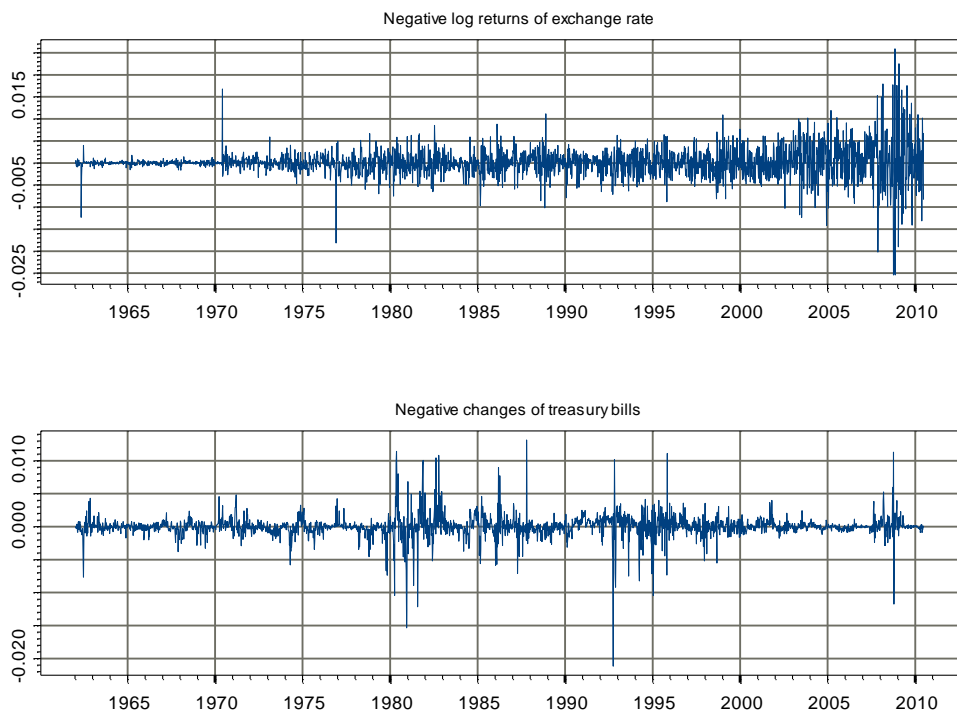


Testing various models suggests that using an AR(2) model removes the autocorrelation of the data and most of the moving average component. Table 10 presents the estimation results

of the AR(2) model for treasury bill negative returns. In other words, $tr_t = c + \alpha tr_{t-1} + \beta tr_{t-2} + \varepsilon_t$, where $c = 0$, $\alpha = 0.2164$, and $\beta = 0.1411$. Graph 11 presents the properties of the residuals of the AR(2) model. The ACF of the residuals of this model suggests that the AR(2) model significantly filters out the autocorrelation in the data.

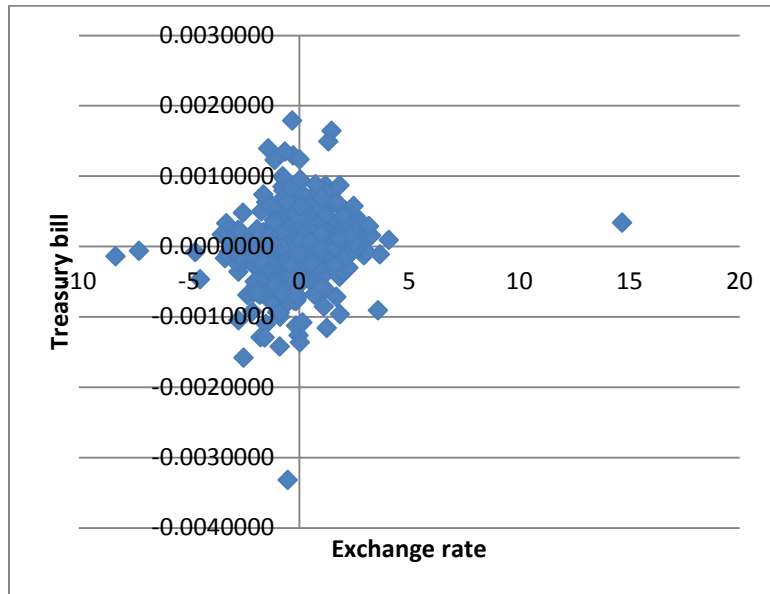
Since we are interested in the bivariate analysis of exchange rate negative returns and treasury bill negative returns, we need to examine the correlation between the two series. Graph 12 plots the two series together. Graph 13 presents the scatter plot between the residuals of the two series. The correlation between the residuals of the two series (and also between the negative changes of the two series) is around 15%. The low correlation and the shape of the scatter plot suggest that we may assume the residuals of the two series are independent.

Graph 12- Comparison of negative returns on US exchange rates and treasury bill rates



In the next chapters, we generate random series for the exchange rate and treasury bill negative returns based on these estimations to analyze their impacts on the banking system. We also show how EVT can be used in the simulation.

Graph 13- Scatter plot of the residuals of the two models



Chapter 4

Simulation results

This Chapter presents the results of the simulation of the banking network. We consider both of the univariate and bivariate cases with exchange rate and treasury bill shocks. The nonparametric simulation is based on reconstructing the exchange rate time series and treasury bill rate time series by resampling from the iid residuals we obtained in Chapters 3, and applying the reconstructed series to the banking network to find out if any failure or contagion occurs in the network. By repeating this process n times and counting the number of failures and contagions in the system, we find the probability of failure and contagion. We assume the reserve value of banks varies from 1% to 10% of their total assets¹. For the bivariate case, we assume the weekly data we used in section 3.3 to model the treasury bill rates has the same distribution as (non-available) daily time series. By using this assumption, we could also simulate daily treasury bill rates.

This process can be explained more precisely as follows. Consider the GARCH(1,1)/ARMA(1,1) model we estimated for the negative returns on US-Canada exchange rates in Section 3.2:

$$\hat{r}_t = c + \alpha \hat{r}_{t-1} + \beta \hat{\varepsilon}_{t-1} + \hat{\varepsilon}_t, \quad (14)$$

where

$$\hat{\varepsilon}_t = \hat{\sigma}_t \hat{z}_t, \quad (15)$$

¹ As mentioned earlier, we are interested in showing the procedure of estimating the probability of failure and contagion, not the actual estimates. Therefore, the probabilities obtained in this section may not be true in reality. Estimating the correct probabilities needs more rigorous estimates of bank assets and considering all shocks to the banking system.

$$\hat{\sigma}_t^2 = a + \lambda \hat{\varepsilon}_{t-1}^2 + \mu \hat{\sigma}_{t-1}^2, \quad (16)$$

$$\hat{z}_t \sim iid(0, 1) \quad (17)$$

with $c = 0.0001$, $\alpha = -0.3526$, $\beta = 0.4205$, $a = 0.0000002$, $\lambda = 0.1207$, and $\mu = 0.8592$. We start from equation (17) and move backward to equation (14) to reconstruct the negative returns on exchange rate.

Suppose we are at time t . First, we sample with replacement from the iid standard residuals we obtained in Section 3.2 and will call it \hat{z}_t . At the same time, we construct $\hat{\sigma}_t^2$ using $\hat{\varepsilon}_{t-1}$ and $\hat{\sigma}_{t-1}$ (equation 16). Then, we rebuild the residual $\hat{\varepsilon}_t$ by using Equation 15. Finally, the negative returns on exchange rates will be constructed by Equation 14. We set the initial values of ε and σ ($\hat{\varepsilon}_0$ and $\hat{\sigma}_0$) as the means of these two variables in the original series evaluated in Section 3.2.

After constructing \hat{r}_t , we calculate the level change in the exchange rate $\Delta \hat{R}_t = \hat{R}_t - \hat{R}_{t-1}$, where $\hat{R}_t = \hat{R}_{t-1} \exp(-\hat{r}_t)$. \hat{R}_0 is set to be the mean of R_t in the original series. At this point, we assess whether this change makes a failure and contagion in the banking system: if $-A_i \Delta \hat{R}_t$ is less than the reserve value bank i keeps to cover the shocks, a failure happens. A_i is the value of foreign asset at bank i . If at least one of the banks fails, we test for the possibility of contagion: first, we update the assets of all banks. This includes two steps. In the first step, we calculate the remaining assets of the failed bank(s). After paying off the personal, business, and government deposits, the remaining assets will be distributed among other banks according to their interbank loan ratios. In the next step, the net asset of all non-failed banks will be calculated with respect to the current exchange rate and interbank losses. If the net asset of one of the banks becomes negative at this time, contagion happens. Since the exchange rate has dropped and the value of

banks' assets is lower, contagion due to default in repaying interbank loans is more probable now.

If no failure happens at time t , we assume the assets will return to the same level as before and the same process will be followed at time $t+1$. However, if a failure happens, we record it and will start a new series. This way, we can bootstrap new replications of counterfactual time series of exchange returns and assess whether there would a failure or not. The banking system structure and estimated assets have been explained in Section 2.2 and Appendix 2.

We do the same for the bivariate case. In Section 3.3, we estimated an AR(2) model for the negative returns on the Bank of Canada treasury bill negative returns as follows,

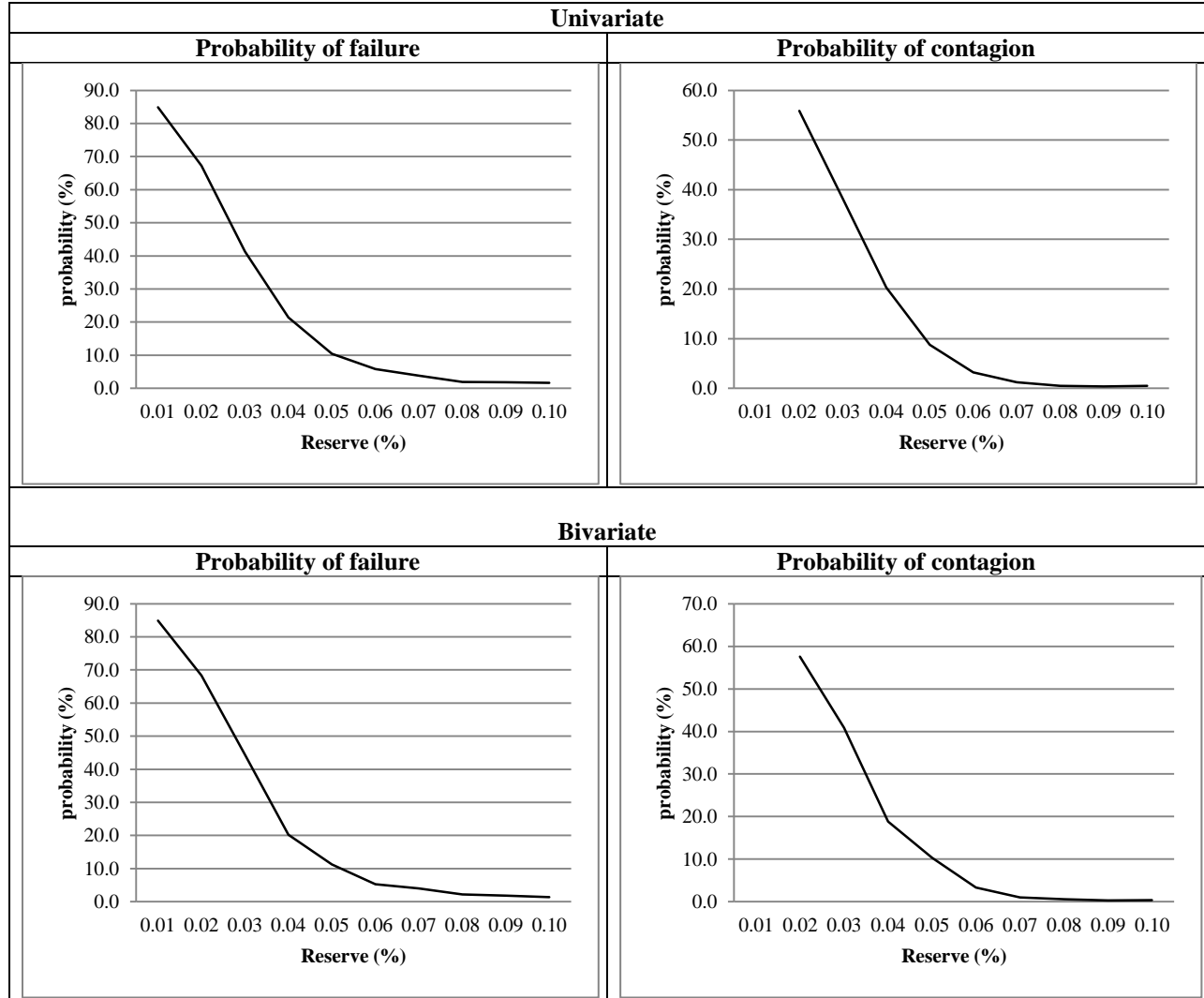
$$\hat{tr}_t = c + \alpha \hat{tr}_{t-1} + \beta \hat{tr}_{t-2} + \hat{\varepsilon}_t, \quad (18)$$

where $c = 0$, $\alpha = 0.2164$, and $\beta = 0.1411$. We set the initial values of \hat{tr}_0 and \hat{tr}_1 to the mean of tr . The change in the treasury bill rate \hat{tr}_t is reconstructed by Equation 18 where $\hat{\varepsilon}_t$ is sampled with replacement from the iid residuals of the AR(2) model estimated in Section 3.3. Then, the change in the treasury bill rate is calculated by $\Delta \widehat{TR}_t = \widehat{TR}_t - \widehat{TR}_{t-1}$, where $\widehat{TR}_t = \widehat{TR}_{t-1}(1 - \hat{tr}_t)$. Since we showed in Section 3.3 that the changes in exchange rates and in treasury bill rates are almost independent, we can reconstruct these two series and assess their impacts on banks' assets at time t independently. The procedure to understand if a failure and contagion happens in the bivariate case is similar to the univariate case, with the assumption that the treasury bill shocks impact the loans part of banks' assets.

Graph 14 presents the probability of failure and contagion for $n = 1000$ replications. As expected, the probability of failure and contagion decreases as banks increase their reserve ratio.

When the reserve rate is as low as 1% of bank assets, we see failures in the network over 80% of time. The probability of failure and contagion reduces significantly as banks keep 10% of their assets in reserve.

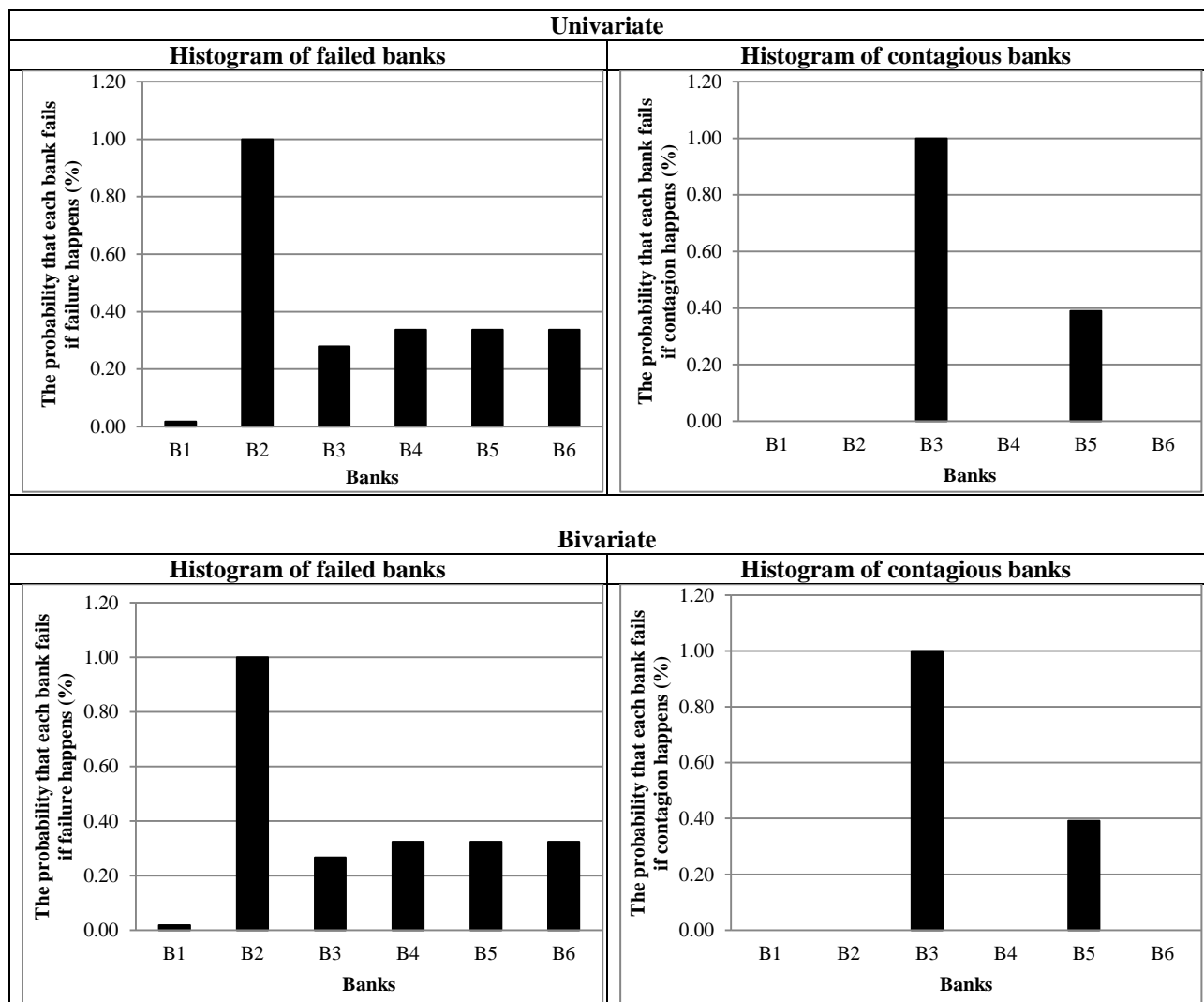
Graph 14- Probability of failure and contagion by reserve value



The bivariate results have been obtained under assumption that exchange rate and treasury bill shocks are independent. Since most of the variation comes from the exchange rate shocks in this model, we do not see a big difference between the univariate and bivariate simulation results in Graph 14. Graph 15 presents the distribution of the failed and contagious

banks. B2 is always the first bank that fails. The reason is that B2 has the largest ratio of international assets to total assets. In fact, there is a negative relationship between a bank's ratio of foreign assets to total assets and the probability of failure of the bank. On the other hand, B3 fails due to contagion whenever there is contagion in the system. B3 has the highest ratio of interbank loans to total assets.

Graph 15- Histogram of failure and contagion by bank



One disadvantage of Monte Carlo simulation in estimating failure in the banking system is that failure and contagion are rare events by nature (reserve levels are kept high enough to

cover the shocks). Therefore, estimating the probability of failure will be very computer intensive when the reserve ratio is high. In the next Chapter, we introduce the extreme value theory which could improve Monte Carlo simulation in this context.

Chapter 5

Theory and application of the extreme value theory

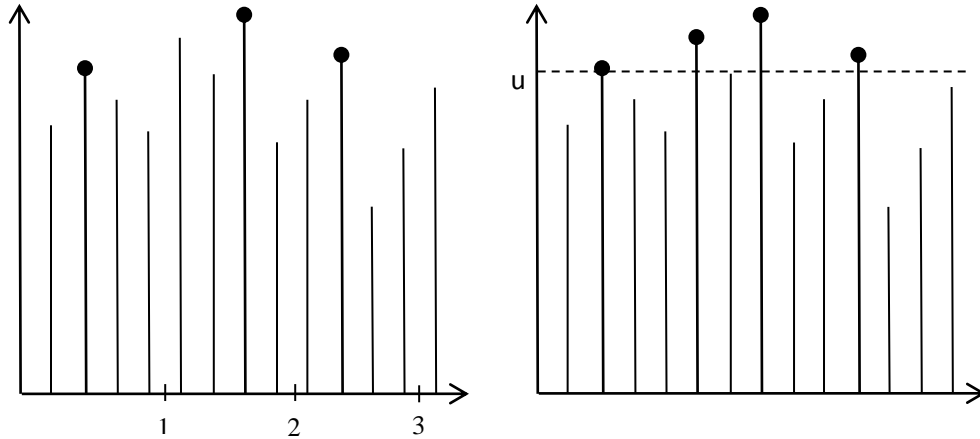
Extreme value theory (EVT) is a framework to analyze the tail behaviour of a distribution. The majority of parametric methods use a normal distribution approximation to model time series data. One of the drawbacks of using a normal approximation to model the extreme values of financial data is that the probability of high quantiles are underestimated since financial data are usually fat-tailed. Using a fat-tailed distribution such as Student-t or log-normal improves the approximation, but may not still fully capture the tail behaviour of the series. On the other hand, nonparametric methods to model extreme values show another disadvantage by being unable to estimate out-of-sample quantiles. EVT attempts to overcome these problems by parameterizing the tail part of the data series. EVT is analogous to the central limit theorem (CLT) in this sense but for the extremes of a distribution.

This Chapter is an introduction to EVT based on Embrechts, Kluppelberg, and Mikosch (1997), Coles (2001), and Smith (2002). Schafgans, Koedijk, and de Vries (1990), Hols and de Vries (1991), Danieelsson and de Vries (1997), Gilli and Kellezi (2006), and Gencay and Selcuk (2006) are examples of studies that used EVT to model the extreme values of exchange rates and financial data. This Chapter also briefly explains random number generation and importance sampling as these techniques will be used in the analysis of the next Chapter.

There are two approaches to model extreme values of a data series. In the first approach, the data is partitioned into successive blocks, where the block maxima represent extreme values. Theorem 1 in Section 5.1 shows the limiting distribution of these maxima. In the second approach, the exceedances over a selected threshold are considered to be extreme values. Theorem

2 in Section 5.2 models the limiting distribution of these exceedances. These two approaches are presented in Graph 16. The second approach is more efficient in modeling exceedances in time series and financial data.

Graph 16 – Block maxima (left panel) and exceedances over a threshold (right panel) approaches



5.1. Block maxima approach

We start with the theories related to the first approach.

Theorem 1 (Fisher and Tippett, 1928; Gnedenko, 1943) - Let $M_n = \text{Max} \{X_1, \dots, X_n\}$, where X_1, \dots, X_n are iid from an unknown distribution F . If there exist sequences of constants $\{a_n > 0\}$ and $\{b_n\}$ such that $p\{(M_n - b_n)/a_n \leq z\} \rightarrow G(z)$ as $n \rightarrow \infty$ for a non-degenerate distribution function G , then G is a member of the “generalized extreme value” (GEV) family,

$$G(z) = \exp \left\{ - \left[1 + \xi \left(\frac{z - \mu}{\sigma} \right) \right]^{\frac{1}{\xi}} \right\}, \quad (19)$$

defined on $\{z: 1 + \xi(z - \mu)/\sigma > 0\}$, where $-\infty < \mu < \infty$, $\sigma > 0$, and $-\infty < \xi < \infty$.

In the above equation, μ and σ are the location and scale parameters, and ξ is called the shape parameter. G belongs to one of the three standard extreme value distributions of Frechet, Weibull, or Gumbel as $\xi > 0$, $\xi < 0$, or $\xi = 0$ after some change of parameters:

$$\text{Frechet: } G(x) = \begin{cases} 0, & x \leq 0, \\ \exp(-x^{-\alpha}), & x > 0, \alpha > 0; \end{cases} \quad (20)$$

$$\text{Weibull: } G(x) = \begin{cases} \exp(-(-x)^\alpha), & x < 0, \alpha > 0, \\ 1, & x \geq 0; \end{cases} \quad (21)$$

$$\text{Gumbell: } G(x) = \exp(-\exp(-x)), \quad -\infty < x < \infty; \quad (22)$$

We should note that Theorem 1 states that if the (normalized) sequence of maxima has a limit, that limit will be a member of the GEV family. However, it does not guarantee the existence of a limit.

The Frechet distribution is heavy-tailed and the one that is usually used in financial modeling as exchange rates and other financial data have usually shown to be fat-tailed. Roughly speaking, the tail distribution function of a light-tailed random variable decays at an exponential rate or faster, while it decays at a slower rate for heavy-tailed random variables. Rare events occur differently in light-tailed and heavy-tailed random variables. It can be shown that the most likely way for the sum of heavy-tailed random variables to become large is by one of the random variables to become large. In contrast in the light-tailed case, all of the random variables in the sum contribute to the sum becoming large¹. That means a large deviation in a light-tailed setting happens most likely due to small occurrences in a specific path instead of having a big event. Studies show that bank failures usually happen due to a single bad event and not a series of consecutive smaller negative returns that may add up to the same highly negative result². The reason is that during a gradual decline of the market or negative returns, banks and other

¹ Juneja and Shahabuddin (2006)

² e.g. in Danfelsson and de Vries (1997)

financial institutes can react and adjust themselves rather than letting the losses accumulate until a failure happens.

We present a sketch of proof for Theorem 1 as follows. Let X_1, \dots, X_n be iid from an unknown distribution F^1 . Let $M_n = \max \{X_1, \dots, X_n\}$. Then,

$$p(M_n \leq x) = p\{X_1 \leq x, \dots, X_n \leq x\} = p\{X_1 \leq x\} * \dots * p\{X_n \leq x\} = (F(x))^n. \quad (23)$$

Since $F(x)^n \rightarrow 0$ if $F(x) < 1$ and $F(x)^n \rightarrow 1$ if $F(x) = 1$ as $n \rightarrow \infty$, i.e. it is a degenerate distribution function, we find the limit of $(M_n - b_n)/a_n$ with appropriate choices of the scale and location constant sequences $\{a_n > 0\}$ and $\{b_n\}$ to avoid this problem.

Definition - The distributions F and F^* are of the “same type” if there exist constants $a > 0$ and b such that $F^*(ax + b) = F(x)$ for all x .

Definition- A distribution is said to be max-stable if

$$G^k(a_k x + b_k) = G(x), k = 1, 2, \dots, \quad (24)$$

for some constants a_k and b_k .

Theorem- A distribution is max stable if and only if it is a GEV distribution.

Now, suppose $(M_n - b_n)/a_n$ has the limit distribution G . That means for large n ,

$$p\{(M_n - b_n)/a_n \leq x\} \approx G(x). \quad (25)$$

Therefore for any k ,

$$p\{(M_{nk} - b_{nk})/a_{nk} \leq x\} \approx G(x). \quad (26)$$

On the other hand,

$$p\{(M_{nk} - b_{nk})/a_{nk} \leq x\} = [p\{(M_k - b_{nk})/a_{nk} \leq x\}]^k. \quad (27)$$

This gives two expressions for the distribution of M_k as follows:

$$p(M_n \leq x) \approx G\left(\frac{x-b_n}{a_n}\right) \text{ and } p(M_n \leq x) \approx G^{1/k}\left(\frac{x-b_{nk}}{a_{nk}}\right). \quad (28)$$

¹ Formal proof of Theorem 1 and the following theorems can be found in Leadbetter et al. (1983) and Resnick (2007). This sketch of proof is from Coles (2001).

That means G and $G^{1/k}$ are asymptotically equal apart from scaling coefficients. Therefore, G is approximately max-stable and GEV.

In practice, we assume that for some $a > 0$ and b , $p\{(M_n - b)/a \leq z\} \approx G(z)$, or equivalently, $p\{M_n \leq z\} \approx G\{(z - b)/a\} = G^*(z)$, where G^* is of the same type as G . That means the GEV family distribution may be fitted directly to a series of observations M_n .

To estimate the parameters of G , for example by maximum likelihood estimation, we need to fit the observations $x_i, i = 1, \dots, N$ into m blocks of length k , for some value of large k , and then find the series of block maxima $M_{k,1}, \dots, M_{k,m}$. Then, the block maxima $z_i = M_{k,i}$'s can be used to estimate the parameters of G . The log-likelihood function will be as follows:

$$l(\mu, \sigma, \xi) = \sum_{i=1}^m \left\{ -\log \sigma - \left(1 + \frac{1}{\xi}\right) \log \left[1 + \xi \left(\frac{z_i - \mu}{\sigma}\right)\right] - \left[1 + \xi \left(\frac{z_i - \mu}{\sigma}\right)\right]^{-1/\xi} \right\}, \quad (29)$$

where $1 + \xi \left(\frac{z_i - \mu}{\sigma}\right) > 0, i = 1, \dots, m$, and $\xi \neq 0$, and,

$$l(\mu, \sigma) = \sum_{i=1}^m \left\{ -\log \sigma - \left(\frac{z_i - \mu}{\sigma}\right) - \exp \left[-\left(\frac{z_i - \mu}{\sigma}\right) \right] \right\}, \quad (30)$$

for $\xi = 0$. Assuming $\theta = (\mu, \sigma, \xi)$, the score and information equations will be as follows:

$$S(\theta) = \frac{\partial l(\theta)}{\partial \theta} = 0, \quad (31)$$

$$J(\theta) = E \left(-\frac{\partial^2 l(\theta)}{\partial \theta \partial \theta^T} \right). \quad (32)$$

The likelihood ratio statistics is $W(\theta) = 2\{l(\hat{\theta}) - l(\theta)\}$, where $\hat{\theta}$ is the maximum likelihood estimator satisfying $l(\hat{\theta}) \geq l(\theta)$ for all θ . Assuming the data were generated from $f(x, \theta^\circ)$, then $\hat{\theta} \sim N_3 \left(\theta^\circ, J(\hat{\theta})^{-1} \right)$ and $W(\theta^\circ) \sim \chi_3^2$ could be used to obtain the confidence intervals and for hypothesis testing. The choice of block size is critical in this method as there is a trade-off between the bias and variance. If the blocks are too small, the approximation will

probably be poor and biased. On the other hand, large blocks generate few block maxima which lead to a large variance.

The blocks may represent time intervals, e.g. a year. In this case, we may find extreme quantiles of the annual maximum distribution by inverting the G :

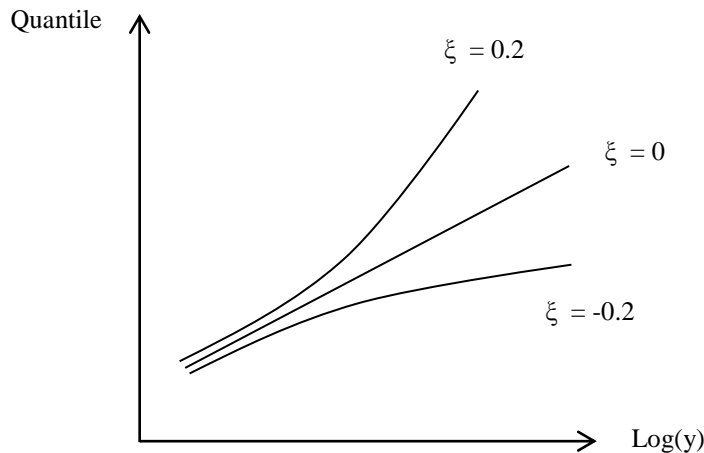
$$z_p = \begin{cases} \mu - \frac{\sigma}{\xi} [1 - \{-\log(1-p)\}^{-\xi}], & \xi \neq 0 \\ \mu - \sigma \log\{-\log(1-p)\}, & \xi = 0 \end{cases} \quad (33)$$

where $G(z_p) = 1-p$. The interpretation is that z_p will be exceeded by the annual maximum in any particular year with probability p . Alternatively, we may say that the level z_p is expected to be exceeded on average once every $1/p$ years. In extreme value terminology, z_p is sometimes called the “return level” associated with the “return period” $1/p$.

Let $y_p = -\log(1-p)$ in the above equation. Then,

$$z_p = \begin{cases} \mu - \frac{\sigma}{\xi} [1 - y_p^{-\xi}], & \xi \neq 0 \\ \mu - \sigma \log y_p, & \xi = 0 \end{cases} \quad (34)$$

Graph 17 – Plot of z_p against $\log y_p$ for different ξ



That means the plot of z_p against $\log y_p$ will be linear if $\xi = 0$. It will be a convex function with the asymptotic limit of $(\mu-\sigma)/\xi$ as $p \rightarrow 0$ if $\xi < 0$. The plot will be concave with no finite bound if $\xi > 0$. Graph 17 presents the plots for $\xi = -0.2, 0$, and 0.2 for certain μ and σ .

According to Smith (1985), the limiting behaviour of the MLE depends on the value of the shape parameter ζ , but the maximum likelihood estimators are consistent as long as $\zeta > -0.5$.

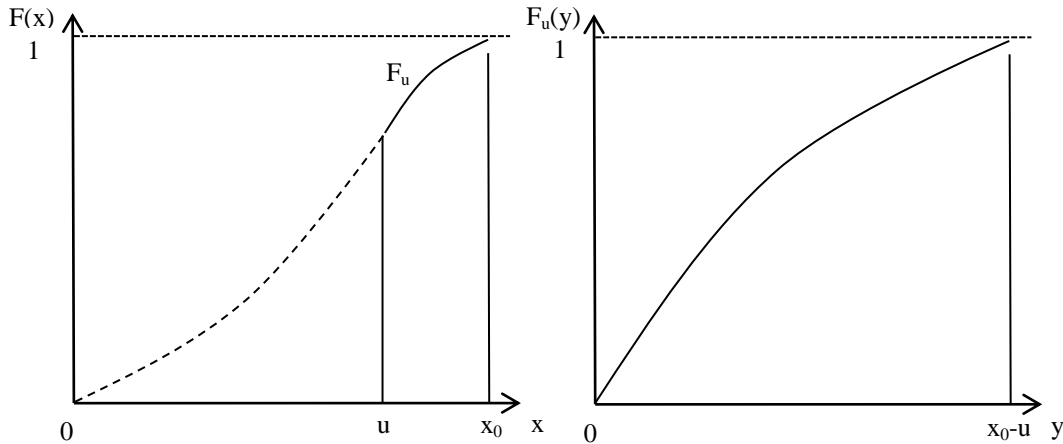
5.2.Exceedances over a threshold

In a different approach, we may model the behaviour of extreme events over a threshold given by the following conditional probability:

$$F_u(y) = p(X-u \leq y | X > u) = \frac{F(y+u)-F(u)}{1-F(u)}, \quad 0 \leq y < x_0. \quad (35)$$

where x_0 is the (finite or infinite) right endpoint of F . Graph 18 presents the relationship between the tail part of a distribution and the original distribution.

Graph 18- Original distribution function and estimation of the tail part



This approach is based on the following theorem:

Theorem 2 (Pickands, 1975; Balkema and de Haan, 1974)- Let X_1, X_2, X_3, \dots be a sequence of independent random variables from F . Let $M_n = \text{Max} \{X_1, \dots, X_n\}$. Suppose F satisfies the EVT, meaning that for large n , $p(M_n \leq z) \approx G(z)$, where G is defined as in Theorem 1. Then, for large enough u , the distribution function of $(X-u)$, conditional on $X > u$, can be approximated by the “generalized Pareto distribution” (GPD),

$$H(y) = \begin{cases} \mathbf{1} - \left(\mathbf{1} + \frac{\xi y}{\tilde{\sigma}} \right)^{-\frac{1}{\xi}}, & \xi \neq \mathbf{0} \\ \mathbf{1} - \exp\left(-\frac{y}{\tilde{\sigma}}\right), & \xi = \mathbf{0} \end{cases} \quad (36)$$

defined on $\{y: y > \mathbf{0} \text{ and } (\mathbf{1} + \xi y/\tilde{\sigma}) > \mathbf{0}\}$, where $\tilde{\sigma} = \sigma + \xi(u - \mu)$.

Note that the shape parameter ξ is the same in G and H .

A justification for Theorem 2 is as follows¹. Assume X has the distribution function F .

By Theorem 1 for large n ,

$$F^n(z) \approx \exp\left\{-\left[1 + \xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right\}, \quad (37)$$

for some parameters μ , $\sigma > 0$, and ξ . By taking the log,

$$n \log F(z) \approx -\left[1 + \xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}. \quad (38)$$

But $\log F(z) \approx -[1 - F(z)]$ by Taylor expansion. Therefore,

$$1 - F(u) \approx \frac{1}{n} \left[1 + \xi\left(\frac{u-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}, \quad (39)$$

for large u . Similarly, for $y > 0$,

$$1 - F(u + y) \approx \frac{1}{n} \left[1 + \xi\left(\frac{u+y-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}. \quad (40)$$

As a result,

$$\begin{aligned} p\{X > u + y | X > u\} &\approx \frac{n^{-1}[1 + \xi(u+y-\mu)/\sigma]^{-1/\xi}}{n^{-1}[1 + \xi(u-\mu)/\sigma]^{-1/\xi}} \\ &= \left[1 + \frac{\xi(u+y-\mu)/\sigma}{1 + \xi(u-\mu)/\sigma}\right]^{-1/\xi} \\ &= \left[1 + \frac{\xi y}{\tilde{\sigma}}\right]^{-1/\xi}, \end{aligned} \quad (41)$$

where $\tilde{\sigma} = \sigma + \xi(u - \mu)$.

¹ Formal proof can be found in Leadbetter et al. (1983). This sketch of proof is from Coles (2001).

One advantage of this method over the block maxima approach is that since each exceedance is associated with a specific event, it is possible to make the parameters ξ and σ depend on covariates (Smith, 2002).

The maximum likelihood estimator could be used to estimate the parameters of GPD. The log-likelihood function with k observations over the threshold u is as follows:

$$l(\sigma, \xi) = -k \log \sigma - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^k \log(1 + \xi y_i / \sigma), \quad (42)$$

for $(1 + \xi y_i / \sigma) > 0$, $i = 1, \dots, k$, and $\xi \neq 0$; and

$$l(\sigma) = -k \log \sigma - \sigma^{-1} \sum_{i=1}^k y_i, \quad (43)$$

for $\xi = 0$.

An important question is how to choose the threshold u . A threshold too high leaves very few observations, which lead to a high variance, while a threshold too low violates the asymptotic properties of the model, which leads to a bias. A useful tool to choose the threshold is the “empirical mean excess function”. Suppose $Y \sim \text{GPD}(\xi, \delta)$. Then, it can be shown that $E(Y) = \sigma / (1 - \xi)$ for $\xi < 1$. $E(Y)$ is infinite for $\xi \geq 1$. This means if the GPD is a valid model for the excesses over an arbitrary threshold u_0 , we have $E(X - u_0 | X > u_0) = \sigma_{u_0} / (1 - \xi)$ for $\xi < 1$. However, the GPD should also be a valid model for all thresholds $u > u_0$ subject to appropriate change of scale parameter σ_u . In this case, it can be shown that

$$E(X - u | X > u) = \frac{\sigma_u}{1 - \xi} = \frac{\sigma_{u_0} + \xi u}{1 - \xi}. \quad (44)$$

$E(X - u | X > u)$ is called the mean excess function. The above equation implies that if the exceedances are correctly modeled by a GPD, its mean excess function is linear in u for $u > u_0$. This suggests that for choosing a threshold, we should find a point after which the empirical

mean excess function is linear with a positive slope. The empirical mean excess function is defined as:

$$e_n(u) = \frac{1}{n_u} \sum_{i=1}^{n_u} (x_{(i)} - u), \quad (45)$$

where $x_{(1)}, \dots, x_{(n)}$ are ordered values of $x_i, 1, \dots, n_u$, such that $x_i > u$. Also, as a tool to check the validity of the model, the estimation could be done with different values of the threshold. If the model is valid for a threshold u_0 , the shape parameter ξ should remain constant for all $u > u_0$. We can show that only the scale parameter may change in this case.

5.3. Extensions and special cases

Analyzing data commonly shows that extreme conditions persist over several consecutive periods. This phenomenon is also observed frequently in financial time series data. Therefore, one extension of the theory is to include cases where dependency exists in the underlying process. It is worth recalling that the EVT theories assume the observations are iid and stationary. Embrechts, Kluppelberg, and Mikosch (1997), Leadbetter et al. (1983) and Danfelson and de Vries (1997) show that the EVT results hold even when there exist serial correlations in the data. This section briefly points out to some methods to overcome this problem. The importance of this section is that it shows even though we obtained iid residuals for our data series in Chapter 3, the results of our EVT estimations in Section 5.5 will hold even when the data series are not iid.

Regarding the dependency in the data, one approach is to assume that the extreme events $X_i > u$ and $X_j > u$ are approximately independent when u is high enough and there is a large separation between times i and j . The following definition and theorems are helpful in this case. More discussion is available in Coles (2001) and Leadbetter et al. (1983).

Definition – A stationary series X_1, X_2, \dots is said to satisfy the $D(u_n)$ condition if for all $i_1 < \dots < i_p < j_1 < \dots < j_q$ with $l_n = j_1 - i_p > 1$,

$$\begin{aligned} & |p(X_{i_1} < u_n, \dots, X_{i_p} < u_n, X_{j_1} < u_n, \dots, X_{j_q} < u_n) - \\ & p(X_{i_1} < u_n, \dots, X_{i_p} < u_n) p(X_{j_1} < u_n, \dots, X_{j_q} < u_n)| \leq \alpha(n, l_n), \end{aligned} \quad (46)$$

where $\alpha(n, l_n) \rightarrow 0$ for some sequence l_n such that $l_n/n \rightarrow 0$ as $n \rightarrow \infty$.

The $D(u_n)$ condition ensures that for a specific sequence of thresholds u_n increasing in n , the sets of variables that are far enough apart are almost independent such that the limit laws for extremes are not affected. Now, we can have an extension of Theorem 1 as follows¹.

Theorem 1.1 – Let X_1, \dots, X_n be a stationary process and define $M_n = \text{Max}\{X_1, \dots, X_n\}$. Then, if $\{a_n > 0\}$ and $\{b_n\}$ are sequences of constants such that $p\{(M_n - b_n)/a_n \leq z\} \rightarrow G(z)$ as $n \rightarrow \infty$ for a non-degenerate distribution function G , and the $D(u)$ condition is satisfied with $u_n = a_n z + b_n$ for every real z , then G is a member of the generalized extreme value family.

However, dependence in the data affects the value of the estimated parameters. The connection between the estimated parameters of an iid series with a dependent series are shown by the following theorem².

Theorem 1.2- Let X_1, \dots, X_n be a stationary process and X_1^*, \dots, X_n^* be a sequence of independent variables with the same marginal distribution. Define $M_n = \text{Max}\{X_1, \dots, X_n\}$ and $M_n^* = \text{Max}\{X_1^*, \dots, X_n^*\}$. Under suitable regularity conditions, $p\{(M_n^* - b_n)/a_n \leq z\} \rightarrow G_1(z)$ as $n \rightarrow \infty$ for normalizing sequences $\{a_n > 0\}$ and $\{b_n\}$,

¹ Formal proof in Leadbetter et al. (1983).

² *ibid.*

where G_I is a non-degenerate distribution function, if and only if, $p\{(M_n - b_n)/a_n \leq z\} \rightarrow G_2(z)$, where $G_2(z) = G_1^\theta(z)$ for a constant θ such that $0 < \theta \leq 1$.

Theorem 1.2 suggests that if the maxima of a stationary series converge provided an appropriate $D(u_n)$ condition, then the limit distribution is related to the limit distribution of an independent series. The precise relationship between the parameters of the model is as follows:

$$\begin{aligned} G_1^\theta(z) &= \exp\left\{-\left[1 + \xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right\}^\theta \\ &= \exp\left\{-\theta\left[1 + \xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right\} \\ &= \exp\left\{-\left[1 + \xi\left(\frac{z-\mu^*}{\sigma^*}\right)\right]^{-\frac{1}{\xi}}\right\} = G(z), \end{aligned} \quad (47)$$

where $\mu^* = \mu - \frac{\delta}{\xi}(1 - \theta^{-\xi})$ and $\sigma^* = \sigma\theta^\xi$.

In practice, we can assume that the $D(u_n)$ condition holds for the block maxima approach since the maxima of blocks can be assumed to be independent. That means that, for example, using a GEV family to model annual maxima is still valid if there is dependency in the data. Similar arguments suggest that exceedances over a threshold could also be approximated by GPD. However, one difficulty for the threshold approach is that exceedances have a tendency to cluster, i.e. extreme events may persist over time. Several treatments have been suggested for this problem such as:

- Identifying clusters of exceedances, choosing the maxima of each cluster as one observation, and assuming these observations are independent. This is the traditional approach to deal with this problem, but has its own limitations (e.g. Coles, 2001).

- Using standardized observations to deal with serial correlations in the data (e.g. Resnick and Starica, 1996). We use this approach in the next section by filtering data through a GARCH model to obtain standardized residuals.
- Model dependence via Markov chains, max-stable processes, and multivariate maxima of moving maxima (M4) processes (e.g. Zhang and Smith, 2008). This approach is under progress specially to deal with multivariate extremes.

With respect to the high volatility and dependency of exchange rate time series data, some authors suggested using a GARCH model to remove the volatility and long-term dependency of the data before using an EVT model (McNeil and Frey, 2000; Smith, 2003, and Davis and Mikosch, 2009). However, we may still need to deal with the short-term dependency in the data. We used a combination of GARCH, ARMA, and AR models in Chapter 3 to remove the long- and short-term dependencies in the data and obtain iid residuals.

Another extension of the EVT is multivariate extreme value theory, in which there exist several variables or processes, and an extreme in any of these variables is of interest. Also, there might be dependency in and between these variables. Multivariate extreme value theory is less widely used in practice, partially due to the difficulty in defining multivariate extremes and higher chance of dependency in the data. Zhang and Smith (2008) is an example of this extension with an application in modeling exchange rates. We also extend our study to a bivariate case in section 5.5, though we take a different approach.

5.4. Random variable generation and importance sampling

This Chapter briefly explains two techniques that will be used in Chapter 6 where the EVT simulation results will be presented: the inverse transform to generate random variables and

importance sampling to expedite the simulation of rare events. The discussions of this Chapter are from Asmussen and Glynn (2000), Robert and Casella (2004), Givens and Hoeting (2005), and Rubino and Tuffin (2009).

The inverse transform is a method to generate a random variable from the distribution f having a random variable generated from the uniform distribution. Suppose the generic probability triple (Ω, F, P) is represented by $([0,1], B, u_{[0,1]})$, where B are the Borel sets on $[0,1]$. In this case, the variable $\omega \in \Omega$ will be equated with that of a uniform variable in $[0,1]$. Then, the random variables X from F are functions of uniform variables transformed by the generalized inverse function defined as follows.

Definition- For a non-decreasing function F on R , the generalized inverse of F , F^- , is the function defined by,

$$F^-(u) = \inf\{x: F(x) \geq u\}. \quad (48)$$

The following lemma gives us a representation of any random variable as a transform of a uniform random variable:

Lemma- If $U \sim u[0,1]$, then the random variable $F^-(u)$ has the distribution F .

Therefore, to generate a random variable $X \sim F$, we need to generate the random variable $U \sim u[0,1]$, and then make the transformation $X = F^-(U)$. That means, for example to generate exponential and GPD random variables in this study, we may first generate a uniform random variable U , and then solve for $X = -1/\lambda \ln(1 - U)$ and $X = -\delta/\xi [(1 - U)^{-\xi} - 1]$ to obtain exponential(λ) and GPD(ξ, σ) random variables.

The next technique we will use in simulation is importance sampling. Importance sampling is a change of measure to increase the probability of events of interest in a Monte Carlo simulation. The probability will then be adjusted to reflect this change of measure.

Importance sampling can be explained as follows. Suppose we want to evaluate the following integral:

$$E_f[h(x)] = \int_{\mathcal{X}} h(x)f(x)dx. \quad (49)$$

Under the normal Monte Carlo approach, we generate a sample (X_1, \dots, X_m) from the density f to approximate (39) by $\bar{h}_m = \frac{1}{m} \sum_{j=1}^m h(x_j)$ since \bar{h}_m converges almost surely to $E_f[h(x)]$ by the SLLN. However, this approach is not efficient if the events of interest happen rarely. To improve the result, we may make the events of interest occur more frequently than it would happen in the normal Monte Carlo method. This can happen by generating random variables from another distribution g that oversamples from the portion of the state space that receives lower probability under f . Later, importance weights correct for this bias. The new estimator will then be $\bar{h}_m^* = \frac{1}{m} \sum_{j=1}^m h(x_j) \omega^*(x_j)$, where $x_j \sim g$, and $\omega^*(x_j) = f(x_j)/g(x_j)$ are importance weights. This approach is called importance sampling and is based on an alternative representation of $E_f[h(x)]$:

$$E_f[h(x)] = \int h(x)f(x)dx = \int h(x) \frac{f(x)}{g(x)} g(x)dx. \quad (50)$$

The application of importance sampling with an exponential distribution can be illustrated as follows. Suppose we are interested in finding the probability of $X > 15$ in *exponential*($\lambda = 1$). The direct calculation shows that $p(X > 15) = 1 - (1 - e^{-10}) = 3 \times$

10^{-7} . However, obtaining this probability by using a Monte Carlo method is very computer intensive as it requires 10^{-7} iterations in average to obtain three events of interest. Using importance sampling method, we may instead sample from $exponential(\lambda = 1/22)$ in which almost half of the sample will be bigger than 15 in a Monte Carlo simulation. The probability of having $X > 15$ will then be corrected by $\omega^*(X) = f(X|\lambda = 1)/f(X|\lambda = 1/22)$, $f(X|\lambda) = \lambda e^{-\lambda x}$, each time this event occurs under $exponential(\lambda = 1/22)$. This way, we will need much less iterations to obtain $p(X > 15)$ than under $exponential(\lambda = 1)$. A sample program of importance sampling in R is presented in appendix 3.

One important consideration when using importance sampling with an exponential distribution is that the second moment of the importance sampling estimator may become infinite and/or the estimator captures the cases that are too unlikely to happen. For example, if we use the rate $\hat{\lambda}$ instead of λ , the second moment of the estimator will be as follows:

$$\hat{E}[(1_A(X)w^*(x))] = \int_T^\infty \left(\frac{\lambda e^{-\lambda x}}{\hat{\lambda} e^{-\hat{\lambda} x}}\right)^2 \hat{\lambda} e^{-\hat{\lambda} x} dx = \frac{\lambda^2}{\hat{\lambda}(2\lambda - \hat{\lambda})} e^{-(2\lambda - \hat{\lambda})T}, \quad (51)$$

where $A = [T, \infty]$ is the area of interest, i.e. $h(X) = P[X \geq T]$. In this case, the second moment is finite if and only if $0 < \hat{\lambda} < 2\lambda$. When $\hat{\lambda} > 2\lambda$, the second moment becomes infinite because the squared likelihood ratio w^* grows exponentially with x at a faster rate than the exponential rate of decrease of the density (Rubino and Tuffin, 2009). Moreover, if $\hat{\lambda}$ becomes too small (i.e. the mean becomes too large), the variance again tends to infinity. That is the reason we cannot use this method to do very large change of measures.

5.5. Application of EVT to exchange rate and treasury bill rate changes

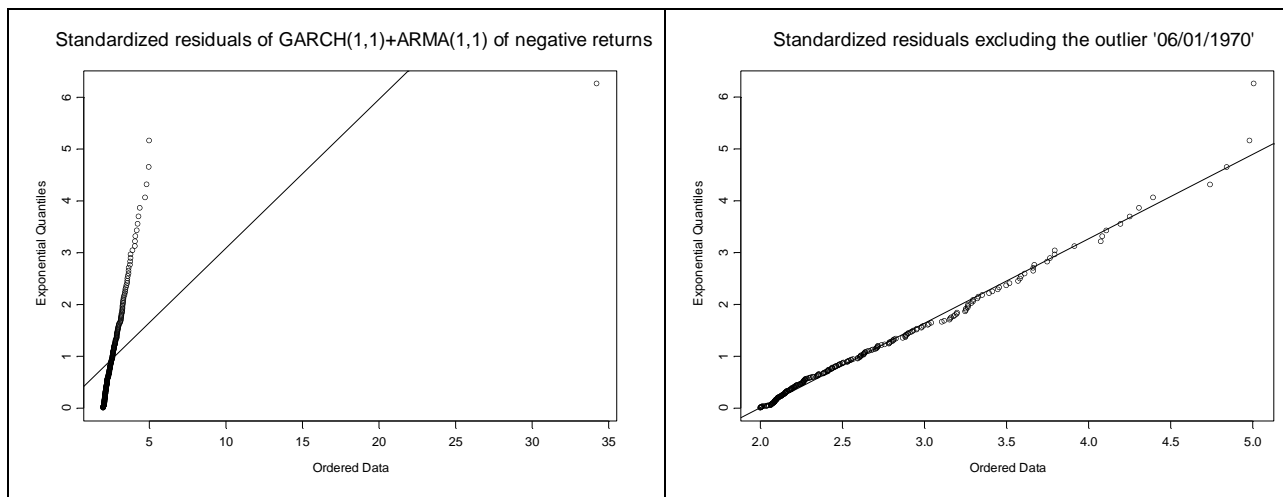
In this Chapter, we estimate the GPD parameters for the exceedances over a threshold on the residuals of the US-Canada exchange rates and the Bank of Canada treasury bill rates we obtained in Chapter 3. A method to understand the tail behaviour of the series is to create a qq-plot against the exponential distribution as a reference distribution. If the exceedances over thresholds are from an exponential distribution, then the shape parameter ζ is 0 and the qq-plot should be linear. Departures from linearity in the qq-plot then indicate either fat-tailed behaviour ($\zeta > 0$) or bounded tails ($\zeta < 0$). Graph 19 presents the qq-plots of the exceedances of z_t under two cases where the outlier on 06/01/1970 is included and excluded. We see a departure from the straight line when the outlier is included, but the qq-plot suggests an exponential distribution when we exclude this outlier (also, leaving this outlier out transform the heavy-tailed residuals to an exponential distribution).

It is worth mentioning that on 06/01/1970 Canada went off fixed exchange rates with the US. After capturing all short and long run dependencies in the data by a GARCH and ARMA model, we observe that this is the largest shock received by the Canadian exchange rate since 1950. The magnitude of this shock is almost seven times bigger than the next biggest shock. We kept this outlier in the non-parametric analysis of Chapter 4. However, we do not include it in the EVT analysis since it does not fit models we estimate in this section.

We need to set a threshold to determine exceedances in the series and fit a model on them. A useful tool to find a threshold is the empirical mean excess function. To determine the threshold level, we need to find an interval where the mean excess function has a positive slope. Also, the estimated shape parameter should be relatively stable with respect to the threshold. As explained before, a right choice of the threshold is critical because if it is too low, we may violate

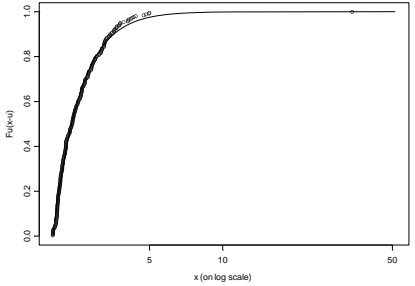
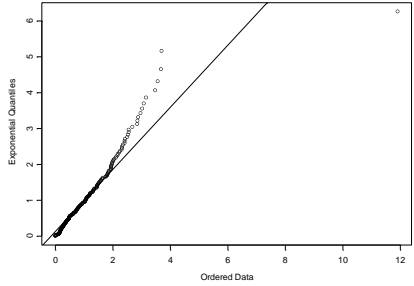
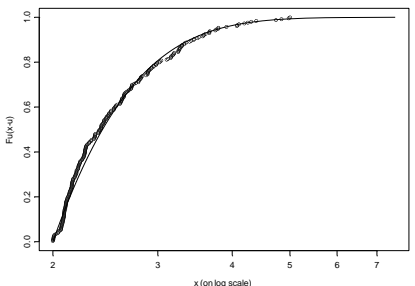
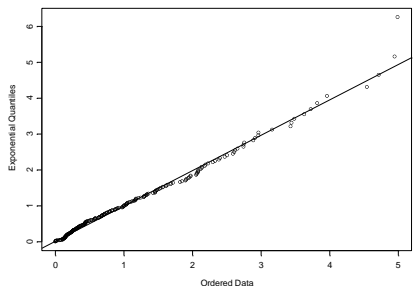
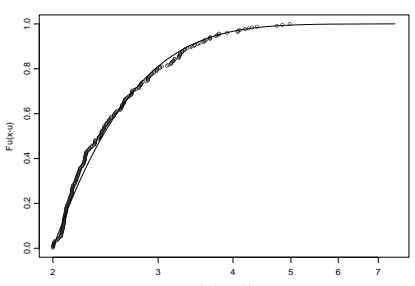
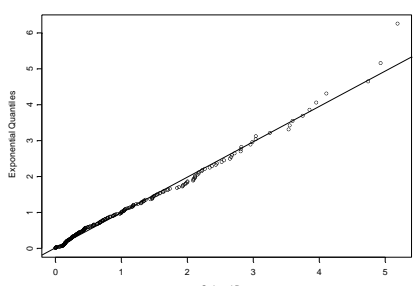
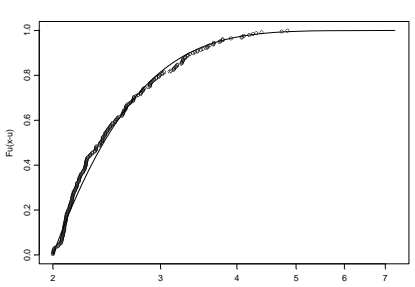
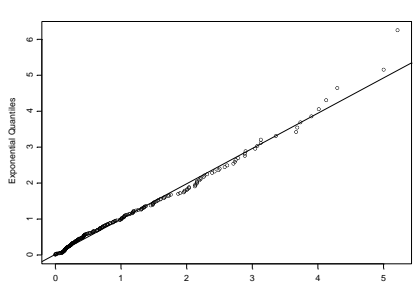
the asymptotic basis of the model which leads to bias in estimation, and if it is too high, few observations will be left which will lead to a high variance. We select two thresholds $u = 2$ and $u = 1.5$ for the estimation. Also, we remove the four right end points of the data one by one to examine the stability of the estimations. Graphs 20 and 21 present the estimated parameters and fitted models for $u = 2$ and $u = 1.5$ respectively. The first column presents the fitted models, and the second columns compare the exceedances with the quantiles of an exponential distribution. As these graphs suggest, the exceedances fit the model quite well when the outlier of 06/01/1970 is excluded. The estimated parameters are also close in both graphs.

Graph 19- qq-Plot of z_t against the exponential distribution

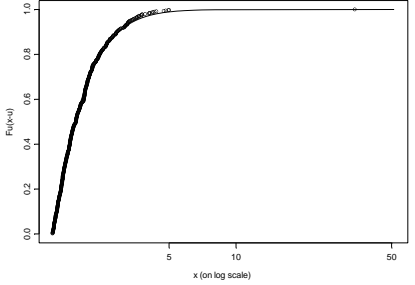
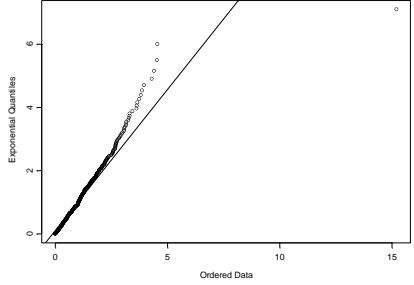
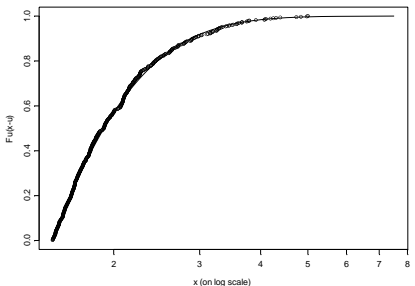
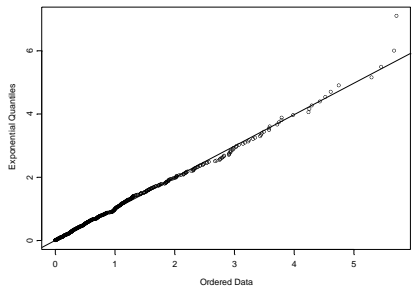
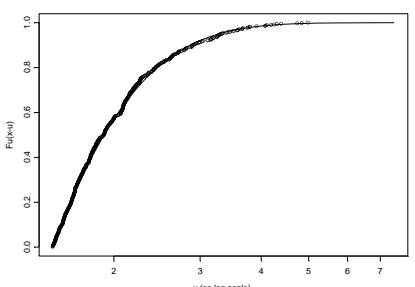
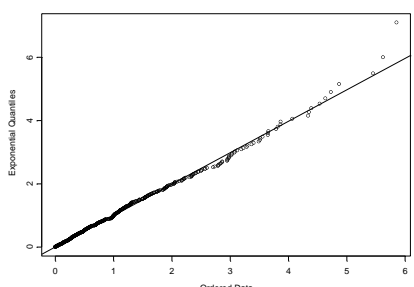
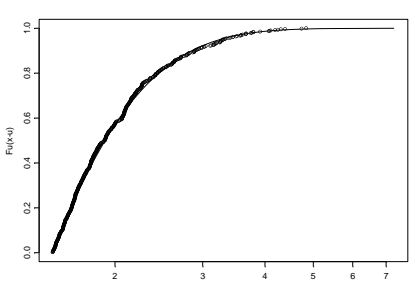
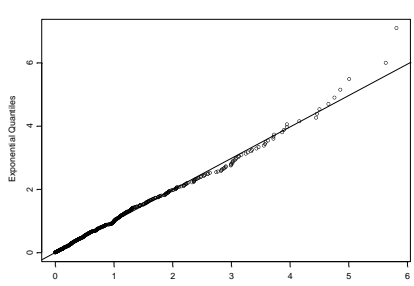


Graph 22 presents the stability of the shape parameter. This plot shows how the MLE of the shape parameter ζ varies with the selected threshold u . We should choose u from a region where the shape parameter ζ remains relatively stable. The upper-right graph belongs to the case where all exceedances are included, and the other three graphs present the cases when we remove the outliers one by one. According to Graph 22, the MLE of the shape parameter ζ is stable when we remove the outlier of 06/01/1970 and we keep over 250 exceedances.

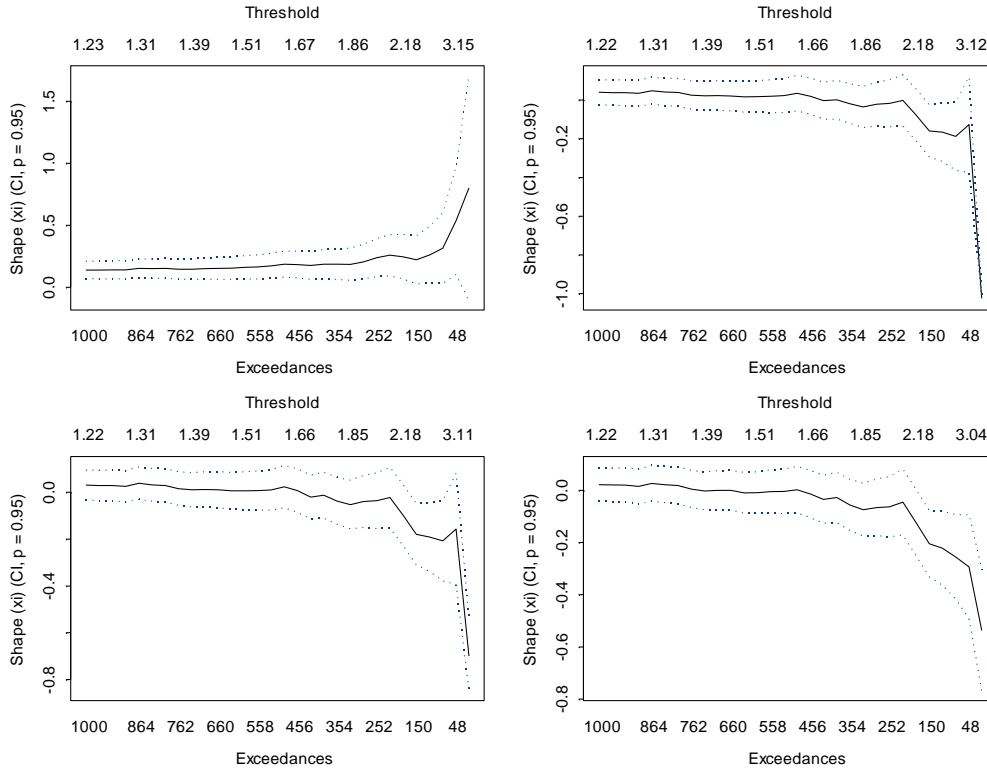
Graph 20- Exceedances over a threshold estimation; $u = 2$

Estimated parameters	Excess Distribution	QQplot of Residuals
<p>Case 1: All data</p> <p>$n = 261$</p> <p>$\xi = 0.2289929$ s.d. = 0.06283206</p> <p>$\beta(u) = 0.5162442$ s.d. = 0.04494105</p>		
<p>Case 2: Excluding 06/01/1970</p> <p>$n = 260$</p> <p>$\xi = -0.005655687$ s.d. = 0.07277196</p> <p>$\beta(u) = 0.6109717$ s.d. = 0.05841683</p>		
<p>Case 3: Excluding 06/01/1970, 12/31/1998</p> <p>$n = 259$</p> <p>$\xi = -0.02569769$ s.d. = 0.07236858</p> <p>$\beta(u) = 0.6136965$ s.d. = 0.05852891</p>		
<p>Case 4: Excluding 06/01/1970, 12/31/1998, 12/29/1999</p> <p>$n = 258$</p> <p>$\xi = -0.04945597$ s.d. = 0.07244067</p> <p>$\beta(u) = 0.6183738$ s.d. = 0.05903489</p>		

Graph 21- Exceedances over a threshold estimation; $u = 1.5$

Estimated parameters	Excess Distribution	QQplot of Residuals
<p>Case 1: All data</p> <p>$n = 604$</p> <p>$\xi = 0.1552308$ s.d. = 0.03734148</p> <p>$\beta(u) = 0.5291026$ s.d. = 0.02902898</p>		
<p>Case 2: Excluding 06/01/1970</p> <p>$n = 603$</p> <p>$\xi = 0.01414949$ s.d. = 0.04459061</p> <p>$\beta(u) = 0.5889857$ s.d. = 0.03556595</p>		
<p>Case 3: Excluding 06/01/1970, 12/31/1998</p> <p>$n = 602$</p> <p>$\xi = 0.001761105$ s.d. = 0.04445515</p> <p>$\beta(u) = 0.5915517$ s.d. = 0.03567908</p>		
<p>Case 4: Excluding 06/01/1970, 12/31/1998, 12/29/1999</p> <p>$n = 601$</p> <p>$\xi = -0.01185767$ s.d. = 0.04468459</p> <p>$\beta(u) = 0.5947643$ s.d. = 0.03598382</p>		

Graph 22- The shape parameter for $u = 1.5$ when removing the outliers



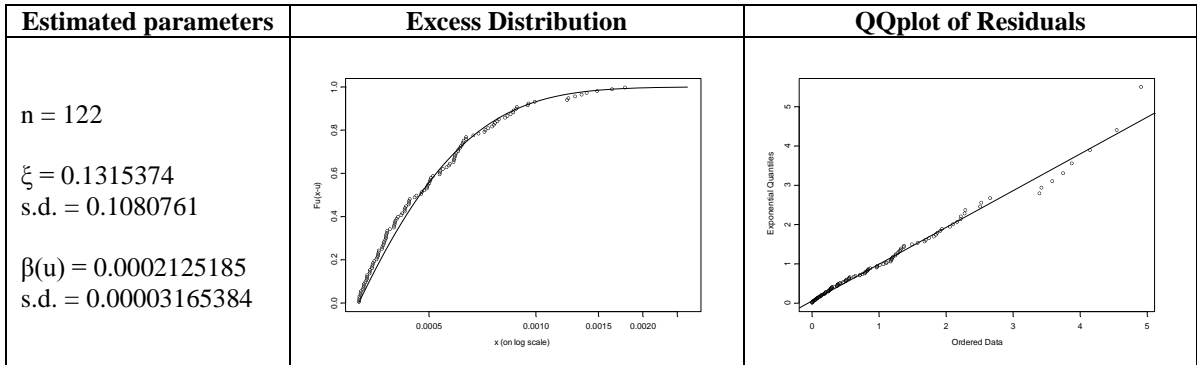
These estimations suggest that when the outlier of 06/01/1970 is taken out, the estimated shape parameter is not statistically different from zero, and the scale parameter is around 0.59 ($u = 1.5$). The models fit the data fairly well. This suggests an exponential distribution with mean 0.59 for the tail part of z_t . A Kolmogorov-Smirnov test also confirms that the exceedances are from an exponential distribution with mean 0.59.

Next, we apply the EVT on the treasury bill residuals. Graph 23 presents the estimation results of EVT on the residuals of the treasury bill after filtering with an AR(2) model as in Section 3.3¹. The estimation suggests a GPD distribution with a shape parameter ξ equal to 0.132 and a scale parameter $\beta(u)$ equal to 0.00021. The model shows an acceptable fit to the data.

¹ The following outlier has been removed to have a better fit to the EVT model: 0.0042910 (08/10/2008).

Graph 23- Exceedances over a threshold estimation of treasury bill negative returns;

$u = 0.000317$ (weekly data)



Chapter 6

Monte Carlo simulation using EVT

This chapter combines the Monte Carlo analysis of Chapter 4 with the EVT and importance sampling techniques introduced in Chapter 5. The idea is to increase the probability of failure in the network, and then adjust the probability with an appropriate change of measure. We tried different approaches to test this method. The conclusion is that even though this method increases the probability of failure in the system, finding an appropriate change of measure would be very difficult. In what follows, we first explain how the EVT and importance sampling could be used in the simulation. Then, we present the results of different methods we used. We do this extension only for the univariate case with the US-Canada exchange rates.

As Chapter 4, the negative exchange rate changes are reconstructed by using equations (52) - (55):

$$\hat{r}_t = c + \alpha \hat{r}_{t-1} + \beta \hat{\varepsilon}_{t-1} + \hat{\varepsilon}_t, \quad (52)$$

$$\hat{\varepsilon}_t = \hat{\sigma}_t \hat{z}_t, \quad (53)$$

$$\hat{\sigma}_t^2 = a + \lambda \hat{\varepsilon}_{t-1}^2 + \mu \hat{\sigma}_{t-1}^2, \quad (54)$$

$$\hat{z}_t \sim iid. \quad (55)$$

where $c = 0.0001$, $\alpha = -0.3526$, $\beta = 0.4205$, $a = 0.0000002$, $\lambda = 0.1207$, and $\mu = 0.8592$. \hat{z}_t 's are sampled again with replacement from the iid standard residuals we obtained in Section 3.2. The only difference with Chapter 4 is that if \hat{z}_t is larger than $u = 1.5$ ¹, we replace it with an

¹ In fact we do this replacement if $1.5 < \hat{z}_t < 34$, since we exclude the outlier from the model.

exponential(λ) random variable. This replacement is justified by the EVT estimation we did in Section 5.5 where we showed the tail part of the residuals exhibits an exponential distribution with rate $\lambda = 1/0.59$. We also propose that with an appropriate change of measure, we could use another exponential distribution with a rate $\hat{\lambda} < 2\lambda$ to increase the probability of failure in the system. The rest of the process is the same as Chapter 4: we calculate the level changes in the exchange rate $\Delta\hat{R}_t = \hat{R}_t - \hat{R}_{t-1}$ and will assess whether a failure and contagion will happen in the system. We did three types of tests on this method as follows:

i) In the first step, we ran some simulations with $\lambda = 1/0.59$ to verify if this method works. The simulation results confirm that this change works quite well and the estimated results under this approach are very close to the non-parametric results obtained in Section 4 for all reserve rates.

ii) In the second step, we wanted to increase the failure rate in the system to have an estimate of the probability of failure and contagion when the reserve ratio is high. The approach we took was to replace $\hat{z}_t > 1.5$ with an *exponential*($\lambda' < 1/0.59$), and then use the likelihood ratio $\omega^*(z) = f(z|\lambda = 1/0.59)/f(z|\lambda')$ to adjust for the change of measure whenever this replacement happens in the reconstructed series. At the end, if a failure happens under this series, these likelihood ratios are multiplied to obtain the probability of this series to occur. Table 24 presents the simulation results of different λ' for the reserve ratio equal to 5% for 1000 iterations.

In table 24, $B = 1/\lambda$ is the mean of the exponential distribution. As table 24 suggests, increasing the mean of the distribution increases the number of failures and contagions in the system. However, when the probability of failure is corrected for this change of measure, two problems happens. On one hand, the probability of failure is usually under-estimated. On the other hand, the estimated probabilities are too volatile even with small changes in the mean.

Table 24- Simulation results of different λ' for the reserve ratio equal to 5%

r = 0.05	B = 0.59	B = 0.61	B = 0.63	B = 0.65	B = 0.67	B = 0.69	B = 0.71
# of failure (in 1000)	98	88	88	118	104	120	106
# of contagion (in 1000)	94	74	78	114	90	102	100
adjusted prob of failure (%)	9.800000	8.300000	4.700000	9.900000	13.359053	2.059788	2.228636
adjusted prob of contagion (%)	9.400000	6.600000	4.300000	8.800000	11.016469	1.222130	1.973480
r = 0.05	B = 0.73	B = 0.75	B = 0.85	B = 1	B = 1.59	B = 2.59	B = 3.59
# of failure (in 1000)	114	104	112	140	250	748	954
# of contagion (in 1000)	98	98	96	124	222	662	796
adjusted prob of failure (%)	2.113183	0.510394	8.717823	0.882176	0.000082	0.001016	2.977956
adjusted prob of contagion (%)	1.807464	0.509049	8.485752	0.882023	0.000081	0.001016	2.977956

iii) We tried to improve the second approach by replacing $\hat{z}_t > 1.5$ with an *exponential*($\lambda' < 1/0.59$) only if the variance of the reconstructed series, $\hat{\sigma}_t^2$, was also high. The idea was that failures usually happen when a large shock is augmented with a large variance. Table 25 presents the results of the simulation for the reserve value equal to 5%, $B = 2.59$, and the variance in 25% quantile. Table 25 suggests that even though this method improved the estimated probability of failure, it is still under-estimated.

**Table 25- Simulation results for the reserve value equal to 5% and $B = 2.59$
when variance is high**

r = 0.05	B = 2.59
# of failure (in 1000)	683
# of contagion (in 1000)	605
adjusted prob of failure (%)	5.67958
adjusted prob of contagion (%)	5.62733

r = 0.05; r > 0	B = 2.59
# of failure (in 1000)	300
# of contagion (in 1000)	279
adjusted prob of failure (%)	27.6217
adjusted prob of contagion (%)	26.5440

We improved this technique one step ahead with making the change only when r_t was negative in the simulation. The rationale was that since the coefficient of r_t is negative in Equation 15, the failure is not likely when r_t is positive. This extra step improved the outcome, but the results are not satisfactory yet.

The conclusion of this section is that combining EVT and importance sampling techniques with the Monte Carlo method can significantly increase the number of failures in the system. However, finding an appropriate change of measure will be difficult.

Conclusion

This study simulated the shocks to a banking network and estimated the probability of failure and contagion in the system. We considered two sources of shocks: external shocks due to market and macroeconomic factors which impact the entire banking system, and idiosyncratic shocks due to failure of a single bank. We used non-parametric simulation and the extreme value theory to model the shocks. In doing so, we also used different techniques such as GARCH and importance sampling. GARCH modeling and other time series techniques have been used to convert the shocks to an iid sequence. Later, we showed that importance sampling could be used to accelerate the probability of failure in the system. We calibrated the model such that the banking network resembles the Canadian banking system.

One of the interesting findings of the empirical Chapter of the paper is that the GARCH/ARMA model could identify large shocks to the US-Canada exchange rates, especially the one that happened on 06/01/1970 when Canada went off fixed exchange rates with the US.

Moreover, the simulation results confirm that the probability of failure and contagion in the system decreases as the reserve ratio increases. A breaking point is around 5% where the rate of failure changes. Moreover, the simulation results show that the failure of banks due to exchange rate shocks has a positive relationship with the ratio of foreign assets to total assets. Similarly, having a high ratio of interbank loans increases the probability of contagion.

Another result is that the exceedances over a threshold can be replaced with the estimated EVT model and obtain the same simulated results. The advantage of using an EVT model is that importance sampling or other techniques could then be used to generate bigger shocks and increase the number of failures in the system to obtain more accurate estimates or reduce the

number of iterations in the Monte Carlo simulation in the case of high reserve ratios (i.e. rare events). Preliminary results show that combining importance sampling and EVT can significantly increase the number of failure and contagion in the system, though the standard deviation of the estimator can be large if the change of measure is naive. Finding a good change of measure could be one of the extensions of this study.

References

- Allen, F., and D. Gale, (2000) “Financial Contagion”, *Journal of Political Economy*, 108, p.1–33.
- Allen, F., and D. Gale, (2007) “An Introduction to Financial Crises”, Chapter 1 in *Financial Crises* edited by F. Allen and D. Gale, The International Library of Critical Writings in Economics.
- Andersen, T.G., R.A. Davis, J.P. Kreiß, and T. Mikosch, (2009) *Handbook of Financial Time Series*, Springer.
- Asmussen, S. and P.W. Glynn, (2008) *Stochastic Simulation: Algorithms and Analysis*, New York: Springer.
- Boss, M., H. Elsinger, M. Summer, and S. Thurner, (2004) “Network topology of the interbank market”, *Quantitative Finance*, 4, p. 677-84.
- Coles S., (2001) “An Introduction to Statistical Modeling of Extreme Values”, Springer.
- Danielsson, J. and C. G. de Vries (1997) “Tail index and quantile estimation with very high frequency data”, *Journal of Empirical Finance*, 4, p. 241-257.
- Davis R.A., and T. Mikosch, “Extreme Value Theory for GARCH Processes” in Andersen T.G., R.A. Davis, J.P. Kreiß, and T. Mikosch, (2009) “Handbook of Financial Time Series”.
- Embrechts, P., C. Kluppelberg, and T. Mikosch, (1997) “Modeling Extremal Events for Insurance and Finance”, Springer, Berlin.
- Elsinger, H., A. Lehar, and M. Summer, (2006) “Risk Assessment for Banking Systems”, *Management Science*, 52 (9), p. 1301-1314.
- Eisenberg, L., and T. Noe, (2001) “Systemic risk in financial systems”, *Management Science*, 47, 236–249.

- Francq, C., and J.M. Zakoïan, (2010) *GARCH models : structure, statistical inference and financial applications*, Wiley.
- Furfine, C., (2003) “Interbank Exposures: Quantifying the Risk of Contagion”, *Journal of Money, Credit and Banking*, 35(1), p.111-128.
- Gai, P., and S. Kapadia, (2010) “Contagion in financial networks”, Bank of England Working Paper 383.
- Gauthier, C., A. Lehar, and M. Souissi, (2010) “Macroprudential Regulation and Systemic Capital Requirements”, Bank of Canada Working Paper 2010-4.
- Gencay, R., and F. Selcuk, (2006) “Overnight borrowing, interest rates and extreme value theory”, *European Economic Review*, 50(3), p. 547-563.
- Givens, G.H., and J.A. Hoeting, (2005) *Computational Statistics*, John Wiley & Sons, New York.
- Gilli, M., and E. Kellezi, (2006) “An Application of Extreme Value Theory for Measuring Financial Risk”, *Computational Economics* 27(1), 2006, p.1-23.
- Hols, M. C., and C. G. de Vries, (1991), “The Limiting Distribution of Extremal Exchange Rate Returns”, *Journal of Applied Econometrics*, 6(3), p. 287-302.
- Illing, M., and Y. Liu (2003) “An Index of Financial Stress for Canada”, Bank of Canada Working Paper 2003-14.
- Leadbetter, M. R., G., Lindgren, and H. Rootzen, (1983) *Extremes and Related Properties of Random Sequences and Series*, Springer Verlag, New York.
- Li, F., (2009) “Testing for Financial Contagion with Applications to the Canadian Banking System”, Bank of Canada Working Paper 2009-14.
- May, R.M., (2010) “Systematic Risk: The Dynamics of Banking System”, The Ockham Lecture.

- McNeil A.J., and R. Frey, (2000) “Estimation of Tail-Related Risk Measures for Heteroscedastic Financial Time Series: an Extreme Value Approach”, *Journal of Empirical Finance*, 7, p. 271-300.
- Resnick, I.R., (2007) “Extreme Values, Regular Variation and Point Processes”, *Springer Series in Operations Research and Financial Engineering*.
- Resnick, I.R., and C. Starica, (1996) “Tail index for Dependent Data”, *School of ORIE, Cornell University, Discussion paper*.
- Robert, C. P. and G. Casella, (2005) *Monte Carlo Statistical Methods*, Second edition. New York: Springer.
- Rubino G., and B. Tuffin, (2009) *Rare Event Simulation using Monte Carlo Methods*, John Wiley & Sons.
- Santor, E., (2003) “Banking Crises and Contagion: Empirical Evidence”, *Bank of Canada Working Paper 2003-1*.
- Schafgans, M., Koedijk, K. G, and C.G. de Vries, (1990) “The tail index of exchange rate returns”, *Journal of international economics*, 29(1-2), p. 93-108.
- Smith, R.L., “Measuring Risk with Extreme Value Theory”, in Dempster, M.A.H., (2002) “Risk Management: Value at Risk and Beyond”, *Cambridge University Press*.
- Smith, R.L., “Statistics of Extremes, with Applications in Environment, Insurance and Finance”, in Finkenstadt B., and H. Rootzén, (2003) “Extreme Values in Finance, Telecommunications and the Environment”, *Chapman and Hall/CRC Press*.
- Upper, C., (2007) “Using counterfactual simulations to assess the danger of contagion in interbank markets”, *Bank for International Settlements Working Paper 234*.
- Zhang Z., and R.S. Smith, (2008) “On the Estimation and Application of Max-Stable Processes”.

Appendix 1- Measures for shocks to the Canadian banking system

This Chapter describes the main shocks to the Canadian banking system according to the study done by Illing and Liu (2003). The measures they proposed for each of the shocks are also presented. To find the shocks to the banking system in this study, we model the extreme values of these measures by applying the EVT on them. The four shocks Illing and Liu describe as the most important shocks to the Canadian banking system are as follows: banking sector stresses, foreign exchange stresses, debt market stresses and equity market stresses.

1- Banking sector:

Illing and Liu refer to Demirgüç-Kunt and Detragiache (1998) to define a banking crisis, where it can be described as a situation in which at least one of the following conditions holds: “(i) the ratio of non-performing assets to total assets is greater than 10 per cent, (ii) the cost of the rescue operation is at least 2 per cent of GDP, (iii) banking problems result in the large-scale nationalization of banks, and (iv) extensive bank runs lead to emergency measures.” Illing and Liu propose the following index to measure banking sector stresses:

$$b = \text{cov}(r,m)/\text{var}(m) \text{ calculated daily,}$$

where,

r = year-over-year percentage change in the TSX Bank & Trust Total Returns Index (formerly the TSE Bank & Trust Total Returns Index; source: Datastream).

m = year-over-year percentage change in the S&P/TSX Total Returns Index (formerly the TSE 300 Total Returns Index; source: Datastream).

2- Foreign exchange market:

Foreign exchange crises may be defined as significant devaluations, losses in reserves, and/or defensive interest rate increases. The following measure has been suggested to estimate foreign exchange stresses:

$$canC6_t / \max[canC6 \text{ in } (canC6_{t-j} \mid j = 0, 1, \dots, T)] \text{ where } T = 365.$$

That is, the daily value of the Canadian effective exchange rate (*canC6*) as a per cent of its maximum value over the preceding 365 calendar days (i.e., the CMAX method). *canC6* is a weighted combination of the U.S. dollar (85.84%), the euro (5.95%), the Japanese yen (5.27%), the U.K. pound (2.17%), the Swiss franc (0.42%), and the Swedish krona (0.35%). Source: Bank of Canada.

3- Debt market:

Illing and Liu refer to Bordo and Schwartz (2000) to describe a debt crisis as “the inability of sovereign nations or the broad private sector to service foreign debts.” They propose the following measure to estimate debt stresses:

$$\text{Canada-US covered interest rate differential } (1 + r_t^*) = F_t / S_t^* (1 + r_t)$$

where,

r_t^* = U.S. Government 90-day Treasury bill rate as at noon on day t

F_t = 90-day forward rate for the US-Canada dollar exchange rate as at noon on day t

S_t = spot rate for the US-Canada dollar exchange rate as at noon on day t

r_t = Government of Canada 90-day Treasury bill rate as at noon on day t

(Source: Bank of Canada)

4- Equity market

Equity crises can be defined as a sharp decrease in the overall market index. The following measure was proposed to measure equity stresses:

$TSX_t / \max[TSX_{t-j} \mid j = 0, 1, \dots, T]$ where $T = 365$.

That is, the S&P/TSX Price Index as a per cent of its maximum value over the preceding 365 calendar days (Source: Toronto Stock Exchange).

Appendix 2- List of the Canadian banks

The banks we consider for this study are the ones that have participated in the Large Value Transfer System (LVTS). The LVTS is a real-time, electronic wire transfer system that processes large-value, time-critical payments throughout the day (Bank of Canada¹). The list of these banks is as follows:

- Alberta Treasury Branches
- Bank of America, National Association
- Bank of Montreal
- The Bank of Nova Scotia
- BNP Paribas (Canada)
- La Caisse centrale Desjardins du Québec
- Canadian Imperial Bank of Commerce
- Credit Union Central of Canada
- HSBC Bank Canada
- Laurentian Bank of Canada
- National Bank of Canada
- Royal Bank of Canada
- State Street Bank and Trust Company
- The Toronto-Dominion Bank

The Royal Bank of Canada (RBC), The Toronto-Dominion Bank (TD), The Bank of Nova Scotia (ScotiaBank), Bank of Montreal (BMO), and Canadian Imperial Bank of Commerce

¹ http://www.bankofcanada.ca/en/financial/financial_pay.html

(CIBC) are called the “Big Five Banks”. These five banks plus the National Bank of Canada (NBC), which stands on the six place, account for 90.3% of all banking assets in Canada (Gauthier, Lehar, and Souissi; 2010). Gauthier, Lehar, and Souissi (2010) use the banks’ monthly balance sheet reports to Office of the Superintendent of Financial Institutions Canada (OSFI) to construct a matrix of Big six banks’ bilateral exposures.

For the purpose of this study and due to data limitations, we also only estimate the interbank borrowing and lending among these six major banks. We will use the matrix form to present interbank linkages. The following tables present the six major banks’ balance sheets according to their financial reports in 2009 and the estimated interbank loans:

a) Royal Bank of Canada (RBC) Balance Sheet (2009)

Assets (millions of dollars)	654,989	Liabilities (millions of dollars)	654,989
Cash and due from banks	8,353	Deposits (personal)	152,328
deposits with other banks	8,923	Deposits (business and government)	220,772
		Deposits (banks)	25,204
securities (available-for-sale)	46,210		
securities (trading)	140,062	other liabilities	256,685
loans (residential mortgage; Canada)	117,292		
loans (personal and credit card; Canada)	68,778		
loans (small business and wholesale; Canada)	52,451		
assets in the US (retail and wholesale)	37,065		
international Assets (retail and wholesale)	17,589		
other assets	158,266		

Source: Royal Bank of Canada Annual Report 2009,
http://www.rbc.com/investorrelations/ir_annual_report.html#2009

b) Toronto-Dominion Bank (TD) Balance Sheet (2009)

Assets (millions of dollars)	557,219		Liabilities (millions of dollars)	557,219
Cash and non-interest-bearing deposits with banks	2,414		Deposits (personal)	223,228
Interest-bearing deposits with banks	19,103		Deposits (business and government)	126,907
			Deposits (banks)	5,480
securities (available-for-sale)	84,841			
securities (trading)	54,320		other liabilities	201,604
loans (residential mortgage)	65,665			
loans (personal and credit card)	102,509			
loans (business and government)	76,176			
other assets	152,191			

Total assets by geography	557,219
Canada	329,454
US	177,593
Other international	50,172

Source: TD Bank Financial Group Annual Report 2009, <http://www.td.com/ar2009/home/>

c) The Bank of Nova Scotia (Scotiabank) Balance Sheet (2009)

Assets (millions of dollars)	496,516		Liabilities (millions of dollars)	496,516
Cash and non-interest-bearing deposits with banks	3,355		Deposits (personal)	123,762
Interest-bearing deposits with banks	34,343		Deposits (business and government)	203,594
			Deposits (banks)	23,063
securities (available-for-sale)	55,699			
securities (trading)	58,067		other liabilities	146,097
loans (residential mortgage; Canada)	88,766			
loans (personal and credit card; Canada)	49,266			
loans (business and government; Canada)	33,540			
assets in the US (retail and wholesale)	20,548			
international Assets (retail and wholesale)	77,052			
other assets	75,880			

Source: 2009 Scotiabank Annual Report, <http://scotiabank.com/annrep2009/ARWeb/shinfo.html>

d) Bank of Montreal (BMO) Balance Sheet (2009)

Assets (millions of dollars)	388,458		Liabilities (millions of dollars)	388,458
Cash and Cash Equivalents	9,955		Deposits (Individuals)	99,445
Interest Bearing Deposits with Banks	3,340		Deposits (business and government)	113,738
			Deposits (banks)	22,973
securities (available-for-sale)	50,303			
securities (trading)	59,071		other liabilities	152,302
loans (residential mortgage)	45,524			
loans (personal and credit card)	48,398			
loans (business and government)	68,169			
other assets	103,698			

Source: BMO Financial Group 192nd Annual Report 2009, www2.bmo.com/ar2009/downloads/bmo_ar2009.pdf

e) Canadian Imperial Bank of Commerce (CIBC) Balance Sheet (2009)

Assets (millions of dollars)	335,944		Liabilities (millions of dollars)	335,944
Cash and non-interest-bearing deposits with banks	1,812		Deposits (personal)	108,324
Interest-bearing deposits with banks	5,195		Deposits (business and government)	107,209
			Deposits (banks)	7,584
securities (available-for-sale)	40,160			
securities (trading)	15,110		other liabilities	112,827
loans (residential mortgage)	86,152			
loans (personal and credit card)	45,677			
loans (business and government)	37,343			
other assets	104,495			

Source: CIBC Annual Accountability Report 2009, www.cibc.com/ca/pdf/about/aar09-en.pdf

f) National Bank of Canada (NBC) Balance Sheet (2009)

Assets (millions of dollars)	132,138		Liabilities (millions of dollars)	132,138
cash	296		Deposits (personal)	34,609
deposits with financial institutes	1,932		Deposits (business and government)	36,698
			Deposits (deposit-taking institutions)	3,638
securities (available-for-sale)	13,281			
securities (Held-for-trading)	36,952		other liabilities	57,193
loans (residential mortgage)	14,961			
loans (personal and credit card)	18,313			
loans (business and government)	20,003			
other assets	26,400			

Source: National Bank of Canada / 2009 Annual Report,
http://www.nbc.ca/bnc/files/bncpdf/en/2/e_ri_NBC_AR2009.pdf

Then, interbank loan estimated matrix based on assets' assumption is as follows (\$millions):

	RBC	TD	Scotiabank	BMO	CIBC	NBC
RBC	0	6634	11623	4099	2160	687
TD	1222	0	2665	940	495	158
Scotiabank	6497	8092	0	5000	2635	838
BMO	4630	5766	10102	0	1878	597
CIBC	1424	1773	3107	1096	0	184
NBC	649	809	1417	500	263	0

Finally, the summary of the bank assets with the estimated interbank loans used in the simulation is as follows (millions of dollars):

Banks Information	RBC	TD	Scotiabank	BMO	CIBC	NBC
Total Assets	654989	557219	496516	388458	335944	132138
Foreign Assets	54654	227765	97600	86392.78	74713.71	29387.4
Loans	238521	244350	171572	162091	169172	53277
Interbank Loans	14421.76	23074.93	28915.1203	11634.84	7432.243	2463.11
Other Assets	340842.3	56456.88	193463.72	124454.8	81266.61	45689.11
Reserve Yi (= 1%)	6549.89	5572.19	4965.16	3884.58	3359.44	1321.38
Total Liability	654989	557219	496516	388458	335944	132138
Interbank Borrowings	25204	5480	23063	22973	7584	3638
Other Deposits	629785	551739	473453	365485	328360	128500

We use the following notation in Chapter 6 to present the banks:

- B1: RBC
- B2: TD
- B3: Scotiabank
- B4: BMO
- B5: CIBC
- B6: NBC

Appendix 3- A sample of importance sampling program in R

```
#####  
# This program uses importance sampling to estimate P(X > 15)  
# where X ~ exponential(lamda = 1). The direct calculation  
# shows that p(X > 15)= 3*10^(-7). In this program, we sample  
# from exponential(lamda = 1/22) and then adjust for the  
# weights.  
#####  
  
n = 10000  
B1 = 1  
B2 = 22  
k = 0  
  
set.seed()  
  
for(i in 1:n) {  
  x = rexp(1, rate = (1/B2))  
  if(x > 15)  
    k = k + ((1/B1)*exp(-(1/B1)*x))/((1/B2)*exp(-(1/B2)*x))  
}  
  
print("P(X > 15):")  
print((k/n))
```

Result:

```
> print("P(X > 15):")  
[1] "P(X > 15):"  
> print((k/n))  
[1] 2.969115e-007
```