

ABSTRACT

DESHPANDE, AMOGH. Value at Risk Comparison of the CreditRisk⁺ and the 2-Stage CreditRisk⁺ Models for a Special Class of Credit portfolios. (Under the direction of Dr. Min Kang.)

We compare two competing credit risk models, namely the CreditRisk⁺ or CR⁺ and the 2-Stage CR⁺ for a special class of portfolio using the theory of Large deviations; more precisely the Cramer's theorem.

Detailed proof of the Cramer's theorem is also provided in the appendix.

© Copyright 2010 by Amogh Deshpande

All Rights Reserved

Value at Risk Comparison of the CreditRisk⁺ and the 2-Stage CreditRisk⁺ Models for a
Special Class of Credit portfolios

by
Amogh Deshpande

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Master of Science

Operations Research

Raleigh, North Carolina

2010

APPROVED BY:

Dr. Tao Pang

Dr. Julie Ivy

Dr. Min Kang
Chair of Advisory Committee

DEDICATION

To my parents.

BIOGRAPHY

The author was born in Pune, India to Meera and Prafulla. He studied Mechanical Engineering as his bachelor's degree followed by a masters degree in Industrial and Management Engineering all in India.

ACKNOWLEDGEMENTS

I would like to thank my benevolent advisor Dr. Min Kang for her kind support and good will. I wish to also thank my committee members for committing their valuable time by agreeing to be a part of my thesis committee.

TABLE OF CONTENTS

List of Tables	vi
Chapter 1 Background and Introduction	1
Chapter 2 The Results	7
References	17
Appendix	18
Appendix A Cramer's Theorem	19

LIST OF TABLES

Table 2.1	VaR comparison for PI	8
Table 2.2	VaR comparison for PII	9
Table 2.3	VaR comparison for PIII	10

Chapter 1

Background and Introduction

Credit risk valuation has been a very active topic of research for quite sometime due to the immense practical implication it creates for financial institution in the time of economic crisis. Hence it is very important for the credit risk models to model the credit risk as realistically as possible and hopefully in an analytically closed form manner. CreditRisk⁺ or CR⁺ credit risk model was developed in 1997 by Credit Suisse. It quickly became widely popular in industry because it could capture portfolio loss distribution analytically in comparison to the contemporary credit risk models like KMV or CreditMetrics which otherwise used expensive monte carlo simulation techniques.

Since its inception, there has been an extensive literature on CreditRisk⁺. The loss distribution in the original CR⁺ model was computed using a recursion scheme due to Panjer and Williams [9]. Gordy [4] showed that this method was numerically unstable for large portfolios. Further, it is difficult to extend this method to more complex models. The standard CR⁺ model apportions the risk of each individual obligor to different sectors which can be thought of as industry sectors. The sector default rates are assumed to follow independent gamma distributions. The choice of the gamma distribution is motivated by the need for having a sparsely parameterized distribution that is concentrated on positive values, is flexible enough and most importantly gives a closed form expressions for the probability generating function of the loss distribution. This allows for easy computation of the loss distribution even for large portfolios with several sectors without taking recourse to time consuming simulations. Correlations in the default rate of the obligors arise due to their common dependence on one or more of the sector default rates. Giese [3] provided a breakthrough by suggesting a method of computing loss distribution that used recursive computation of exponential and logarithmic polynomials instead of the numerically unstable Panjer recursion. This method in Giese allows us to evaluate the portfolio loss distribution for a wider range of risk factor distributions including the possibility that the risk factors can be modeled as stochastic processes.

Burgisser et al. [1] was the first to introduce correlations among the sector default rates. This was done by adjusting the standard deviation of the portfolio default rate to account for the sector correlations and the carrying out a single sector analysis. It was shown by Giese [3] that the Burgisser model cannot adequately capture concentration risks arising from sectors with large exposures and large factor variance. In Giese [3] correlation is induced among the sectoral default rates via a single scaling variable that follows a gamma distribution. This introduces a *uniform* level of covariance. A recent paper of Deshpande and Iyer [7] addresses some drawbacks in the compound gamma model of Giese [3]. It was noticed that though Giese’s model performs better than the Burgisser et al. model, it distorts concentration risks by enhancing low levels of correlations and suppressing higher levels of correlations. Further, estimation of this uniform covariance from the more complex set of observed sectoral default rate covariance matrix can lead to certain inconsistency.

The approach of Deshpande and Iyer in modeling the correlated defaults between the sectoral default rates is more straightforward. In Giese, the sectors are identified with systematic risk factors that represent forecasts on macro-economic variables. The 2-stage model built by Deshpande and Iyer separates the two notions of industry sectors and macro-economic risk factors. The sectors in our model refer to industry sectors or countries as in the original CR⁺. In the second level, the sector default rates, in turn, are assumed to depend on risk factors that can be thought of as macro economic factors or factors that explain the default rates. We therefore refer to this model as the 2-stage CR⁺ model. We now describe in brief the CR⁺, Giese’s compound gamma model and the 2-stage CR⁺ model.

1.2 Problem formulation

1.2.1 Description of CreditRisk⁺ model

A suitable base unit of currency ΔL is chosen. A generic obligor of a portfolio is denoted by “ A ”. The adjusted exposure E_A of obligor A is replaced by $\nu_A = \lfloor E_A/\Delta L \rfloor$ where $\lfloor \cdot \rfloor$ denotes integer part of real number. Let p_A be average default probability for obligor A . Let N_A denote integer valued default of obligor A . Therefore portfolio loss of obligors is represented by the integer random variable

$$L = \sum_A \nu_A N_A \tag{1.0.1}$$

The risk of each obligor is apportioned among a set of K sectors (industry) by choosing g_k^A such that $\sum_{k=1}^K g_k^A = 1$. In the standard CR⁺ model, the default rate of each sector is represented by non-negative variable γ_k with gamma distribution $\Gamma(\frac{1}{\sigma_k}, \sigma_k)$ that satisfies $E[\gamma_k] = 1 \quad \forall k = 1, \dots, K$ and covariance matrix $Cov(\gamma_k, \gamma_l) = 0$, for all $k, l \in \{1, \dots, K\}$

The CR⁺ framework assumes that the default rates of the obligors depend on the sector default rates via the linear relationship

$$X_A(\gamma) = p_A \sum_{k=1}^K g_k^A \gamma_k \quad (1.0.2)$$

where $X_A(\gamma)$ is the default rate of obligor A conditional on the sector default rates $\gamma = (\gamma_1, \dots, \gamma_K)$. Conditional on γ the default variables N_A are assumed as independent and Poisson distributed with intensity $X_A(\gamma)$. Note that

$$E[X_A(\gamma)] = p_A \sum_{k=1}^K g_k^A E[\gamma_k] = p_A.$$

Thus the above structure preserves expectations regarding default by individual obligors, while enabling correlated movements among default realizations. Under these assumptions the conditional probability generating function (pgf) of the portfolio loss given $\gamma = (\gamma_1, \dots, \gamma_K)$ is calculated as

$$G_\gamma(z) = \exp\left(\sum_{k=1}^K \gamma_k P_k(z)\right). \quad (1.0.3)$$

where

$$P_k(z) = \sum_A g_k^A p_A (z^{\nu_A} - 1). \quad (1.0.4)$$

Default correlations between obligors arise only through their dependence on the common set of sector default rates. We get the unconditional pgf of the loss distribution by averaging $G_\gamma(z)$ over the distribution of the sector default rates.

$$G(z) = E^\gamma[\exp(\sum_{k=1}^K \gamma_k P_k(z))] = M_\gamma(\mathbf{P}(z)) \quad (1.0.5)$$

where $\mathbf{P}(z) = (P_1(z), \dots, P_K(z))$. Since the moment generating function (mgf) of the univariate gamma distribution with mean 1 and variance σ_k is $(1 - \sigma_k t_k)^{-\frac{1}{\sigma_k}}$, the pgf of the loss distribution can be written as

$$G^{CR^+}(z) = \exp\left\{-\sum_{k=1}^K \frac{1}{\sigma_k} \log(1 - \sigma_k P_k(z))\right\} \quad (1.0.6)$$

1.2.2 The compound gamma CR⁺ model

In this model correlations are introduced between variables γ_k via a common scaling factor S as follows: Conditional on S , each γ_k is independently gamma distributed with shape parameter

$$\hat{\alpha}_k(S) = S\alpha_k, \quad \alpha_k > 0,$$

and constant scale parameter β_k . The common scaling factor S is itself modeled as gamma distribution with $\alpha = \frac{1}{\hat{\sigma}^2}$ and $\beta = \hat{\sigma}^2$ such that $E[S] = 1$ and $\text{Var}(S) = \hat{\sigma}^2$. Under these assumptions, it follows that $E[\gamma_k] = \alpha_k\beta_k$, which together $E[\gamma_k] = 1$ yields $\alpha_k = \frac{1}{\beta_k}$. The cross covariances $\sigma_{kl} = \delta_{kl}\beta_k + \hat{\sigma}^2$, where δ_{kl} is Kronecker's delta. The pgf of the compound gamma model is given in Giese [3] by

$$G^{CG}(z) = M_\gamma^{CG}(\mathbf{P}(z)), \tag{1.0.7}$$

where $\mathbf{P}(z) = (P_1(z), \dots, P_K(z))$, and

$$M_\gamma^{CG}(T) = \exp \left\{ -\frac{1}{\hat{\sigma}^2} \log \left[1 + \hat{\sigma}^2 \sum_{k=1}^K \frac{1}{\beta_k} \log(1 - \beta_k t_k) \right] \right\} \tag{1.0.8}$$

with $T = (t_1, \dots, t_K)$. The cross covariances $\hat{\sigma}^2$ is estimated from the observed time series of sectoral default rates as follows:

$$\hat{\sigma}^2 = \frac{\sum_{k \neq l} EL_k V_{kl} EL_l}{\sum_{k \neq l} EL_k EL_l},$$

where V_{kl} is an estimate of the covariance in the default rates among sectors k and l and EL_k is the expected loss of sector k . This factor, in principle can become negative in which case the model will become ill-defined.

1.2.3 2-Stage CR⁺ model

The 2-stage CR⁺ model proposed recently explains default risk at two levels. In the first level as in case of standard CR⁺ model, the risk of each obligor is apportioned to a common set of industry sectors using the parameters g_k^A . In the second level the default rates of the industry sectors $\gamma_k, k = 1, \dots, K$ are assumed to depend linearly on a set of common independent risk factors Y_1, \dots, Y_M i.e.

$$\gamma_k = \sum_{i=1}^M a_{ki} Y_i \tag{1.0.9}$$

where Y_i are independent gamma with mean 1 and variance $\tilde{\sigma}_i$ and $\sum_{i=1}^M a_{ki} = 1$ for all $k = 1, \dots, K$ and $|a_{ki}| < 1$. The variables Y_i are the risk factors that drive correlated changes in the sector default rates. These variables can be obtained from a principal component analysis of macro-economic variables that influence the sector default rates. It is not necessary to model the sector default rates using macro-economic variables. A factor analysis based on the observed default rate correlation matrix would suffice to obtain an estimate of the parameters a_{ki} .

Remark I: Note that γ_k in (1.0.9) loses its identity as a Gamma distributed random variable. γ_k may not even be uncorrelated now. It is merely a construction of sum of i.i.d. gamma random variables Y_i 's. However, in comparing the CR^+ and the 2-Stage CR^+ through large deviations perspective, one can assume that the sum of these Y_i 's manages to induce high variance σ on each γ_k , the same one as that of the variance of original gamma random variable γ_k in the CR^+ model.

Remark II: The portfolio case dealt in Giese [3] is of the form $\sum_A g_k^A = G$ for all $k \in \{1, \dots, K\}$. where G is some positive constant number. In addition to this it is assumed in the portfolio example constructed in Deshpande and Iyer [7] that $\sum_k a_{ki} = 1$ for all $i \in \{1, \dots, M\}$. In this thesis we will therefore analyze a "special" portfolio where $\sum_A g_k^A = G$ and $a_{ki} = c$ where c is some constant $\forall i$ and k such that $\sum_k a_{ki} = A = 1$.

Define $Q_i(z) := \sum_{k=1}^K a_{ki} P_k(z)$. The portfolio loss distribution as obtained by Deshpande and Iyer [7] to be re-derived in chapter 2 is given by

$$G(z) = \exp \left\{ - \sum_{i=1}^M \frac{1}{\tilde{\sigma}_i} \log(1 - \tilde{\sigma}_i Q_i(z)) \right\}. \quad (1.0.10)$$

1.2.4 Connection of Credit risk analysis to large deviations

Credit portfolios are typically large and include exposures of many obligors whose default probabilities are extremely small. Rare but large loss events that occur puts emphasis on obtaining small probability of large losses that are relevant in computing Value at Risk or VaR. We define VaR as follows.

Definition 1 (Value at Risk) For some confidence level $\alpha \in (0, 1)$, the Value at Risk or VaR of the portfolio is given by the smallest number ' l ' such that the probability that the loss exceeds ' l ' is not larger than $1 - \alpha$. Thus mathematically,

$$VaR_\alpha = \inf \{ l \in \mathbb{R} : P(L > l) \leq (1 - \alpha) \}$$

To understand better the occurrence of large loss events, large deviations is widely gaining importance in computing probability of such rare events. See for an introduction Pham [10]. The early work of applying large deviations to credit risk dates back to Dembo et al. [5] who suggest that generically, rare events are exponentially rare in the dimension of the portfolio in other words, the tail of loss distribution of the portfolio decays at the rate of $e^{-\lambda N}$ where N is the number of assets in the portfolio. Same conclusion was drawn upon by recent work of Glasserman et al. [8]. This was observed as unrealistic by Sowers [11]. For example he argued that some standard credit default indices typically consist of 125 names, and hence via the previous results of Dembo et al.[5] or Glasserman [3], the tail will be of the form of e^{-125} . This is of the order of 10^{-55} . This value is so small to be meaningless. In this emerging paradigm of applying large deviations to credit risk analysis we mention the following contribution that we intend to add,

- 1.** we compare two competing credit risk models, viz. the standard CR^+ and its new variant the 2-Stage CR^+ model in the context of VaR for a special portfolio setup.
- 2.** instead of computing the tail decay rate of the loss distribution; in a unique way, we provide large deviations result for the total portfolio default rate of the two models and then comment on large losses and subsequent VaR results.

To the best of our knowledge no such comparison has been done before between any known competing portfolio credit risk models via application of large deviations theory.

As mentioned earlier, in the next chapter we derive the loss distribution for the 2-stage CR^+ model followed by its comparison with its new variant the 2-Stage CR^+ model in terms of Value at Risk or (VaR) using large deviations techniques.

Chapter 2

The Results

We divide this chapter into two parts. The first part discusses numerical credit loss computations for various portfolio cases. In the second part we analyze VaR performance for both the CR⁺ and the 2-stage CR⁺ using Large deviations theory. In the first part we subject both the credit risk models to three types of portfolios and generate VaR values. We next compare them and summarize our strategy that should be adopted while analyzing VaR performance using Large deviations. Throughout this chapter we will be citing the following assumptions.

Assumptions

A1) $\nu := \max\{\nu_A, \dots, \nu_N\}$, then $\nu_i = \nu$ for all $i \in \{A, \dots, N\}$ and $p^* := \max\{p_A, \dots, p_N\}$, then $p_i = p^*$ for all $i \in \{A, \dots, N\}$. This entails a *uniform* portfolio.

A2) γ_k is i.i.d. Gamma distributed with parameters $(\frac{1}{\sigma}, \sigma)$. Namely $\Gamma(\frac{1}{\sigma}, \sigma)$ where $\sigma = \max\{\sigma_1, \dots, \sigma_k\} \forall k \in \{1, \dots, K\}$ and Y_i is i.i.d. Gamma distributed with parameters $(\frac{1}{\tilde{\sigma}}, \tilde{\sigma})$ namely $\Gamma(\frac{1}{\tilde{\sigma}}, \tilde{\sigma})$ where $\tilde{\sigma} = \max\{\tilde{\sigma}_1, \dots, \tilde{\sigma}_i\}$ for all $i \in \{1, \dots, M\}$. This entails a *homogeneous* portfolio.

We first tweak a version of the original portfolio (PO) mentioned in the paper of Deshpande and Iyer [7] by applying assumptions A1 and A2 on the same. This will entail a uniform but highly stressed portfolio characterized by high exposure and high average obligor probability of default in addition to homogeneity of sectoral default and risk factor variances via assumption A2. We name this portfolio as (PI). We will then compare VaR performance for both the credit risk models on this portfolio and tabulate our results in Table(2.1). Next we shall consider an alternative randomized portfolio (PII). In it we shall sample σ_k from a uniform distribution and provide VaR results in Table(2.2). In the third case (PIII) we remove assumption A1, in which case the portfolio is not uniform. We tabulate the VaR results on the same in Table(2.3) keeping the same sampled variances from P(II). Suitable interpretations follow the computations. We

begin with the modified version of the portfolio mentioned in Deshpande and Iyer [7]. All the portfolio cases considered have $\sum_A g_k^A = 3000$

2.1 Numerical computations

Portfolio(PO). The test portfolio is made up of $K = 12$ sectors, each containing 3000 obligors. Obligor in sectors 3-10 belong in equal parts to one of three classes with adjusted exposures $E_1 = 1, E_2 = 2.5$ and $E_3 = 5$ monetary units and respective default probabilities are $p_1 = 0.55\%, p_2 = 0.08\%, p_3 = 0.02\%$. For the three obligor classes in sectors 1,2,11 and 12, we assume the same default rates but twice as large exposures ($E_1 = 2, E_2 = 5$ and $E_3 = 10$.) We consider the sector default rate variances to be sampled from a uniform $U(0, 1)$ distribution. We now apply assumptions A1 and A2 to this portfolio that leads us to portfolio (PI).

Portfolio (PI). The test portfolio is again made up of $K = 12$ sectors, each containing 3000 obligors. Obligor in sectors 1-12 belong in equal parts to one of three classes with adjusted exposures $E_1 = 10, E_2 = 10$ and $E_3 = 10$ monetary units and respective default probabilities are $p_1 = 0.55\%, p_2 = 0.55\%, p_3 = 0.55\%$. We consider the sector default rate variances to be sampled from uniform $U(0, 1)$ distribution. In the case of the 2-stage model, correlation between sector default rates would in principle be the outcome of dependence on common set of risk factors in other words $\gamma_k = 0.5Y_k + 0.5Y_{k+1}$ for $k \in \{1, \dots, 11\}$ while $\gamma_{12} = 0.5Y_{12} + 0.5Y_1$. VaR computations on this uniform and homogeneous portfolio for the three cases lead us to the following results.

Table 2.1: VaR comparison for PI

σ	VaR ₉₀ -CR ⁺	VaR ₉₀ -2 Stage CR ⁺
0.2644	0.6694	0.7112
0.63544	0.7250	0.8029
0.99	0.7667	0.9267

Next, we consider a modified version of (PI) obtained and compute VaR for these three cases of varying values of σ sampled from a uniform distribution.

Portfolio (PII). The test portfolio is again made up of $K = 12$ sectors, each containing 3000 obligors. Obligor in sectors 1-12 belong in equal parts to one of three classes with adjusted

exposures $E_1 = 10, E_2 = 10$ and $E_3 = 10$ monetary units and respective default probabilities are $p_1 = 0.55\%, p_2 = 0.55\%, p_3 = 0.55\%$. We consider the sector default rate variances σ_i for all $i \in \{1, \dots, 12\}$ to be sampled from uniform $U(0.1, 0.2643), U(0.263, 0.6353)$ and $U(0.6353, 0.99)$ distribution. We pick the maximum out of these. In the case of the 2-stage model, correlation between sector default rates would in principle be the outcome of dependence on common set of risk factors in other words $\gamma_k = 0.5Y_k + 0.5Y_{k+1}$ for $k \in \{1, \dots, 11\}$ while $\gamma_{12} = 0.5Y_{12} + 0.5Y_1$. VaR computations on this uniform portfolio for these three cases yield us to the following results.

Table 2.2: VaR comparison for PII

σ	VaR ₉₀ -CR ⁺	VaR ₉₀ -2 Stage CR ⁺
$\sigma_i \sim U(0.1, 0.2643)$	0.6500	0.6833
$\sigma_i \sim U(0.2643, 0.6353)$	0.7083	0.7722
$\sigma_i \sim U(0.6353, 0.99)$	0.7528	0.8778

Finally we consider a modified version of the portfolio (PII) and denote it as (PIII) wherein we relax assumption A1. This is the most general portfolio case that we shall deal with.

Portfolio (PIII). The test portfolio is made up of $K = 12$ sectors, each containing 3000 obligors. Obligor in sectors 3-10 belong in equal parts to one of three classes with adjusted exposures $E_1 = 1, E_2 = 2.5$ and $E_3 = 5$ monetary units and respective default probabilities are $p_1 = 0.55\%, p_2 = 0.08\%, p_3 = 0.02\%$. For the three obligor classes in sectors 1,2,11 and 12, we assume the same default rates but twice as large exposures ($E_1 = 2, E_2 = 5$ and $E_3 = 10$.) We consider the same sector default rate variances σ_i for all $i \in \{1, \dots, 12\}$ that were sampled in (PII). In the case of the 2-stage model, correlation between sector default rates would in principle be the outcome of dependence on common set of risk factors in other words $\gamma_k = 0.5Y_k + 0.5Y_{k+1}$ for $k \in \{1, \dots, 11\}$ while $\gamma_{12} = 0.5Y_{12} + 0.5Y_1$. VaR computations on this uniform portfolio for these three cases yield us to the following results.

Based on the results obtained in Tables (2.1),(2.2) and (2.3), we summarize the computations.

2.2 Numerical Results and Observations

- (PI) is a stress tested version of the original portfolio. We test it for three different values of σ which are 0.2634, 0.6353 and 0.99. In (PII) to make explicit the relationship be-

Table 2.3: VaR comparison for PIII

σ	VaR ₉₀ -CR ⁺	VaR ₉₀ -2 Stage CR ⁺
$\sigma_i \sim U(0.1, 0.2643)$	0.1250	0.1290
$\sigma_i \sim U(0.2643, 0.6353)$	0.1315	0.1419
$\sigma_i \sim U(0.6353, 0.99)$	0.1387	0.1484

tween portfolio VaR and σ , we carefully sample σ_i uniformly between these three variance values considered in (PI) keeping exposures and average default probability high and uniform. In (PIII) we also relax assumption A1 while maintaining same sampling scheme of σ_i as in (PII) and note down VaR values. Thus this portfolio is no longer “stressed”.

2. Results show that as expected stress tested portfolio show highest VaR followed by (PII) and (PIII) in descending order for both the CR⁺ and the 2-stage model.

3. In all these three cases, VaR obtained from 2-stage CR⁺ model is higher than CR⁺ and sectoral default and risk factor variance play a dominant role in loss computations and that all these cases had one common feature that $\sum_{k=1}^K \sigma_k < \sum_{i=1}^M \tilde{\sigma}_i$

We conclude this first part by making the following observations.

1. We observe that the VaR for stress tested case is highest than its non-standard counterpart. Hence for comparison, it makes sense to compare performance of CR⁺ and 2-stage CR⁺ for the stress tested portfolio and extrapolate for any arbitrary case through some arguments to be detailed later supported by encouraging results obtained in in Table(2.3). More rigorously one could argue the same VaR inference as that for the stress tested portfolio using Gärtner-Ellis theorem for the case (PII). This is part of our future work.

2. As observed from Table(2.2) and Table(2.3), the values of only σ and $\tilde{\sigma}$ are responsible for differing VaR results. We will define a random variable say X whose tail decays as a function of risk variance. We will connect its large deviations behaviour with large portfolio losses. If X for CR⁺ decays faster than X for 2-stage CR⁺ then we get higher portfolio loss for the 2-Stage credit risk model. This is the technique we shall utilize in the second part of this chapter. For the sake of completeness, the X in our case is the portfolio default rate which decays exponentially as a function of σ and $\tilde{\sigma}$ for CR⁺ and the 2-stage CR⁺ respectively. We shall argue that it decays slowly for the 2-stage CR⁺ than CR⁺ as $\tilde{\sigma} \geq \sigma$. Hence the 2-stage model will produce more VaR than its primitive i.e. the CR⁺. We elaborate on this further later.

2.3 Further Analysis

We first synopsise the analysis done in attacking this problem. We derieve the portfolio loss distribution in Theorem 1 for the case of the 2-stage model as done in Deshpande and Iyer [7]. We conclude the discussion on the specifics of the 2-stage model by a short note on its Exact risk contributions which more informally is known as $q\%$ portfolio loss quantile.

We now come to attacking the problem of understanding why the 2-stage credit risk model gives equal or higher Value at Risk than the CR^+ for the homogenous portfolio case. One straightforward approach is to determine the tail decay rate of the portfolio loss distributions for both models. The one that decays slowly will produce more VaR. This technique though apparently straightforward is not practical since it is hard if not impossible to determine the exact rate of decay of the portfolio loss distribution in a realistic setup. As mentioned earlier in the concluding remarks of the first part, we counter this problem , instead by working on the large deviation asymptotic of a random variable called the “Portfolio default rate”, for both the CR^+ and the 2-stage CR^+ model. We define the same in Definitions 2 and 3. It is through the result in Lemma 4 that we will connect large value of portfolio default rate to large losses and hence higher VaR. Next we introduce the key theorem in large deviations attributed to Cramer, well-known as the Cramer’s theorem. We detail the large deviation asymptotics of our interest and present result concerning it in Lemma 6. Knowing the portfolio default decay rate from Lemma 6, we now compare the same via Theorem 7. It is through this theorem and Lemma 4 that we can comment on why the 2-stage credit risk model will give us higher VaR than its primitive, the CR^+ model. Through Theorem 8 we extrapolate Value at Risk (VaR) comparison for a non-uniform portfolio. As usual the discussion will end with concluding remarks.

Theorem 1 *The portfolio loss distribution in the case of the 2-Stage CR^+ model is given as*

$$G(z) = \exp \left\{ - \sum_{i=1}^M \frac{1}{\tilde{\sigma}_i} \log(1 - \tilde{\sigma}_i Q_i(z)) \right\}. \quad (2.0.1)$$

Proof. From $G(z)$ computed in (0.1.5), we have

$$\begin{aligned}
G(z) &= E^\gamma[\exp(\sum_{k=1}^K \gamma_k P_k(z))] \\
&= E^Y[\exp(\sum_{k=1}^K (\sum_{i=1}^M a_{ki} Y_i) P_k(z))] \\
&= E^Y[\exp(\sum_{i=1}^M (\sum_{k=1}^K a_{ki} P_k(z)) Y_i)] \\
&= E^Y[\exp(\sum_{i=1}^M Y_i Q_i(z))] = M_Y(Q(z)),
\end{aligned}$$

where $Q(z) = (Q_1(z), \dots, Q_M(z))$ with $Q_i(z) = \sum_{k=1}^K a_{ki} P_k(z)$ ■

Before we state the next important theorem that provides large deviations understanding of our problem, we first need to define some crucial terms, explain our assumptions and give important lemmas that will help us achieve our goal.

As in Giese [3], we also now give the exact formula for the risk contribution for each individual obligor. The risk or the quantile contribution of obligor A , QC_A is defined as the expected individual loss conditional on the portfolio loss being equal to l_q , the $q\%$ portfolio loss quantile, hence

$$QC_A = \nu_A E(N_A | L = l_q) = p_A \nu_A \frac{\sum_{k=1}^K g_k^A D^{(l_q - \nu_A)} G_k(z)}{D^{l_q} G(z)}$$

where the moment generating function $M_\gamma(t)$, $t = (t_1, t_2, \dots, t_k)$, denote $M_{\gamma,k}(t) = \frac{\partial}{\partial t_k} M_\gamma(t)$, $G_k(z) = M_{\gamma,k}(\mathbf{P}(z))$, and D^k is an operator that returns the coefficient of z^k of a power series. Evaluating $G_k(z)$ is quite straightforward and yields,

$$G_k(z) = G(z) \sum_{i=1}^N \frac{a_{ki}}{1 - \sigma_i Q_i(z)}$$

The above representation for $G_k(z)$ allows us to calculate the risk contributions QC_A using the recursive approximation of exponential and logarithm of polynomials as described earlier. The $G_k(z)$ expression for the CR^+ model is given by

$$G_k^{CR^+}(z) = \frac{G^{CR^+}(z)}{1 - \sigma_k P_k(z)}.$$

This completes our discussion for the 2-stage CR^+ model. We now turn our attention to answering the question as to why the 2-stage CR^+ model gives us equal or higher VaR than CR^+ for the simple portfolio setup. Let us start by defining some key terms.

Definition 2 *The portfolio default rate for the CR^+ using equation (1.0.2) is defined as*

$$\begin{aligned} T^{CR^+}(\gamma) &= \sum_A X_A(\gamma) = \sum_A p_A \sum_{k=1}^K g_k^A \gamma_k \\ &= \sum_{k=1}^K \sum_A (g_k^A p_A) \gamma_k \\ &= \sum_{k=1}^K \tilde{g}_k \gamma_k \end{aligned}$$

where $\tilde{g}_k = \sum_A g_k^A p_A = Gp^*$ for all $k = 1, \dots, K$. from Remark II and A1.

Definition 3 *Similarly the portfolio default rate for the 2-stage CR^+ using Definition 2 is defined as*

$$\begin{aligned} T^{2-stageCR^+}(\gamma) &= \sum_{k=1}^K \tilde{g}_k \gamma_k = \sum_{k=1}^K \tilde{g}_k \sum_{i=1}^M a_{ki} Y_i \\ &= \sum_{i=1}^M \left(\sum_{k=1}^K \tilde{g}_k a_{ki} \right) Y_i \\ &= \sum_{i=1}^M g^i Y_i \end{aligned}$$

where $g^i = \sum_{k=1}^K \tilde{g}_k a_{ki}$ for all $i = 1, \dots, M$. From Remark II and Definition 2, $g^i = Gp^*$ for all $i \in \{1, \dots, M\}$.

Our intention is to compare VaR values for the two competing credit risk models for a stress tested portfolio. In that spirit we simplify the problem set up as discussed before by making assumptions A1 and A2.

Lemma 4 *The larger the value of $T^{CR^+}(\gamma)$ or $T^{2-stageCR^+}(\gamma)$ is, the higher the portfolio Value at risk (VaR) is.*

Proof. By assumption A1 on obligor exposure, the portfolio loss is approximated as

$$\begin{aligned} L &= \sum_A \nu_A N_A \\ &= \nu \sum N_A = \nu N \end{aligned}$$

Since the default variables N_A are independent Poisson distributed with intensity rate γ then N is also Poisson distributed with intensity rate $T(\gamma)$. Hence the probability that the portfolio is subject to more than k defaults is given by,

$$P(N > k) = 1 - \sum_{j=0}^k \frac{e^{-T(\gamma)} (T(\gamma))^j}{j!},$$

Thus for fixed number of defaults k , if the likelihood of $T(\gamma)$ taking higher value increases then $P(N > k)$ correspondingly increases. Thus there is high probability of occurrence of large loss and hence leading to higher VaR as seen from its definition. Thus our main goal is to determine how probable $T^{CR^+}(\gamma)$ or $T^{2-stageCR^+}(\gamma)$ takes large values when the credit risk model is subject to same “k” number of defaults or more. ■

Theorem 5 (Cramer’s Theorem) Denote Λ for cumulant moment generating function. Let $D_\Lambda = \{\lambda : \Lambda(\lambda) < \infty\}$. Let X_i be i.i.d. real valued random variables. Let the empirical mean $\hat{S}_n = \frac{1}{n} \sum_{j=1}^n X_j$ be distributed as μ_n . The sequence of measures $\{\mu_n\}$ is said to satisfy the large deviations principle with the convex rate function $\Lambda^*(\cdot)$ which is the Fenchel-Legendre transform of $\Lambda(\lambda)$. Then:

(a) For any closed set $F \subset \mathbb{R}$

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(F) \leq - \inf_{x \in F} \Lambda^*(x).$$

(b) For any open set $G \subset \mathbb{R}$

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(G) \geq - \inf_{x \in G} \Lambda^*(x).$$

where, $\Lambda^*(x) = \sup_{\lambda \in \mathbb{R}} [\lambda x - \Lambda(\lambda)]$

Proof. Refer Appendix-A and for details refer [6]. ■

Now under the assumptions made earlier, by the law of large numbers $\frac{T^{CR^+}(\gamma)}{K}$ converges to Gp^* . Hence in order that the default event $\{T^{CR^+}(\gamma) > Kq\}$ become rare without being trivially

impossible we let $\frac{T(\gamma)}{K}$ approach $q \in (Gp^*, \infty)$. We are interested in large deviation asymptotic of $P_K(q) = P(\frac{T^{CR^+}(\gamma)}{K} > q)$ as $K \nearrow \infty$. Analogously we are also interested in large deviation asymptotic of $P_M(q) = P(\frac{T^{2-StageCR^+}(\gamma)}{M} > q)$ as $M \nearrow \infty$.

Lemma 6 $P_K(q)$ satisfies large deviation principle with rate “K” and good rate function $\Lambda_{CR^+}^*(q) = \frac{1}{\sigma}(\frac{q}{Gp^*} - 1 + \log \frac{Gp^*}{q})$. Similarly, $P_M(q)$ satisfies large deviation principle with rate “M” and good rate function $\Lambda_{2-Stage}^*(q) = \frac{1}{\tilde{\sigma}}(\frac{q}{Gp^*} - 1 + \log \frac{Gp^*}{q})$.

Proof. Let $\bar{\gamma}_k$ have the same distribution as $Gp^*\gamma_k$ i.e. $\bar{\gamma}_k =^d Gp^*\gamma_k$. By Cramer’s theorem, $\lim_{K \rightarrow \infty} \frac{1}{K} \log P_K(q) = \Lambda_{CR^+}^*(q) = \sup_{\lambda > 0} [\lambda q - \Lambda(\lambda)]$ where, $\Lambda(\lambda) = \log E(e^{\lambda \bar{\gamma}_1}) = -\frac{1}{\sigma} \log(1 - \lambda Gp^* \sigma)$. The function inside the supremum is concave and has a unique optimum λ^* . One can easily check that it is $\lambda^* = \frac{q - Gp^*}{Gp^* q \sigma}$. Substituting it in the objective function we get $\Lambda_{CR^+}^*(q) = \frac{1}{\sigma}(\frac{q}{Gp^*} - 1 + \log \frac{Gp^*}{q})$. In a similar way, since Y_i is distributed as $\Gamma(\frac{1}{\tilde{\sigma}}, \tilde{\sigma})$ we will get the rate function as $\Lambda_{2-Stage}^*(q) = \frac{1}{\tilde{\sigma}}(\frac{q}{Gp^*} - 1 + \log \frac{Gp^*}{q})$ as $A=1$. ■

Theorem 7 There exists a constant $p^* \in (0, 1)$ such that for all $q > Gp^*$, $\Lambda_{2-stage}^*(q) \leq \Lambda_{CR^+}^*(q)$.

Proof. Note here $\sum_{i=1}^M a_{ki} = 1$ for all $k = 1, \dots, K$. Therefore from (1.0.9) and the Remark I in chapter 1, $\sigma = \sum_{i=1}^M (a_{ki})^2 \tilde{\sigma}$. Thus we have $\sigma \leq \tilde{\sigma}$. Now the result will follow from Lemma 6. ■

Theorem 8 The order of VaR comparison between the CR^+ and the 2-Stage CR^+ is the same for both the uniform portfolio case and the non-uniform portfolio case.

Proof. It is clear that the “uniform” portfolio shall lead to higher VaR values in comparison to a realistic portfolio as we chose highest exposure and average default probability. The quantum of portfolio loss and hence the VaR is determined by the value of exposure and the average obligor probability of default. The γ_k and Y_i are independent gamma distributed. Hence the VaR values differ between the CR^+ and the 2-Stage CR^+ based only on the differing values of the sectoral default rate variance and risk factor variances since the exposures and the average obligor default probability values are the same across both the two credit risk models. Hence even upon relaxing the “uniformity” assumption A1, the VaR values just get scaled down but the order of VaR comparison between the CR^+ and the 2-Stage CR^+ remains the same. Therefore in brief the only factor that contributes to differing VaR values between the two models are the risk factor variances $\tilde{\sigma}_i$ and σ_k , more correctly the collation of them i.e. $\sum_k \sigma_k$ and $\sum_i \tilde{\sigma}_i$. ■

2.4 Conclusion

In practice $\sigma_k < 1$ and $K \geq M$ as usually $K \geq 50$. $P(\frac{T^{2-stageCR^+}(\gamma)}{M} > q)$ is approximated as $e^{-M\Lambda_{2-stage}^*(q)}$. Similarly $P(\frac{T^{CR^+}(\gamma)}{K} > q)$ is approximated by $e^{-K\Lambda_{CR^+}^*(q)}$. From Theorem 7 we observe that the probability of portfolio default rate taking larger values in the 2-stage credit risk model is equal or higher compared to that of the CR^+ for the homogenous case when $\sigma \leq \tilde{\sigma}$. Hence from Lemma 4, the 2-stage CR^+ may produce higher VaR than the CR^+ . If we remove assumption A1 based on uniformity of the portfolio, then the portfolio under both the credit risk models will suitably be scaled down keeping intact the order of VaR between them. This can be seen numerically from the test portfolio example (PIII) and also through Theorem 8. From (1.0.9) we saw that $\sigma_k \leq \max(\tilde{\sigma}_i)$ and $\sigma_k \geq \min(\tilde{\sigma}_i)$ for all $i \in \{1, \dots, M\}$ and for all $k \in \{1, \dots, K\}$. Hence nothing concrete can be said when one needs to discuss VaR comparison results for the non-homogeneous case. We observed that the inequality $\sum_{k=1}^K \sigma_k < \sum_{i=1}^M \tilde{\sigma}_i$ held true for all the examples we considered in (PI), (PII) and (PIII). We probe this angle further by an application of the Gärtner-Ellis theorem. This is a part of the future work.

Thus in brief the case that we dealt with in this thesis is comparing VaR performance of competing credit risk models on a homogeneous portfolio wherein $\sum_{k=1}^K \sigma_k \leq \sum_{i=1}^M \tilde{\sigma}_i$ for $K \geq M$. A numerical example implemented in [7] supports 2-stage CR^+ model vis-a-vis CR^+ . The 2-stage model computed higher Value at Risk and was shown to be better off in allocating sectoral risk contributions in comparison to CR^+ . However while working on it the authors of [7] observed that the 2-stage CR^+ model also produced higher VaR than the CR^+ . This thesis intends to provide justification for this phenomenon using the theory of large deviations for a special case of a homogeneous portfolio. We did so on a stress tested portfolio and justified this approach through numerical computations. We further provided rigorous analysis for this VaR phenomenon using the Cramer's theorem in large deviations. Consequently in a unique way, we inferred higher tail losses for large Portfolio default rate which was seen to be proxy for the portfolio losses. Thus our analysis was rendered simple and tractable. Enthused by the numerical results for the VaR computations on an non-homogenous portfolio (PIII) , we gave a possible reason behind this result by comparing $\sum_{k=1}^K \sigma_k$ and $\sum_{i=1}^M \tilde{\sigma}_i$. By Theorem 8, we show that the order relation between VaR values of the two credit risk models will remain preserved even if we play around i.e apply or remove assumption A1 since they (the order relation between VaR values) depend only on the collective behaviour of sector default rate variances and risk factor variances.

REFERENCES

- [1] Burgisser P., Kurt A., Wolf M. (1999) Integrating correlations, *Risk*, 12(7):57-60
- [2] Credit Suisse Financial Products (1997) CreditRisk⁺, A credit risk management framework. <http://www.csfb.com/creditrisk>
- [3] Giese G. (2003) Enhancing CreditRisk⁺, *Risk*, 16(4):73-77.
- [4] Gordy M. (2002) Saddlepoint approximation of CreditRisk⁺, *J Banking Finance*, 26(7):1337-1355.
- [5] Dembo, A., Deuschel J-D. and Duffie D. (2004) Large portfolio losses, *Finance and Stochastics*, 8(1):3-16.
- [6] Dembo A., Zeitouni O.(1998) *Large Deviations Techniques and Applications*, Second Edition, Springer-Verlag, New York.
- [7] Deshpande A., Iyer S.K. (2009) Credit risk+ model with general sector correlations, *Central European Journal of Operations Research*, 17(2): 219-228.
- [8] Glasserman P., Kang W. and Shahabuddin P. (2007) Large deviations in multifactor portfolio credit risk, *Mathematical Finance*, 17(3):345-379.
- [9] Panjer H., Williams G. (1992) *Insurance risk models*. Society of Actuaries, Schaumburg.
- [10] Pham H. (2007) Some applications and methods of large deviations in finance and insurance, Lectures from a Bachelier course at Institut Henri Poincar é, Paris.
- [11] Sowers, R. (2010) Large correlation structures in pools of defaults: A Large deviations analysis. Preprint.

APPENDIX

Appendix A

Cramer's Theorem

Cramer's theorem is about Large deviations associated with the empirical mean of i.i.d random variables taking values in a finite set. Specifically, consider the empirical means $\hat{S}_n = \frac{1}{n} \sum_{j=1}^n X_j$ for i.i.d random variables $X_1, \dots, X_n \stackrel{\Delta}{=} \mu \in M_1(\mathbb{R}^d)$. The logarithmic moment generating function associated with the law μ is defined as $\Lambda(\lambda) = \log M(\lambda) = \log E[e^{\langle \lambda, X_1 \rangle}]$ where $\langle \lambda, x \rangle = \sum_{j=1}^d \lambda^j x^j$ is the usual scalar product in \mathbb{R}^d and x^j is the j^{th} coordinate of x . Another common name for $\Lambda(\cdot)$ is the cumulant generating function. In what follows $|x| = \sqrt{\langle x, x \rangle}$ is the usual Euclidian norm. Let μ_n denote the law of \hat{S}_n and $\bar{x} = E[X_1]$. When \bar{x} exists and is finite and $E[|X_1 - \bar{x}|^2] < \infty$ then $\frac{\hat{S}_n}{n} \rightarrow \bar{x}$ in probability. Since $E[|X_1 - \bar{x}|^2] = \frac{1}{n} E[|X_1 - \bar{x}|^2]$ converges to 0. In this situation $\mu_n(F) \rightarrow 0$ for any closed set F . Cramer's theorem characterizes the logarithmic rate of this convergence by the following defined rate function. We consider here the case where $d = 1$.

Definition 1.1 *The Fenchel-Legendre transform of $\Lambda(\lambda)$ is*

$$\Lambda^*(x) := \sup_{\lambda \in \mathbb{R}} [\langle \lambda, x \rangle - \Lambda(\lambda)].$$

Theorem 1.2 *As defined above Λ is the cumulant moment generating function. Let $D_\Lambda = \{\lambda : \Lambda(\lambda) < \infty\}$. Let $X_i \in \mathbb{R}$, then sequence of measures $\{\mu_n\}$ satisfies the LDP with the convex rate function $\Lambda^*(\cdot)$ namely:*

(a) *For any closed set $F \subset \mathbb{R}$*

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(F) \leq - \inf_{x \in F} \Lambda^*(x). \tag{A.0.1}$$

(b) For any open set $G \subset \mathbb{R}$

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(G) \geq - \inf_{x \in G} \Lambda^*(x).$$

The following lemma states the various properties of $\Lambda^*(\cdot)$ and $\Lambda(\cdot)$ that are needed to prove Theorem 1.2. We only mention it here and proceed directly with derivation of Theorem 1.2. For a detailed proof on Lemma 1.3 refer Dembo and Zeitouni [6].

Lemma 1.3

(a) Λ is a convex function and Λ^* is also a convex rate function.

(b) If $D_\Lambda = \{0\}$, then Λ^* is identically zero. If $\Lambda(\lambda) < \infty$ for some $\lambda > 0$ then $\bar{x} < 0$ (possibly $\bar{x} = -\infty$) for all $x \geq \bar{x}$,

$$\Lambda^*(x) = \sup_{\lambda \geq 0} [\lambda x - \Lambda(\lambda)] \tag{A.0.2}$$

is for $x > \bar{x}$, a nondecreasing function. Similarly if $\Lambda(\lambda) < \infty$ for some $\lambda < 0$ then $\bar{x} > -\infty$ (possibly $\bar{x} = \infty$) and for all $x \leq \bar{x}$

$$\Lambda^*(x) = \sup_{\lambda \leq 0} [\lambda x - \Lambda(\lambda)] \tag{A.0.3}$$

is for $x < \bar{x}$, a nonincreasing function when $\bar{x} < \infty$, $\Lambda^*(\bar{x}) = 0$ and always

$$\inf_{x \in \mathbb{R}} \Lambda^*(x) = 0 \tag{A.0.4}$$

(c) $\Lambda(\cdot)$ is differentiable in interior of \mathcal{D}_* with

$$\Lambda'(\eta) = \frac{E[X_1 e^{\eta X_1}]}{M(\eta)} \tag{A.0.5}$$

and

$$\Lambda'(\eta) = y \Rightarrow \Lambda^*(y) = \eta y - \Lambda(\eta) \tag{A.0.6}$$

Proof of Theorem 1.2:

(a) Let F be a non-empty closed set. Note that (A.0.1) will hold trivially when $I_F = \inf_{x \in F} \Lambda^*(x) = 0$. Assume $I_F > 0$. It follows from part (b) of Lemma 1.3 that \bar{x} exists, possibly as an extended

real number. For all x and every $\lambda \geq 0$ an application of Chebycheff's inequality,

$$\begin{aligned} \mu_n([x, \infty)) &= E[1_{\hat{S}_n - x \geq 0}] \leq E[e^{n\lambda(\hat{S}_n - x)}] \\ &= e^{-n\lambda x} \prod_{i=1}^n E[e^{\lambda X_i}] \\ &= e^{-n\lambda x} (E[e^{\lambda \sum_{i=1}^n X_i}]) \end{aligned}$$

Since X_i 's are i.i.d

$$\begin{aligned} &= e^{-n\lambda x} E[e^{n\lambda X_1}] \\ &= e^{-n\lambda x} e^{n\Lambda(\lambda)} \\ &= e^{-n[\lambda x - \Lambda(\lambda)]} \end{aligned} \tag{A.0.7}$$

Therefore, if $\bar{x} < \infty$ then by (A.0.2) for every $x > \bar{x}$

$$\mu_n([x, \infty)) \leq e^{-n\Lambda^*(x)} \tag{A.0.8}$$

By the similar arguments if $\bar{x} > -\infty$ and $x < \bar{x}$, then

$$\mu_n((-\infty, x]) \leq e^{-n\Lambda^*(x)} \tag{A.0.9}$$

Now first consider the case of \bar{x} is finite. Then $\Lambda^*(\bar{x}) = 0$ and because by assumption $I_F > 0$, \bar{x} must be contained in the open set F^c . Let (x_-, x_+) be the union of all the open intervals $(a, b) \in F^c$ that contain \bar{x} . Note that $x_- < x_+$ and that either x_- or x_+ must be finite since F is non-empty. If $x_- < \infty$ then $x_- \in F$ and consequently $\Lambda^*(x_-) \geq I_F$. Likewise $\Lambda^*(x_+) \geq I_F$ whenever x_+ is finite. Applying (A.0.8) for $x = x_+$ and (A.0.9) for $x = x_-$, the union of events bound ensures that $\mu_n(F) \leq \mu_n((-\infty, x_-]) + \mu_n([x_+, \infty)) \leq 2e^{-nI_F}$ and the upper bound follows when the normalized logarithmic limit as $n \rightarrow \infty$ is considered.

Suppose now that $\bar{x} = -\infty$. Then, since Λ^* is nondecreasing, it follows from (A.0.4) that $\lim_{x \rightarrow -\infty} \Lambda^*(x) = 0$ and hence $x_+ = \inf\{x : x \in F\}$ is finite for otherwise $I_F = 0$. Since F is a closed set, $x_+ \in F$ and consequently $\Lambda^*(x_+) \geq I_F$. Moreover $F \subset [x_+, \infty)$ and therefore, the Large deviations upper bound follows by applying (A.0.8) for $x = x_+$.

The case of $\bar{x} = \infty$ is handled analogously.

(b) We prove next that for every $\delta > 0$ and every marginal law $\mu \in M_1(\mathbb{R})$,

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(-\delta, \delta) \geq \inf_{\lambda \in \mathbb{R}} \Lambda(\lambda) \geq -\Lambda^*(0) \tag{A.0.10}$$

Since the transformation $Y = X - x$ results with $\wedge_Y(\lambda) = \Lambda(\lambda) - \lambda x$, and hence with $\Lambda_Y^*(\cdot) = \Lambda(\cdot + x)$, it follows from the preceding inequality that for every x and every $\delta > 0$

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n((x - \delta, x + \delta)) \geq -\Lambda^*(x) \quad (\text{A.0.11})$$

For every open set G , any $x \in G$ and any $\delta > 0$ small enough, $(x - \delta, x + \delta) \subset G$. Thus the large deviation lower bound follows from (A.0.11).

Turning to the proof of the key inequality (A.0.10), first suppose that $\mu((-\infty, 0)) > 0, \mu((0, \infty)) > 0$ and that μ is supported on a bounded subset of \mathbb{R} . By the former assumption $\Lambda(\lambda) \rightarrow \infty$ as $|\lambda| \rightarrow \infty$ and by the latter assumption, $\Lambda(\cdot)$ is finite everywhere. Accordingly, $\Lambda(\cdot)$ is a continuous, differentiable function (see Lemma 1.3 (c)) and hence $\exists \eta < \infty \ni \Lambda(\eta) = \inf_{\lambda \in \mathbb{R}} \Lambda(\lambda)$ and $\Lambda'(\eta) = 0$. Define a new probability measure $\tilde{\mu}$ in terms of μ via $\frac{d\tilde{\mu}}{d\mu}(x) = e^{\eta x - \Lambda(\eta)}$ and observe that $\tilde{\mu}$ is a probability measure because

$$\int_{\mathbb{R}} d\tilde{\mu} = \frac{1}{M(\eta)} \int_{\mathbb{R}} e^{\eta x} d\mu = 1$$

Let $\tilde{\mu}_n$ be the law governing \hat{S}_n when X_i are i.i.d random variables of law $\tilde{\mu}$. Note that for every $\epsilon > 0$

$$\begin{aligned} \mu_n((-\epsilon, \epsilon)) &= \int_{|\sum_{i=1}^n x_i| < n\epsilon} \mu(dx_1) \dots \mu(dx_n) \\ &\geq e^{-n\epsilon|\eta|} \int_{|\sum_{i=1}^n x_i| < n\epsilon} \exp(\eta \sum_{i=1}^n x_i) \mu(dx_1) \dots \mu(dx_n) \\ &= e^{-n\epsilon|\eta|} e^{n\Lambda(\eta)} \tilde{\mu}_n((-\epsilon, \epsilon)) \end{aligned} \quad (\text{A.0.12})$$

By (A.0.5) and the choice of η

$$E_{\tilde{\mu}}[X_1] = \frac{1}{M(\eta)} \int_{\mathbb{R}} x e^{\eta x} d\mu = \Lambda'(\eta) = 0$$

Hence, by the Law of Large Numbers,

$$\lim_{n \rightarrow \infty} \tilde{\mu}_n((-\epsilon, \epsilon)) = 1 \quad (\text{A.0.13})$$

It follows now from (A.0.12) that for every $0 < \epsilon < \delta$,

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n((-\delta, \delta)) \geq \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n((-\epsilon, \epsilon)) \geq \Lambda(\eta) - \epsilon|\eta|$$

and (A.0.10) follows by considering the limit as $\epsilon \rightarrow 0$. Suppose now that μ is of unbounded support, while both $\mu((-\infty, 0)) > 0$ and $\mu((0, \infty)) > 0$ and let

$$\wedge_M(\lambda) = \log \int_{-M}^M e^{\lambda x} d\mu$$

Let ν denote the law of X_1 conditioned on $\{|X_1| \leq M\}$ and let ν_n be the law of \hat{S}_n conditioned on $\{|X_i| \leq M, i = 1, \dots, n\}$. Then $\forall n$ and every $\delta > 0$,

$$\mu_n(\delta, \delta) \geq \nu_n((-\delta, \delta))\mu([-M, M])^n$$

Observe that by the preceding proof (A.0.10) holds for ν_n .

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n((-\delta, \delta)) \geq \log \mu([-M, M]) + \liminf_{n \rightarrow \infty} \frac{1}{n} \log \nu_n((-\delta, \delta)) \geq \inf_{\lambda \in \mathbb{R}} \wedge_M(\lambda)$$

with $I_M = -\inf_{\lambda \in \mathbb{R}} \wedge_M(\lambda)$ and $I^* = \limsup_{M \rightarrow \infty} I_M$, it follows that

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n((-\delta, \delta)) \geq -I^* \tag{A.0.14}$$

Note that $\wedge_M(\cdot)$ is non-decreasing in M , and thus so is $-I_M$. Moreover, $-I_M \leq \wedge_M(0) \leq \Lambda(0) = 0$ and hence $-I^* \leq 0$. Now, since $-I_M$ is finite $\forall M$ large enough, $-I^* > -\infty$. Therefore the level sets $\{\lambda : \wedge_M(\lambda) \leq -I^*\}$ are non-empty, compact sets (i.e bounded because $I^* < \infty$ and closed because I^* is lower semicontinuous) that are nested with respect to M , and hence \exists at least one point, denoted by λ_0 , in their intersection. Since $\wedge_M(\lambda)$ is monotone increasing with M one can by Lebesgue's Monotone convergence theorem conclude that, $\Lambda(\lambda_0) = \lim_{M \rightarrow \infty} \wedge_M(\lambda_0) \leq -I^*$, and consequently from bound (A.0.14) yields (A.0.10) holds, now for μ of unbounded support.

The proof of (A.0.10) for an arbitrary probability law μ is completed by observing that if either $\mu((-\infty, 0)) = 0$ or $\mu((0, \infty)) = 0$, from the definition, $\Lambda(\cdot)$ is a monotone function with $\inf_{\lambda \in \mathbb{R}} \Lambda(\lambda) = \log \mu(\{0\})$. Hence, in this case since X_i 's are i.i.d., (A.0.10) follows as, $\mu_n((-\delta, \delta)) \geq \mu_n(\{0\}) = \mu(\{0\})^n$. This completes the proof of the theorem.