2011

Essays on risk and volatility

Junyue Xu
Louisiana State University and Agricultural and Mechanical College, junyuex@hotmail.com

ESSAYS ON RISK AND VOLATILITY

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Economics

by

Junyue Xu
B.S., Fudan University, Shanghai, China, 2004
M.S., Louisiana State University, 2008
December 2011

Acknowledgments

I want to thank my parents who have made great sacrifices to make sure I can receive the education I want. It is them who helped me to understand the value of knowledge and encouraged me to pursue my dreams. They were also the first ones who stimulated my research interest in financial economics. Throughout the years, they have provided me with unconditional love and support. Words cannot express how grateful I am.

I am also deeply indebted to my advisors Dr. Eric Hillebrand and Dr. Ambar Sengupta. It has been a most rewarding experience for me to work with them and to learn from them. This dissertation would not have been possible without their guidance and inspirations.

I would like to express my gratitude to Dr. Carter Hill and Dr. Douglas McMillin. They have offered me constant support and many invaluable advices throughout the years. It is a pleasure also to thank the Department of Economics at Louisiana State University for providing me with a pleasant working environment.

Table of Contents

List of Tables

List of Figures

Abstract

In this work, we begin with an investigation into the temporal correlation in default risk. We first establish a link between the dynamics of house price changes and the dynamics of default rates in the Gaussian copula framework by specifying a time series model for a common risk factor. We show that the serial correlation propagates from the common risk factor to default rates. In the second essay, we specify a model where the default correlation is stochastic. We find the distribution of expected value of cash flows received by securitized investment vehicles is distorted by the dynamics of default correlation. The third essay provides an empirical study on variance risk premium, which is defined as the difference between implied variance and ex post realized variance. We show that an individual stock's variance risk premium and its two components can be used to predict future equity premium. In the fourth essay, we derive asymptotic properties of the quasi maximum likelihood estimator of smooth transition regressions. We show that the estimator converges at the usual $\sqrt{T}$-rate and has an asymptotically normal distribution.

Introduction

This dissertation includes four essays on credit risk and volatility. The first two parts examine the dynamics of default correlation and its implication for risk management. The third part provides an empirical study using variance risk premia to predict equity excess returns. The fourth part focuses on smooth transition regressions, whose application can be found in credit risk and stochastic volatility models.

The subprime crisis in 2007 started the worst global economic crisis since the Great Depression. The trigger of the crisis was a large number of defaults in subprime mortgages that arrived in a serially correlated manner. The first part of this dissertation examines the stochastic properties of the default behavior of mortgages and that of cash flows received by mortgage backed securities (MBS). We introduce the concept of vintage correlation and establish a link between the dynamics of house price changes and the dynamics of default rates in the Gaussian copula framework by specifying a time series model for a common risk factor. We show analytically and in simulations that serial correlation propagates from the common risk factor to default rates. We simulate cash flows received by MBS, which are securitized from pools of mortgages using a waterfall structure. We find that subsequent vintages of these securities inherit temporal correlation from the common risk factor.

In the second essay, we analyze the distribution of the expected value of cash flows in securitized investment vehicles implied by a model in which default correlation is stochastic. We impose different structures on the dynamics of default correlation including a regime-switching model and a logistic transition model. We find that the dynamics of default correlation distort the distribution of tranche prices. The distortion is affected by the sensitivity of tranche prices to change in default correlation as well as the smoothness of the transition in default correlation.

The third essay provides an empirical study on the variance risk premia' predictive power for equity excess returns. We construct variance risk premium for individual stocks as the difference between implied variance and ex post realized variance. We find that this variable has good predictive power for an individual stock's equity premium. We then decompose variance risk premia into two parts by using a Log-Linear Realized GARCH model. We show that this decomposition significantly increases the return predictive power of variance risk premia.

An important innovation in the second essay of this dissertation is the usage of a logistic transition function in modeling the dynamics of default correlation. The logistic transition function belongs to the group of smooth transition functions, which have been widely used in time series applications. In the fourth essay, we derive the asymptotic properties of the quasi maximum likelihood estimator of logistic transition regressions when time is the transition variable. The consistency of the estimator and its asymptotic distribution are examined. It is shown that the estimator converges at the usual $\sqrt{T}$-rate and has an asymptotically normal distribution.

Chapter 2. Temporal Correlation of Defaults in Subprime Securitization

2.1 Introduction

From its beginning in the summer of 2007, the subprime crisis has plunged the world into one of the worst recessions in history. At the center of the crisis is the subprime mortgage market, where lenders provide mortgages to borrowers with poor credit standing. During the crisis, subprime mortgages created at different times have defaulted one after another. The default arrivals of these mortgages were serially correlated. Figure 2.1, lower panel, shows the time series of serious delinquency rates of subprime mortgages from 2002 to 2009.[1] This series obviously displays very high serial correlation. Defaults of subprime mortgages are closely connected to house price fluctuations, as suggested, among others, by Gorton (2008).[2] Most subprime mortgages are Adjustable-Rate Mortgages (ARM). This means that the interest rate on a subprime mortgage is fixed at a relatively low level for a "teaser" period, usually two to three years, after which it increases substantially. Gorton (2008) points out that the interest rate usually resets to such a high level that it "essentially forces" a mortgage borrower to refinance or default after the teaser period. Therefore, whether the mortgage defaults or not is largely determined by the borrower's access to refinancing. At the end of the teaser period, if the value of the house is much greater than the outstanding principal of the loan, the borrower is likely to be approved for a new loan since the house serves as collateral. On the other hand, if the value of the house is less than the outstanding principal of the loan, the borrower is unlikely to be able to refinance and has to default.

Gorton's view is supported by data. Figure 2.1 displays two-year changes in the Case-Shiller index (upper panel) and subprime ARM serious delinquency rates (lower panel) using quarterly data from 2002 to 2009. It is clear that from 2002 to 2006, subprime delinquency rates declined as home prices climbed steadily. The delinquency series reached its trough around the same time the home price peaked. When the house price index started to drop in 2006, delinquency rates began to increase significantly, which triggered the subprime crisis.

Therefore, the hypothesis that this paper examines is that the dynamics of defaults are inherited from the dynamics of house prices. The aim of this paper is to formalize this relationship using the industry-standard framework of a Gaussian copula, which was routinely used to price derivatives constructed from subprime mortgages. To this end, we introduce the notion of *vintage correlation*, which captures the correlation of default rates in mortgage pools issued at different times.

[1] By definition of the Mortgage Banker Association, seriously delinquent mortgages refer to mortgages that have either been delinquent for more than 90 days or are in the process of foreclosure.

[2] For example, see also Bajari, Chu, and Park (2008), Daglish (2009), Hayre, Saraf, Young, and Chen (2008).

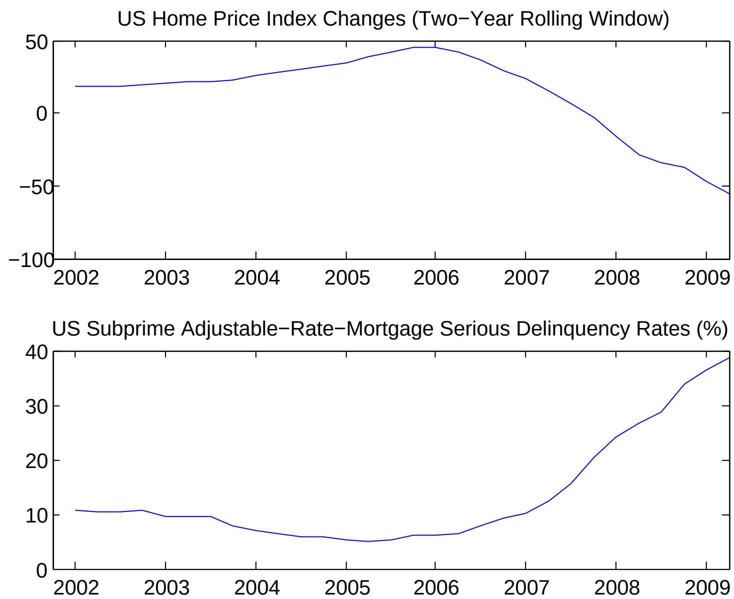

Figure 2.1: **Two-Year Changes in U.S. House Price and Subprime ARM Serious Delinquency Rates**

"U.S. home price two-year rolling changes" are two-year overlapping changes in the S&P Case-Shiller U.S. National Home Price index. "Subprime ARM Serious Delinquency Rates" are obtained from the Mortgage Banker Association. Both series cover the first quarter in 2002 to the second quarter in 2009.

Under certain assumptions, vintage correlation is the same as serial correlation. After showing that changes in a housing index can be regarded as a common risk factor of individual subprime mortgages, we specify a time series model for the common risk factor in the Gaussian copula framework. We show analytically and in simulations that the serial correlation of the common risk factor introduces vintage correlation into default rates of pools of subprime mortgages of subsequent vintages. In this sense, serial correlation propagates from the common risk factor to default rates. In simulations of the price behavior of Mortgage-Backed Securities (MBS) over different cohorts, we find that the price of MBS also exhibits vintage correlation, which is inherited from the common risk factor of individual mortgages.

The main point of this paper is to provide a formal examination of one of the important causes of the current crisis.[3] Vintage correlation in default rates and MBS prices also has implications for asset pricing. To price some derivatives, for example forward starting Collateralized Debt Obligations (CDO), it is necessary to predict default rates of credit assets created at some future time. Knowing the serial correlation of default probabilities can improve the quality of prediction. For risk management in general, some credit asset portfolios may consist of credit derivatives of different cohorts. Vintage correlation of credit asset performance affects these portfolios' risks. For instance, suppose there is a portfolio consisting of two subsequent vintages of the same-type MBS. If the vintage correlation of the MBS price is close to one, for example, the payoff of the

[3]For different perspectives on the causes and effects of the subprime crisis, see also Caballero and Krishnamurthy (2009), Crouhy, Jarrow, and Turnbull (2008), Figlewski (2009), Gorton (2009), Murphy (2008), Reinhart and Rogoff (2008), and Reinhart and Rogoff (2009).

portfolio has a variance almost twice as big as if there were no vintage correlation.

The outline of the paper is as follows. In Section 2.2, we introduce the concept of vintage correlation and give some examples to provide intuition. We then briefly describe the Gaussian copula model. We show that changes in a house price index can be seen as a common risk factor in the copula framework. Section 2.3 contains the main analytical results. It shows the link between the serial correlation of the common risk factor and vintage correlation in default rates. Section 2.4 explores this link in two sets of simulations: First, a series of mortgage pools is simulated to confirm our analytical results. Second, a waterfall structure is simulated to study the propagation of serial correlation in MBS. In Section 2.5 we summarize the main conclusions.

2.2 Modeling Temporal Correlation in Subprime Securitization

In this section, we introduce the concept of vintage correlation and give some examples to provide intuition. We outline the Gaussian copula model. We show that changes in a house price index can be seen as a common risk factor in the copula framework.

DEFINITION 1 (Vintage Correlation). *Suppose we have a pool of mortgages created at each time* $v = 1, 2, \cdots, V$. *Denote the default rates of each vintage observed at a fixed time* $T > V$ *as* $p_1, p_2, \cdots, p_V$, *respectively. We define* vintage correlation $\phi_j := Corr(p_1, p_j)$ *for* $j = 2, 3, \cdots, V$ *as the default correlation between the* $j - th$ *vintage and the first vintage.*

As an example of vintage correlation, consider wines of different vintages. Suppose there are several wine producers that have produced wines of ten vintages from 2011 to 2020. The wines are packaged according to vintages and producers, that is, one box contains one vintage by one producer. In the year 2022, all boxes are opened and the percentage of wines that have gone bad is obtained for each box. Consider the correlation of fractions of bad wines between the first vintage and subsequent vintages. This correlation is what we call vintage correlation.

The definition of vintage correlation can be extended easily to the case where the base vintage is not the first vintage but any one of the other vintages. Obviously, vintage correlation is very similar to serial correlation. There are two main differences. First, the consideration is at a specific time in the future. Second, in calculating the correlation between any two vintages, the expected values are averages over the cross-section. That is, in the wine example, expected values are averages over producers. In mortgage pools, they are averages over different mortgage pools. Only if we assume the same stochastic structure for the cross-section and for the time series of default rates, are vintage correlation and serial correlation equivalent. We do not have to make this assumption to obtain our main results. Making this assumption, however, does not invalidate any of the results either. Therefore, we use the terms “vintage correlation” and “serial correlation” interchangeably in our paper.

To model vintage correlation in subprime securitization, we use the copula approach. The Gaussian copula approach is widely used in industry to model default correlation across names. A copula is a function that takes the marginal distribution functions of a set of variables as arguments and returns the joint distribution of the variables. Thus, the copula approach provides a general way to link univariate marginal distribution functions to their multivariate distribution function.

This feature makes it very useful for modeling multivariate correlations. Frees and Valdez (1998) explain in detail how to specify a copula, and how to simulate a multivariate distribution once its copula form is known.

The credit industry standard copula model was introduced by Li (2000) and is called the default-time (or survival-time) Gaussian copula. This model is applicable to all types of CDO, MBS, and almost all other credit derivatives that are derived from multiple assets with credit risk. The idea behind Li's model is that each credit asset has a default time (or survival time), after which the mortgage defaults. Instead of modeling the correlation between default events of mortgages, Li proposes a copula approach to capture the joint distribution of default times. A copula in this case takes the marginal distribution of default times and returns their joint distribution.

The literature on credit risk pricing with copulas and other models has grown substantially in recent years and an exhaustive review is beyond the scope of his paper. Bluhm, Overbeck, and Wagner (2002), Schönbucher (2003), Duffie and Singleton (2003), and Lando (2004) are standard monographs. Some modifications of the standard Gaussian copula model are discussed. For example, Servigny and Renault (2002) and Das, Freed, Geng, and Kapadia (2006) provide empirical evidence that asset correlation may be stochastic. Andersen and Sidenius (2005), Hull, Predescu, and White (2009), and Berd, Engle, and Voronov (2007) allow default correlation to vary over time. Copula models using distribution functions other than Gaussians have also been suggested. For example, Andersen, Sidenius, and Basu (2003) and Frey and McNeil (2003) consider the student t-copula. Schönbucher and Schubert (2001) and Laurent and Gregory (2005) discuss the Clayton copula. The Marshall-Olkin copula has been considered by Lindskog and McNeil (2003) and Giesecke (2003). A recent study by Beare (2010) explicitly addresses the temporal correlation problem in copulas.

There are approaches to model default correlation other than default-time copulas. One method relies on the so-called structural model, which goes back to Merton's (1974) work on pricing corporate debt. An essential point of the structural model is that it links the default event to some observable economic variables. Hull and White (2001) extend the model to a multi-issuer scenario, which can be applied to the pricing of corporate debt CDO. It is assumed that a firm defaults if its credit index hits a certain barrier. Therefore, correlation between credit indices determines the correlation of default events. The advantage of a structural model is that it gives economic meaning to underlying variables. Other approaches to CDO pricing are found, for example, in Graziano and Rogers (2009) and in Sidenius, Piterbarg, and Andersen (2008). Burtschell, Gregory, and Laurent (2009) provide a comparison of common CDO pricing models.

In this paper, we adopt Li's (2000) default time copula approach and extend it by adding a time series model for a common risk factor. Each mortgage i of vintage v has a default time $\tau_{v,i}$, which is a random variable representing the time at which the mortgage defaults. If the mortgage never defaults, this value is infinity. If we assume that the distribution of $\tau_{v,i}$ is the same across all mortgages of vintage v, we have

$$F_v(s) = \mathbb{P}[\tau_{v,i} < s],\ \forall i = 1, 2, ..., N, \tag{2.1}$$

where the index i denotes individual mortgages and the index v denotes vintages. We assume that F_v is continuous and strictly increasing. Given this information, for each vintage v the Gaussian copula approach provides a way to obtain the joint distribution of the $\tau_{v,i}$ across i. Generally, a copula is a joint distribution function

$$C\left(u_1, u_2, ..., u_N\right) = \mathbb{P}\left(U_1 \leq u_1, U_2 \leq u_2, ..., U_N \leq u_N\right),$$

where $u_1, u_2, ..., u_N$ are N uniformly distributed random variables that may be correlated. It can be easily verified that the function

$$C\left[F_1(x_1), F_2(x_2), ..., F_N(x_N)\right] = G(x_1, x_2, ..., x_N) \tag{2.2}$$

is a multivariate distribution function with marginal distribution functions $F_1(x_1)$, $F_2(x_2)$,...,$F_N(x_N)$. Sklar (1959) proved the converse. He showed that for an arbitrary multivariate distribution function $G(x_1, x_2, ..., x_N)$ with continuous marginal distributions functions $F_1(x_1)$, $F_2(x_2)$,...,$F_N(x_N)$, there exists a unique C such that equation (2.2) holds. Therefore, in the case of default times, there is a C_v for each vintage v such that

$$C_v\left[F_v(\tau_{v,1}), F_v(\tau_{v,2}), ..., F_v(\tau_{v,N})\right] = G_v(\tau_{v,1}, \tau_{v,2}, ..., \tau_{v,N}). \tag{2.3}$$

The joint distribution function G_v on the right-hand side of equation (2.3) is the object we want to obtain. Since we assume F_v to be continuous and strictly increasing, we can find a standard Gaussian random variable $X_{v,i}$ such that

$$\Phi(X_{v,i}) = F_v(\tau_{v,i}) \quad \forall v = 1, 2, ..., V;\ i = 1, 2, ..., N, \tag{2.4}$$

or equivalently,

$$\tau_{v,i} = F_v^{-1}(\Phi(X_{v,i})) \quad \forall v = 1, 2, ..., V;\ i = 1, 2, ..., N, \tag{2.5}$$

where Φ is the standard normal distribution function. To see that this is correct, observe that

$$\begin{aligned}
\mathbb{P}[\tau_{v,i} \leq s] &= \mathbb{P}\left[\Phi(X_{v,i}) \leq F_v(s)\right] \\
&= \mathbb{P}\left[X_{v,i} \leq \Phi^{-1}\left(F_v(s)\right)\right] \\
&= \Phi\left[\Phi^{-1}\left(F_v(s)\right)\right] \\
&= F_v(s).
\end{aligned}$$

Substituting equation (2.4) into the left-hand side of equation (2.3), we have

$$C_v\left[\Phi(X_{v,1}), \Phi(X_{v,2}), ..., \Phi(X_{v,N})\right] = G_v(\tau_{v,1}, \tau_{v,2}, ..., \tau_{v,N}). \tag{2.6}$$

Since $\Phi(\cdot)$ is the marginal distribution function for all $X_{v,i}$, the left-hand side of equation (2.6) is equal to the joint distribution of $X_{v,i}$. The Gaussian copula approach assumes that this joint distribution has a multivariate normal distribution function Φ_N,

$$G_v(\tau_{v,1}, \tau_{v,2}, ..., \tau_{v,N}) = \Phi_N\left(X_{v,1}, X_{v,2}, ..., X_{v,N}\right). \tag{2.7}$$

Thus the joint distribution function of default times $\tau_{v,i}$ is obtained once the correlation matrix of the $X_{v,i}$ is known. A standard simplification in practice is to assume that the pairwise correlations between different $X_{v,i}$ are the same across i. Suppose that the value of this correlation is ρ_v for each vintage v. Consider the following definition

$$X_{v,i} := \sqrt{\rho_v} Z_v + \sqrt{1-\rho_v}\varepsilon_i \quad \forall i = 1, 2, \ldots, N;\ v = 1, 2, \ldots, V, \tag{2.8}$$

where $\varepsilon_{v,i}$ are i.i.d. standard Gaussian random variables and Z_v is a Gaussian random variable independent of the $\varepsilon_{v,i}$. It can be shown easily that in each vintage v, the variables $X_{v,i}$ defined in this way have the exact joint distribution function Φ_N.

Using the information above, for each vintage v, the Gaussian copula approach obtains the joint distribution function G_v for default times as follows. First, N Gaussian random variables $X_{v,i}$ are generated according to equation (2.8). Second, from equation (2.5) a set of N default times $\tau_{v,i}$ is obtained, which has the desired joint distribution function G_v. In equation (2.8), the common factor Z_v can be viewed as a latent variable that captures the default risk in the economy, and ε_i is the idiosyncratic risk for each mortgage. The variable $X_{v,i}$ can be viewed as a state variable for each mortgage. The parameter ρ_v is the correlation between any two individual state variables. It is obvious that the higher the value of ρ_v, the greater the correlation between the default times of different mortgages.

In pricing derivatives created on subprime mortgages, Monte Carlo simulations are employed to study the default behavior of mortgages by the method described above. In each simulation, the default times for all mortgages are generated with the joint distribution G_v. Mortgage i is said to default before time T, if its simulated default time $\tau_{v,i}$ is less than T. The value of ρ_v can then be calibrated to market data. The market-implied ρ_v may vary over time. Indeed, this is supported by empirical evidence provided by Servigny and Renault (2002) and Das, Freed, Geng, and Kapadia (2006). This time dependence may capture dynamic correlation between default events not explicitly captured in the default time copula approach. One way to explicitly model the dynamics of defaults is to specify a stochastic process for default correlation. This approach is called stochastic correlation (see for example Andersen and Sidenius (2005) and Hull, Predescu, and White (2009)). Das, Freed, Geng, and Kapadia (2006) propose a model where the default intensity is determined by the state of the economy, which follows a Markov process.

In this paper, we propose a time series model for the common risk factor Z_v in the copula framework and show that its serial correlation propagates to the default rates. To illustrate the intuition behind our approach, we first give a structural interpretation for the common risk factor Z_v of subprime ARM.

Assume that we have a pool of N mortgages $i = 1, \ldots, N$ for each vintage $v = 1, \ldots, V$. Each individual mortgage within a pool has the same initiation date v and interest adjustment date $v' > v$. Let $Y_{v,i}$ be the change in the logarithm of the price $P_{v,i}$ of borrower i's (of vintage v) house during the teaser period $[v, v']$. Consider

$$Y_{v,i} := \log P_{v',i} - \log P_{v,i} = H_v + e_{v,i}, \tag{2.9}$$

where $H_v := \log I_{v'} - \log I_v$ is the change in the logarithm of a housing market index I_v, and $e_{v,i}$ are i.i.d. normal random variables for all $i = 1, 2, ..., N$, and $v = 1, 2, ..., V$. As outlined in the introduction, default rates of subprime ARM depend on house price changes during the teaser period. If the house price fails to increase substantially or even declines, the mortgage borrower cannot refinance, absent other substantial improvements in income or asset position. They have to default shortly after the interest rate is reset to a high level. We assume that the default, if it happens, occurs at time v'. Therefore, we assume that a mortgage defaults if and only if $Y_{v,i} < Y^*$, where Y^* is a predetermined threshold. For example, if we set $Y^* = 0$, we are implicitly assuming that if the house price increases, the mortgage borrower is able to refinance. Otherwise, they cannot be approved for a new loan and have to default. Suppose we have a portfolio of N mortgages, which satisfy the assumptions above. Now, if we assume a form of stationary stochastic process for H_v, say an ARMA(p,q) process, we can simulate the default rates over time in the portfolio by Monte Carlo simulation.

We can now give a structural interpretation of the common risk factor Z_v in the Gaussian copula framework. Define

$$Z_v' := \frac{H_v}{\sigma_H}, \tag{2.10}$$

where σ_H is the unconditional standard deviation of H_v. Then we have

$$Y_{v,i} = Z_v' \sigma_H + e_{v,i}.$$

Further standardizing $Y_{v,i}$, we have

$$\begin{aligned} X_{v,i}' &:= \frac{Y_{v,i}}{\sigma_Y} \\ &= \frac{Z_v' \sigma_H + e_{v,i}}{\sigma_Y} \\ &= \frac{\sigma_H}{\sqrt{\sigma_H^2 + \sigma_e^2}} Z_v' + \frac{\sigma_e}{\sqrt{\sigma_H^2 + \sigma_e^2}} \varepsilon_{v,i}' \end{aligned}$$

where σ_e is the standard deviation of $e_{v,i}$, and $\varepsilon_{v,i}' := e_{v,i}/\sigma_e$. The third equality follows from the fact that

$$\sigma_Y = \sqrt{\sigma_H^2 + \sigma_e^2}.$$

Define

$$\rho' := \frac{\sigma_H^2}{\sigma_H^2 + \sigma_e^2}.$$

Then

$$X_{v,i}' = \sqrt{\rho'} Z_v' + \sqrt{1-\rho'} \varepsilon_{v,i}' \quad \forall i = 1, 2, \ldots, N; \quad t = 1, 2, \ldots, T \tag{2.11}$$

Note that equation (2.11) has exactly the same form as equation (2.8). The default event is defined as $X_{v,i}' < X^{*\prime}$ where

$$X^{*\prime} := \frac{Y^*}{\sqrt{\sigma_H^2 + \sigma_e^2}}.$$

Let

$$\tau_{v,i}' := F_v^{-1}\left(\Phi(X_{v,i}')\right),$$

and

$$\tau_v^{*\prime} := F_v^{-1}\left(\Phi(X^{*\prime})\right),$$

then the default event can be defined equivalently as $\tau_{v,i}' \leq \tau^{*\prime}$. The comparison between equation (2.11) and (2.8) shows that the common risk factor Z_v in the Gaussian copula model for subprime mortgages can be interpreted as a standardized change in a house price index.

In light of this structural interpretation, the common risk factor Z_v is very likely to be serially correlated across subsequent vintages. More specifically, we find that Z_v' is proportional to a moving average of monthly log changes in a housing price index. To see this, let v be the time of origination and v' be the end of the teaser period. Then,

$$H_v = \int_v^{v'} \mathrm{d}\log I_\tau,$$

where I is the house price index. For example, if we measure house price index changes quarterly, as in the case of the Case-Shiller housing index, we have

$$H_v = \sum_{\tau\in[v,v']} (\log I_\tau - \log I_{\tau-1}), \tag{2.12}$$

where the unit of τ is a quarter. If we model this index by some random shock arriving each quarter, equation (2.12) is a moving average process. Therefore, from equation (2.10) we know that Z'_v has positive serial correlation.

2.3 Main Theorems - Vintage Correlation in Default Rates

Since the common risk factor is likely to be serially correlated, we examine the implications for the stochastic properties of mortgage default rates. Most subprime ARM have a teaser period of two years, therefore equation (2.12) suggests that a two-year house index change can be used as a common risk factor for these mortgages. Figure 2.1 compares two-year changes in the Case-Shiller index with subprime ARM serious delinquency rates. The two variables are highly and negatively correlated with each other. To study this observation analytically, we specify a time series model for the common risk factor in the Gaussian copula approach. We then determine the relationship between the serial correlation of the default rates and that of the common risk factor.

PROPOSITION 1 (Default Probabilities and Numbers of Defaults). *Let* $k = 1, 2, ..., N$,

$$X_k = \sqrt{\rho}Z + \sqrt{1-\rho}\,\varepsilon_k, \qquad \textit{and} \qquad X'_k = \sqrt{\rho'}Z' + \sqrt{1-\rho'}\,\varepsilon'_k \tag{2.13}$$

with

$$Z' = \phi Z + \sqrt{1-\phi^2}\,u, \tag{2.14}$$

where $\rho, \rho' \in (0,1)$, $\phi \in (-1,1)$, *and* Z, $\varepsilon_1, ..., \varepsilon_N$, $\varepsilon'_1, ..., \varepsilon'_N$, u *are mutually independent standard Gaussians. Consider next the number of* X_k *that fall below some threshold* X_*, *and the number of* X'_k *below* X'_*:

$$A = \sum_{k=1}^N \mathbb{1}_{\{X_k \le X_*\}}, \qquad \textit{and} \qquad A' = \sum_{k=1}^N \mathbb{1}_{\{X'_k \le X'_*\}}, \tag{2.15}$$

where X_* *and* X'_* *are constants. Then*

$$\mathit{Cov}(A, A') = N^2\,\mathit{Cov}(p, p'), \tag{2.16}$$

where

$$p = p(Z) := \mathbb{P}[X_k \le X_* \,|\, Z] = \Phi\left(\frac{X_* - \sqrt{\rho}Z}{\sqrt{1-\rho}}\right), \qquad \textit{and} \qquad p' = \mathbb{P}[X'_k \le X'_* \,|\, Z'] = p'(Z'). \tag{2.17}$$

Moreover, the correlation between A *and* A' *equals the correlation between* p *and* p', *in the limit as* $N \to \infty$.

Proof. We first show that

$$\mathbb{E}[AA'] = \mathbb{E}\left[\mathbb{E}[A \mid Z]\mathbb{E}[A' \mid Z']\right]. \tag{2.18}$$

Note that A is a function of Z and $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_N)$, and A' is a function (indeed, the same function as it happens) of Z' and $\boldsymbol{\varepsilon'} = (\varepsilon'_1, \ldots, \varepsilon'_N)$. Now for any non-negative bounded Borel functions f and g on $\mathbb{R}^N$, and any non-negative bounded Borel functions F and G on $\mathbb{R} \times \mathbb{R}^N$, we have, on using self-evident notation,

$$\begin{aligned}
&\mathbb{E}[f(Z)g(Z')F(Z,\boldsymbol{\epsilon})G(Z',\boldsymbol{\epsilon'})] \\
&\quad= \int f(z)g(\underbrace{\phi z + \sqrt{1-\phi^2}x}_{z'})F(z,y_1,\ldots,y_N)G(z',y'_1,\ldots,y'_N)\ \mathrm{d}\Phi(z,x,\boldsymbol{y},\boldsymbol{y'}) \\
&\quad= \int f(z)g(z')\left[\left\{\int F(z,y_1,\ldots,y_N)\ \mathrm{d}\Phi(\boldsymbol{y})\right\}\left\{\int G(z',y'_1,\ldots,y'_N)\ \mathrm{d}\Phi(\boldsymbol{y'})\right\}\right]\ \mathrm{d}\Phi(z,x) \\
&\quad= \mathbb{E}\left[f(Z)g(Z')\mathbb{E}[F(Z,\boldsymbol{\epsilon}) \mid Z]\mathbb{E}[G(Z',\boldsymbol{\epsilon'}) \mid Z']\right].
\end{aligned} \tag{2.19}$$

This says that

$$\mathbb{E}\left[F(Z,\boldsymbol{\varepsilon})G(Z',\boldsymbol{\varepsilon'}) \mid Z, Z'\right] = \mathbb{E}[F(Z,\boldsymbol{\varepsilon}) \mid Z]\mathbb{E}[G(Z',\boldsymbol{\varepsilon'}) \mid Z']. \tag{2.20}$$

Taking expectation on both sides of equation (2.20) with respect to Z and Z', we obtain

$$\mathbb{E}\left[F(Z,\boldsymbol{\varepsilon})G(Z',\boldsymbol{\varepsilon'})\right] = \mathbb{E}\left[\mathbb{E}[F(Z,\boldsymbol{\varepsilon}) \mid Z]\mathbb{E}[G(Z',\boldsymbol{\varepsilon'}) \mid Z']\right]. \tag{2.21}$$

Substituting $F(Z,\boldsymbol{\varepsilon}) = A$, and $G(Z',\boldsymbol{\varepsilon}') = A'$, we have equation (2.18) and

$$\begin{aligned}
\mathbb{E}[AA'] &= \mathbb{E}\left[\mathbb{E}[A \mid Z]\mathbb{E}[A' \mid Z']\right] \\
&= \mathbb{E}[NpNp'] = N^2\mathbb{E}[pp'],
\end{aligned} \tag{2.22}$$

The last line is due to the fact that conditional on Z, A is a sum of N independent indicator variables and follows a binomial distribution with parameters N and $\mathbb{E}p$. Applying (2.21) again with $F(Z,\boldsymbol{\varepsilon}) = A$, and $G(Z',\boldsymbol{\varepsilon}') = 1$, or indeed, much more directly by repeated expectations, we have

$$\mathbb{E}[A] = N\mathbb{E}[p], \qquad \text{and} \qquad \mathbb{E}[A'] = N\mathbb{E}[p']. \tag{2.23}$$

Hence we conclude that

$$\begin{aligned}
\mathrm{Cov}(A,A') &= \mathbb{E}(AA') - \mathbb{E}[A]\mathbb{E}[A'] \\
&= N^2\mathbb{E}[pp'] - N^2\mathbb{E}[p]\mathbb{E}[p'] \\
&= N^2\,\mathrm{Cov}(p,p').
\end{aligned}$$

We have

$$\begin{aligned}
\mathrm{Var}(A) &= \mathbb{E}\left[\mathbb{E}[A^2 \mid Z]\right] - N^2(\mathbb{E}[p])^2 \\
&= \mathbb{E}[Np + N(N-1)p^2] - N^2(\mathbb{E}[p])^2 \\
&= N\mathbb{E}[p(1-p)] + N^2\,\mathrm{Var}(p).
\end{aligned} \tag{2.24}$$

Similarly,

$$\text{Var}(A') = N\mathbb{E}[p'(1-p')] + N^2 \text{Var}(p').$$

Putting everything together, we have for the correlations:

$$\begin{aligned}
\text{Corr}(A, A') &= \frac{N^2 \text{Cov}(p, p')}{\sqrt{N\mathbb{E}[p(1-p)] + N^2 \text{Var}(p)}\sqrt{N\mathbb{E}[p'(1-p')] + N^2 \text{Var}(p')}} \\
&= \frac{\text{Corr}(p, p')}{\sqrt{1 + \frac{\mathbb{E}[p(1-p)]}{N \text{Var}(p')}}\sqrt{1 + \frac{\mathbb{E}[p'(1-p')]}{N \text{Var}(p)}}} \\
&= \text{Corr}(p, p') \quad \text{as} \quad N \to \infty.
\end{aligned} \tag{2.25}$$

□

THEOREM 1 (Vintage Correlation in Default Rates). *Consider a pool of N mortgages created at each time v, where N is fixed. Suppose within each vintage v, defaults are governed by a Gaussian copula model as in equations* (2.1), (2.5), (2.7), *and* (2.8) *with common risk factor Z_v being a zero-mean stationary Gaussian process. Assume further that $\rho_v = Corr(X_{v,i}, X_{v,j})$, the correlation parameter for state variables $X_{v,i}$ of individual mortgages of vintage v, is positive. Then, A_v and $A_{v'}$, the numbers of defaults observed at time T within mortgage vintages v and v' are correlated if and only if $\phi_{v,v'} = Corr(Z_v, Z_{v'}) \neq 0$, where Z_v is the common Gaussian risk factor process. Moreover, in the large portfolio limit, $Corr(A_v, A_{v'})$ approaches a limiting value determined by $\phi_{v,v'}$, ρ_v, and $\rho_{v'}$.*

Proof. Conditional on the common risk factor Z_v, the number of defaults A_v is a sum of N independent indicator variables and follows a binomial distribution. More specifically,

$$\mathbb{P}(A_v = k | Z_v) = \binom{N}{k} p_v^k (1 - p_v)^{N-k} \tag{2.26}$$

where p_v is the default probability conditional on Z_v, i.e.,

$$p_v = \mathbb{P}(\tau_{v,i} \leq \tau^* | Z_v) = \mathbb{P}(X_{v,i} \leq X_v^* | Z_v),$$

with

$$X_v^* = \Phi^{-1}(F_v(T)),$$

where $F_v(T)$ is the probability of default before the time T. Then

$$\begin{aligned}
p_v &= \mathbb{P}\left(X_{v,i} \leq X_v^* | Z_v\right) \\
&= \mathbb{P}\left(\varepsilon_i \leq \frac{X_v^* - \sqrt{\rho_t} Z_v}{\sqrt{1 - \rho_v}}\right) \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{X_v^* - \sqrt{\rho_t} Z_v}{\sqrt{1-\rho_v}}} \exp\left(-\frac{x^2}{2}\right) \mathrm{d}x \\
&= \Phi\left(Z_v^*\right),
\end{aligned} \tag{2.27}$$

where

$$Z_v^* = \frac{X_v^* - \sqrt{\rho_v} Z_v}{\sqrt{1 - \rho_v}}. \tag{2.28}$$

Similarly,

$$p_{v'} = \Phi\left(Z_{v'}^*\right), \tag{2.29}$$

where

$$Z_{v'}^* = \frac{X_{v'}^* - \sqrt{\rho_{v'}} Z_{v'}}{\sqrt{1-\rho_{v'}}}. \tag{2.30}$$

Note that if Z_v and $Z_{v'}$ are jointly Gaussian with correlation coefficient $\phi_{v,v'}$, we can write

$$Z_v = \phi_{v,v'} Z_{v'} + \sqrt{1-\phi_{v,v'}^2}\, u_{v,v'} \quad \text{for } t > j, \tag{2.31}$$

where $u_{v,v'}$ are standard Gaussians that are independent of $Z_{v'}$. Combining equation (2.28), (2.30) and (2.31), we have

$$Z_v^* = a\phi_{v,v'} Z_{v'}^* + \frac{X_t^* - b\phi_j X_{v'}^*}{\sqrt{1-\rho_t}} - \frac{\sqrt{\rho_v(1-\phi_{v,v'}^2)}}{\sqrt{1-\rho_v}} u_{v,v'}, \tag{2.32}$$

where

$$a = \sqrt{\frac{\rho_v(1-\rho_{v'})}{\rho_{v'}(1-\rho_v)}}, \quad b = \sqrt{\frac{\rho_v}{\rho_{v'}}}.$$

Therefore,

$$\begin{aligned} \text{Cov}(p_v, p_{v'}) =& \text{ Cov}\left(\Phi(Z_v^*), \Phi(Z_{v'}^*)\right) \\ =& \text{ Cov}\left(\Phi\left(a\phi_{v,v'} Z_{v'}^* + \frac{X_v^* - b\phi_{v,v'} X_{v'}^*}{\sqrt{1-\rho_v}} - \frac{\sqrt{\rho_v(1-\phi_{v,v'}^2)}}{\sqrt{1-\rho_v}} u_{v,v'}\right), \Phi\left(Z_{v'}^*\right)\right). \end{aligned} \tag{2.33}$$

Since $a > 0$ as $\rho_v \in (0,1)$, we know that the covariance and the correlation between p_v and $p_{v'}$ are determined by $\phi_{v,v'}$, ρ_v, and $\rho_{v'}$. They are nonzero if and only if $\phi_{v,v'} \neq 0$. Applying Proposition 1, we know that

$$\text{Corr}(A_v, A_{v'}) = \frac{\text{Corr}(p_v, p_{v'})}{\sqrt{1 + \frac{\mathbb{E}[p_v(1-p_v)]}{N\,\text{Var}(p_{v'})}}\sqrt{1 + \frac{\mathbb{E}[p_{v'}(1-p_{v'})]}{N\,\text{Var}(p_v)}}} \quad \forall v \neq v'. \tag{2.34}$$

Therefore, A_v and $A_{v'}$ have nonzero correlation as long as p_v and $p_{v'}$ do. □

Equations (2.33) and (2.34) provide closed-form expressions for the serial correlation of default rates p_v of different vintages and the number of defaults A_v. However, we cannot directly read from equation (2.33) how the vintage correlation of default rates depends on the serial correlation parameter $\phi_{v,v'}$ in the common risk factor. The theorem below shows that this dependence is always positive.

THEOREM 2 (Dependence on Common Risk Factor). *Under the same settings as in Theorem 1, assume that both the serial correlation $\phi_{v,v'}$ of the common risk factor and the individual state variable correlation ρ_v are always positive. Then the number A_v of defaults in the vintage-v cohort by time T is positively correlated with the number $A_{v'}$ in the vintage-(v') cohort. Moreover, this correlation is an increasing function of the serial correlation parameter $\phi_{v,v'}$ in the common risk factor.*

Proof. We will use the notation established in Proposition 1. We can assume that $v \neq v'$. Recall that in the Gaussian copula model, name i in the vintage-v cohort defaults by time T if the standard Gaussian variable $X_{v,i}$ falls below a threshold X_v^*. The unconditional default probability is

$$\mathbb{P}[X_{v,i} \leq X_v^*] = \Phi(X_v^*).$$

For the covariance, we have

$$\begin{aligned} \mathrm{Cov}(A_v, A_{v'}) &= \sum_{k,l=1}^{N} \mathrm{Cov}(\mathbb{1}_{[X_{v,k} \leq X_v^*]}, \mathbb{1}_{[X_{v',l} \leq X_{v'}^*]}) \\ &= N^2 \mathrm{Cov}(\mathbb{1}_{[X \leq X_v^*]}, \mathbb{1}_{[X' \leq X_{v'}^*]}), \end{aligned} \tag{2.35}$$

where X, X' are jointly Gaussian, each standard Gaussian, with mean zero and covariance $\mathbb{E}[XX'] = \mathbb{E}[X_{v,k}X_{v',l}]$, which is the same for all pairs k, l, since $v \neq v'$. This common value of the covariance arises from the covariance between Z_v and $Z_{v'}$ along with the covariance between any $X_{v,k}$ with Z_v; it is

$$\mathrm{Cov}(X, X') = \phi_j \sqrt{\rho_v \rho_{v'}}. \tag{2.36}$$

Now since X, X' are jointly Gaussian, we can express them in terms of two independent standard Gaussians:

$$\begin{aligned} W_1 &:= X, \\ W_2 &:= \frac{1}{\sqrt{1 - \rho_v \rho_{v'} \phi_{v,v'}^2}} [X' - \phi_{v,v'} \sqrt{\rho_v \rho_{v'}} X]. \end{aligned} \tag{2.37}$$

We can check readily that these are standard Gaussians with *zero* covariance, and

$$\begin{aligned} X &= W_1, \\ X' &= \phi_{v,v'} \sqrt{\rho_v \rho_{v'}} W_1 + \sqrt{1 - \rho_v \rho_{v'} \phi_{v,v'}^2}\, W_2. \end{aligned} \tag{2.38}$$

Let

$$\alpha = \phi_{v,v'} \sqrt{\rho_v \rho_{v'}}.$$

The assumption that ρ and $\phi_{v,v'}$ are positive (and, of course, less than 1) implies that

$$0 < \alpha < 1.$$

Note that the covariance between p_v and $p_{v'}$ can be expressed as

$$\begin{aligned} \mathrm{Cov}(p_v, p_{v'}) &= \mathbb{E}(p_v p_{v'}) - \mathbb{E}(p_v)\mathbb{E}(p_{v'}) \\ &= \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{\{X_{v,i} \leq X_v^*\}} \middle| Z_v\right] \mathbb{E}\left[\mathbb{1}_{\left\{X_{v',i} \leq X_{v'}^*\right\}} \middle| Z_{v'}\right]\right] - \mathbb{E}(p_v)\mathbb{E}(p_{v'}) \\ &= \mathbb{E}\left[\mathbb{1}_{\{X_{v,i} \leq X_v^*\}} \mathbb{1}_{\left\{X_{v',i} \leq X_{v'}^*\right\}}\right] - \mathbb{E}(p_v)\mathbb{E}(p_{v'}) \\ &= \mathbb{P}\left[X_{v,i} \leq X_v^*,\, X_{v',i} \leq X_v^*\right] - \mathbb{E}(p_v)\mathbb{E}(p_{v'}) \\ &= \mathbb{P}\left[W_1 \leq X_v^*,\, \alpha W_1 + \sqrt{1-\alpha^2}\, W_2 \leq X_{v'}^*\right] - \mathbb{E}(p_v)\mathbb{E}(p_{v'}) \\ &= \int_{-\infty}^{X_v^*} \Phi\left(\frac{X_{v'}^* - \alpha w_1}{\sqrt{1-\alpha^2}}\right) \varphi(w_1)\, \mathrm{d}w_1 - \mathbb{E}(p_v)\mathbb{E}(p_{v'}), \end{aligned}$$

where $\varphi(\cdot)$ is the probability density function of the standard normal distribution. The third equality follows from equation (2.21). The fifth equality follows from equation (2.38). The unconditional expectation of p_v is independent of α, because

$$\begin{aligned}\mathbb{E}(p_v) &= \mathbb{E}\left(\mathbb{P}(X_{v,i} \le X_v^* | Z_v)\right) \\ &= \mathbb{P}(X_{v,i} \le X_v^*) \\ &= \Phi(X_v^*).\end{aligned} \tag{2.39}$$

It follows that

$$\begin{aligned}\frac{\partial}{\partial\alpha}\operatorname{Cov}(p_v, p_{v'}) &= \int_{-\infty}^{X_v^*} \varphi\left(\frac{X_{v'}^* - \alpha w_1}{\sqrt{1-\alpha^2}}\right)\varphi(w_1)\frac{\partial\left(\frac{X_{v'}^* - \alpha w_1}{\sqrt{1-\alpha^2}}\right)}{\partial\alpha}\,\mathrm{d}w_1 \\ &= \int_{-\infty}^{X_v^*} \varphi\left(\frac{X_{v'}^* - \alpha w_1}{\sqrt{1-\alpha^2}}\right)\varphi(w_1)\frac{-w_1 + \alpha X_{v'}^*}{(1-\alpha^2)^{\frac{3}{2}}}\,\mathrm{d}w_1 \\ &= -\frac{1}{(1-\alpha^2)^{\frac{3}{2}}}\int_{-\infty}^{X_v^*} (w_1 - \alpha X_{v'}^*)\varphi\left(\frac{X_{v'}^* - \alpha w_1}{\sqrt{1-\alpha^2}}\right)\varphi(w_1)\,\mathrm{d}w_1.\end{aligned} \tag{2.40}$$

The last two terms in the integrand can be rewritten as

$$\begin{aligned}\varphi\left(\frac{X_{v'}^* - \alpha w_1}{\sqrt{1-\alpha^2}}\right)\varphi(w_1) &= \frac{1}{2\pi}\exp\left[-\frac{(X_{v'}^* - \alpha w_1)^2}{2(1-\alpha^2)} - \frac{w_1^2}{2}\right] \\ &= \frac{1}{2\pi}\exp\left[-\frac{{X_{v'}^*}^2 - 2\alpha X_{v'}^* w_1 + w_1^2\alpha^2 + w_1^2(1-\alpha^2)}{2(1-\alpha^2)}\right] \\ &= \frac{1}{2\pi}\exp\left[-\frac{w_1^2 - 2\alpha X_{v'}^* w_1 + \alpha^2 {X_{v'}^*}^2 + {X_{v'}^*}^2(1-\alpha^2)}{2(1-\alpha^2)}\right] \\ &= \frac{1}{2\pi}\exp\left[-\frac{(w_1 - \alpha X_{v'}^*)^2 + {X_{v'}^*}^2(1-\alpha^2)}{2(1-\alpha^2)}\right].\end{aligned} \tag{2.41}$$

Substituting equation (2.41) into (2.40), we have

$$\frac{\partial}{\partial\alpha}\operatorname{Cov}(p_v, p_{v'}) = -\frac{\exp\left(-\frac{{X_{v'}^*}^2}{2}\right)}{2\pi\,(1-\alpha^2)^{\frac{3}{2}}}\int_{-\infty}^{X_v^*}(w_1 - \alpha X_{v'}^*)\exp\left[-\frac{(w_1 - \alpha X_{v'}^*)^2}{2(1-\alpha^2)}\right]\mathrm{d}w_1.$$

Make a change of variable and let

$$y := \frac{w_1 - \alpha X_{v'}^*}{\sqrt{1-\alpha^2}}.$$

It follows that

$$\begin{aligned}\frac{\partial}{\partial\alpha}\operatorname{Cov}(p_v, p_{v'}) &= -\frac{\exp\left(-\frac{{X_{v'}^*}^2}{2}\right)}{2\pi\sqrt{1-\alpha^2}}\int_{-\infty}^{\frac{X_v^* - \alpha X_{v'}^*}{\sqrt{1-\alpha^2}}} y\exp\left(-\frac{y^2}{2}\right)\,\mathrm{d}y \\ &= \frac{\exp\left(-\frac{{X_{v'}^*}^2}{2}\right)}{2\pi\sqrt{1-\alpha^2}}\exp\left[-\frac{(X_v^* - \alpha X_{v'}^*)^2}{2(1-\alpha^2)}\right] \\ &= \frac{1}{2\pi\sqrt{1-\alpha^2}}\exp\left(-\frac{X_v^{*2} - 2\alpha X_v^* X_{v'}^* + {X_{v'}^*}^2}{2(1-\alpha^2)}\right) > 0.\end{aligned}$$

Thus, we have shown that the partial derivative of the covariance with respect to α is positive. Since

$$\alpha = \sqrt{\rho_v \rho_{v'}} \phi_{v,v'},$$

with ρ_s and $\phi_{v,v'}$ assumed to be positive, we know that the partial derivatives of the covariance with respect to $\phi_{v,v'}$, ρ_v and $\rho_{v'}$ are also positive everywhere. Note that the unconditional variance of p_v is independent of $\phi_{v,v'}$ (although dependent of ρ_s), which can be seen from equation (2.27). It follows that the serial correlation of p_v has positive partial derivative with respect to $\phi_{v,v'}$. Recall equation (2.33), which shows that the covariance of p_v and $p_{v'}$ is zero for any value of ρ_s when $\phi_{v,v'} = 0$. This result together with the positive partial derivatives of the covariance with respect to $\phi_{v,v'}$ ensure that the covariance and thus the vintage correlation of p_v and $p_{v'}$ is always positive. From equation (2.34), noticing the fact that both the expectation and variance of p_v are independent of $\phi_{v,v'}$, we know that the correlation between A_v and $A_{v'}$ must also be positive everywhere and monotonically increasing in $\phi_{v,v'}$. $\square$

So far we have shown that the vintage correlation of p_v and thus A_t is positive and increasing in $\phi_{v,v'}$. Due to the complexity of the analytical form of the vintage correlation, we resort to numerical methods to study its exact magnitude. We generate plots of the vintage correlation of p_v as a function of $\phi_{v,v'}$. For simplicity, we assume $\rho_v = \rho_{v'} = \rho$, and $X_v^* = X_{v'}^* = X^*$. Different values of X_v^* ranging from -3 to 3 are implemented. Two of these plots can be seen in Figures 2.2 and 2.3. Others are omitted for brevity as they look very similar. In all these plots, vintage correlation is always positive for $\phi_{v,v'} \in (0, 1)$ and monotonically increasing in $\phi_{v,v'}$, matching our theoretical findings. Moreover, the magnitude of vintage correlation is always close to $\phi_{v,v'}$, although when the absolute value of X_v^* increases, the function becomes more convex.

2.4 Monte Carlo Simulations

In this section, we study the link between serial correlation in a common risk factor and vintage correlation in pools of mortgages in two sets of simulations: First, a series of mortgage pools is simulated to confirm the analytical results of Section 2.3. Second, a waterfall structure is simulated to study temporal correlation in MBS.

2.4.1 Vintage Correlation in Mortgage Pools

We conduct a Monte Carlo simulation to study how serial correlation of a common risk factor propagates into vintage correlation in default rates. We simulate default times for individual mortgages according to equations (2.1), (2.5), (2.7), and (2.8). From the simulated default times, the default rate of a pool of mortgages is calculated.

In each simulation, we construct a cohort of $N = 100$ homogeneous mortgages in every month $v = 1, 2, \ldots, 120$. We simulate a monthly time series of the common risk factor Z_v, which is assumed to have an AR(1) structure with unconditional mean zero and variance one,

$$Z_v = \phi Z_{v-1} + \sqrt{1-\phi^2}\, u_v \quad \forall v = 2, 3, \ldots, 120. \tag{2.42}$$

The errors u_v are i.i.d. standard Gaussian. The initial observation Z_1 is a standard normal random variable. We report the case where $\phi = 0.95$. We choose ϕ close to one due to the observation

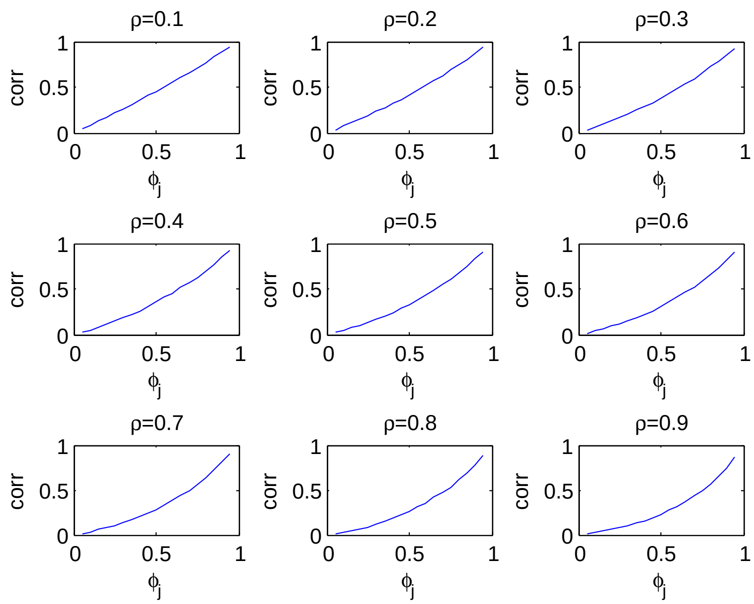

Figure 2.2: **Correlation of p_v and $p_{v'}$ when $X^* = -2$**

Correlation between p_v and $p_{v'}$ as a function of ϕ for different scenarios of the cross-mortgage correlation parameter $\rho \in (0, 1)$. X^* is set to -2 in all scenarios.

that the autocorrelation of the housing index is high.[4] Each mortgage i issued at time v has a state variable $X_{v,i}$ assigned to it that determines its default time. The time series properties of $X_{v,i}$ follow equation (2.8). The error ε_i in equation (2.8) is independent of u_v.

To simulate the actual default rates of mortgages, we need to specify the marginal distribution functions of default times $F(\cdot)$ as in equation (2.1). We define a function $\mathcal{F}(\cdot)$, which takes a time period as argument and returns the default probability of a mortgage within that time period since its initiation. We assume that this $\mathcal{F}(\cdot)$ is fixed across different vintages, which means that mortgages of different cohorts have a same unconditional default probability in the next S periods from their initiation, where $S = 1, 2, \ldots$. It is easy to verify that $F_v(T) = \mathcal{F}(T - v)$. The values of the function $\mathcal{F}(\cdot)$ are specified in Table 2.1, for both subprime and prime mortgages. Intermediate values of $\mathcal{F}(\cdot)$ are linearly interpolated from this table. While these values are in the same range as actual default rates of subprime and prime mortgages in the last ten years, their specification is rather arbitrary as it has little impact on the stochastic structure of the simulated default rates. We set the observation time T to be 144, which is two years after the creation of the last vintage, as we need to give the last vintage some time window to have possible default events. For example, in each month from January 1998 to December 2007, 100 mortgages are created. Then in December 2009, we examine the default rates of these mortgages within each vintage.

We need to consider two cases, subprime and prime. For the subprime case, every vintage is given a two-year window to default, so the unconditional default probability is constant across vintages. On the other hand, prime mortgages have decreasing default probability through subsequent vintages. For example, in our simulation, the first vintage has a time window of 144 months to

[4] Other values of ϕ are also tried, but not reported here. The results are all consistent with our theoretical findings.

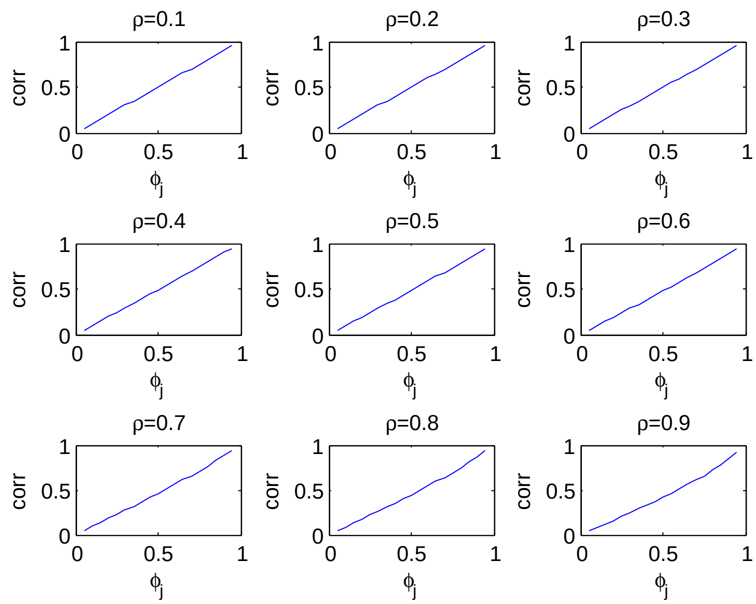

Figure 2.3: **Correlation of p_v and $p_{v'}$ when $X^* = 0$**

Correlation between p_v and $p_{v'}$ as a function of ϕ for different scenarios of the cross-mortgage correlation parameter $\rho \in (0, 1)$. X^* is set to 0 in all scenarios.

default, the second vintage has 143 months, the third has 142 months, and so on. Therefore, older vintages are more likely to default by observation time T than newer vintages. This is why the fixed ex-post observation time of defaults is one difference that distinguishes vintage correlation from serial correlation.

We construct a time series $\tau_{v,i}$ of default times of mortgage i issued at time v according to equation (2.5).[5] Time series of default rates $\bar{A}_v$ are computed as:

$$\bar{A}_v(\tau_v^*) = \frac{\#\{\text{mortgages for which } \tau_{v,i} \leq \tau_v^*\}}{N}.$$

In the subprime case, $\tau_v^* = 24$ is set to be constant. In the prime case, $\tau_v^* = T - v$ varies across vintages.

The simulation is repeated 1000 times. For the subprime case, the average simulated default rates are plotted in Figure 2.4. For the prime case, average simulated default rates are plotted in Figure 2.5. Note that because of the decreasing time window to default, the default rates in Figure 2.5 have a decreasing trend.

In the subprime case, we can use the sample autocorrelation and partial autocorrelation functions to estimate vintage correlation, because the unconditional default probability is constant across vintages, so that averaging over different vintages and averaging over different pools is the same. In the prime case, we have to calculate vintage correlation proper. Since we have 1000 Monte Carlo observations of default rates for each vintage, we can calculate the correlation

[5]Note that this is not a time series of default times for a single mortgage, since a single mortgage defaults only once or never. Rather, the index i is a placeholder for a position in a mortgage pool. In this sense, $\tau_{v,i}$ is the time series of default times of mortgages in position i in the pool over vintages v.

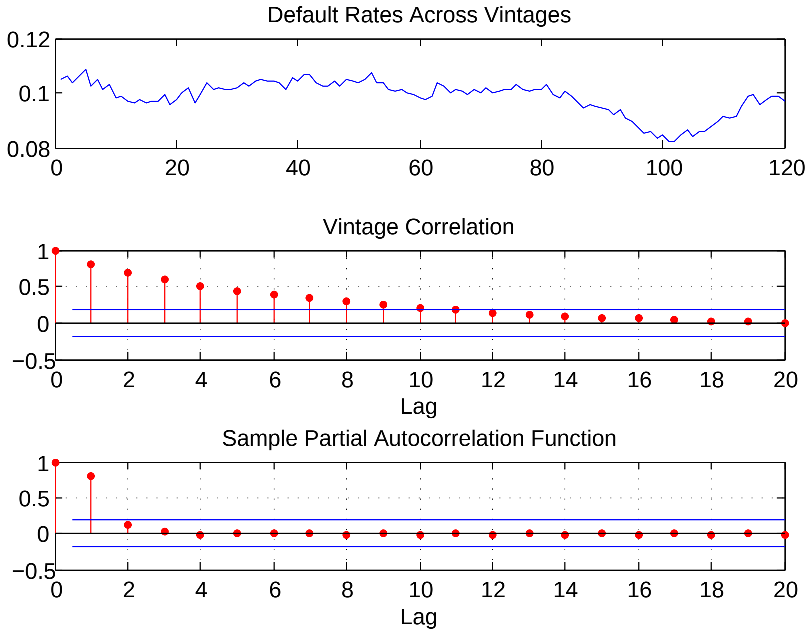

Figure 2.4: **Serial Correlation in Default Rates of Subprime Mortgages**

between two vintages using those samples. For the partial autocorrelation function, we simply demean the series of default rates and obtain the usual partial autocorrelation function.

We plot the estimated vintage correlation in the second rows of Figure 2.4 and 2.5 for subprime and prime cases, respectively. As can be seen, the correlation of the default rates of the first vintage with older vintages decreases geometrically. In both cases, the estimated first-order coefficient of default rates is close to but less than $\phi = 0.95$, the AR(1) coefficient of the common risk factor. The partial autocorrelation functions are plotted in the third rows of Figures 2.4 and 2.5. They are significant only at lag one. This phenomenon is also observed when we set ϕ to other values. Both the sample autocorrelation and partial autocorrelation functions indicate that the default rates follow a first-order autoregressive process, similar to the specification of the common risk factor. However, compared with the subprime case, the default rates of prime mortgages seem to have longer memory.

The similarity between the magnitude of the autocorrelation coefficient of default rates and common risk factor can be explained by the following Taylor expansion. Taylor-expanding equation (2.27) at $Z_v^* = 0$ to first order, we have

$$p_v \approx \frac{1}{2} + \frac{1}{\sqrt{2\pi}} Z_v^* \tag{2.43}$$

Since p_v is in this sense approximately a linear transformation of Z_v^*, which is a linear transformation of Z_t, it approximately follows a stochastic process that has the same serial correlation as Z_t.

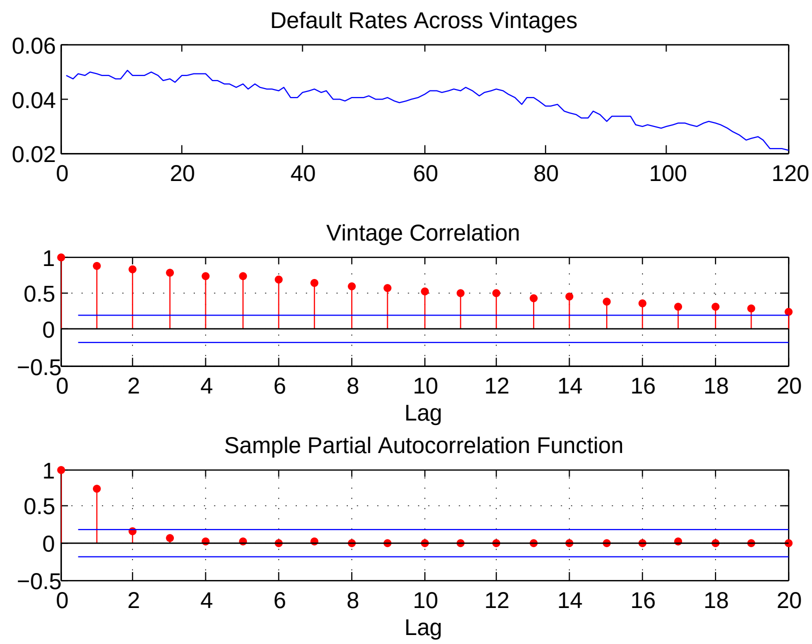

Figure 2.5: **Serial Correlation in Default Rates of Prime Mortgages**

2.4.2 Vintage Correlation in Waterfall Structures

Using the Gaussian copula approach, we have already shown that the time series of default rates in mortgage pools inherits vintage correlation from the serial correlation of the common risk factor. We now study how this affects the performance of assets such as MBS that are securitized from the mortgage pool in a so-called waterfall.

The basic elements of the simulation are:

1. a time line of 120 months and an observation time $T = 144$;
2. a mortgage contract with a principal of $\$1$ and a maturity of 15 years. The annual interest rate on the mortgage loan is 9%. Fixed monthly payments are received until the mortgage defaults or is paid in full. In each month, a pool of 100 such mortgages is created.
3. A pool of 100 units of MBS is securitized from the mortgage cohort in each month. Every unit of MBS has a principal of $\$1$. There are four tranches in our structure: the senior tranche, the mezzanine tranche, the subordinate tranche, and the equity tranche. The senior tranche consists of the top 70% of the face value of all mortgages created in each month (that is, there are 70 units of senior MBS); the mezzanine tranche consists of the next 25%; the subordinate tranche consist of the next 4%; the equity tranche has the bottom 1%. Each senior MBS pays an annual interest rate of 6%; each mezzanine MBS pays 15%; each subordinate MBS pays 20%. The equity tranche does not pay interest but retains residual profits, if any.

The basic setup of the simulation is illustrated in Figure 2.6. For a cohort of mortgages issued at time v and the MBS derived from it, the securitization process works as follows. At the end of each month, each mortgage either defaults or makes a fixed monthly payment. The method to determine

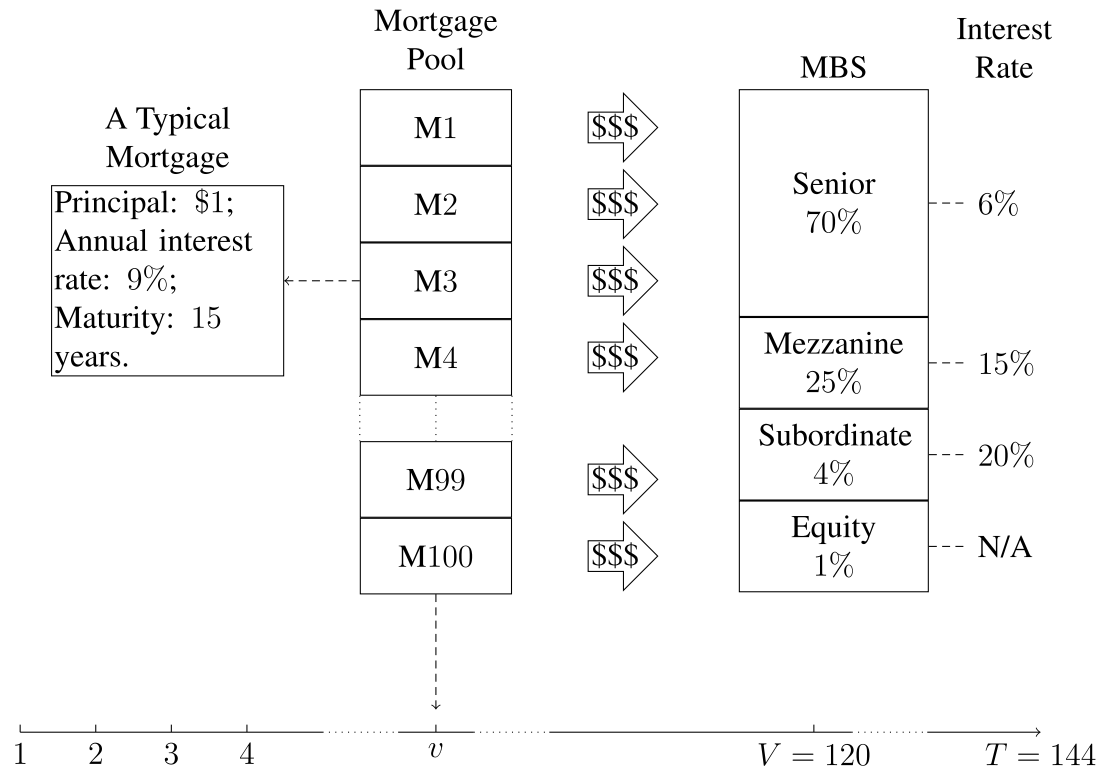

Figure 2.6: **Simulated Mortgages and MBS**

default is the same that we have used before: mortgage i issued at time v defaults at $\tau_{v,i}$, which is generated by the Gaussian copula approach according to equations (2.1), (2.5), (2.7), and (2.8). We consider both subprime and prime scenarios, as in the case of default rates. For subprime mortgages, we assume that each individual mortgage receives a prepayment of the outstanding principal at the end of the teaser period if it has not defaulted, so the default events and cash flows only happen within the teaser period. For the prime case, there is no such restriction. Again, we assume the common risk factor to follow an AR(1) process with first-order autocorrelation coefficient $\phi = 0.95$. The cross-name correlation coefficient ρ is set to be 0.5. The unconditional default probabilities over time are obtained from Table 2.1.

If a mortgage has not defaulted, the interest payments received from it are used to pay the interest specified on the MBS from top to bottom. Thus, the cash inflow is used to pay the senior tranche first (6% of the remaining principal of the senior tranche at the beginning of the month). The residual amount, if any, is used to pay the mezzanine tranche, after that the subordinate tranche, and any still remaining funds are collected in the equity tranche. If the cash inflow passes a tranche threshold but does not cover the following tranche, it is prorated to the following tranche. Any residual funds after all the non-equity tranches have been paid add to the principal of the equity tranche. Principal payments are processed analogously. We assume a recovery rate of 50% on the outstanding principal for defaulted mortgages. The 50% loss of principal is deducted from the principal of the lowest ranked outstanding MBS. This means that the equity tranche covers the principal loss first. If there is no principal left in the equity tranche, the subordinate tranche

Table 2.1: **Default Probabilities Through Time ($F(\tau)$).**

			Subprime		
Time (Month)	12	24	36	72	144
Default Probability	0.04	0.10	0.12	0.13	0.14
			Prime		
Time (Month)	12	24	36	72	144
Default Probability	0.01	0.02	0.03	0.04	0.05

covers the remaining loss and so on upwards. In order to capture MBS price performance through time, we calculate the present value of all cash flows received from each MBS tranche and average across 1000 simulations. This results in a time series of expected present values and can be viewed as a proxy for tranche price evolution.

Before we examine the vintage correlation of the present value of MBS tranches, we look at the time series of total principal loss across MBS tranches. In our simulations, no loss of principal occurred for the senior tranche. The series of expected principal losses of other tranches and their sample autocorrelation and sample partial autocorrelation are plotted in Figures 2.7 and 2.8 for subprime and prime scenarios respectively. We use the same method to obtain the autocorrelation functions for prime mortgages as in the case of default rates. The correlograms show that the expected loss of principal for each tranche follows an AR(1) process, although the estimated coefficients are smaller than $\phi = 0.95$, the first-order autocorrelation coefficient of Z_v, in all cases.

The series of present values of cash flows for each tranche and their sample autocorrelation and partial autocorrelation functions are plotted in Figures 2.9 and 2.10 for subprime and prime scenarios, respectively. The senior tranche displays a significant first-order autocorrelation coefficient due to losses in interest payments although there are no losses in principal. The partial autocorrelation functions, which have significant positive values for more than one lag, suggest that the cash flows may not follow an AR(1) process due to the high non-linearity. However, the estimated vintage correlation still decreases over vintages, same as in an AR(1) process, which indicates that our findings for default rates can be extended to cash flows.

2.5 Conclusions

Default rates of subprime mortgages exhibit temporal correlation. Default events of subprime-mortgages depend on house price changes that are serially correlated, and this serial correlation is inherited by the sequence of defaults. We model a house price index by a common risk factor in a Gaussian copula model. Analytical findings and simulations show that credit assets inherit correlation of defaults across different vintages of initiation from serial correlation in the common risk factor. The same is true in waterfall structures, used in mortgage-backed securities and collateralized debt obligations. Simulation results demonstrate that when the common risk factor of different cohorts of individual mortgages is serially correlated, the price of the MBS securitized from these mortgages also displays serial correlation. These findings are consistent with the view that the current financial crisis, which was triggered by a large and serially correlated arrival of

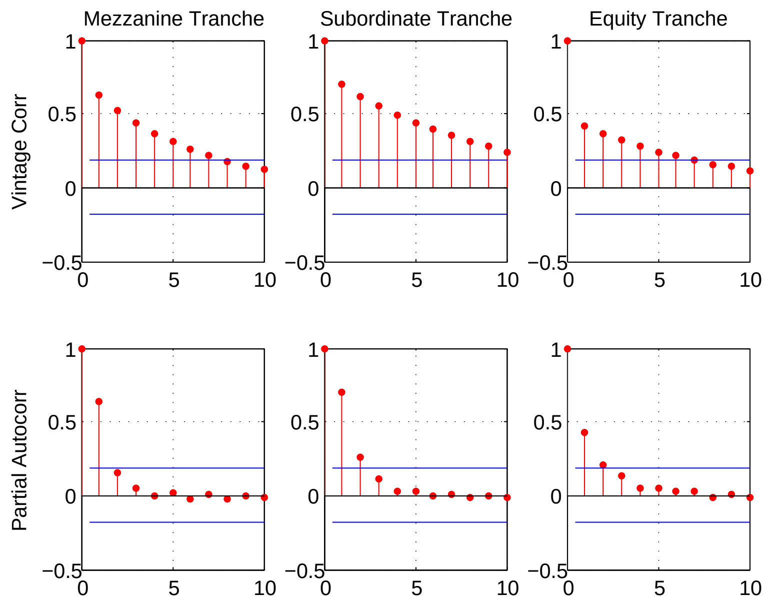

Figure 2.7: **Serial Correlation in Principal Losses of Subprime MBS**

The first row plots the vintage correlation of the principal loss of each tranche. The correlation is estimated using the sample autocorrelation function. The second row plots the partial autocorrelation functions.

subprime mortgage defaults, has its origin in the decline of the housing market.

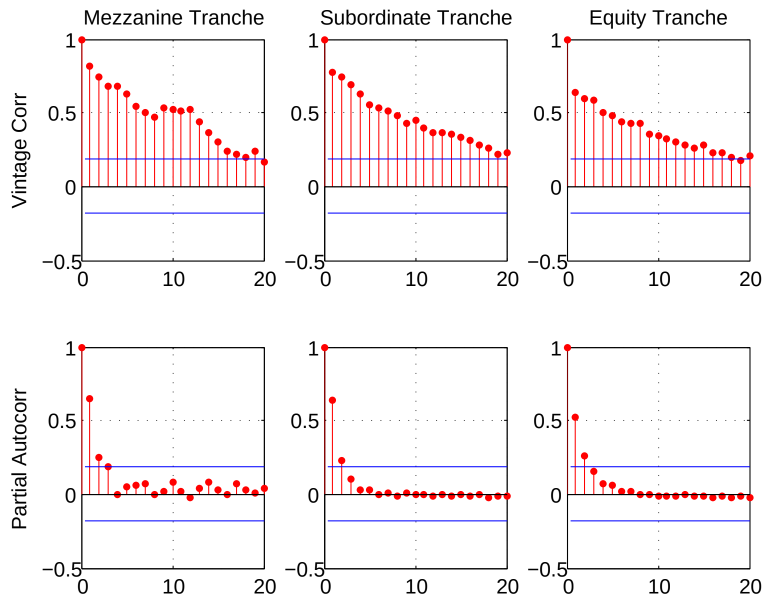

Figure 2.8: **Serial Correlation in Principal Losses of Prime MBS**

The first row plots the vintage correlation of the principal loss of each tranche. The correlation is estimated using the correlation between the first and subsequent vintages, each of which has a Monte Carlo sample size of 1000. The second row plots the partial autocorrelation functions of the demeaned series of principal losses.

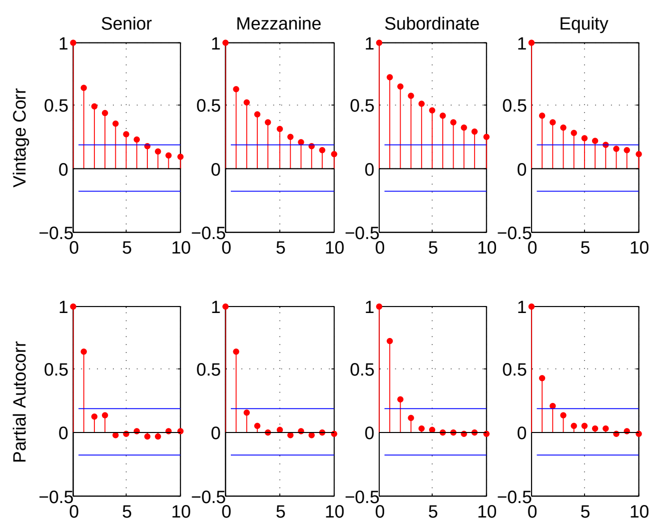

Figure 2.9: **Cash Flows from Subprime MBS**

The first row plots the vintage correlation of the cash flow received by each tranche. The correlation is estimated using the sample autocorrelation function. The second rows plot the partial autocorrelation functions.

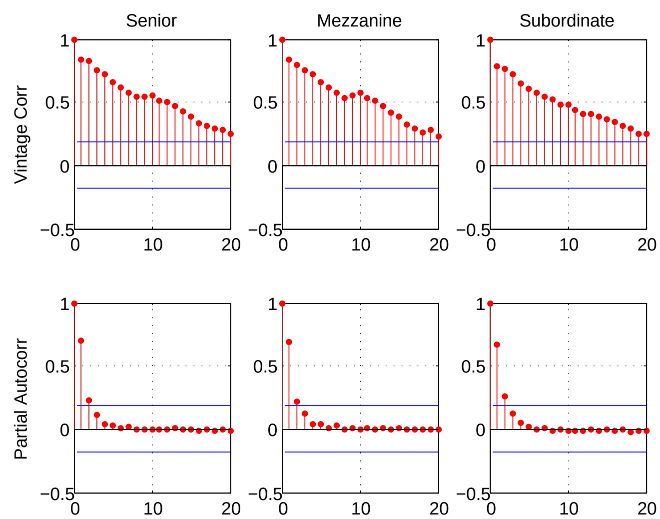

Figure 2.10: **Cash Flows from Prime MBS**

The first row plots the vintage correlation of the cash flow received by each tranche. The correlation is estimated using the correlation between the first and subsequent vintages, each of which has a Monte Carlo sample size of 1000. The second row plots the partial autocorrelation functions of the demeaned series of cash flow.

Chapter 3. Impact of Correlation Fluctuations on Securitized Structures

3.1 Introduction

The financial crisis precipitated by the subprime mortgage fiasco has focused attention on the use of Gaussian copula methods in pricing and risk managing CDOs involving subprime mortgages. Gorton (2008) has analyzed the role of structured mortgage backed securities (MBS) vehicles in the subprime crisis. In this paper we study, both theoretically and numerically, Gaussian default modeling and its sensitivity to changes in default correlation over time. Our method avoids some typical pitfalls of using copulas over extended time periods by using the notion of *vintage* of a portfolio, rather than its age, an important distinction. In brief, this means we examine, at a fixed time, a sequence of portfolios which have been issued at different initiation times, hence of different *vintages*, in the past. The different vintages across time are connected through a serial correlation parameter, which we hold fixed for this study. An extensive study of vintage correlation is carried out in Hillebrand, Sengupta, and Xu (2010).

One concern about the simple Gaussian copula model of Li (2000), which assumes a static correlation of default between different assets in the portfolio, is that the value of default correlation migrates over time (for example, see Servigny and Renault (2002) and Das, Freed, Geng, and Kapadia (2006)). To address this concern, Andersen and Sidenius (2005), Berd, Engle, and Voronov (2007), and Hull, Predescu, and White (2009) allow default correlation to vary. The dynamics of default correlation have important implications for risk management. As is shown by Andersen and Sidenius (2005), introducing dynamics into default correlation changes the loss distribution of CDO tranches. It is natural to ask then, how exactly default correlation dynamics affect the distribution of CDO tranche prices. It is of interest to know if the manner in which default correlation changes matters. We explore a model where the default correlation changes more smoothly than the simple regime switching scenario of Andersen and Sidenius (2005). We also study how tranche price distributions are affected when we introduce dynamics into both default correlation and the state variable.

As mentioned before, we consider MBS tranches of many subsequent vintages. Both the state variable and default correlation are assumed to be stochastic across vintages. We examine the impact of the dynamics of the state variable and default correlation on tranche prices, as measured roughly through the expected value of cash flows. We also allow the smoothness of the change in default correlation to vary and inspect its impact on the tranche price distribution. We find that some of our results depend on the sensitivity of the tranche price to the change in default correlation, which in turn depends on the seniority of the MBS tranche. To study the impact of a change in default correlation, we slice the full investing spectrum of an MBS into a large number of thin tranches, which we refer to as 'high-frequency' tranching. The presence of such thin tranches in subprime MBS portfolios has been noted by Gorton (2008). Through a set of Monte Carlo simulations, we are able to study the impact of the dynamics of default correlation on the prices of such tranches.

This paper is organized as follows. Section 3.2 describes a typical MBS vehicle and our models. We conduct a set of Monte Carlo simulations in Section 3.3 to examine the prices of a two-tranche MBS. We assume a regime-switching model for the default correlation and study the effects on the distribution of tranche prices. We then impose a more general logistic transition structure on the dynamics of default correlation. This model nests, as one limit point, the constant default correlation model and, as another limit point, the regime-switching default correlation model. In Section 3.4, we study an MBS with high-frequency tranching. We numerically estimate the sensitivities of the tranche prices to a change in default correlation in a static model. We explore how dynamic default correlation affects the serial correlation and overall distributions of tranche prices.

3.2 Description of the Products and Models

The function of an MBS vehicle is to allocate capital from investors with a spectrum of risk tolerance to borrowers. Suppose there are investors labeled by interest rates $r_1 < \ldots < r_I$ that they seek, and there are borrowers whose risk levels qualify them for loan rates of $r'_1 < \ldots < r'_B$. An MBS portfolio pools together the funds from the investors and issues mortgage loans to the borrowers from this pool (of course, the same considerations apply to other asset-backed loans). More accurately, the interest rates r_i label *tranches* of the portfolio, which the investors may purchase. The allocation of risk of defaults in the loans to the different investors/tranches is a critically important task in the structuring of such an investment vehicle. An active market place in the securities (investments) would generate market-implied rates r_i, or, more precisely, *tranche prices*, but pricing any derivative product for such tranches would require a good model of the correlated default behavior of the mortgages. Errors in the modeling of such default behavior would show up either as arbitrage opportunities or, more seriously, as market instabilities. Copula models for default correlation have proved to be most useful in practice, despite serious theoretical drawbacks. More theoretically sound models may require a large number of parameters to be estimated, and each such parameter would itself be a source of possible error in risk management.

In this paper we consider V portfolios, issued at time instants

$$T_1 < T_2 < \ldots < T_V,$$

with the j-th portfolio consisting of N mortgages. We assume, as an idealized situation, that each mortgage has a maturity of only one period. If the mortgage defaults during that period, all principal is lost. Otherwise, the mortgage receives full principal value at maturity. We construct two or more tranches of MBS from each portfolio of mortgages.

In principle, the labels T_v might indicate something other than time (geographic or industry sector, for instance). However, we will refer to v, as the *vintage* of the portfolio, and for the purpose of this paper the vintage interpretation is more appropriate than are other interpretations of v. The performance of the MBS are examined at a certain time, say time T. The expected value of the cash flow received by each tranche of MBS of vintage v is calculated conditional on the information available at vintage v. This conditional expected value of cash flows can be regarded, in this simplified setup, as the price of the MBS.

The default behavior of mortgage i in the portfolio of vintage v is governed by a random variable $X_{v,i}$. If $X_{v,i}$ is below a threshold c_v, then mortgage i in vintage v has defaulted. Therefore, the default probability of a mortgage is

$$\mathbb{P}(X_{v,i} < c_v) = F_{v,i}(c_v), \tag{3.44}$$

where $F_{v,i}$ is the standard cumulative distribution function of $X_{v,i}$. There is no model yet.

In the one-factor Gaussian copula model, there are independent standard Gaussian variables Z_v and $\epsilon_{v,1}, \ldots, \epsilon_{v,N}$, such that

$$X_{v,j} = \sqrt{\rho_v} Z_v + \sqrt{1-\rho_v}\, \epsilon_{v,j}. \tag{3.45}$$

Thus the assumption is that the variables $X_{v,j}$ are jointly Gaussian, with each being standard Gaussian, and have the following specific correlation structure:

$$\mathbb{E}[X_{v,i} X_{v,j}] = \rho_v \in [0, 1]. \tag{3.46}$$

Recall that for jointly Gaussian variables, each of mean 0, the joint distribution is completely determined by the pairwise covariances.

Following Anderson and Sidenius (2005), we will allow the possibility that the correlation parameter ρ_v is dependent on the stochastic state variable Z_v, and then take

$$X_{v,j} = \sqrt{\rho_v} Z_v + \kappa \epsilon_{v,j} + m, \tag{3.47}$$

where κ and m are parameters that ensure $X_{v,j}$ has mean 0 and variance 1. Intuitively, Z_v can be viewed as a state variable that determines the conditional default probability for mortgages of vintage v. For these variables, we assume the following *serial correlation* behavior: the variables Z_1,..., Z_V are also jointly Gaussian with correlations

$$\mathbb{E}[Z_v Z_{v'}] = \phi^2_{v,v'} \in [0, 1], \tag{3.48}$$

where $\phi_{v,v'} \geq 0$.

In the case of continuous-time vintage $v \in [0, V]$, the process

$$v \mapsto Z_v$$

is Gaussian, with Z_v having mean 0 and variance 1 for each v. For instance, $Z_v = a(v) B_{b(v)}$, for a Brownian motion $v \mapsto B_v$, and $a(\cdot)$ and $b(\cdot)$ are suitable functions which determine the correlation $\phi_{v,v'}$.

In this paper, for simplicity, we always assume that Z_v follows an $AR(1)$ process of the form:

$$Z_v = \phi Z_{v-1} + \sqrt{1-\phi^2} u_v, \;\; v = 0, 1, \cdots, V \tag{3.49}$$

where Z_0 and u_v are all standard Gaussian variables.

Note that unlike the ‘traditional’ use of the copula model, we do not apply the copula model to one portfolio over different time instants; instead we are considering *different portfolios* issued at different times (thereby of different vintages).

We have studied this model extensively in Hillebrand, Sengupta, and Xu (2010), the main results of which may be summarized as:

1. Default rates exhibit vintage correlation if and only if the state variable Z_v has serial correlation. Moreover, in the large portfolio limit, the vintage correlation approaches a limiting value determined by the magnitude of serial correlation and the value of default correlation.

2. Vintage correlation of default rates is positively correlated with serial correlation in the state variable Z_v.

In the present paper we explore situations where there may be stochastic fluctuation in the default correlation parameter ρ_v with respect to the vintage parameter v in particular, a transition between a high correlation regime and a low correlation regime. We then examine the impact on the tranche prices, as measured through the expected value of cash flows.

3.3 Impact of Dynamics of Default Correlation on Low-Frequency Tranches

3.3.1 Constant Default Correlation

In this section, we study the impact of dynamics in the state variable on MBS tranche price distributions. For this purpose, we hold the default correlation constant and let the state variable vary across vintages.

We construct 2500 vintages by simulating one path of the state variables Z_v, where $v = 1, 2, \ldots, 2500$. For each vintage, we simulate a cohort of $N = 100$ homogenous mortgages. Each mortgage has a principal of $\$1$. To reduce computational burden, these mortgages are assumed to have a life time of one period. A mortgage is assumed to receive no cash flow if a default event happens during its life time and to receive full principal otherwise.

From each cohort of mortgages, we construct 100 units of mortgage backed securities (MBS). Each unit of the MBS has a principal of $\$1$. There are two tranches of our simplified MBS. The senior tranche consists of the top 50% of the face value of all mortgages created in each vintage. The equity tranche consists of the bottom 50%. The prices of each tranche of vintage v are simulated as the expected value of the cash flow received by each tranche conditional on the value of Z_v. The underlying procedure of the simulation is:

1. A series of 2500 state variables Z_v is simulated with i.i.d standard Gaussian distribution.

2. For each Z_v, the default behavior of a cohort of $N = 100$ mortgages are simulated according to equations (3.44) and (3.47). The unconditional default probability for each mortgage is fixed at 50%, meaning we keep c_v in equation (3.44) fixed at 0. At this stage, we keep the default correlation parameter $\rho_v = 0.5$ constant. The cash flows received by each tranche of MBS at each vintage are calculated according to the simulated mortgage default behavior.

3. Step (2) is repeated 100000 times, and the expected values of the cash flows received by each tranche of MBS for each vintage are estimated. These expected values can be viewed as the tranche prices.

Note that in our simulation, we only simulate one series of Z_v. This implies that the simulated series of prices follows one realization of the path of tranche prices observed at a later date T.

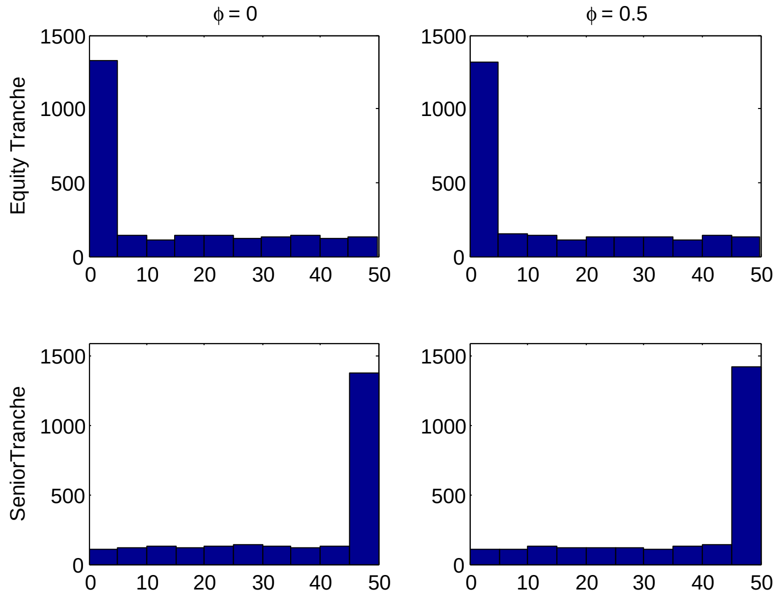

Figure 3.11: **Histograms of Tranche Prices when Z_v Has Different Values of Serial Correlation**

In these figures, we plot the histogram of tranche prices across vintages. The tranche price is simulated as the expected value of cash flow received by a certain tranche of a certain vintage. In our simulations, we construct 2500 vintages, so there are 2500 prices for each tranche. In the first column, the state variable Z_v is white noise. In the second column, Z_v follows an $AR(1)$ process described in equation (3.49) with $\phi = 0.5$. In both columns, $\rho_v = 0.5$ assumes to be constant across vintages. The unconditional default probability is assumed to be 50%.

Therefore, the unconditional distribution of these prices is comparable to the distribution of a series of observed historical prices of MBS of subsequent vintages. The histograms of tranche prices are shown in the first column of Figure 3.11.

We then let Z_v be serially correlated according to equation (3.49), where $\phi = 0.5$. The histogram of the simulated tranche prices in this case are shown in the second column of Figure 3.11. Comparing the two columns of the figure, it appears that there is no significant difference in the two cases. This finding is verified by the Quantile-Quantile (QQ) plot of the price distributions shown in Figure 3.12. In the first column of this figure, we display the QQ plot of the distribution of tranche prices when Z_v is not serially correlated against the tranche price distribution when Z_v is serially correlated. It can be seen from the figure that the two distributions are very similar. This result indicates that the dynamics in Z_v do not affect the unconditional distribution of tranche prices.

3.3.2 Regime-Switching Default Correlation

We now allow the default correlation parameter ρ_v to vary stochastically across vintages. To model the dynamics of ρ_v, we follow Andersen and Sidenius (2005) and use a regime-switching model. Specifically, we set

$$\rho_v = \rho_l \cdot \mathbb{1}_{\{Z_v \geq Z^*\}} + \rho_u \cdot \mathbb{1}_{\{Z_v < Z^*\}}, \tag{3.50}$$

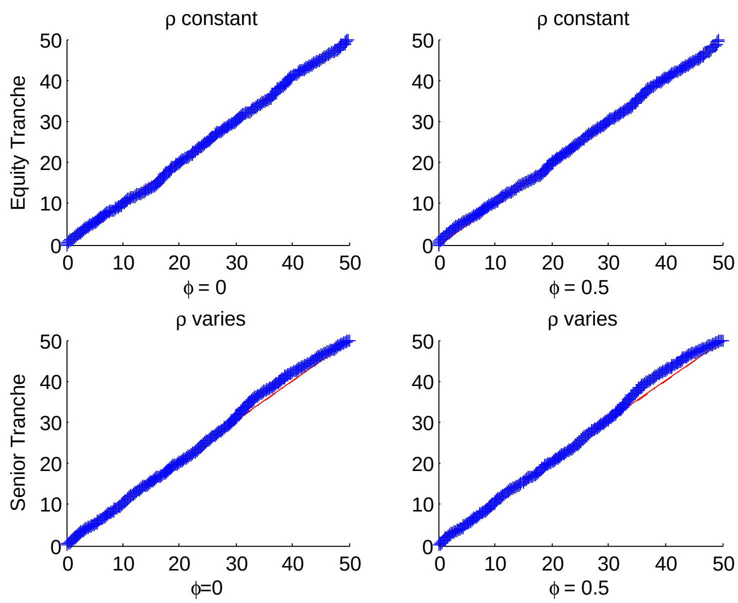

Figure 3.12: **QQ Plot:** $\phi = 0$ **Versus** $\phi = 0.5$

In these figures, we display the quantile-quantile plot of tranche prices when Z_v has no serial correlation (horizontal axis) versus tranche prices when Z_v follows an $AR(1)$ process where the first order coefficient $\phi = 0.5$ (vertical axis). In the first column, the value of ρ_v is constant across all vintages. In the second column, the value of ρ_v varies across vintages according to equation (3.50), where $\rho_l = 0.3$, $\rho_u = 0.7$. The unconditional default probability is assumed to be 50%.

where $\rho_l = 0.3$, $\rho_u = 0.7$, $Z^* = 0$. This means that the value of ρ_v assumes a lower value 0.3 when the state variable Z_v is positive and assumes a higher value 0.7 when the state variable is negative. This is consistent with the empirical finding that default correlation tends to be higher when the overall economy is in a bad state. Compared with our previous model, the mean of ρ_v remains the same. In doing this, we insulate the impact on default behavior caused by the dynamics of ρ_v from that caused by a change in the absolute level of ρ_v.

Note that the tranche price can be positively or negatively correlated with the value of default correlation according to the seniority of the tranche. (For example, Meng and Sengupta (2011) provide a detailed analysis of the sensitivities of tranche prices to change in default correlation.) This means that the dynamics of ρ_v described above have different impact on prices of different tranches. Since the price of a tranche is generally positively correlated with the value of Z_v, the dynamics of ρ_v exaggerate the effect of Z_v on tranches whose prices are negatively correlated with default correlation. For example, the price of a senior tranche is generally negatively correlated with the value of default correlation. When the value of Z_v is high, not only the conditional default probability is low but the default correlation is also low if ρ_v follows equation (3.50). Thus in this case, a senior tranche benefits from a small ρ_v as well as a low conditional default probability. Similarly, the dynamics of ρ_v alleviate the effect of Z_v on tranches whose prices are positively correlated with default correlation. This may in turn affect the overall distribution of tranche prices.

Since ρ_v is now assumed to be a random variable that is dependent on Z_v, the state variable for

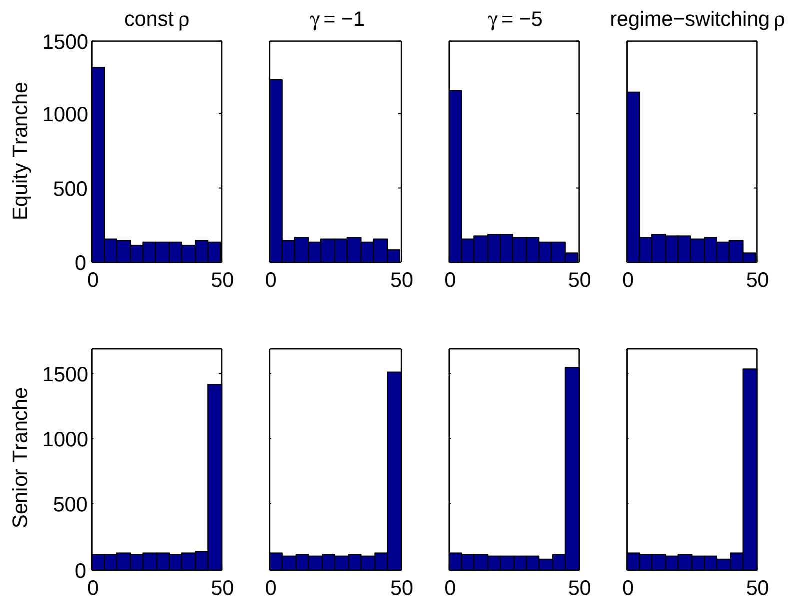

Figure 3.13: **The Histograms of Tranche Prices Over Vintages**

In these figures, we plot the histogram of tranche prices across vintages, where the state variable Z_v is assumed to follow an $AR(1)$ process with the first order coefficient $\phi = 0.5$. The tranche price is simulated as the expected value of cash flow received by a certain tranche of a certain vintage. In our simulations, we construct 2500 vintages, so there are 2500 prices for each tranche. In the first column, the value of ρ_v is fixed at 0.5. In the second and third columns, the value of ρ_v is changing between 0.3 and 0.7 according to equation (3.53). In the fourth column, the value of ρ_v assumes a regime-switching model specified by equation (3.50), where $\rho_l = 0.3$, $\rho_u = 0.7$. The unconditional default probability is assumed to be 50%.

each vintage, $X_{v,j}$ in equation (3.45) no longer has an unconditional standard Gaussian distribution. In the model for the state variable in equation (3.47), which we repeat here:

$$X_{v,j} = \sqrt{\rho_v} Z_v + \kappa \epsilon_{v,j} + m,$$

where Z_0 and $\varepsilon_{v,j}$ are standard Gaussian variables, we now use the parameters (see Andersen and Sidenius (2005))

$$\kappa = \sqrt{1 - \text{Var}(\sqrt{\rho_v} Z_v)}, \tag{3.51}$$

and

$$m = -\mathbb{E}(\sqrt{\rho_v} Z_v). \tag{3.52}$$

The first and fourth columns in Figure 3.13 compare the histograms of tranche prices in the case where ρ_v is constant and in the case where ρ_v varies according to equation (3.50). As can be seen, for the senior tranche, when ρ_v varies, the likelihood of either receiving full payment or receiving nothing increases. In other words, the distribution of senior tranche prices has fatter tails when ρ_v varies than when it is constant. For the equity tranche however, the exact opposite happens. The likelihood of either getting full payment or receiving no payment decreases. These results are intuitive. Since the price of a senior tranche is negatively correlated with ρ_v, the dynamics of ρ_v, modeled by equation (3.50), amplified the effect of Z_v on tranche prices, increasing the

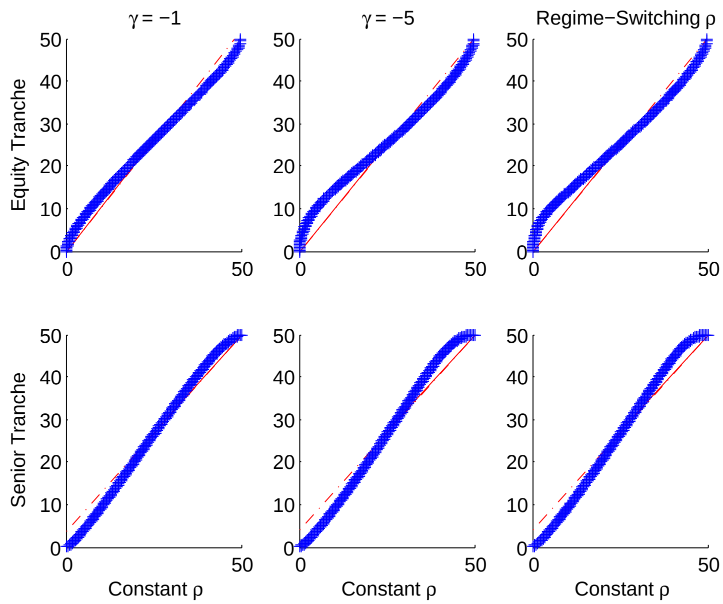

Figure 3.14: **QQ Plot of Constant ρ_v vs Varying ρ_v**

In these figures, we display the quantile-quantile plot of tranche prices when default correlation is constant versus tranche prices when default correlation is stochastic. In all figures, the horizontal axis denotes the case where ρ_v is constant. The vertical axis denote the case where ρ_v varies across vintages. In the first column, the value of ρ_v has a logistic transition model according to equation (3.53), where $\gamma = -1$. In the second column, the value of ρ_v has a logistic transition model according to equation (3.53), where $\gamma = -5$. In the third column, the value of ρ_v has a regime-switching model and varies between 0.3 and 0.7 according to equation (3.50). The state variable Z_v is assumed to follow an $AR(1)$ process where the first order coefficient $\phi = 0.5$. The unconditional default probability is assumed to be 50%.

probabilities of either a very high or very low default rate. For the equity tranche, its price is positively correlated with default correlation. Thus, the assumed dynamics of default correlation alleviate the effect of Z_v, making the tranche less likely to receive either zero or full payment. These findings are also supported by Figure 3.14. In the third column of Figure 3.14, we display the QQ plot of the distribution of tranche prices when ρ_v is constant against the distribution of tranche prices when ρ_v assumes a regime-switching model. It is clear from this figure that when ρ_v is stochastic relative to when ρ_v is constant, the equity tranche price has a distribution with thinner tails while the senior tranche price has a distribution of fatter tails. We also notice that the same can be concluded if Z_v does not have serial correlation. This can be seen from the first and fourth columns in Figures 3.15 and the third column in Figure 3.16, where we set Z_v to have no serial correlation.

Again, the dynamics of Z_v do not seem to affect the unconditional distribution of prices even when ρ_v is stochastic. This can be seen in Figure 3.17. In the first column of the figure, we fix the autocorrelation parameter of Z_v to be zero, so Z_v is a series of white noises. In the second column, Z_v follows an $AR(1)$ process with $\phi = 0.5$. In both columns, we allow ρ_v to change according to equation (3.50). Note that even though the state variable Z_v has different serial correlations, the distributions of tranche prices are almost identical between the two columns. The QQ plot of the two distributions, which can be seen in the second column of Figure 3.12, confirms that these two

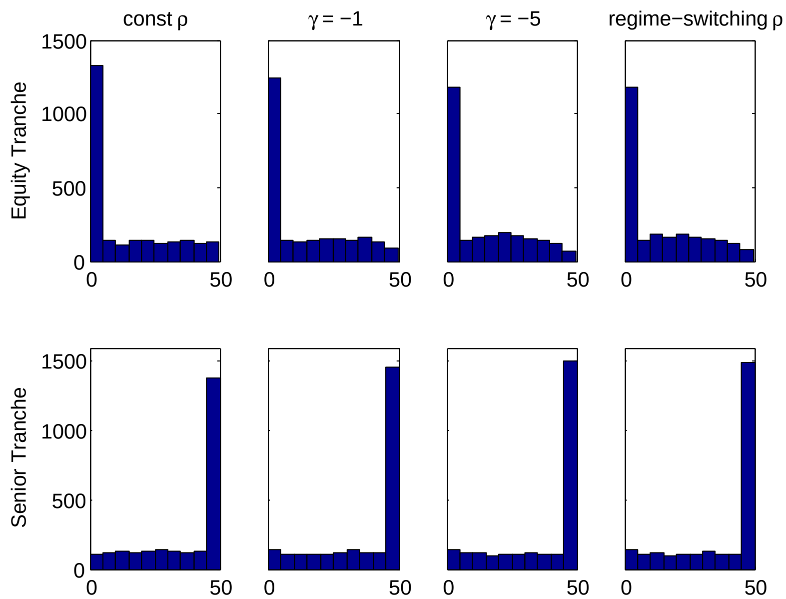

Figure 3.15: **The Histograms of Tranche Prices Across Vintages**

The state variable Z_v is assumed to have no serial correlation. The tranche price is simulated as the expected value of cash flow received by a certain tranche of a certain vintage. In our simulations, we construct 2500 vintages, so there are 2500 prices for each tranche. In the first column, the value of ρ_v is fixed at 0.5. In the second and third columns, the value of ρ_v is changing between 0.3 and 0.7 according to equation (3.53). In the fourth column, the value of ρ_v assumes a regime-switching model specified by equation (3.50), where $\rho_l = 0.3$, $\rho_u = 0.7$. The unconditional default probability is assumed to be 50%.

distributions are indeed the same. This is likely due to the fact that a change in serial correlation only affects the conditional distribution of Z_v but not its unconditional distribution. Therefore, only the conditional distribution of tranche prices changes while the unconditional distribution remains the same.

3.3.3 Logistic Transitional Default Correlation

A regime-switching model for ρ_v as described in the last section is intuitive, consistent with empirical findings, and allows efficient calibration to market prices. (See for example, Andersen and Sidenius (2005).) However, it does have one major drawback in that the true default correlation is unlikely to assume just two values. To address this issue, we allow the default correlation parameter to change between two values "smoothly", meaning that ρ_v is a smooth function of Z_v. Specifically, we assume a logistic transition model for the default correlation parameter ρ_v,

$$\rho_v = \rho_l + (\rho_u - \rho_l)\frac{1}{1 + \exp(-\gamma(Z_v - c))}. \tag{3.53}$$

With this setup, ρ_v can take any values between ρ_l and ρ_u according to the value of Z_v. When the value of γ is negative, the value of ρ_v decreases smoothly towards ρ_l as Z_v increases. As Z_v decreases, ρ_v increases smoothly towards ρ_u. Note that γ determines the smoothness of the

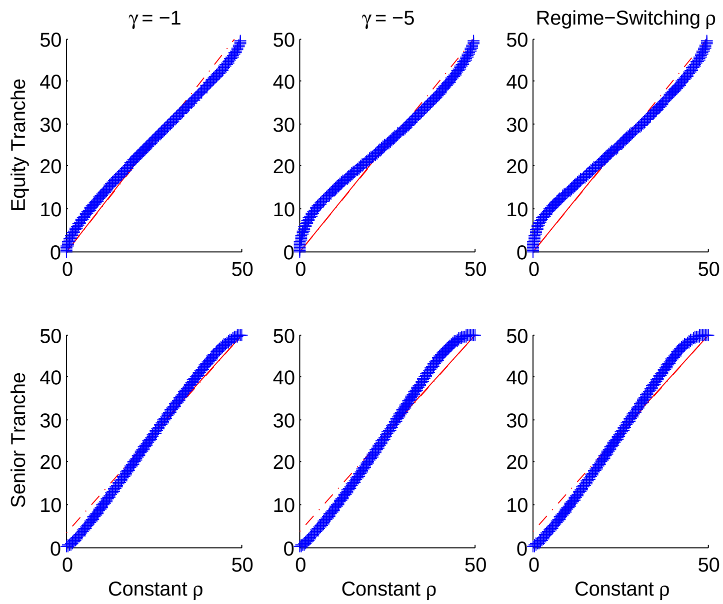

Figure 3.16: **QQ Plot of Constant ρ_v vs Varying ρ_v**

In these figures, we display the quantile-quantile plot of tranche prices when default correlation is constant versus tranche prices when default correlation is stochastic. In all figures, the horizontal axis denotes the case where ρ_v is constant. The vertical axis denotes the case where ρ_v varies across vintages. In the first column, the value of ρ_v has a logistic transition model according to equation (3.53), where $\gamma = -1$. In the second column, the value of ρ_v has a logistic transition model according to equation (3.53), where $\gamma = -5$. In the third column, the value of ρ_v has a regime-switching model and varies between 0.3 and 0.7 according to equation (3.50). The state variable Z_v is assumed to have no serial correlation. The unconditional default probability is assumed to be 50%.

transition of ρ_v. The smaller the absolute value of γ, the smoother is the transition. For $\gamma \to -\infty$, this model converges to the regime-switching model described in the sections above. On the other hand, when γ is zero, it degenerates into the constant ρ_v model.

In our paper, we fix $\rho_l = 0.3$, $\rho_u = 0.7$, and $c = 0$. Note that by using this specification, we manage to keep the mean value of $\rho_v = 0.5$. The relationship between ρ_v and Z_v when $\gamma = -1, -5$ can be seen in Figure 3.18. We follow the same steps as in Sections 3.3.2 except that we replace the regime-switching model for ρ_v with the logistic transition model. We let Z_v follow an $AR(1)$ process according to equation (3.49) with $\phi = 0.5$.

The histograms of the tranche prices are displayed in the second and third columns of Figure 3.13. The QQ plots of the distributions of tranche prices when ρ_v assumes logistic transition models against the distribution of tranche prices when ρ_v is constant are displayed in the first and second columns of Figure 3.14. As can be seen, compared with the case where ρ_v is constant, the equity tranche prices exhibit thinner tails while the senior tranche prices exhibit fatter tails. This is consistent with our findings in Section 3.3.2. Also note that the greater the absolute value of γ in equation (3.53), or in other words, the less smooth the transition of ρ_v, the thinner the tails of the distribution of equity tranche prices, and the fatter the tails of the distribution of senior tranche prices. As the absolute value of γ increases to infinity, the tranche prices converge in distribution to those when ρ_v follows a regime-switching model. As the absolute value of γ decreases to 0, the tranche prices converge in distribution to the those when ρ_v is constant.

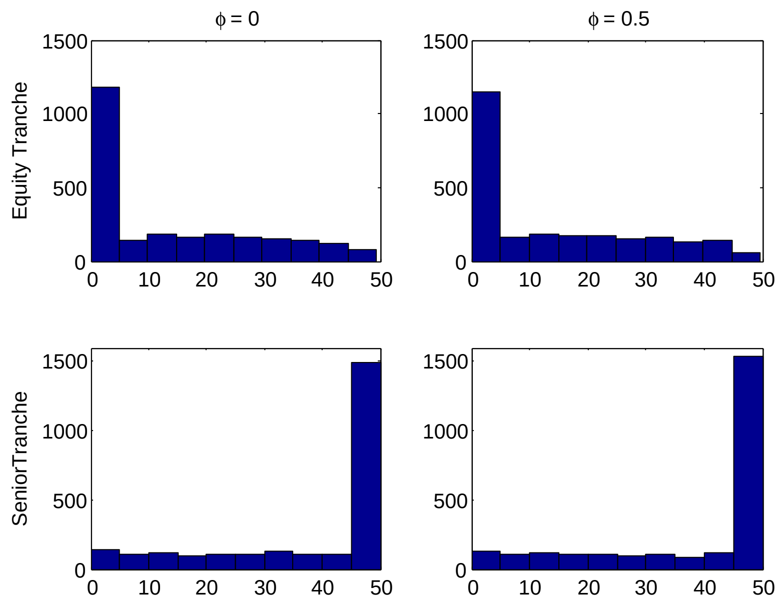

Figure 3.17: **Histograms of Tranche Prices when Z_v Has Different Values of Serial Correlation**

In these figures, we plot the histogram of tranche prices across vintages. The tranche price is simulated as the expected value of cash flow received by a certain tranche of a certain vintage. In our simulations, we construct 2500 vintages, so there are 2500 prices for each tranche. In the first column, the state variable Z_v is white noise. In the second column, Z_v follows an $AR(1)$ process described in equation (3.49) with $\phi = 0.5$. In both columns, ρ_v assumes a regime switching model as is described in equation (3.50). The unconditional default probability is assumed to be 50%.

We also consider the case where Z_v does not have serial correlation. The results are shown in the second and third columns of Figure 3.15. Their QQ plots are shown in the first and second columns of Figure 3.16. As can be seen, the pattern of the histograms of tranche prices when Z_v are white noise is very similar to the case where Z_v has serial correlation. This is also consistent with our findings in Section 3.3.2.

3.4 Impact of Dynamics of Default Correlation on High-Frequency Tranches

We study now the situation where the full risk spectrum for investors is subdivided into a large number tranches each of thin width. We call such a situation high frequency tranching. As pointed out by Gorton (2008), subprime MBS portfolios included very thin tranches.

3.4.1 Sensitivity of High-Frequency Tranche Prices to Default Correlation

As illustrated in the section above, imposing dynamics on default correlation has different effects on the price distribution of different tranches. This is because the sensitivity of tranche prices to a change in default correlation is different for each tranche. For example, holding everything else

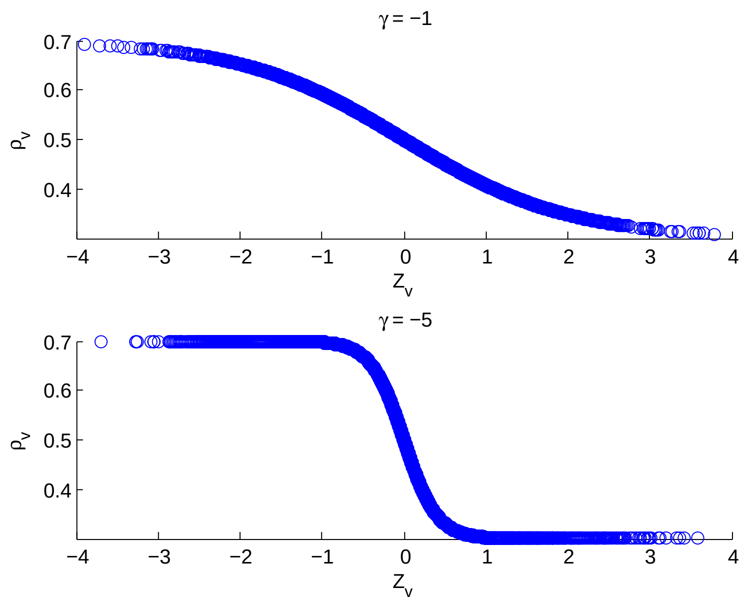

Figure 3.18: **Relationship Between ρ_v and Z_v**

In this figure, 10000 Z_v are generated with i.i.d standard Gaussian distribution. The corresponding values of ρ_v are calculated according to equation (3.53), where $\rho_l = 0.3$, $\rho_u = 0.7$, $c = 0$.

constant, the value of an equity tranche increases when default correlation increases, while the value of a senior tranche decreases.

Since the sensitivity of price with respect to default correlation of a tranche with arbitrary seniority and size is unclear, we cannot deduce from Section 3.3 the impact of the dynamics of default correlation on an arbitrary tranche. To better understand how a change in default correlation affects tranche prices, we slice each MBS portfolio into many thin tranches instead of just two tranches. We estimate numerically the prices of all these high-frequency tranches of a fixed vintage for different values of default correlation. These prices can be used to calculate the price of any larger tranche that consists of a set of small tranches by simply summing up the prices of these smaller tranches.

Specifically, we slice each MBS into 100 equal-sized small tranches. Each tranche has a principal of $\$1$ at initiation. For different values of ρ_v, the expected values of cash flows received by these tranches are estimated using Monte Carlo simulation. Figure 3.19 shows the case where we fix the unconditional default probability at 50%. The figure shows that the prices of all the tranches of relatively high seniority decrease as ρ_v increases and the prices of tranches of relatively low seniority increase as ρ_v increases.

We also estimated the high-frequency tranche prices when the unconditional default probability assumes other values. These can be found in Figure 3.20. In all cases, the tranche price is positively correlated with default correlation when the tranche seniority is below a certain level, which we call transition level. The tranche price is negatively correlated with default correlation when the seniority is above that level. The position of the transition level (its distance from the bottom tranche) is determined by the unconditional default probability. The higher the unconditional default probability, the higher the transition level. In other words, the price as a function of

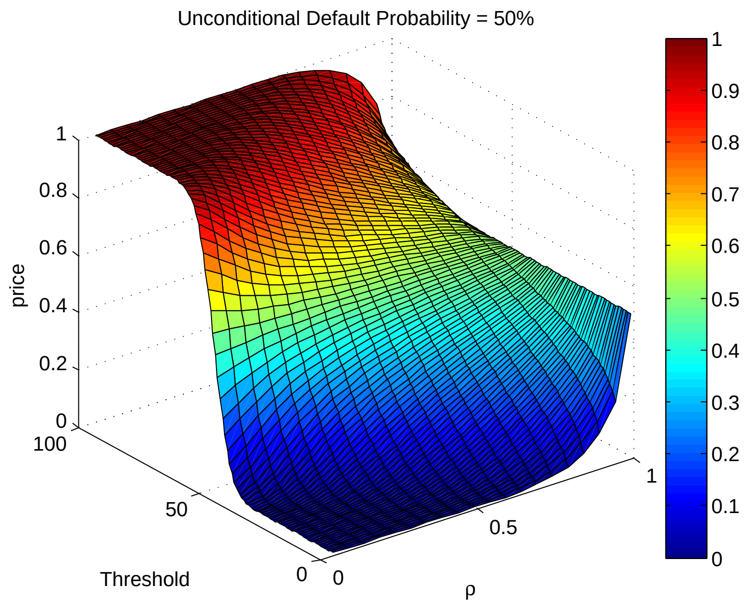

Figure 3.19: **The Sensitivity of Tranche Price to ρ**

seniority has a transition with locus at a certain threshold whose position is positively correlated with the value of unconditional default probability.

3.4.2 Sensitivity of High Frequency Tranche Prices to Dynamic Default Correlation

In the previous section, we have fixed the vintage v and analyzed the relationship between high-frequency tranche prices and default correlation in a static manner. Now we introduce the vintage dimension into the analysis. extending results in Sections 3.3.1 and 3.3.2 to the case of 100 high-frequency tranches of MBS rather than just two tranches.

Again, we generate 2500 vintages of mortgages whose unconditional default probability is fixed at 50%. We assume the state variable Z_v to be serially correlated. We assume that Z_v follows an $AR(1)$ process described in equation (3.49), where the $AR(1)$ coefficient $\phi = 0.5$. We first let the default correlation parameter $\rho_v = 0.5$. The expected values of the cash flow received by each tranche at each vintage $v = 1, 2, \cdots, 2500$ are estimated using Monte Carlo simulations. We present here series of prices of five representative tranches (out of 100 tranches in total) — the tranches at 10% level from the bottom of the portfolio, at 30% level, at 50% level, at 70% level, and at 90% level.

We then let ρ_v change across vintages and examine how the dynamics of ρ_v affect the distribution of tranche prices across vintages.In Figure 3.21, we compare the histograms of tranche prices when the default correlation parameter ρ_v is fixed at 0.5 and when ρ_v switches between 0.3 and 0.7 according to equation (3.50) or (3.53). The corresponding QQ plots are shown in Figure 3.22. We can see from this figure that the distributions of tranche prices are obviously different for a constant ρ_v and for a varying ρ_v. Furthermore, the differences in distribution vary across tranches. For example, for a thin slice of MBS tranche created at the 50% level from the bottom of a mortgage pool,

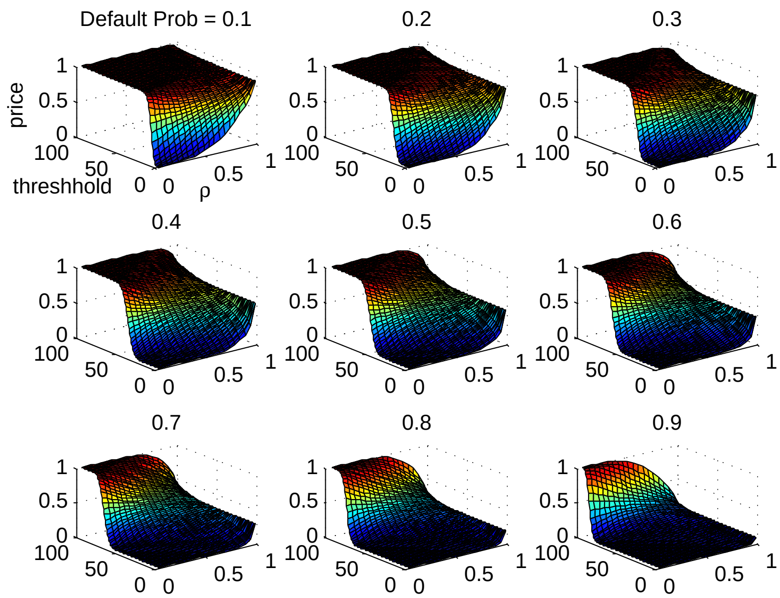

Figure 3.20: **The Sensitivity of Tranche Price to ρ**

the distribution of its tranche prices across vintages have thinner left tail when ρ_v is time-varying compared with when ρ_v is constant. On the other hand, for a thin MBS tranche at 90% level of a mortgage pool, the distribution of its tranche prices has fatter tails when ρ_v is dynamic relative to when ρ_v is constant. Also notice that when ρ_v has a logistic transition structure, the resulting tranche prices converge in distribution to constant-ρ_v tranche prices as γ approaches zero. They converge in distribution to regime-switching-ρ_v tranche prices as γ decreases to negative infinity.

A noteworthy observation is that for certain tranches, such as the 90% tranche in Figure 3.21, the distortion of the price distribution caused by the dynamics of ρ_v is relatively small. This fact suggests that for certain tranches, the simple Gaussian copula model is good enough to model tranche prices as long as it correctly specifies the mean value of default correlation. On the other hand, for certain tranches such as the 50% tranche, the distortion of price distribution from the baseline model is relatively big, suggesting that a simple Gaussian Copula model is inappropriate to capture the distribution of tranche prices when default correlation is indeed varying across vintages.

This observation also holds when Z_v does not exhibit serial correlation. This can be seen in Figure 3.23 and 3.24 where we set Z_v to be a series of i.i.d. standard Gaussian variables.

3.5 Conclusion

We find that introducing dynamics into the state variable does not affect the distribution of tranche prices across vintages. On the other hand, the dynamics of default correlation do influence the distribution of tranche prices. The smoothness of the change of default correlation also matters. If the default correlation changes smoothly according to the value of Z_v, the distortion tends to be small, and vice versa.

The distortions of distributions caused by the dynamics of default correlation parameter ρ_v is

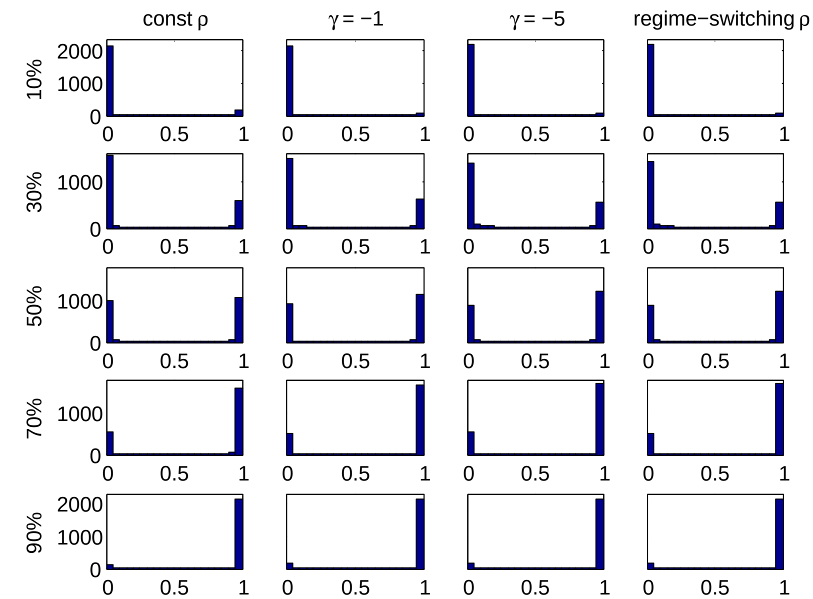

Figure 3.21: **The Histograms of High-Frequency Tranche Prices Over Vintages**

In the first column, the value of ρ_v is fixed at 0.5. In the second and third column, the value of ρ_v changes between 0.3 and 0.7 according to equation (3.53), with $\gamma = -1$ in the second column and $\gamma = -5$ in the third column. In the fourth column, the value of ρ_v switches between 0.3 and 0.7 according to equation (3.50). The state variable Z_v is assumed to follow an $AR(1)$ process where the first order coefficient $\phi = 0.5$. The unconditional default probability is assumed to be 50%.

determined by seniority of the tranche. In general, if the tranche price is positively correlated with default correlation, the distribution of the price of the tranche tends to have fatter tails when ρ_v is negatively correlated with the conditional default probability. The exact opposite is true for a tranche whose price is negatively correlated with default correlation.

The findings above prompt us to study the sensitivity of high-frequency tranching prices to change in default correlation. We find that the tranche price as a function of the seniority is a transition with locus at a threshold whose distance from the bottom of the MBS portfolio is positively correlated the unconditional default probability.

The results in this paper have important implications for risk management. If the default correlation changes across vintages either in a regime-switching model or in a logistic transition, our findings suggest that a Gaussian copula model that assumes a constant default correlation ρ_v underestimates the default risk of a senior tranche while overestimating the default risk for an equity tranche, even if it correctly characterizes the mean value of ρ_v and the true unconditional default probability.

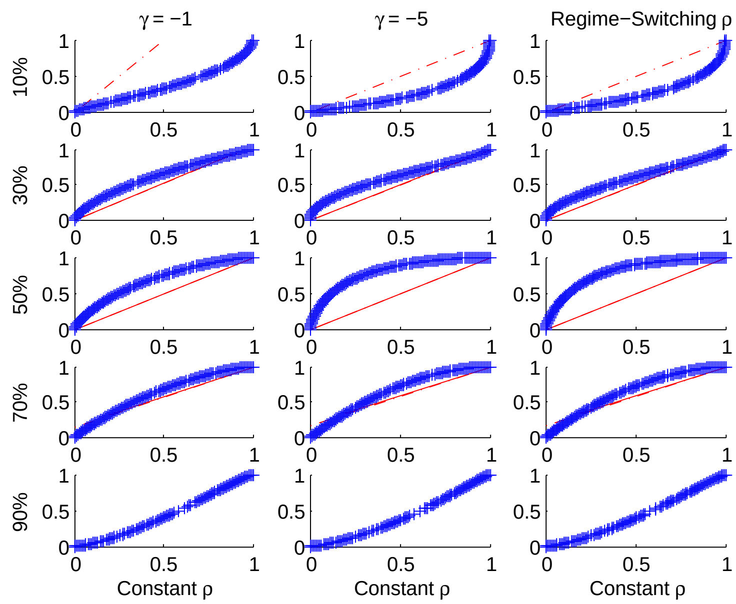

Figure 3.22: **QQ Plot of Constant ρ_v vs Varying ρ_v**

In these figures, we display the quantile-quantile plot of high-frequency tranche prices when default correlation is constant versus tranche prices when default correlation is stochastic. In all figures, the horizontal axis denotes the case where ρ_v is constant. The vertical axis denotes the cases where ρ_v varies across vintage. In the first column, the value of ρ_v has a logistic transition model according to equation (3.53), where $\gamma = -1$. In the second column, the value of ρ_v has a logistic transition model according to equation (3.53), where $\gamma = -5$. In the third column, the value of ρ_v has a regime-switching model and varies between 0.3 and 0.7 according to equation (3.50). The number in percentage in front of each row indicates the position of that tranche. For example, the number 10% on the first row means that the tranche of that row is located at the 10% position from the bottom of the MBS portfolio. The state variable Z_v is assumed to follow an $AR(1)$ process where the first order coefficient $\phi = 0.5$. The unconditional default probability is assumed to be 50%.

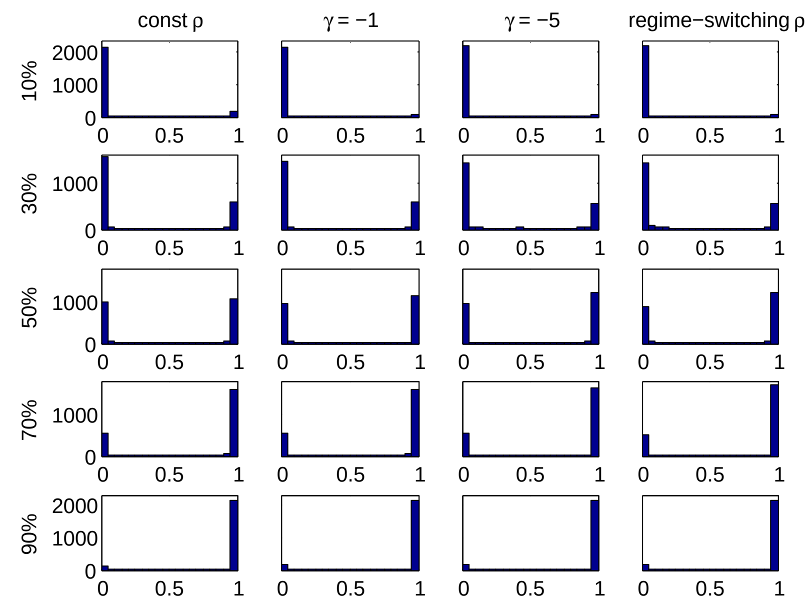

Figure 3.23: **The Histograms of High-Frequency Tranche Prices Over Vintages**

In the first column, the value of ρ_v is fixed at 0.5. In the second and third column, the value of ρ_v changes between 0.3 and 0.7 according to equation (3.53), with $\gamma = -1$ in the second column and $\gamma = -5$ in the third column. In the fourth column, the value of ρ_v switches between 0.3 and 0.7 according to equation (3.50). The state variable Z_v is assumed to have no serial correlation. The unconditional default probability is assumed to be 50%.

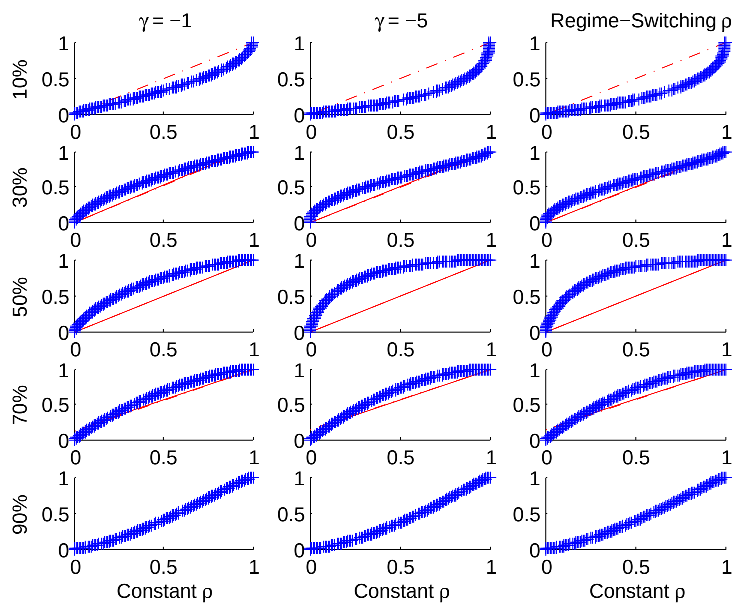

Figure 3.24: **QQ Plot of Constant ρ_v vs Varying ρ_v**

In these figures, we display the quantile-quantile plot of high-frequency tranche prices when default correlation is constant versus tranche prices when default correlation is stochastic. In all figures, the horizontal axis denotes the case where ρ_v is constant. The vertical axis denotes the cases where ρ_v varies across vintage. In the first column, the value of ρ_v has a logistic transition model according to equation (3.53), where $\gamma = -1$. In the second column, the value of ρ_v has a logistic transition model according to equation (3.53), where $\gamma = -5$. In the third column, the value of ρ_v has a regime-switching model and varies between 0.3 and 0.7 according to equation (3.50). The number in percentage in front of each row indicates the position of that tranche. For example, the number 10% on the first row means that the tranche of that row is located at the 10% position from the bottom of the MBS portfolio. The state variable Z_v is assumed to have no serial correlation. The unconditional default probability is assumed to be 50%.

Chapter 4. Predicting Equity Premia with Variance Risk Premia

4.1 Introduction

Whether equity premia are predictable is an everlasting topic in finance. While many return predictors such as price-earning ratio, dividend-price ratio, and book-to-market ratio have been proposed in the past, the predictive power of these variables are usually low. The prediction regression's in-sample R-squares are often below 1% when monthly data are used (for example, see Welch and Goyal (2008)). A recent paper by Bollerslev, Tauchen, and Zhou (2009) however, finds that the variance risk premium of the S&P500 can explain a large portion of variations in excess market returns. The good predictive power of the market variance risk premium for return prompts us to test the forecast power of individual stock's variance risk premia for monthly equity premia.

Similar to the definition of equity premia, variance risk premium is the difference between risk neutral and objective expectation of future variance. Bollerslev, Tauchen, and Zhou (2009) construct variance risk premium as the difference between model-free implied variance and ex post realized variance. Model-free implied variance was introduced by Britten-Jones and Neuberger (2000). It is the variance of the underlying asset implied by observed option prices under no-arbitrage conditions. Unlike Black-Scholes implied variance, the derivation of model-free implied variance requires no assumption of the option pricing model. Jiang and Tian (2005) propose a simple method to numerically calculate model-free implied variance from observed option prices.

Bollerslev, Tauchen, and Zhou (2009) use the sum of five-minute intraday squared returns as the measure for ex post realized variance. While model-free implied variance can be viewed as a good approximation of risk neutral expectation of future variance (for example, see Jiang and Tian (2005) and Bollerslev, Gibson, and Zhou (2008)), the ex post realized variance is likely not a good proxy for objective expectation of future variance. This prompts us to decompose their measure of variance risk premium into two components. The first component is defined as the difference between implied and expected value of realized variance. The second component represents the difference between expected value and ex post realized variance.

In this paper, we test the predictive power of variance risk premia for returns using individual stock data. The usage of individual stock data has advantages and disadvantages. The major disadvantage is that we do not have access to the data of model-free implied variances for individual stocks. We have to use Black-Scholes implied variance as a substitute, which is an biased estimate of risk neutral expectation of future variance and contains less information than model-free implied variance (Day and Lewis (1992), Lamoureux and Lastrapes (1993), and Jiang and Tian (2005)). The main advantage is that the sample size increases significantly as the number of stocks increases. This helps to avoid spurious results that can easily results from a small sample size.

Despite the fact that we use the Black-Scholes implied variance instead of the model-free implied variance, we find that the variance risk premia constructed as the difference between implied variance and ex-post realized variance of individual stocks have predictive powers for an individual stock's monthly excess returns. Also, the forecast accuracy increases greatly when the first two lags of the variance risk premia are included as additional predictors. Moreover, each component of the variance risk premium serves as a predictor by itself. When both components and their first two lags are used as explanatory variables, they can explain about 8% of the variations in the monthly equity premium.

The outline of the paper is as follows. In Section 4.2, we describe the data we use in the chapter. We examine the predictive power of variance risk premia and theirs lags in Section 4.3. In Section 4.4, we decompose the variance risk premium into two parts and show their predictive power for returns. Section 4.5 explores a method to estimate the impact of the two components on future equity premia without suffering measurement errors introduced in the decomposition process. In Section 4.6 we summarize the main conclusions.

4.2 Data Description

We use non-overlapping monthly data. The whole sample period is from January 1995 to July 2009. The sample period in which we study the predictive power of variance premium on equity return is from January 2000 to July 2009. We focus on our study on stocks which were/are in the Dow Jones Industrial Average (DJIA) index during the sample period. We leave out the stocks whose NYSE ticker symbol had changed during the period.[6] The ticker symbols and names of the selected stocks are listed in Table 4.2. To ensure the non-overlapping property, we only use the implied variance data on the last trading day of each month. We use realized kernels computed from two-minute intraday return data as daily realized variance. Our choice of kernel is Tukey-Hanning with order 2 introduced by Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008). The average daily value is used for the realized variance in each month. We list the description and source of each variable below. Summary statistics of these variables are given in Table 4.3.

1. Book-to-Market Ratio:

   The book value is the difference between total assets and total liabilities of a firm. The data of total assets and total liabilities are extracted from the Standard $\&$ Poor's Compustat North America database. These are quarterly data. We linearly interpolate to obtain monthly values.

   The market value is the product of number of outstanding common shares and current stock price. The data of number on the outstanding common shares are extracted from the Compustat database. These are quarterly data. We linearly interpolate to obtain monthly values.

   The book-to-market ratio is the ratio of book value to market value for each stock. Following Kothari and Shanken (1997) and Welch and Goyal (2008), for the months from March to

[6] A change in a stock's ticker symbol implies that the corresponding company has merged with another company. When this happens, the stock's realized variance series exhibits a structure break at the time of the mergence, which poses difficulty for us to estimate LRGARCH model.

Table 4.2: **List of Stocks.**

Ticker	Name	Implied Variance	Realized Variance	Variance Premium	Equity Premium (%)
AA	Alcoa Inc	7.302	5.495	1.794	−1.364
AXP	American Express Co	6.019	4.843	1.160	−0.416
BA	Boeing Co	3.921	2.723	1.193	0.073
CAT	Caterpillar Inc	4.308	3.324	0.965	0.268
DD	DuPont Co	3.124	2.596	0.529	−0.450
DIS	Disney Walt Co	4.457	2.950	1.522	0.041
GE	General Electric Co	4.196	3.428	0.746	−1.269
HD	Home Depot Inc	4.490	3.407	1.097	−0.900
IBM	Int'l Business Machines	2.896	1.913	0.992	−0.259
INTC	Intel Corp	5.801	3.866	1.969	−0.779
JNJ	Johnson & Johnson	1.798	1.328	0.479	−0.053
KO	Coca Cola Co	1.866	1.421	0.451	0.032
MCD	McDonalds Corp	3.251	2.251	1.008	0.848
MMM	Minnesota Mining & Mfg Co	2.401	1.709	0.697	0.018
MO	Philip Morris Cos Inc	3.012	1.901	1.123	0.723
MRK	Merck & Co Inc	3.360	2.637	0.694	−0.637
MSFT	Microsoft Corp	3.642	2.440	1.209	−0.346
PFE	Pfizer Inc	2.944	2.107	0.836	−0.977
PG	Procter & Gamble Co	1.694	1.267	0.427	0.276
UTX	United Technologies Corp	3.102	2.158	0.951	0.484
WMT	Wal Mart Stores Inc	2.546	1.881	0.675	−0.278

December, this is computed by dividing book value at the end of the previous year by market value at the end of the current month. For the months of January and February, this is computed by dividing book value at the end of two years ago by market value at the end of the current month.

2. Dividend-Price Ratio:

   The dividend data are extracted from the Compustat database. We use a one-year moving average of these data as our variable for dividends. The dividend price ratio is the log difference between dividends and current price. The dividend-yield ratio is the log difference between dividends and lagged price.

3. Dow Jones Industrial Average Index Log Return:

   The DJIA log-return variable is constructed using DJIA index monthly close price data, which are obtained from Compustat North America database. The DJIA log-return variable is scaled by $\sqrt{T}$, where T is the number of trading days in the month. It is scaled this way to match the scaled version of the realized variance variable.

4. Earning-Price Ratio:

   The earning data are extracted from the Compustat database. We use a one-year moving average of these data as our variable for earning. The earning price ratio is the ratio between earning and current price.

5. Equity Premium:

Table 4.3: **Summary Statistics.**

Variable	Sample Start	Sample End	Mean	Std. Dev.	Min	Max
Book-to-mkt ratio	Jan. 2000	Jul. 2009	0.266	0.171	−0.052	2.377
Dividend-price ratio	Jan. 2000	Jul. 2009	−4.156	1.041	−10.000	−1.917
DJIA Log Return	Jan. 2000	Jul. 2009	−0.197	0.460	−15.153	10.079
Earning-price ratio	Jan. 2000	Jul. 2009	0.054	0.028	−0.217	0.215
Equity premium	Jan. 1995	Jul. 2009	−0.065	1.891	−14.073	13.583
Implied variance	Jan. 2000	Jul. 2009	4.177	4.036	0.431	45.875
Realized variance	Jan. 1995	Jul. 2009	2.862	3.474	0.163	62.730
VP	Jan. 2000	Jul. 2009	1.087	2.201	−29.083	21.979
VP^a	Jan. 2000	Jul. 2009	1.187	1.936	−8.305	27.440
VP^b	Jan. 2000	Jul. 2009	−0.101	1.531	−29.387	6.056

Equity premium is the difference between monthly stock log-return and market risk-free rate. The risk-free rate is the one-month average U.S. treasure bill rate. Data are collected from the Center for Research in Security Prices (CRSP) database.

The equity premium variable is scaled by $\sqrt{T}$, where T is the number of trading days in the month. It is scaled this way to match the scaled version of realized variance variable.

6. Implied Variance:

   Original data of implied volatility are obtained from Bloomberg under the variable name "History_PUT_IMP_VOL". The description of the data from Bloomberg is:

   > RK075 - HIST_PUT_IMP_VOL (Hist. Put Implied Volatility) Implied volatility calculated from a weighted average of the volatilities of the two closest options. For all securities, the contract used is the closest pricing contract month that is expiring at least 20 business days out from today.

   These daily data are collapsed to monthly data by keeping only the last entry of the previous month as the value for last month. The raw data represent the implied volatility for the coming year. To use these data in the paper, we first square them to obtain implied variance variable. We then divide this variable by 252 to get average daily implied variance.

7. Realized Variance:

   We use trade and quote data to construct daily realized variance according to Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008). The trade and quote data are obtained from the Trade and Quote (TAQ) database through Wharton Research Data Service (WRDS). The monthly variable of realized variance represents the sample average of daily realized variance.

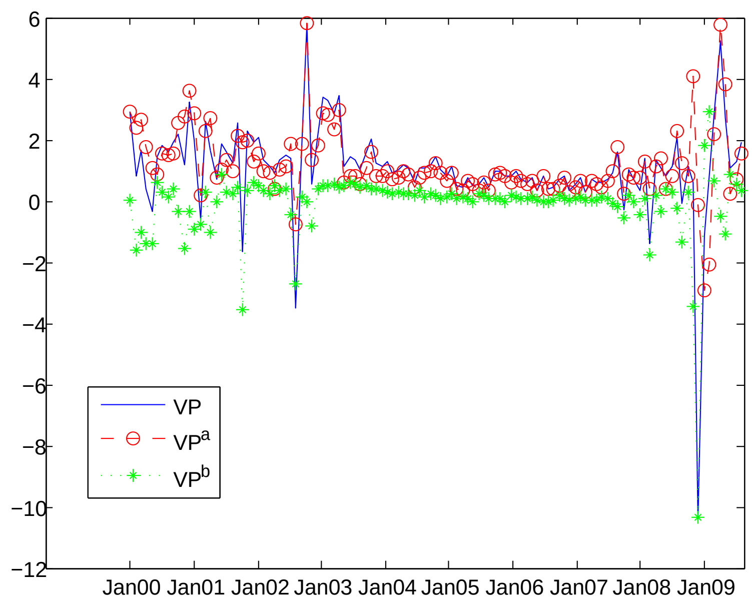

Figure 4.25: **Variance Risk Premium and Its Components**

The figure shows variance risk premium variable we constructed according to equation (4.54) and its two components defined in equation (4.55). The sample period extends from January 2000 to July 2009.

4.3 Variance Risk Premium

Following Bollerslev, Tauchen, and Zhou (2009), we start by constructing a series of variance premia for each stock i according to the equation:

$$VP_{i,t} = IV_{i,t} - RV_{i,t-1}, \tag{4.54}$$

where $IV_{i,t}$ is the implied variance of stock i at the last trading day of month $t-1$, normalized to daily value; $RV_{i,t-1}$ is the average realized variance of stock i in month $t-1$. This definition ensures that the predictor is observable at time t. Summary statistics of this variable are given in Table 4.3. The cross sectional mean of the series over time is plotted in Figure 4.25. The time series mean of the variance risk premium for each individual stock is listed in Table 4.2. It is noteworthy that all of the stocks in our sample have positive variance risk premia on average.

We run a pooled linear regression using $VP_{i,t}$ as explanatory variable to forecast individual stock excess return. The result is given in Column A of Table 4.4. Robust clustered standard errors (Rogers (1993)) are reported in all of our pooled OLS regressions[7]. The estimated coefficient of variance risk premium is positive and significant at 5% level. This implies that a positive (negative) variance risk premium leads to higher (lower) excess return for individual stocks. The changes in variance premium explain about 0.7% of the variations in future monthly returns. This finding is consistent with Bollerslev, Tauchen, and Zhou (2009) who also report a positive coefficient of the variance risk premium when using monthly data of $S\&P500$ index.

[7]Bollerslev, Tauchen, and Zhou (2009) show that, with finance panel data set like the one we have for this paper, clustered standard errors are unbiased as they account for the residual dependence created by the firm effect, while standard and panel data version of Newey-West (Newey and West (1987)) error estimators are biased.

Table 4.4: **Predictive Power of Variance Risk Premium.**

Column	A	B	C	D	E
R^2	0.007	0.038	0.005	0.011	0.055
Coeff.(S.E.)					
VP_t	$0.073(0.027)$	$0.041(0.026)$	-	-	$0.061(0.024)$
VP_{t-1}	-	$0.125(0.028)$	-	-	$0.140(0.026)$
VP_{t-2}	-	$0.061(0.034)$	-	-	$0.078(0.034)$
VP'_t	-	-	-	$0.001(0.009)$	-
VP'_{t-1}	-	-	-	$0.037(0.016)$	-
VP'_{t-2}	-	-	-	$0.007(0.012)$	-
Book-to-mkt ratio	-	-	$-0.270(0.234)$	-	$-0.755(0.212)$
Dividend-price ratio	-	-	$0.127(0.055)$	-	$0.204(0.051)$
Earning-price ratio	-	-	$0.675(2.384)$	-	$2.679(1.623)$
Sample size	2145	2373	2394	2331	2373

This table reports estimated coefficients from pooled OLS regressions. The dependent variable is the equity premium of individual stocks. The sample period extends from January 2000 to July 2009. Robust-clustered standard errors are reported in the parentheses.[8] Variable VP is constructed according to equation (4.54). Variable VP' is the difference between implied variance and ex-post squared return. Other explanatory variables are defined in Section 4.2.

The current value of the variance risk premium by itself is a good predictive factor for individual stock returns. For comparison, we use more conventional predictive variables such as book-to-market ratio, earning-price ratio, and dividend-price ratio to forecast equity premium. The result can be seen in Column C of Table 4.4. Variance risk premia explain more variations in individual stock excess returns than all those three traditionally popular predictors combined.

Moreover, when we include lagged values of variance risk premia as explanatory variables, the forecast accuracy increases significantly. The regression result is given in Column B of Table 4.4. The R^2 equals 3.8% in this case, indicating a better monthly stock return predictive power for variance risk premium than most other predictors in the literature. (For example, see Welch and Goyal (2008).) The predictive power of lagged values of variance risk premia implies it has both short run and long run impacts on excess return. Similar findings have been proposed by Adrian and Rosenberg (2008) who find that volatility risk can be broken into short run and long run components to predict equity premia. It is noteworthy that most of the predictive power of variance risk premium stems from its first order lag. In this regression, neither the current value of variance risk premium nor its second order lag has estimated coefficients significant at 5% level, while the first lag of variance risk premium has a coefficient significant at 1% level.

Note that the usage of realized variance constructed by high frequency data is crucial in forecasting excess return as is pointed out by Bollerslev, Tauchen, and Zhou (2009). To test the robustness of our finding, we create another measure of variance risk premium with name VP', using monthly sums of squared returns to replace $RV_{i,t-1}$ in equation (4.54). The regression result is reported in Column D of Table 4.4. While $VP'_{i,t}$ and its first two lags still have predictive power, it is much lower than that of $VP_{i,t}$ and its lags. Moreover, the estimated coefficients before $VP'_{i,t}$ and $VP'_{i,t-2}$ are not statistically significant.

We also conduct the "kitchen-sink" regression where we use both variance risk premia and conventional predictive variables as explanatory variables. The result is reported in Column E of Table 4.4. We find that adding conventional predictors into the regression does not reduce the predictive power of variance risk premia. On the contrary, the predictive power of variance risk premia appears to increase in this case. The coefficients of current variance risk premium and its two lags are all significant at 5% level and their magnitudes are greater than before.

4.4 Variance Risk Premium Decomposition

While we have created a variance risk premium variable that serves as a good return predictor, the variable we construct according to equation (4.54) is slightly different from the conventional definition of variance premium. Similar to the definition of the equity premium, the variance risk premium should be defined as the difference between risk neutral and objective expectations of future variance. While the implied variance in equation (4.54) can be viewed as a proxy for risk neutral expectation of the future stock return variance, the lagged value of realized variance is past variance rather than an expected value of future variance. In light of this, we decompose our variance risk premium variable in the following way:

$$VP_{i,t} = IV_{i,t} - RV_{i,t-1} = (IV_{i,t} - \widehat{RV_{i,t}}) + (\widehat{RV_{i,t}} - RV_{i,t-1}), \tag{4.55}$$

where $\widehat{RV_{i,t}}$ is the conditional expectation of realized variance for stock i in month t:

$$\widehat{RV_{i,t}} = \mathbb{E}(RV_{i,t}|\mathfrak{F}_{t-1}). \tag{4.56}$$

where $\mathfrak{F}_{t-1}$ denotes the collection of information up to time $t-1$. We define

$$VP^a_{i,t} := IV_{i,t} - \widehat{RV_{i,t}};\ VP^b_{i,t} := \widehat{RV_{i,t}} - RV_{i,t-1}. \tag{4.57}$$

This decomposition is quite straightforward. The first component, VP^a is by definition the actual variance risk premium if implied variance can be viewed as risk neutral expectation of future variance. The second component, VP^b is the expected change in future variance. While the first part clearly has the potential to predict stock excess returns, the second part has also been found to have return predictability in the literature (for example, see Adrian and Rosenberg (2008)).

In order to test the predictive power of each part, an essential task is to have a measure of the expectation of future realized variance. If we set the expectation as the realized variance in the past month so that

$$\widehat{RV_{i,t}} = RV_{i,t-1}, \tag{4.58}$$

our decomposition in equation (4.55) degenerates back to equation (4.54). However, a single lagged value of realized variance is likely a very noisy measure of expectation for future variance. It is better to utilize all the historical information available. There are many existing models for predicting future variance using historic stock price data, such as ARCH and GARCH. For example Day and Lewis (1992) and Lamoureux and Lastrapes (1993) used GARCH to predict future variances and compare them to implied variances. In this paper, we apply a model called Log Realized GARCH (LRGARCH). It is a model introduced by Hansen, Huang, and Shek (2011), which is used to jointly model return and realized measures of variance. One of the advantages of this model over ARCH and GARCH is that it can predict future realized variance by incorporating the information of both stock returns and historical realized variance.

We use a LRGARCH$(1,1)$ model with the following structure:

$$r_{i,t} = \mu + \sqrt{h_{i,t}}z_{i,t}, \tag{4.59}$$

$$\log(h_{i,t}) = \omega + \beta \log(h_{i,t-1}) + \gamma \log(RV_{i,t-1}), \tag{4.60}$$

$$\log(RV_{i,t}) = \xi + \phi \log(h_{i,t}) + \tau_1 z_{i,t} + \tau_2(z_{i,t}^2 - 1) + \sigma_u e_{i,t}. \tag{4.61}$$

where $r_{i,t}$ is stock i's excess return in month t, and $e_{i,t}$ are assumed to be i.i.d standard Gaussian variables. The numbers $\mu, \omega, \beta, \gamma, \xi, \phi, \tau_1, \tau_2, \sigma_u$ are model parameters. The realized variance in the next period can be predicted as follows:

$$\widehat{RV_{i,t}} = \mathbb{E}(RV_{i,t}|\mathfrak{F}_{t-1}) = \frac{\exp(\xi + \phi \log(h_{i,t}) + 0.5\sigma_u^2 + 0.5\tau_1^2 - \tau_2)}{\sqrt{(1 - 2\tau_2)}}. \tag{4.62}$$

To obtain $\widehat{RV_t}$, we need to estimate the LRGARCH model for each stock. In order to have enough degrees of freedom to estimate the LRGARCH parameters properly, we assume that these parameters are constant across stocks. We use a rolling window of 5 years to estimate the model at each month and obtain $\widehat{RV_t}$.[9] Once we have $\widehat{RV_t}$, we use them to decompose VP into VP^a and VP^b according to equation (4.55). The plots of the cross-sectional means of the two series are shown in Figure 4.25. Table 4.3 shows that VP^a is positive on average while VP^b has a mean close to zero. The predictive powers of VP^a and VP^b are tested by running pooled OLS regressions. The results are given in Table 4.5.

When VP_t^a is used as the only predictor, it can explain 0.5% of the variations in individual stock excess return (Column C of Table 4.5). Correspondingly, VP_t^b explains 0.1% of future return variations. Both of them have positive coefficient. However, their coefficient are not significant at 5% level. Jointly, VP_t^a and VP_t^b explain 0.7% of variations in equity premium (Column G of Table 4.5), indicating a similar predictive power as VP_t.

Even though the current values of VP^a and VP^b lack significant predictive power, when lagged values of them are added as additional predictors, their predictive power increases sharply. Variable VP_t^a and its two lags explain 5.3% of changes in future excess return (Column D of Table 4.5), which is better than VP and its two lags. This implies that constructing variance risk premium variable using predicted realized variance instead of ex-post realized variance improves its predictive power. Variable VP_t^b and its two lags explain 2% of changes in future excess return (Column D of Table 4.5). Put together, these two components and their lags exhibit very good predictive power. This is shown in Column H of Table 4.5. As can be seen, after we decompose VP_t and its lagged value into two parts, the R^2 increases from 3.7% to 8.4%.

Examining the estimated coefficients before the explanatory variables, we notice that the reason behind the jump in forecast accuracy is that the lags of VP_t^a and VP_t^b have opposite effects on future excess return. For example, the first lag of VP_t^a has a large positive impact on future return while the first lag of VP_t^a has a small negative impact. When they are summed together, the two forces offset each other, resulting in a small positive impact on future return from the first lag of VP_t. Therefore, the decomposition of VP is crucial for a better forecast of the equity premium.

4.5 Modified Log-linear Realized GARCH Model

While we have found outstanding return predictive power from the two components of the variance risk premium, there is one concern in the approach above. Since we need to estimate an

[9] We have data for stock excess return and realized variance back to January 1995.

LRGARCH first to decompose VP into two parts, additional estimation errors are introduced into the process. Simply speaking, the objective expectation of future variance $\widehat{RV_{i,t}}$ obtained from the two-step procedure contains significant measure errors. This may cause our estimation of the predictive power of each component of the variance risk premium to be biased. To address this issue, we design a procedure that combines estimation and prediction in one step.

We modify the LRGARCH model to include predictors in the return function. Specifically, the model we propose is the following:

$$r_t = \mu + \mathbf{X}\mathbf{a} + \sqrt{h_t} z_t, \tag{4.63}$$

$$\log(h_{i,t}) = \omega + \beta \log(h_{i,t-1}) + \gamma \log(RV_{i,t-1}), \tag{4.64}$$

$$\log(RV_{i,t}) = \xi + \phi \log(h_{i,t}) + \sigma_u e_{i,t}.^{10} \tag{4.65}$$

where $\mathbf{X}$ is a row vector of return predictors and $\mathbf{a}$ is a column vector of parameters. When we use two components of variance premium and their lags as forecast variables, the return function becomes

$$r_t = \mu + \sum_{j=1}^{3} a_j(VP^a_{t-j+1}) + \sum_{j=1}^{3} b_j(VP^b_{t-j+1}) + \sqrt{h_t} z_t. \tag{4.66}$$

where

$$VP^a_{i,t} = IV_{i,t} - \widehat{RV_{i,t}};\ VP^b_{i,t} = \widehat{RV_{i,t}} - RV_{i,t-1}, \tag{4.67}$$

where

$$\widehat{RV_{i,t}} = \mathbb{E}(RV_{i,t}|\mathfrak{F}_{t-1}) = \frac{\exp(\xi + \phi \log(h_{i,t}) + 0.5\sigma_u^2}{\sqrt{(1-2\tau_2)}}. \tag{4.68}$$

Estimating the vector $\mathbf{a}$ this way, we eliminate the measurement error problem associated with the predicted value of $\widehat{RV_{i,t}}$.

Using different variables as predictors, we estimate our model by maximum likelihood. The results are reported in Table 4.6. In Column A, we can see that VP_t and its first lags have a positive impact on excess return. Now the current value of VP has a negative but small and insignificant coefficient. The magnitude of the estimated coefficient of its first lag is positive and significant and slightly larger than that in the pooled OLS regression. Overall, VP_t and its lags explain 3.1% of the variations in excess return. The R^2 is only slightly lower than that in the pooled OLS regression. Note that the pooled OLS regression achieves the highest R^2 possible as it minimizes the sum of squared residuals.

When VP^a_t and its lags are used as regressors in the return function (Column B in Table 4.6), the result is also similar to that of the pooled OLS regression. VP^a_t now has a negative and significant coefficient. In the pooled OLS regression, this coefficient is negative and insignificant. Overall, VP^a_t is estimated to explain 4.2% of the variations in equity premium.

The estimated impact of VP^b_t and its lags on excess return is somewhat different from what the pooled OLS regression shows. In this case, VP^b_t has a positive but insignificant coefficient while in the pooled OLS the estimated coefficient is positive and significant. The coefficient of

[10] We have removed the leverage function from the measurement funciton (equation (4.65)). This is because when variance risk premia enter into the return function, the measurement function contains $\log(RV_{i,t})$ on the left-hand side and expected value of $RV_{i,t}$ on the right-hand side implicitly if the leverage function remains. This renders the underlying data generation process impossible.

VP^b_{t-1} is now positive but small and statistically insignificant while that coefficient is negative and significant in the pooled OLS. Despite those discrepancy, the main predictive power of VP^b still comes from its second order lag which has a large positive coefficient. This is the same as in the OLS regression. Overall, VP^b_t is estimated to explain 1.8% of variations in future excess return.

When both components of VP_t and their lags are used as predictors, the forecast accuracy again increase sharply. In Column D of Table 4.6, we see that those variables can explain 7.3% of variations in excess return. According to the best of our knowledge, no other variables have been found in the literature to have such a good predictive power for monthly stock excess returns.

According to Bollerslev, Tauchen, and Zhou (2009) and Bollerslev, Gibson, and Zhou (2008), the variance risk premium can predict stock returns because it is likely a proxy for time-varying risk aversion. While we use individual stock variance risk premia as predictors in this paper, it is possible that individual variance risk premia only serve as an approximation of the market variance risk premium. Therefore, they are also proxies for market risk aversion. They can predict individual stock equity premia because they can affect variations in the market risk premium.

To test the theory above, we form a new variable by subtracting Dow Jones Industrial Average index monthly log-returns from individual stock monthly log-returns. The resulting variable can be viewed as a proxy for idiosyncratic excess returns. If individual stocks' variance risk premia can predict equity premia only because they serve as proxies for market risk aversion, they should have no predictive power on idiosyncratic excess returns. We use pooled OLS to test the forecast power of variance risk premium variables on idiosyncratic excess returns. The results are reported in Table 4.7. It can be seen that the predictive power of variance premia has decreased sharply. VP_t, VP^a_t, and VP^b_t each alone have no predictive power for idiosyncratic excess returns. However, the second order lag of VP^a still has significant predictive power. When we use VP^a_t, VP^b_t, and their first two lags as predictors, they can explain about 2% of the variations in idiosyncratic excess return.

The findings above have two important implications. First, an individual stock's variance risk premium can predict excess returns mainly because it is correlated with market risk aversion or maybe some other market risk factors. Second, an individual stock's variance risk premium seems to also represent certain idiosyncratic characteristics of each stock. Thus, it appears to have some forecast power for the part of equity premium that is in excess of market risk premium.

4.6 Conclusion

In sum, we find that an individual stock's variance risk premium constructed as the difference between implied variance and ex post realized variance has good predictive power for the equity premium. The forecast accuracy can be greatly improved if this variable is properly decomposed into two parts. The predictive power of the variance risk premium mainly stems from its correlation with certain market risk factors such as time-varying risk aversion. Nevertheless, an individual stock's variance risk premium also represents certain idiosyncratic characteristics of that stock, and can explain future variations in the equity premium in excess of the market premium. Also, forecast accuracy increases when lags of the variance risk premium are added as additional predictors. This implies that there exist both short-run and long-run effects of variance risk premia on excess returns.

Table 4.5: **Predictive Power of Variance Premium Components.**

Column	A	B	C	D	E	F	G	H
R^2	0.007	0.037	0.005	0.053	0.001	0.020	0.007	0.084
Coeff.(S.E.)								
VP_t	0.070(0.026)	0.035(0.026)	-	-	-	-	-	-
VP_{t-1}	-	0.127(0.028)	-	-	-	-	-	-
VP_{t-2}	-	0.056(0.035)	-	-	-	-	-	-
VP^a_t	-	-	0.069(0.035)	$-0.028(0.036)$	-	-	0.079(0.036)	$-0.021(0.038)$
VP^a_{t-1}	-	-	-	0.259(0.027)	-	-	-	0.315(0.039)
VP^a_{t-2}	-	-	-	$-0.087(0.030)$	-	-	-	$-0.103(0.042)$
VP^b_t	-	-	-	-	0.041(0.019)	0.061(0.020)	0.062(0.020)	0.210(0.039)
VP^b_{t-1}	-	-	-	-	-	$-0.066(0.028)$	-	$-0.052(0.031)$
VP^b_{t-2}	-	-	-	-	-	0.168(0.036)	-	0.124(0.032)
Sample size	2415	2373	2415	2373	2415	2373	2415	2373

This table reports estimated coefficients from pooled OLS regressions. The dependent variable is the equity premium of individual stocks. The sample period extends from January 2000 to July 2009. Robust-clustered standard errors are reported in the parentheses. Variable VP is constructed according to equation (4.54). Variables VP^a and VP^b are defined in equation (4.55). They are estimated using an LRGARCH model described in equations (4.59) to (4.61). The estimated coefficients in the LRGARCH model are omitted for brevity.

Table 4.6: **Modified LRGARCH Model.**

Column	A	B	C	D
R^2	0.031	0.043	0.018	0.073
Coeff.(S.E.)				
VP_t	$-0.024(0.027)$	-	-	-
VP_{t-1}	$0.157(0.027)$	-	-	-
VP_{t-2}	$0.004(0.025)$	-	-	-
VP^a_t	-	$-0.062(0.032)$	-	$-0.053(0.032)$
VP^a_{t-1}	-	$0.247(0.035)$	-	$0.272(0.036)$
VP^a_{t-2}	-	$-0.093(0.031)$	-	$-0.106(0.032)$
VP^b_t	-	-	$0.035(0.053)$	$0.163(0.055)$
VP^b_{t-1}	-	-	$0.043(0.052)$	$-0.016(0.055)$
VP^b_{t-2}	-	-	$0.147(0.045)$	$0.152(0.045)$
μ	$-0.114(0.043)$	$-0.060(0.041)$	$-0.017(0.030)$	$-0.109(0.043)$
ω	$0.270(0.029)$	$0.266(0.029)$	$0.273(0.029)$	$0.276(0.029)$
β	$0.269(0.024)$	$0.276(0.023)$	$0.272(0.024)$	$0.269(0.024)$
γ	$0.626(0.030)$	$0.615(0.030)$	$0.627(0.030)$	$0.602(0.030)$
ξ	$-0.325(0.051)$	$-0.330(0.052)$	$-0.330(0.051)$	$-0.349(0.054)$
ϕ	$1.045(0.043)$	$1.059(0.044)$	$1.040(0.043)$	$1.085(0.047)$
σ^2_u	$0.212(0.006)$	$0.212(0.006)$	$0.212(0.006)$	$0.211(0.006)$
Sample size	2415	2415	2415	2415

This table reports estimated coefficients of modified LRGARCH model defined in equations (4.63) to (4.66). The sample period extends from January 2000 to July 2009. The coefficients are estimated by maximum likelihood. Hessian matrix are calculated numerically to obtain the standard errors for the estimated coefficients.

Table 4.7: **Predictive Power of Variance Premium Components on Idiosyncratic Excess Return.**

Column	A	B	C	D	E	F	G	H
R^2	0.002	0.011	0.002	0.014	0.000	0.004	0.002	0.020
Coeff.(S.E.)								
VP_t	0.027(0.025)	0.010(0.024)	-	-	-	-	-	-
VP_{t-1}	-	0.051(0.026)	-	-	-	-	-	-
VP_{t-2}	-	0.038(0.031)	-	-	-	-	-	-
VP_t^a	-	-	0.036(0.031)	$-0.013(0.030)$	-	-	0.037(0.033)	$-0.011(0.032)$
VP_{t-1}^a	-	-	-	0.106(0.033)	-	-	-	0.123(0.043)
VP_{t-2}^a	-	-	-	$-0.018(0.032)$	-	-	-	$-0.020(0.042)$
VP_t^b	-	-	-	-	$-0.001(0.019)$	0.012(0.020)	0.009(0.021)	0.068(0.037)
VP_{t-1}^b	-	-	-	-	-	$-0.033(0.029)$	-	$-0.017(0.030)$
VP_{t-2}^b	-	-	-	-	-	0.066(0.040)	-	0.052(0.039)
Sample size	2415	2373	2415	2373	2415	2373	2415	2373

This table reports estimated coefficients from pooled OLS regressions. The dependent variable is the difference between individual stock monthly log-returns and Dow Jones Industrial Average index monthly log-returns. The sample period extends from January 2000 to July 2009. Robust-clustered standard errors are reported in the parentheses. Variable VP is constructed according to equation (4.54). Variables VP^a and VP^b are defined in equation (4.55). They are estimated using an LRGARCH model described in equations (4.59) to (4.61). The estimated coefficients in the LRGARCH model are omitted for brevity.

Chapter 5. Asymptotic Theory for Regressions with Smoothly Changing Parameters

5.1 Introduction

In this paper, we derive the asymptotic properties of the quasi maximum likelihood estimator (QMLE) of smooth transition regressions (STR) when time is the transition variable and the regressors are stationary. The consistency of the estimator and its asymptotic distribution are examined.

Nonlinear regression models have been widely used in practice for a variety of time series applications; see Granger and Teräsvirta (1993) for some examples in economics. In particular, STR models, initially proposed in its univariate form by Chan and Tong (1986), and further developed in Luukkonen, Saikkonen, and Teräsvirta (1988) and Teräsvirta (1994,1998), have been shown to be very useful for representing asymmetric behavior.[11] A comprehensive review of time series STR models is presented in van Dijk, Teräsvirta, and Franses (2002).

In most applications, stationarity, weak exogeneity,[12] and homoskedasticity have been imposed. In these cases, the standard method of estimation is nonlinear least squares (NLS), which is equivalent to quasi-maximum likelihood or, when the errors are Gaussian, to conditional maximum likelihood. The asymptotic properties of the NLS are discussed in Mira and Escribano (2000), Suarez-Fariñas, Pedreira, and Medeiros (2004), and Medeiros and Veiga (2005). Lundbergh and Teräsvirta (1998) and Li, Ling, and McAleer (2002) consider STR models with heteroskedastic errors. Chan, McAleer, and Medeiros (2005) study the properties of the QMLE when the errors follow a GARCH (Generalized Autoregressive Conditional Heteroskedasticity) model. Saikkonen and Choi (2004) consider the case of STR models with cointegrated variables when the transition variable is integrated of order one, and Medeiros, Mendes, and Oxley (2009) analyze a similar model but with stationary transition variables. The case with endogenous regressors is considered in Areosa, McAleer, and Medeiros (in press).

An important case to consider is time as transition variable in STR models. Lin and Teräsvirta (1994) and Medeiros and Veiga (2003) consider this type of specification to construct models with parameters that change smoothly over time. Strikholm (2006) use this transition variable to determine the number of breaks in regression models. However, the asymptotic properties of the QMLE in this case have not been fully understood. If time is the transition variable, asymptotic

[11]The term "smooth transition" in its present meaning first appeared in Bacon and Watts (1971). They presented their smooth transition model as a generalization of models of two intersecting lines with an abrupt change from one linear regression to another at some unknown change point. Goldfeld and Quandt (1972, pp. 263–264) generalized the so-called two-regime switching regression model using the same idea.

[12]With respect to the parameters of interest.

theory of the QML estimator cannot be achieved in the standard way, because as the sample size T goes to infinity, the proportion of finite sub-samples goes to zero. Our solution to this problem is to scale the transition variable t so that the location of the transition is a certain fraction of the total sample rather than a fixed sample point. This modification allows asymptotic theory of the QML estimator. Andrews and McDermott (1995) and Saikkonen and Choi (2004) use similar transformations.

The outline of this paper is as follows. Section 5.2 describes the model and asymptotic properties of the QMLE. Monte Carlo simulations are presented in Section 5.3. Section 5.4 concludes the paper. All proofs are presented in the Appendix.

5.2 Model Definition and Estimation

5.2.1 The Model

We consider the following time series regression with time-varying parameters

$$y_t = \boldsymbol{x}_t'\boldsymbol{\beta}_0 + \sum_{m=1}^{M} \boldsymbol{x}_t'\boldsymbol{\beta}_m f[\gamma_m(t - c_m)] + \varepsilon_t,\ t = 1, 2, \ldots, T, \tag{5.69}$$

where ε_t is a martingale difference sequence with variance σ_ε^2. $\boldsymbol{x}_t$ is a vector of pre-determined regressors. The function f is the logistic transition function which has the form

$$f[\gamma(t-c)] = \frac{1}{1 + e^{-\gamma(t-c)}}, t = 1, 2, \ldots, T. \tag{5.70}$$

where $\gamma > 0$ controls the smoothness of the transition and $c \in \{1, 2, \ldots, T\}$ is a location parameter. The loci $c_m \in \{1, 2, \ldots, T\}$ in (5.69) are change-points. Note that when $\gamma_m \longrightarrow \infty$, $m = 1, \ldots, M$, model (5.69) becomes a linear regression with M structural breaks occurring at the c_m.

5.2.2 Embedding the Model in a Triangular Array

Asymptotic theory for the QML estimator of the model defined above cannot be achieved the standard way. Consider model (5.69) with $M = 1$. As $T \to \infty$, the proportion of observations in the first regime goes to zero. Since for T large,

$$f\left[\gamma(t-c)\right] = f\left[T\gamma(T^{-1}t - T^{-1}c)\right] \approx \mathbb{1}_{\{T^{-1}t>0\}},$$

the parameter vector $\boldsymbol{\beta}_0$ that governs the first regime as well as the transition parameters γ and c vanish from the model and become unidentified. Figure 5.26 illustrates this. In the simulation, γ is set to be 0.2, c is equal to 50. In the upper plot of the figure, c is in the middle of the sample; in the lower plot ($T = 1000$), the second regime dominates. QML estimation of model (5.69) will be dominated by the second regime as the sample size increases. As the sample size goes to infinity, the first regime vanishes and its parameters become unidentified in the estimation. In order to obtain asymptotic theory for the estimator, the proportion of sub-samples in two regimes (before

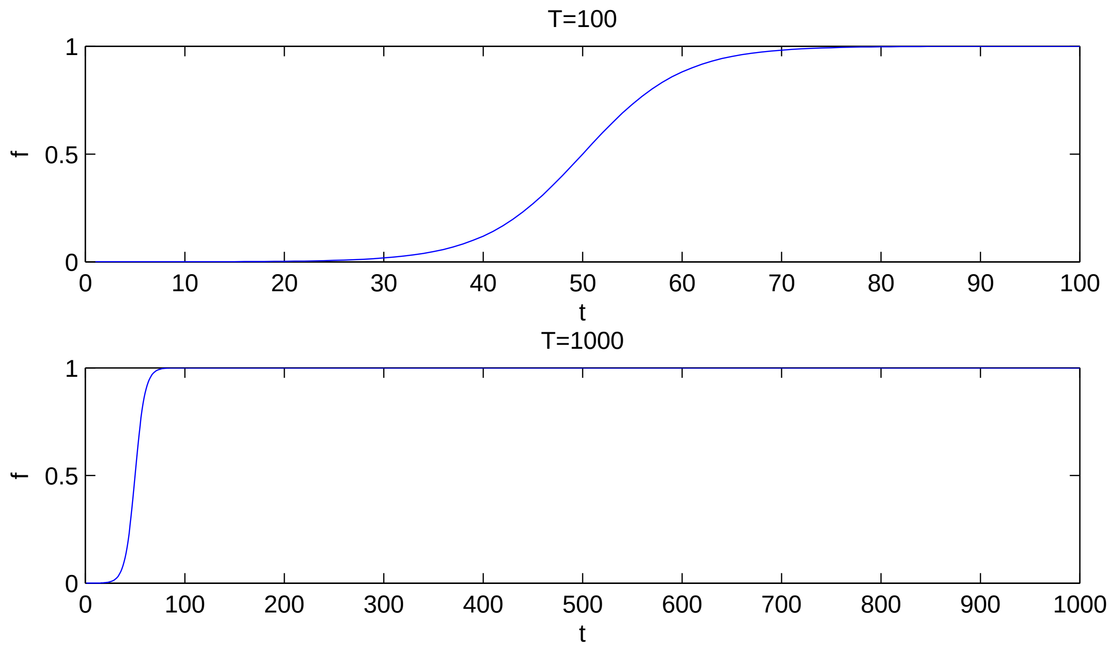

Figure 5.26: **Same unscaled logistic transition functions with different sample sizes** $T = 100$ **&** 1000**.** $\gamma = 0.2; c = 50$**.**

and after the transition) should remain constant as T goes to infinity. In other words, the shape of the plot of the time series should remain qualitatively the same as T grows. For this purpose, we scale the logistic transition function as

$$f\left[\gamma\left(\frac{T_0}{T}t - c\right)\right] = f\left[T^{-1}\gamma\left(T_0 t - Tc\right)\right];\ t = 1, \ldots, T;\ c \in \left[\frac{T_0}{T}, T_0\right]. \tag{5.71}$$

where T_0 is the actual sample size in any given data situation. Accordingly,

$$y_t = \boldsymbol{x}_t'\boldsymbol{\beta}_0 + \sum_{m=1}^{M} \boldsymbol{x}_t'\boldsymbol{\beta}_m f\left[\gamma_m\left(\frac{T_0}{T}t - c_m\right)\right] + \varepsilon_t. \tag{5.72}$$

Note that a given small-sample situation is embedded in this sequence of models at $T = T_0$. As can be seen in (5.71), with this scaling the slope of the logistic function is decreasing with T while the locus of the transition is increasing with T. The scaling of the time counter, T_0, remains constant. Therefore, the proportions of observations in the first regime, during the transition, and in the last regime remain the same as T grows, and the parameters in these groups of observations remain identified.

5.2.3 Assumptions

We denote the data-generating parameter vector as

$$\boldsymbol{\theta}_0 = (\boldsymbol{\beta}_{0,0}', \boldsymbol{\beta}_{1,0}', \ldots, \boldsymbol{\beta}_{M,0}', \gamma_{1,0}, \ldots, \gamma_{M,0}, c_{1,0}, \ldots, c_{M,0}, \sigma_{\varepsilon,0}^2)',$$

where the (second) 0-subscript indicates the data-generating character.

We write $\varepsilon_t(\boldsymbol{\theta})$ such that the notation can be used for both the residuals from the estimation and the data-generating errors:

$$\varepsilon_t\left(\boldsymbol{\theta}\right) = y_t - g(\boldsymbol{x}_t; \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{c})$$

where $\boldsymbol{\beta} = (\boldsymbol{\beta}_0, \ldots, \boldsymbol{\beta}_M)'$; $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_M)'$; $\boldsymbol{c} = (c_1, \ldots, c_M)'$ and

$$g(\boldsymbol{x}_t; \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{c}) = \boldsymbol{x}_t'\boldsymbol{\beta}_0 + \sum_{m=1}^{M} \boldsymbol{x}_t'\boldsymbol{\beta}_m f\left[\gamma_m\left(\frac{T_0}{T}t - c_m\right)\right].$$

We use the shorthand notation $\varepsilon_{t,0} := \varepsilon_t(\boldsymbol{\theta}_0)$, for the data-generating errors and $\varepsilon_t = \varepsilon_t(\boldsymbol{\theta})$ for the residual evaluated at any $\boldsymbol{\theta}$.

We consider the following assumptions.

ASSUMPTION 1 (Parameter Space). *The parameter vector $\boldsymbol{\theta}_0$ is an interior point of $\boldsymbol{\Theta}$, a compact real parameter space.*

ASSUMPTION 2 (Errors).

1. *$\varepsilon_{t,0}$ is a martingale difference sequence with constant variance $\sigma_\varepsilon^2 > c > 0$.*
2. *$\mathbb{E}|\varepsilon_{t,0}|^q < \infty$ for $q \leq 4$.*
3. *$\boldsymbol{x}_t$ and $\varepsilon_{t,0}$ are independent.*

ASSUMPTION 3 (Stationarity and Moments).

1. *$\boldsymbol{x}_t = (\boldsymbol{x}_{A,t}, \boldsymbol{x}_{B,t})'$, where $\boldsymbol{x}_{A,t}$ consists of stationary and ergodic exogenous variables and $\boldsymbol{x}_{B,t}$ is a set of lagged values of y_t. The autoregressive polynomial in each regime associated to $\boldsymbol{x}_{B,t}$ has all roots outside the unit circle.*
2. *$\mathbb{E}\left\|\boldsymbol{x}_{A,t}\right\|^q < \infty$ for $q \leq 4$, where $\|\cdot\|$ is the Euclidean vector norm.*
3. *$\frac{1}{T}\sum_{t=1}^{T}\left(\boldsymbol{x}_t\boldsymbol{x}_t'\right)$ converges in probability to $\Omega = \mathbb{E}\left(\boldsymbol{x}_t\boldsymbol{x}_t'\right)$, which is symmetric positive definite.*

ASSUMPTION 4 (Transition Function). *$g(\boldsymbol{x}_t; \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{c})$ is parameterized such that the parameters are well defined.*

Assumption 1 is standard in the literature and is not too restrictive in the present case as we expect $\boldsymbol{\beta}_0$ to be finite, $\boldsymbol{\gamma}_0$ is positive and finite, and $\boldsymbol{c}_0 \in [0, 1]$. Assumption 2 is also standard.

5.2.4 Quasi Maximum Likelihood Estimator

The quasi log-likelihood function is given by

$$\mathcal{L}_T(\boldsymbol{\theta}) = \frac{1}{T}\sum_{t=1}^{T} \ell_t(\boldsymbol{\theta}),$$

where

$$\ell_t(\boldsymbol{\theta}) = -\frac{1}{2}\left(\log 2\pi + \log \sigma_\varepsilon^2 + \varepsilon_t^2 \sigma_\varepsilon^{-2}\right).$$

The parameter vector is estimated by quasi maximum likelihood as

$$\widehat{\boldsymbol{\theta}}_T = \operatorname*{argmax}_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \mathcal{L}_T(\boldsymbol{\theta}), \tag{5.73}$$

where $\boldsymbol{\Theta}$ is the parameter space.

THEOREM 3 (Consistency). *Under Assumptions 1 through 4, the quasi maximum likelihood estimator $\widehat{\boldsymbol{\theta}}_T$ is consistent:*

$$\widehat{\boldsymbol{\theta}}_T \xrightarrow{p} \boldsymbol{\theta}_0.$$

The proof is provided in the Appendix.

THEOREM 4 (Asymptotic Normality). *Under Assumptions 1 through 4, the quasi maximum likelihood estimator $\widehat{\boldsymbol{\theta}}_T$ is asymptotically normally distributed:*

$$\sqrt{T}\left(\widehat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0\right) \xrightarrow{d} \mathcal{N}\left[0,\ \boldsymbol{A}(\boldsymbol{\theta}_0)^{-1}\boldsymbol{B}(\boldsymbol{\theta}_0)\boldsymbol{A}(\boldsymbol{\theta}_0)^{-1}\right], \tag{5.74}$$

where

$$\boldsymbol{A}(\boldsymbol{\theta}_0) = -\mathbb{E}\left(\left.\frac{\partial^2 \ell_t}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right|_{\boldsymbol{\theta}_0}\right),$$

$$\boldsymbol{B}(\boldsymbol{\theta}_0) = \mathbb{E}\left(\left.\frac{\partial \ell_t}{\partial\boldsymbol{\theta}}\right|_{\boldsymbol{\theta}_0} \left.\frac{\partial \ell_t}{\partial\boldsymbol{\theta}'}\right|_{\boldsymbol{\theta}_0}\right).$$

PROPOSITION 2 (Covariance Matrix Estimation). *Under Assumptions 1 through 4,*

$$\boldsymbol{A}_T \xrightarrow{p} \boldsymbol{A}, \quad \boldsymbol{B}_T \xrightarrow{p} \boldsymbol{B},$$

where

$$\boldsymbol{A}_T(\boldsymbol{\theta}) = -\frac{1}{T}\sum_{t=1}^{T} \frac{\partial^2 \ell_t}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'},$$

and

$$\boldsymbol{B}_T(\boldsymbol{\theta}) = \frac{1}{T}\sum_{t=1}^{T} \frac{\partial \ell_t}{\partial\boldsymbol{\theta}}\frac{\partial \ell_t}{\partial\boldsymbol{\theta}'},$$

and $\boldsymbol{A}$, $\boldsymbol{B}$ as defined in Theorem 4.

5.3 Small Sample Simulations

We conduct a set of Monte Carlo simulations in order to evaluate both the small-sample properties and the asymptotic behavior of the QMLE. In particular, we consider the following models with three limiting regimes:

Model A – Independent and identically distributed (IID) regressors:

$$y_t = \boldsymbol{x}_t'\boldsymbol{\beta}_0 + \sum_{m=1}^{2} \boldsymbol{x}_t'\boldsymbol{\beta}_m f\left[\gamma_m\left(\frac{t}{T} - c_m\right)\right] + \varepsilon_t,$$
$$y_t = 1 + x + (-1 - 2x) f\left[30\left(\frac{t}{T} - \frac{1}{3}\right)\right] + (1 + 3x) f\left[30\left(\frac{t}{T} - \frac{2}{3}\right)\right] + \varepsilon_t,$$

where $\{x_t\}$ is a sequence of independent and normally distributed random variables with zero mean and unit variance, $x_t \sim \mathsf{NID}(0,1)$, and $\{\varepsilon_t\}$ is either a sequence of $\mathsf{NID}(0,1)$ or $\mathsf{Uniform}(-2,2)$ random variables.

Model B – Dependent regressors:

$$y_t = \boldsymbol{x}_t'\boldsymbol{\beta}_0 + \sum_{m=1}^{2} \boldsymbol{x}_t'\boldsymbol{\beta}_m f\left[\gamma_m\left(\frac{t}{T} - c_m\right)\right] + \varepsilon_t,$$
$$\begin{aligned} y_t = 0.5 + 0.4y_{t-1} &+ (-0.5 + 0.5y_{t-1}) f\left[30\left(\frac{t}{T} - \frac{1}{3}\right)\right] \\ &+ (0.5 - 1.7y_{t-1}) f\left[30\left(\frac{t}{T} - \frac{2}{3}\right)\right] + \varepsilon_t, \end{aligned}$$

where $\{\varepsilon_t\}$ is either a sequence of $\mathsf{NID}(0,1)$ or $\mathsf{Uniform}(-2,2)$ random variables.

Different values of T are used, ranging from 100 to 5000 observations. For each value of T, 1000 Monte Carlo simulations are repeated. When the errors are normally distributed, the estimators are maximum likelihood estimators. On the other hand, when the errors are uniformly distributed, the error distribution is misspecified and we have a quasi maximum likelihood estimation setup. For sample sizes up to 300 observations, the estimation procedure did not converge in less than 5% of the replications. These cases were discarded. The parameters $\boldsymbol{\gamma}$ are chosen in order to keep the transitions neither too smooth nor too sharp; see Figure 5.27.

The results are presented in Figures 5.28–5.39. Figures 5.28–5.31 show the average bias and the mean squared error (MSE) as a function of the sample size. Apart from the slope parameter, the average biases are rather small for all sample sizes, models, and error distributions. Furthermore, the MSE, as expected, converges to zero as the sample size increases. With respect to the slope parameter, the MSE is quite high for very small samples (100–300 observations) but also goes to zero as the sample size increases. The bias is also large in small sample, but turns to be negligible for larger sample sizes. The large biases and MSE are mainly caused by few very large estimates (less than 1% of the cases). For example, for Model A with Gaussian errors and 100 observations, the average bias and MSE for the first slope parameter ($\widehat{\gamma}_1$) are, respectively 908.82 and

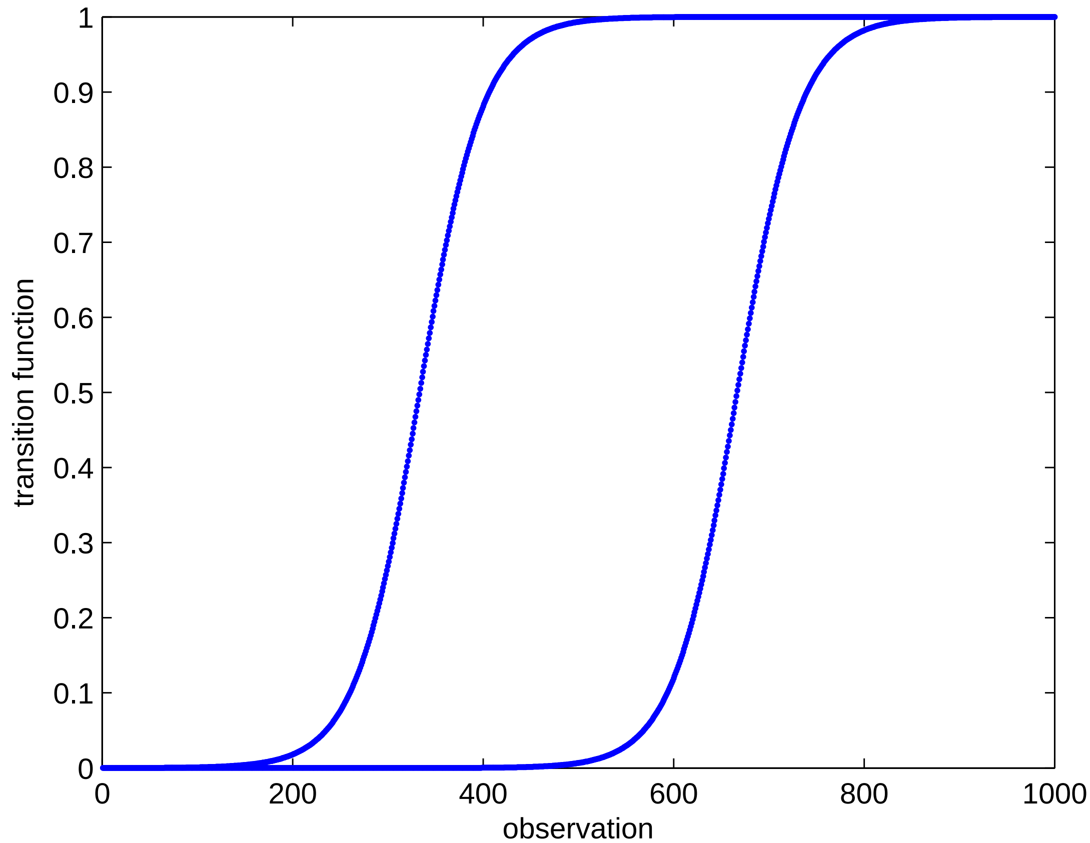

Figure 5.27: **Transition function for Models A and B with 1000 observations.**

$106,447,280.55$. On the other hand, the median bias is just 13.00. For 500 observations and the same model, the average bias and MSE are 19.28 and $155,859.76$, respectively. The median bias is just 0.66 when $T = 500$. This pattern is somehow expected, as it is quite difficult to estimate the slope parameters in small samples. On the other hand, the location ($\boldsymbol{c}$) and the linear parameters ($\boldsymbol{\beta}$) are estimated quite precisely.

Figures 5.32–5.35 present the distribution the standardized QMLE of the linear parameters of the model ($\boldsymbol{\beta}$). Some interesting facts emerge from the graphs. First, even in very small samples, the estimate $\widehat{\boldsymbol{\beta}}_0$ has a distribution close to normal for all models and error distributions. Second, the distributions of $\widehat{\boldsymbol{\beta}}_1$ and $\widehat{\boldsymbol{\beta}}_2$ have some outliers in small samples, but, as expected, they are close to normal for very large samples ($T = 5,000$).

Turning to the location parameter, Figures 5.36–5.39 show the distribution of the standardized QMLE for $\boldsymbol{c}$. It is quite remarkable that even for $T = 100$, the empirical distributions are close to normal.

5.4 Conclusion

In this paper, we propose asymptotic theory for the QML estimator of a logistic smooth transition regression model with time as the transition variable. Although asymptotic theory cannot be achieved in the standard way as the transition variable is not stationary, after proper scaling, we show that the QML estimator is consistent and asymptotically normal. The estimator is shown to converge to the true value of the parameter at the speed of $\sqrt{T}$. We explore the small sample behavior in simulations.

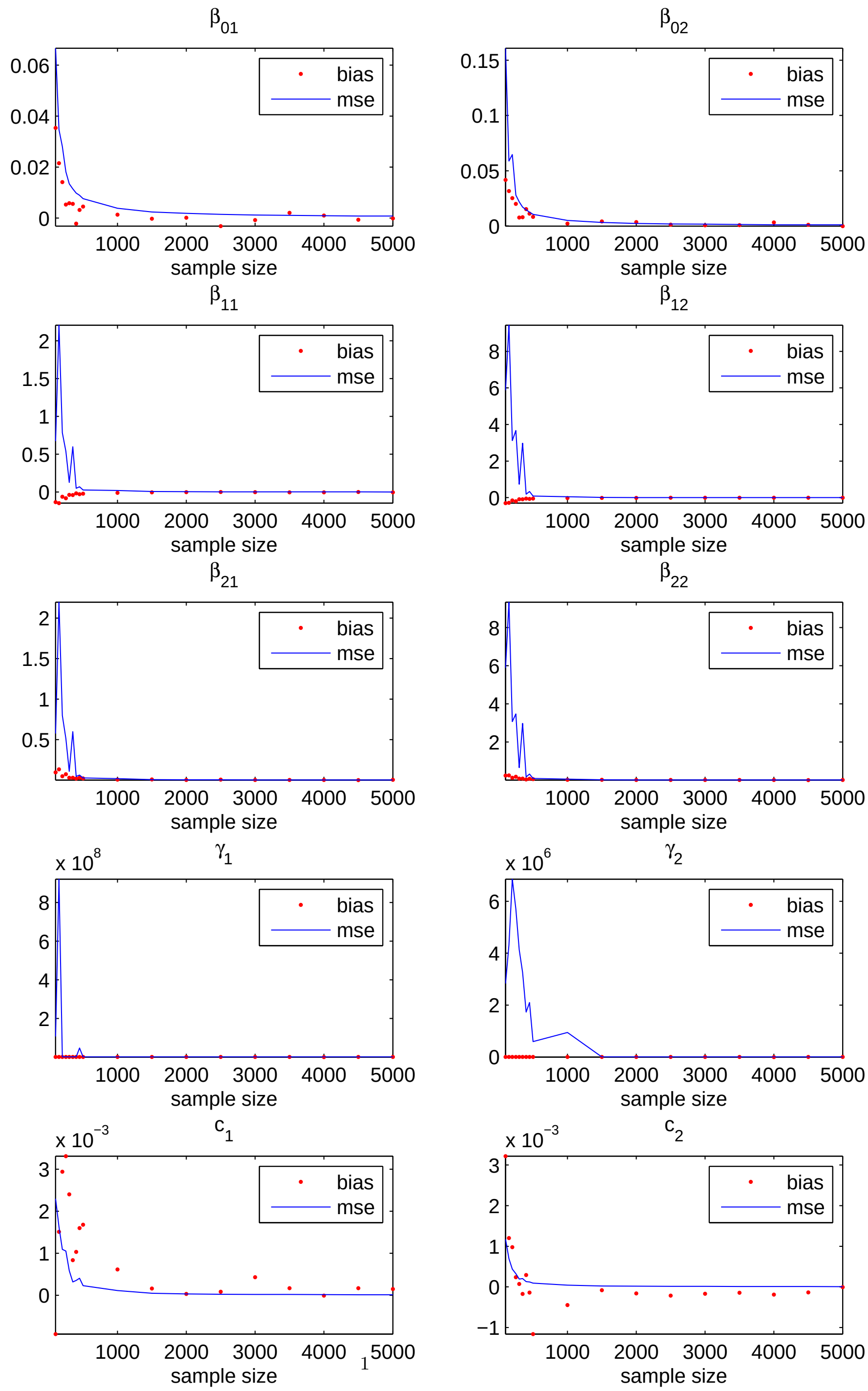

Figure 5.28: **Bias and mean squared error (MSE) of the QMLE of the parameters of Model A with gaussian errors.**

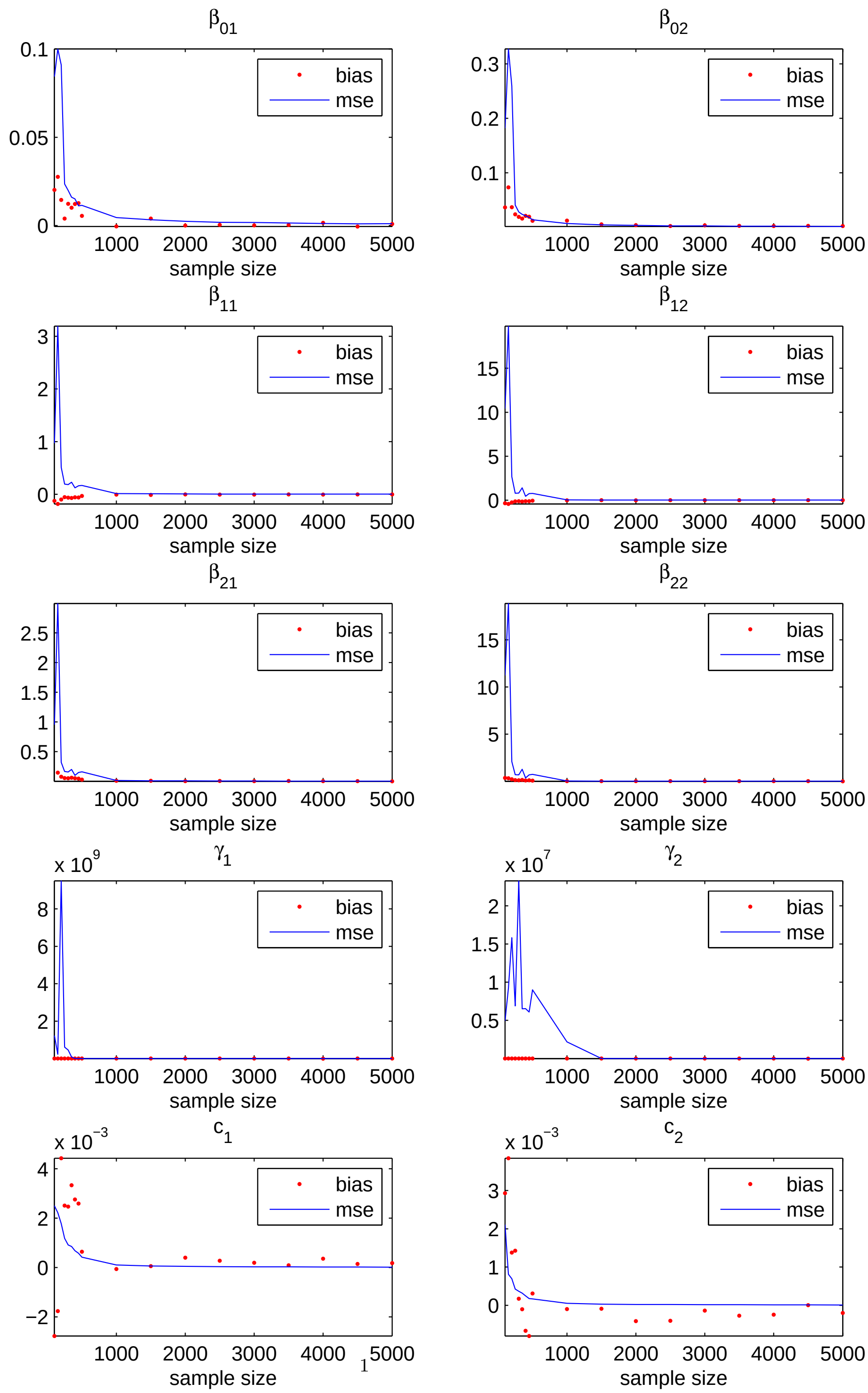

Figure 5.29: **Bias and mean squared error (MSE) of the QMLE of the parameters of Model A with uniform errors.**

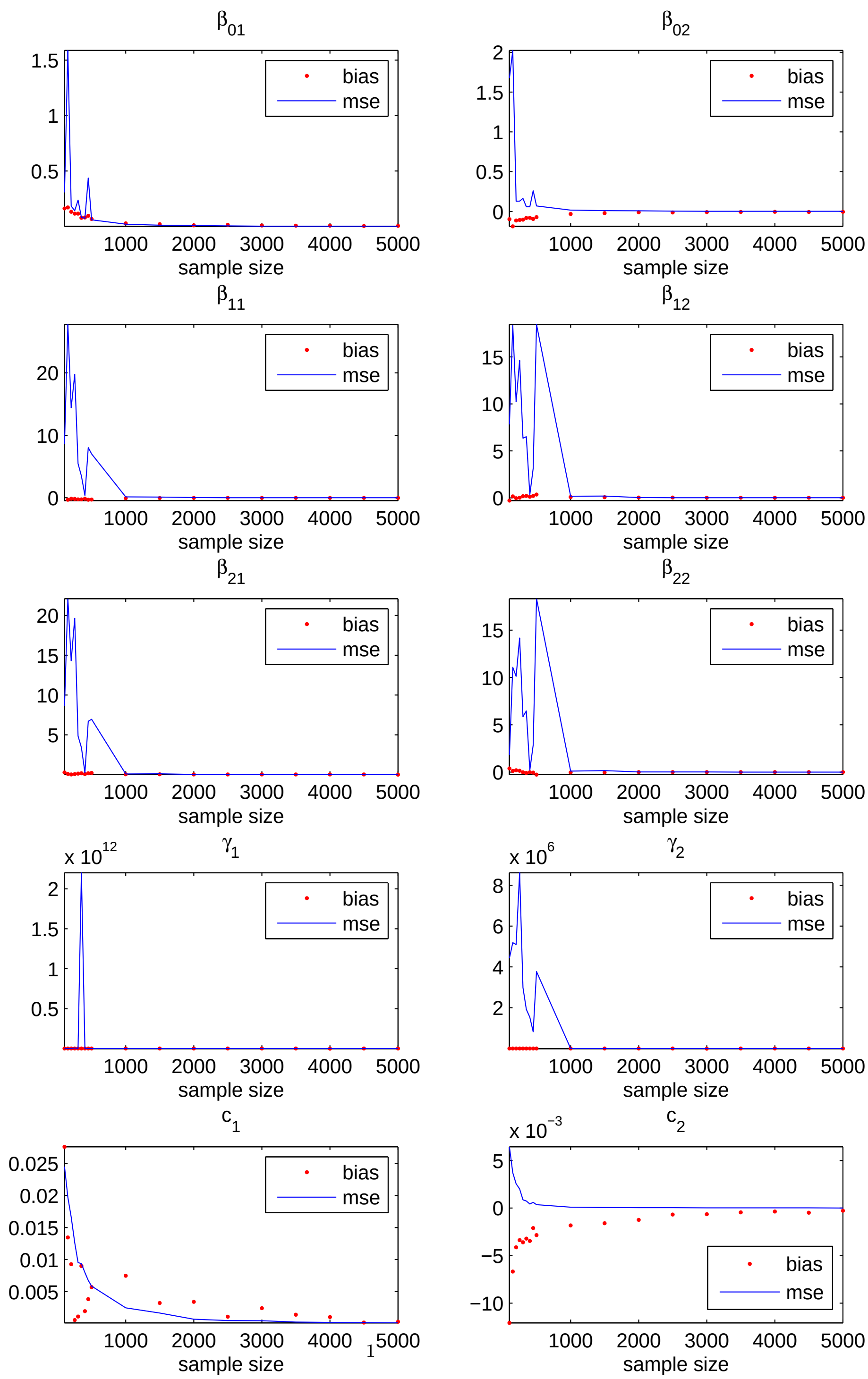

Figure 5.30: **Bias and mean squared error (MSE) of the QMLE of the parameters of Model B with gaussian errors.**

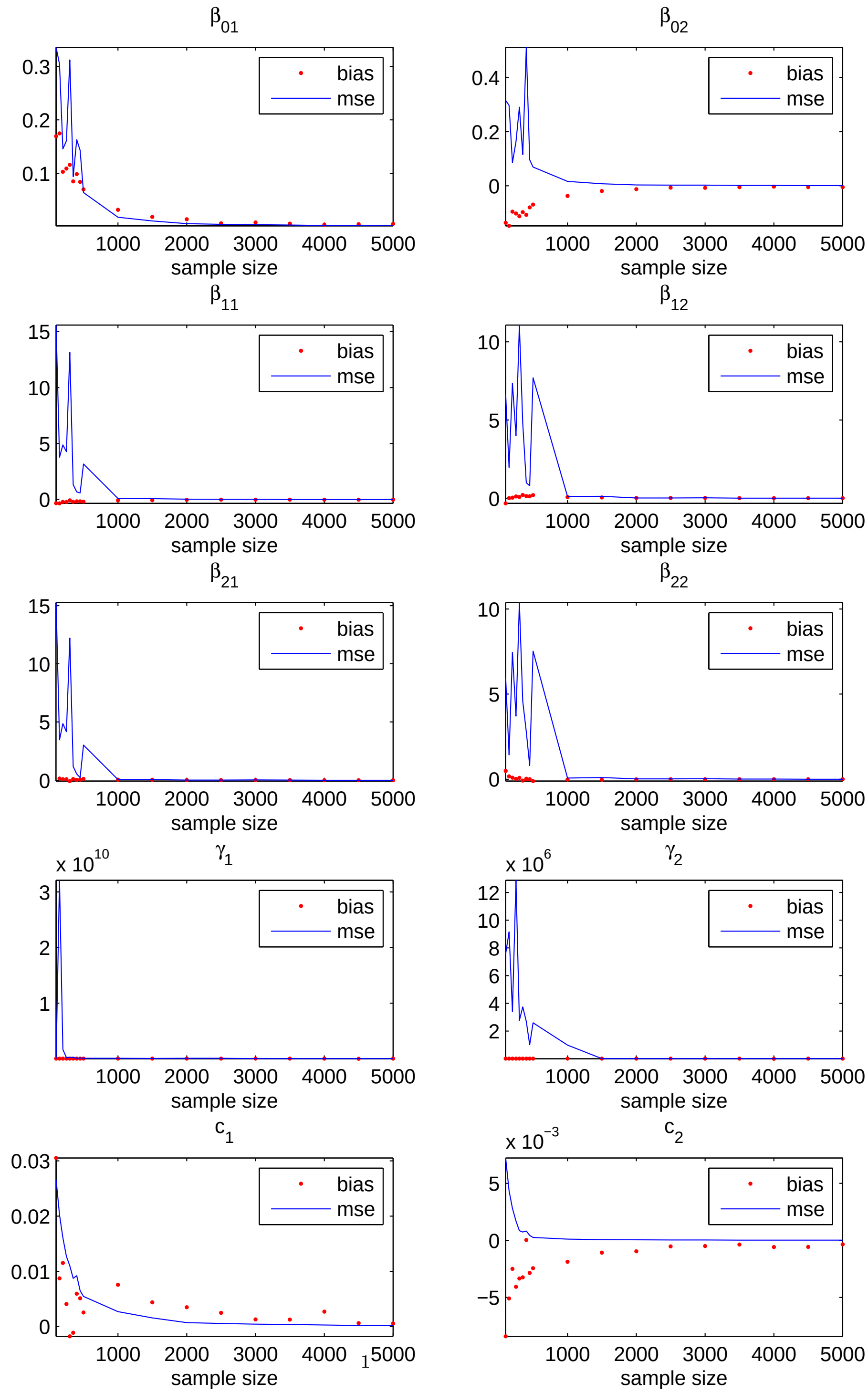

Figure 5.31: **Bias and mean squared error (MSE) of the QMLE of the parameters of Model B with uniform errors.**

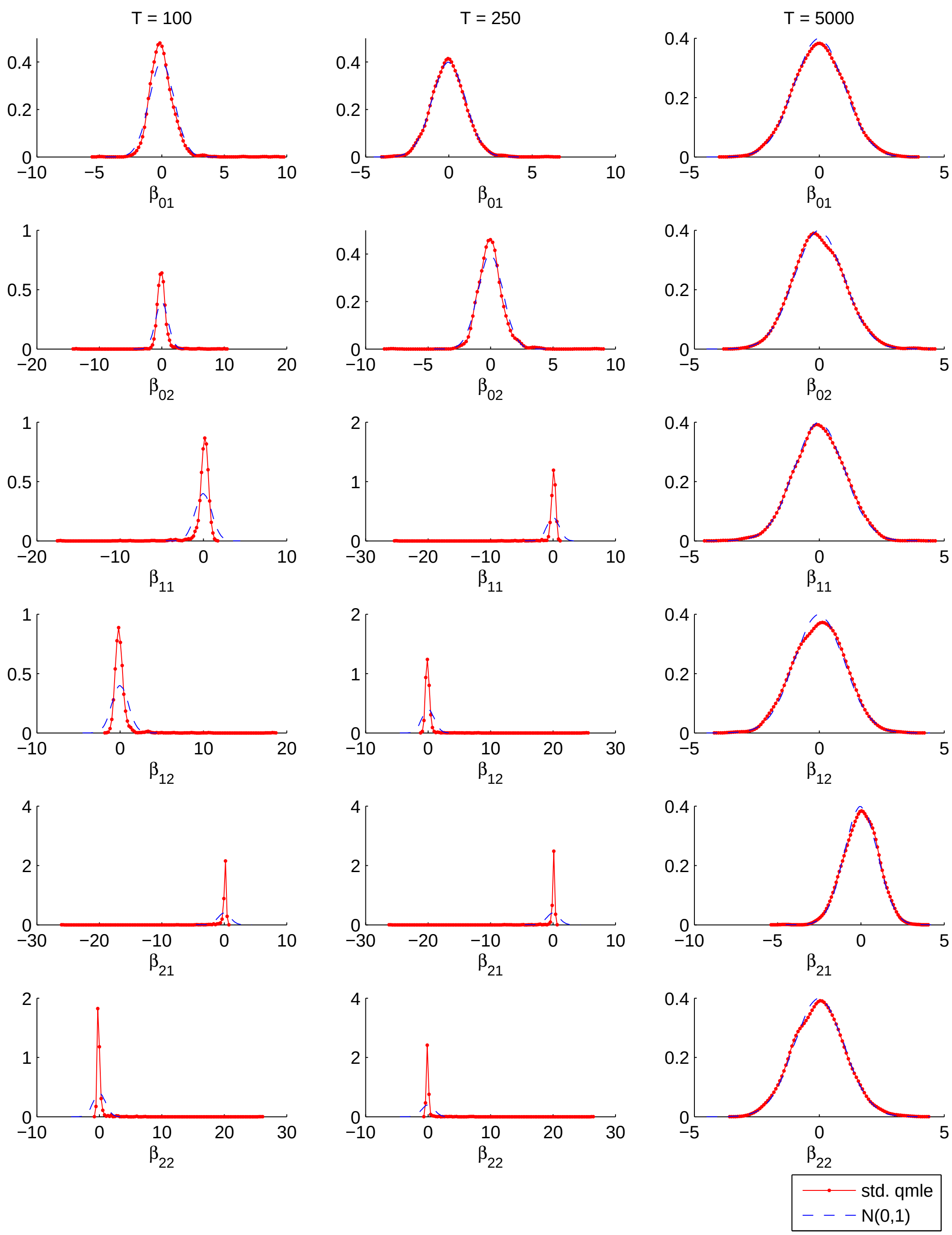

Figure 5.32: **Distribution of the standardized QMLE of the linear parameters of Model A with gaussian errors.**

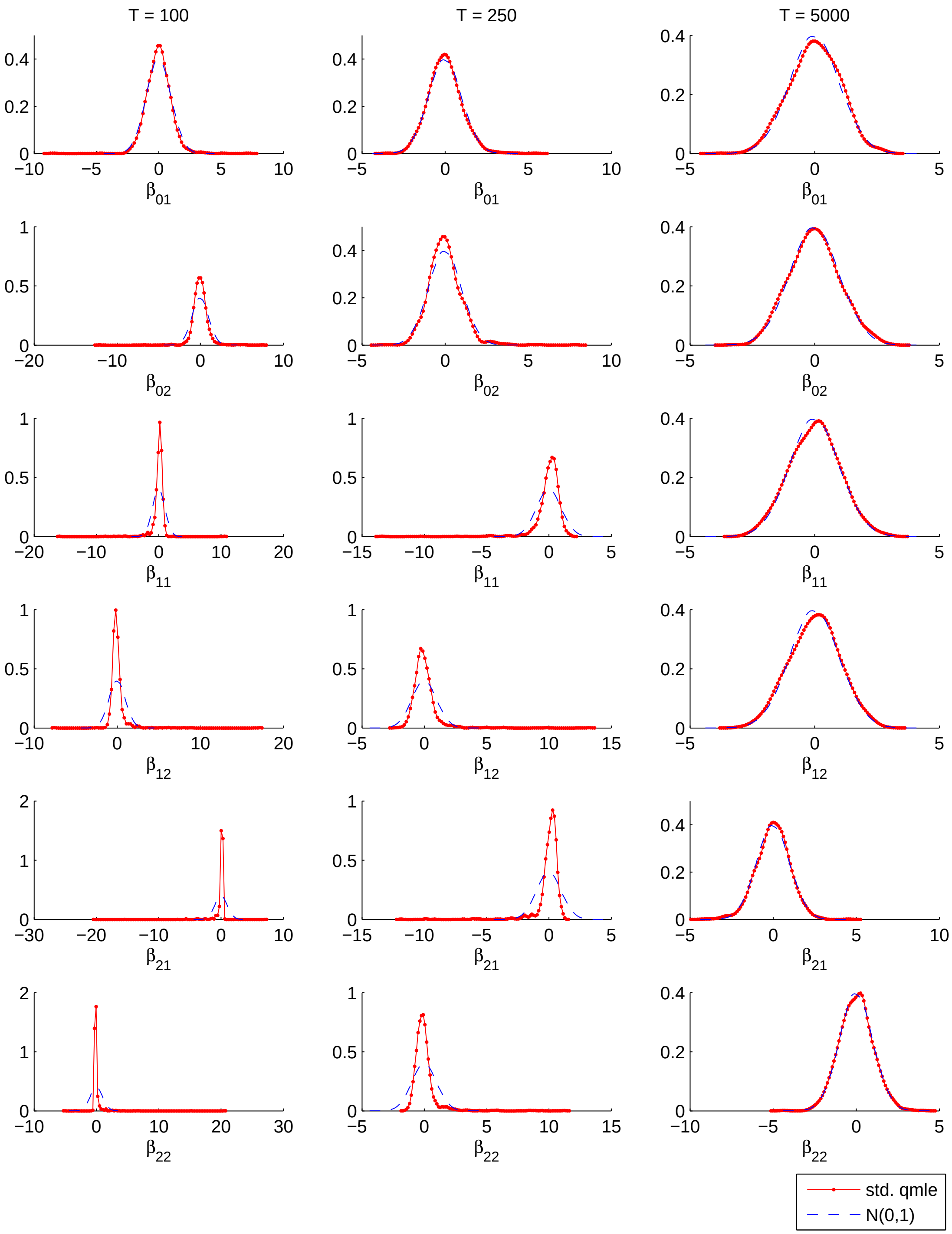

Figure 5.33: **Distribution of the standardized QMLE of the linear parameters of Model A with uniform errors.**

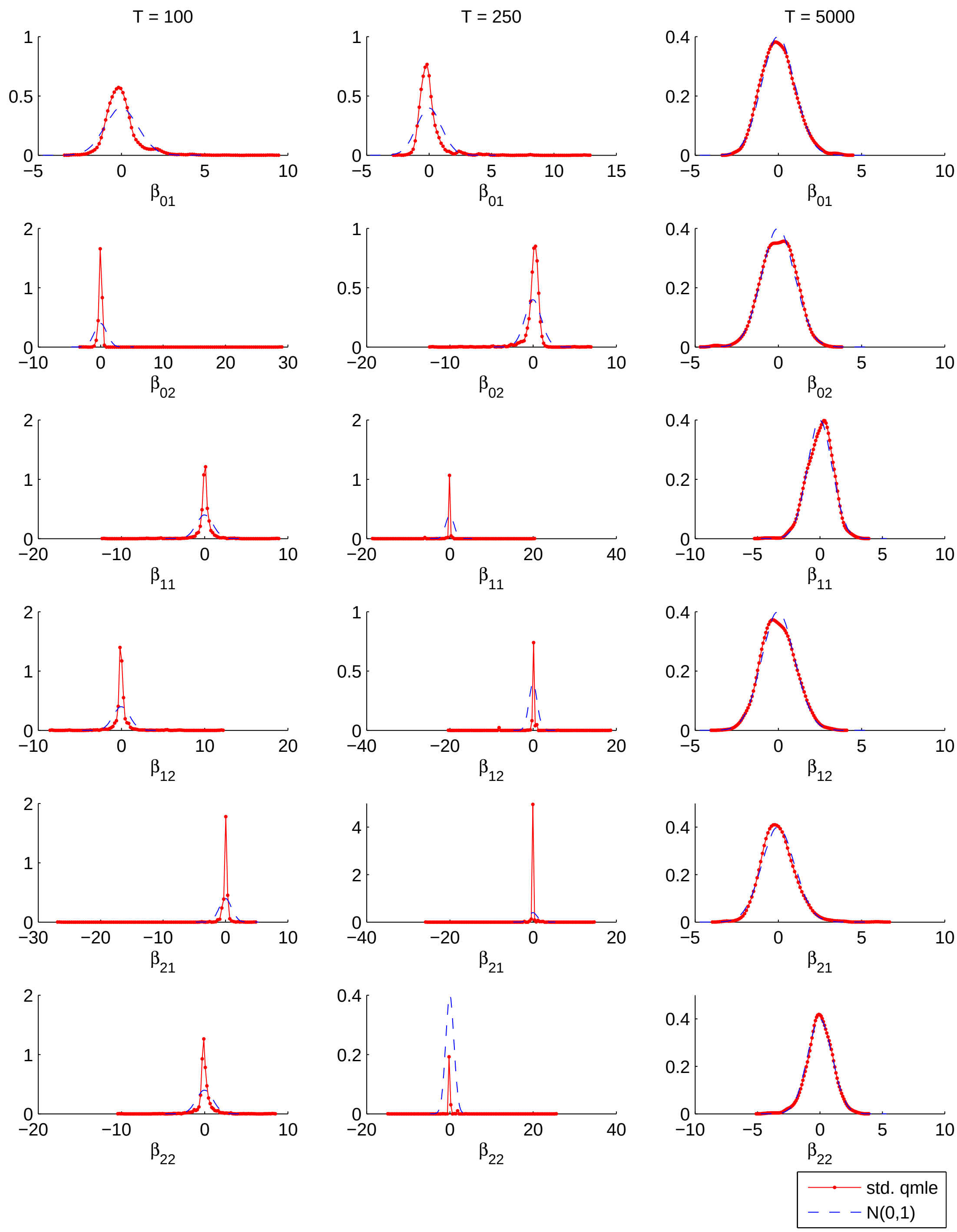

Figure 5.34: **Distribution of the standardized QMLE of the linear parameters of Model B with gaussian errors.**

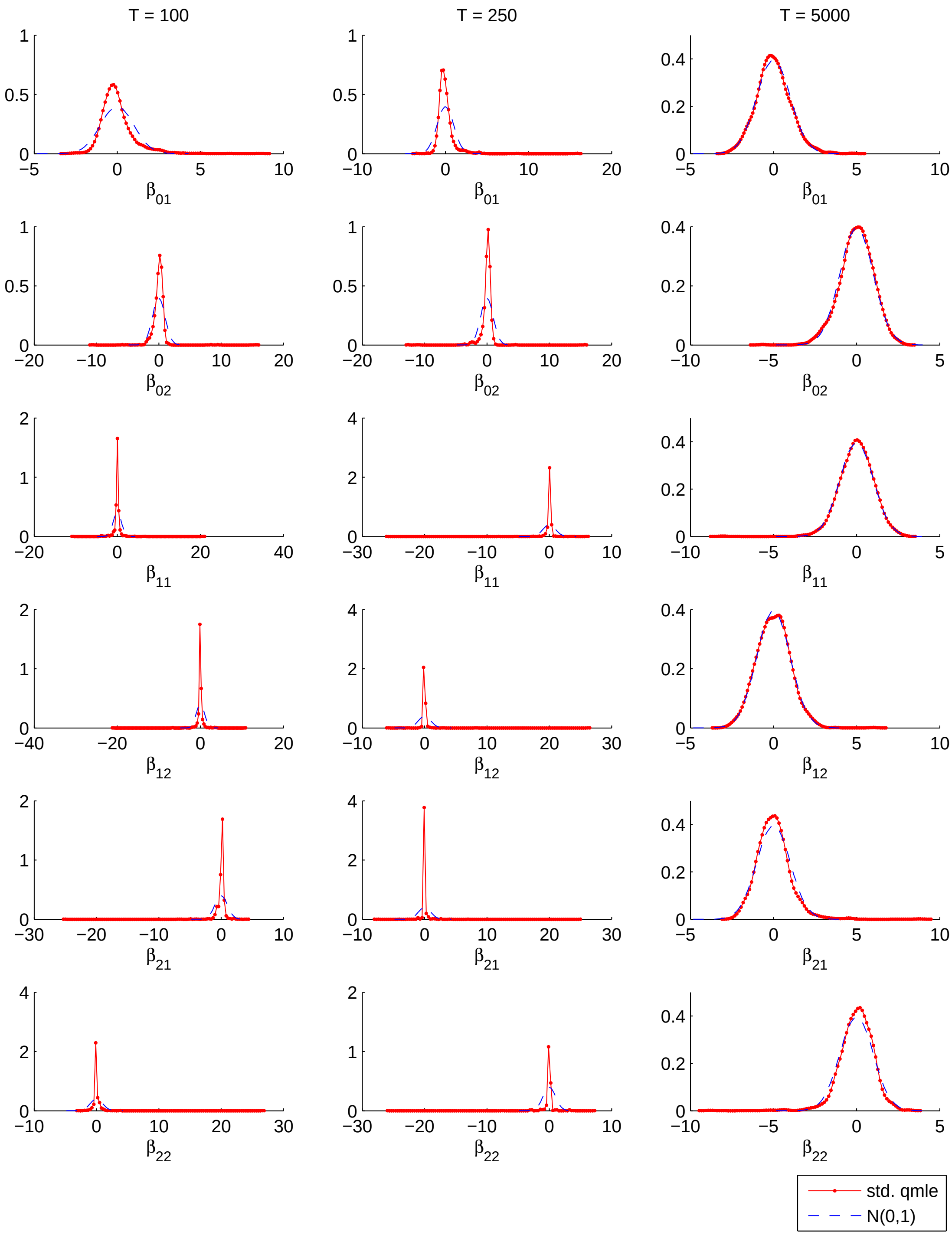

Figure 5.35: **Distribution of the standardized QMLE of the linear parameters of Model B with uniform errors.**

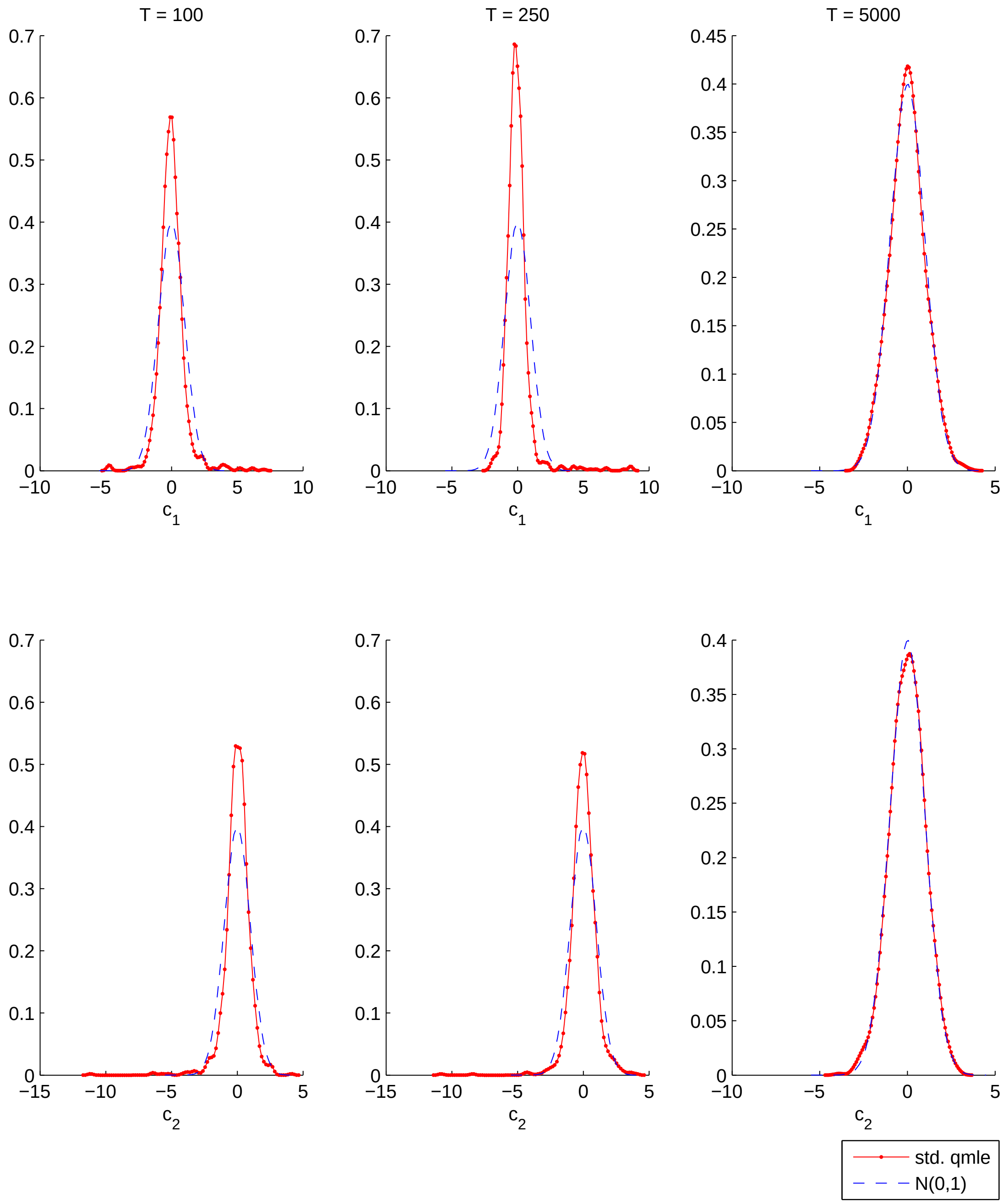

Figure 5.36: **Distribution of the standardized QMLE of the location parameters for Model A with gaussian errors.**

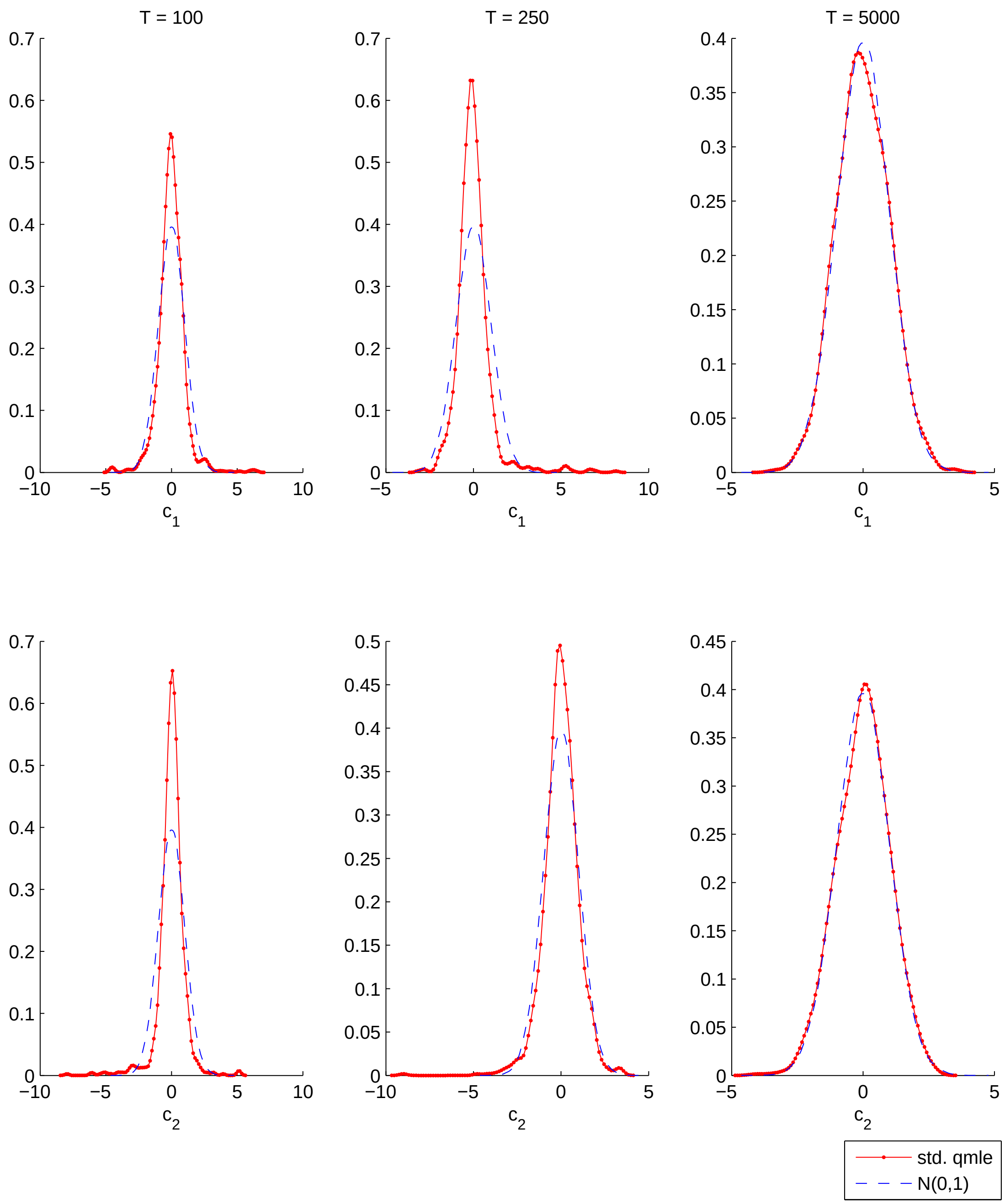

Figure 5.37: **Distribution of the standardized QMLE of the location parameters for Model A with uniform errors.**

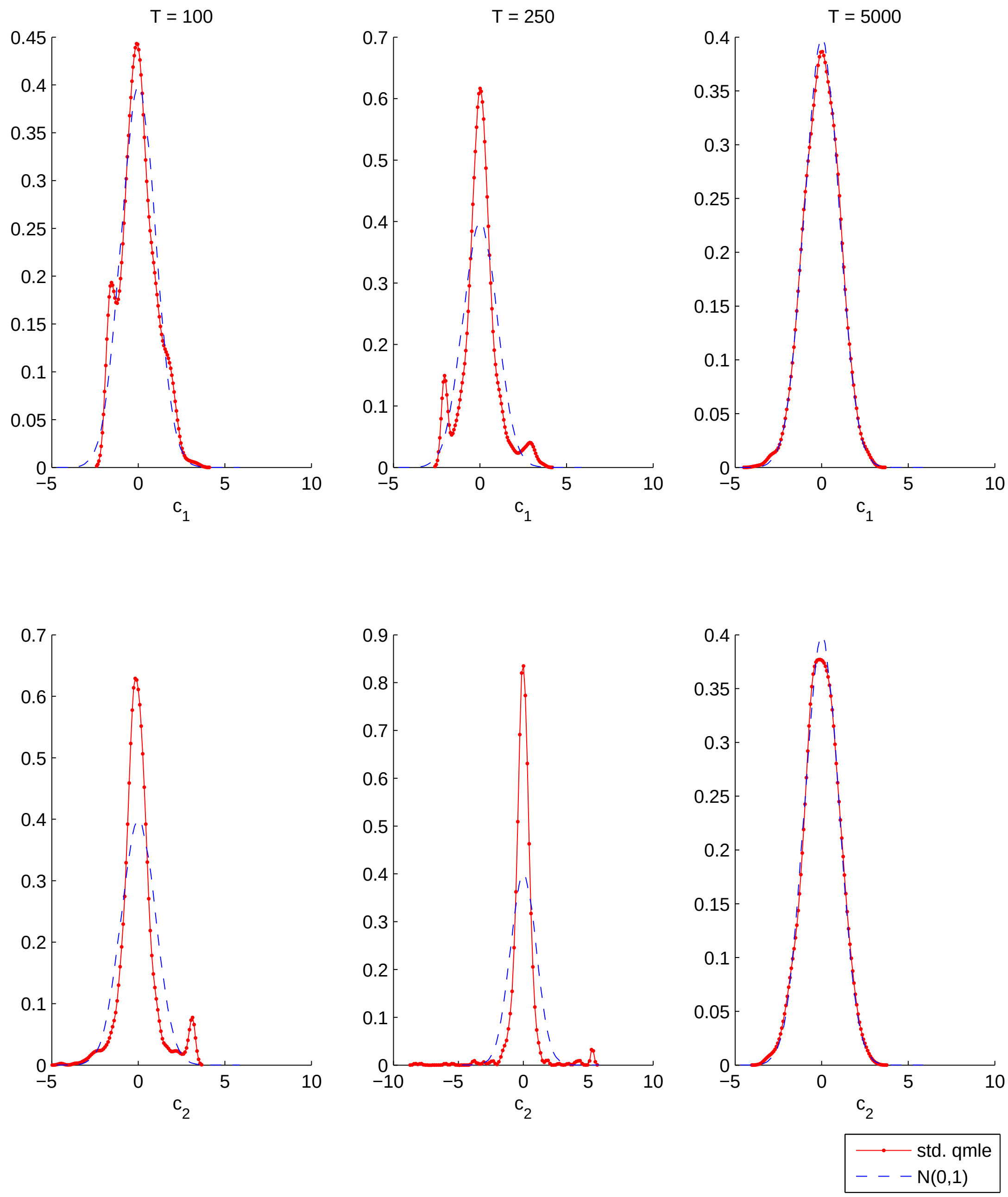

Figure 5.38: **Distribution of the standardized QMLE of the location parameters for Model B with gaussian errors.**

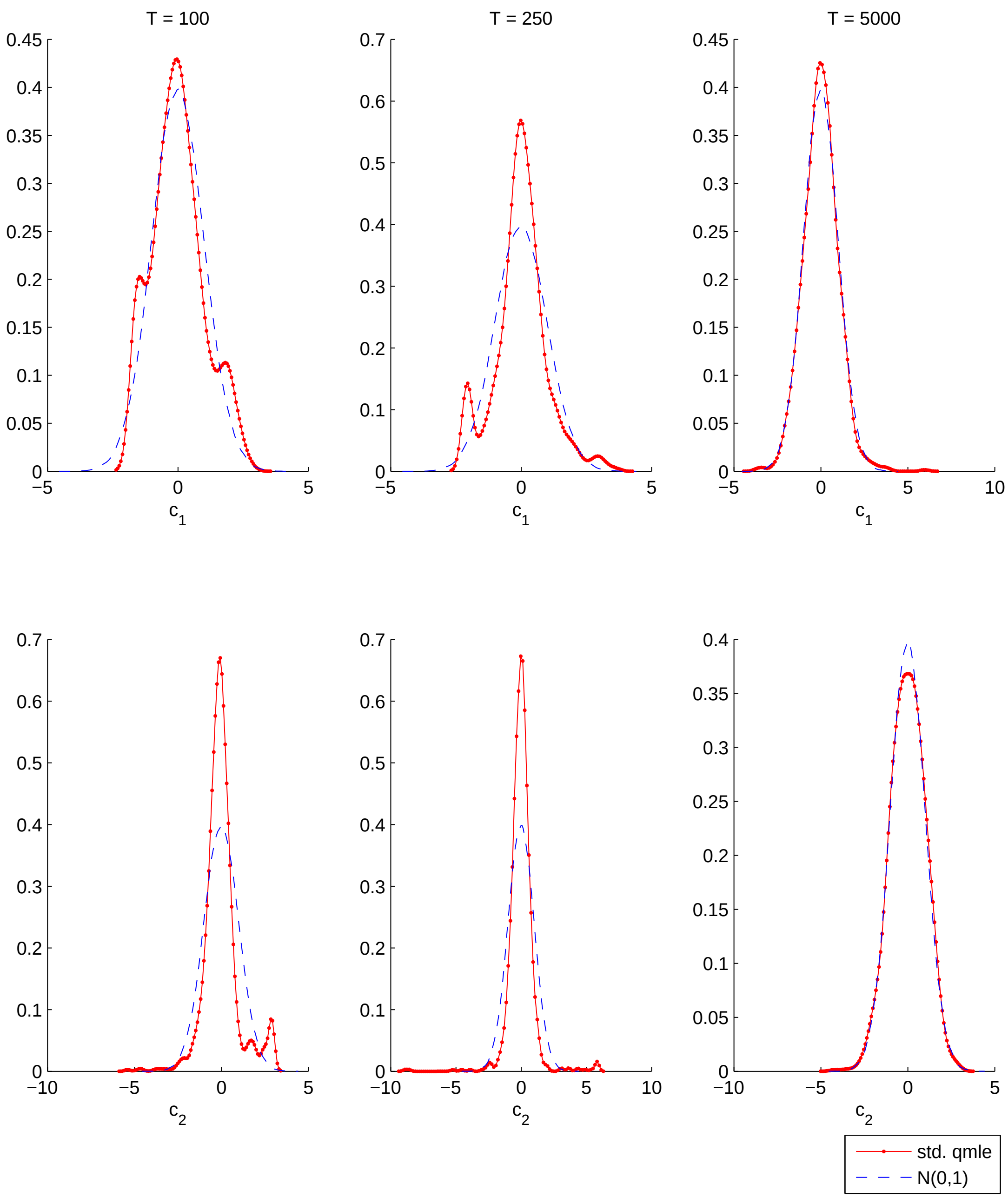

Figure 5.39: **Distribution of the standardized QMLE of the location parameters for Model B with uniform errors.**

Chapter 6. Summary and Conclusion

The four essays included in my dissertation contribute to the literature of econometric modeling of risk and volatility. The first two essays study the dynamics involved in credit default risk. The third essay explores the return predictive power of variance risk premia. The fourth essay derives asymptotic properties of the QMLE of a specific type of logistic transition regressions.

In the first essay, we establish a link between the dynamics of default rates and that of the common risk factor. We find that default rates inherit inter-vintage correlation from the common risk factor. For example, if the common risk factor is an AR(1) time series, default rates of mortgages also display AR(1) structure over different vintages. This result is shown both analytically and in simulations under a Gaussian copula framework. Furthermore, the expect values of cash flows received by securities structured from the defaultable asset (such as MBS) are shown to have similar temporal correlation.

In the subprime market, we identify the monthly common risk factor to be a moving average of monthly log changes in a housing price index. This explains why default rates of subprime mortgages of subsequent vintages display a very high value of correlation, which we define as vintage correlation. The introduction of vintage correlation is important in the literature, not only because it illustrates the source of the subprime crisis, but because it has practical implications for risk management in general. Traditionally, the modeling of temporal correlation in default rates focuses on the serial correlation of default probability of a single asset. This is often done by modeling the dynamics of default intensity. However, the correlation of default rates between two different vintages of assets has been largely ignored in the literature. Neglecting vintage correlation can lead to underestimation of the portfolio risk, if the portfolio consists of credit risk of different vintages.

In the second essay, we examine how the dynamics of default correlation affect the distribution of tranche prices in securitized investment vehicles. We assume that the default correlation exhibits vintage correlation, which is affected by the conditional default probability. We construct regime switching models as well as logistic transition models for the dynamics of default correlation. We find that the distributions of tranche prices in securitized investment vehicles are distorted by the dynamics of default correlation. The shape of distortion is determined by seniority of the tranche. In general, if the tranche price is positively correlated with default correlation, the distribution of the price of the tranche tends to have fatter tails when default correlation is negatively correlated with the conditional default probability. The exact opposite is true for a tranche whose price is negatively correlated with default correlation.

The results in this essay also have important implications for risk management. If the default correlation changes across vintages either in a regime-switching model or in a logistic transition one, our findings suggest that a Gaussian copula model that assumes a constant default correlation underestimates the default risk of a senior tranche while overestimating the default risk for an equity tranche.

In the third essay, we construct monthly series of variance risk premia for 21 Dow Jones stocks during the period from 2000 to 2009. We find that variance risk premia can be used to predict individual stocks' equity premia. Their predictive power is greater than those of conventional return predictors such as price earning ratio and book to market ratio. Furthermore, we apply a Log-linear Realized GARCH model to forecast future realized variance and use it to decompose the variance risk premium into two components. Both of these two components are shown to have predictive powers for equity premia. The combination of these two components and their first two lags explain about 8% of variations in monthly equity premia.

In the literature, the predictive power of variance risk premium is attributed to its correlation with a stochastic macro economic factor such as the degree of risk aversion. While we confirm that the correlation to macro economic factor is likely the most important source of variance risk premia' predictive power, we find evidences that variance risk premium also represents certain idiosyncratic characteristics of a stock, which can explain part of future variations in equity premium in excess of market premium.

In the fourth essay, we propose asymptotic theories for the QML estimator of a logistic transition regression model with time as the transition variable. Logistic transition functions have been widely applied in models where a smooth transition rather than an abrupt change in a coefficient or variable is required. An important case is when time is used as the transition variable. However, if this is the case, asymptotic theory of the QML estimator cannot be achieved in the standard way, because as the sample size T goes to infinity, the proportion of finite sub-samples goes to zero. Our solution to this problem is to scale the transition variable t so that the location of the transition is a certain fraction of the total sample rather than a fixed sample point. we show that the QML estimator is consistent and asymptotically normal. The estimator is shown to converge to the true value of the parameter at the speed of $\sqrt{T}$.

Bibliography

ADRIAN, T., AND J. ROSENBERG (2008): "Stock Returns and Volatility: Pricing the Short-Run and Long-Run Components of Market Risk," *The Journal of Finance*, 63(6), 2997–3030.

AMEMIYA, T. (1985): *Advanced Econometrics*. Harvard University Press, Cambridge.

ANDERSEN, L., AND J. SIDENIUS (2005): "Extensions to the Gaussian Copula: Random Recovery and Random Factor Loadings," *Journal of Credit Risk*, 1(1), 29–70.

ANDERSEN, L., J. SIDENIUS, AND S. BASU (2003): "All your Hedges in One Basket," *Risk*, November, 67–72.

ANDREWS, D. (1992): "Generic Uniform Convergence," *Econometric Theory*, 8, 241–257.

ANDREWS, D. W. K., AND C. J. MCDERMOTT (1995): "Nonlinear Econometric Models with Deterministically Trending Variables," *Review of Economic Studies*, 62, 343–360.

AREOSA, W., M. MCALEER, AND M. MEDEIROS (in press): "Moment-Based Estimation of Smooth Transition Regression Models with Endogenous Variables," *Journal of Econometrics*.

BACON, D. W., AND D. G. WATTS (1971): "Estimating the Transition Between Two Intersecting Lines," *Biometrika*, 58, 525–534.

BAJARI, P., C. CHU, AND M. PARK (2008): "An Empirical Model of Subprime Mortgage Default From 2000 to 2007," Working Paper 14358, NBER.

BARNDORFF-NIELSEN, O., P. HANSEN, A. LUNDE, AND N. SHEPHARD (2008): "Designing Realized Kernels to Measure the Ex Post Variation of Equity Prices in the Presence of Noise," *Econometrica*, 76(6), 1481–1536.

BEARE, B. (2010): "Copulas and Temporal Dependence," *Econometrica*, 78(1), 395–410.

BERD, A., R. ENGLE, AND A. VORONOV (2007): "The Underlying Dynamics of Credit Correlations," *Journal of Credit Risk*, 3(2), 27–62.

BILLINGSLEY, P. (1999): *Convergence of Probability Measures*. Wiley, New York, second edn.

BLUHM, C., L. OVERBECK, AND C. WAGNER (2002): *An Introduction to Credit Risk Modeling*, Financial Mathematics Series. Chapman & Hall/CRC, London, UK.

BOLLERSLEV, T., M. GIBSON, AND H. ZHOU (2008): "Dynamic Estimation of Volatility Risk Premia and Investor Risk Aversion from Option-Implied and Realized Volatilities," Working paper.

BOLLERSLEV, T., G. TAUCHEN, AND H. ZHOU (2009): "Expected Stock Returns and Variance Risk Premia," *Review of Financial Studies*, 22(11), 4463–4492.

BRITTEN-JONES, M., AND A. NEUBERGER (2000): "Option Prices, Implied Price Processes, and Stochastic Volatility," *Journal of Finance*, 55(2), 839–866.

BURTSCHELL, X., J. GREGORY, AND J.-P. LAURENT (2009): "A Comparative Analysis of CDO Pricing Models," *Journal of Derivatives*, 16(4), 9–37.

CABALLERO, R. J., AND A. KRISHNAMURTHY (2009): "Global Imbalances and Financial Fragility," *American Economic Review*, 99(2), 584–588.

CHAN, F., M. MCALEER, AND M. C. MEDEIROS (2005): "Structure and Asymptotic Theory for Nonlinear Models with GARCH Errors," Discussion Paper 506, Department of Economics, Pontifical Catholic University of Rio de Janeiro.

CHAN, K. S., AND H. TONG (1986): "On Estimating Thresholds in Autoregressive Models," *Journal of Time Series Analysis*, 7, 179–190.

CROUHY, M. G., R. A. JARROW, AND S. M. TURNBULL (2008): "The Subprime Credit Crisis of 2007," *Journal of Derivatives*, 16(1), 81–110.

DAGLISH, T. (2009): "What Motivates a Subprime Borrower to Default?," *Journal of Banking and Finance*, 33(4), 681–693.

DAS, S. R., L. FREED, G. GENG, AND N. KAPADIA (2006): "Correlated Default Risk," *Journal of Fixed Income*, 16(2), 7–32.

DAVIDSON, J. (1994): *Stochastic Limit Theory*. Oxford University Press, Oxford.

DAY, T., AND C. LEWIS (1992): "Stock market volatility and the information content of stock index options," *Journal of Econometrics*, 52, 267–287.

DUFFIE, D., AND K. SINGLETON (2003): *Credit Risk: Pricing, Measurement, and Management*, Princeton Series in Finance. Princeton University Press, Princeton, NJ.

FIGLEWSKI, S. (2009): "Viewing the Crisis from 20,000 Feet Up," *Journal of Derivatives*, 16(3), 53–61.

FREES, E. W., AND E. A. VALDEZ (1998): "Understanding Relationships Using Copulas," *North American Actuarial Journal*, 2(1), 1–25.

FREY, R., AND A. MCNEIL (2003): "Dependent Defaults in Models of Portfolio Credit Risk," *Journal of Risk*, 6(1), 59–92.

GIESECKE, K. (2003): "A Simple Exponential Model for Dependent Defaults," *Journal of Fixed Income*, 13(3), 74–83.

GOLDFELD, S. M., AND R. QUANDT (1972): *Nonlinear Methods in Econometrics*. North Holland, Amsterdam.

GORTON, G. (2008): "The Panic of 2007," Working Paper 14358, NBER.

——— (2009): "Information, Liquidity, and the (Ongoing) Panic of 2007," *American Economic Review*, 99(2), 567–572.

GRANGER, C. W. J., AND T. TERÄSVIRTA (1993): *Modelling Nonlinear Economic Relationships*. Oxford University Press, Oxford.

GRAZIANO, G. D., AND L. C. G. ROGERS (2009): "A Dynamic Approach to the Modeling of Correlation Credit Derivatives Using Markov Chains," *International Journal of Theoretical and Applied Finance*, 12(1), 45–62.

HANSEN, P., Z. HUANG, AND H. SHEK (2011): "Realized GARCH: A Joint Model for Returns and Realized Measures of Volatility," *Journal of Applied Econometrics*, forth coming.

HAYRE, L. S., M. SARAF, R. YOUNG, AND J. CHEN (2008): "Modeling of Mortgage Defaults," *Journal of Fixed Income*, 17(4), 6–30.

HILLEBRAND, E., A. SENGUPTA, AND J. XU (2010): "Temporal Correlation of Defaults in Subprime Securitization," Working paper.

HULL, J., M. PREDESCU, AND A. WHITE (2009): "The Valuation of Correlation-Dependent Credit Derivatives Using a Structural Model," Working paper.

HULL, J., AND A. WHITE (2001): "Valuing Credit Default Swaps II: Modeling Default Correlations," *Journal of Derivatives*, 8(3), 12–22.

JIANG, G., AND Y. TIAN (2005): "Model-Free Implied Volatility and Its Information Content," *Review of Financial Studies*, 18(4), 1305–1342.

KOTHARI, S. P., AND J. SHANKEN (1997): "Book-to-market, dividend yield, and expected market returns: A time-series analysis," *Journal of Financial Economics*, 44(2), 169–203.

LAMOUREUX, C., AND W. LASTRAPES (1993): "Forecasting Stock-Return Variance: Toward an Understanding of Stochastic Implied Volatilities," *Review of Financial Studies*, 6(2), 293–326.

LANDO, D. (2004): *Credit Risk Modeling: Theory and Applications*, Princeton Series in Finance. Princeton University Press, Princeton, NJ.

LAURENT, J., AND J. GREGORY (2005): "Basket Default Swaps, CDOs and Factor Copulas," *Journal of Risk*, 7(4), 103–122.

LI, D. (2000): "On Default Correlation: A Copula Function Approach," *Journal of Fixed Income*, 9(4), 43–54.

LI, W. K., S. LING, AND M. MCALEER (2002): "Recent Theoretical Results for Time Series Models with GARCH Errors," *Journal of Economic Surveys*, 16, 245–269.

LIN, C. F. J., AND T. TERÄSVIRTA (1994): "Testing the Constancy of Regression Parameters Against Continuous Structural Change," *Journal of Econometrics*, 62, 211–228.

LINDSKOG, F., AND A. MCNEIL (2003): "Common Poisson Shock Models: Applications to Insurance and Credit Risk Modelling," *ASTIN Bulletin*, 33(2), 209–238.

LING, S., AND M. MCALEER (2003): "Asymptotic Theory for a Vector ARMA-GARCH Model," *Econometric Theory*, 19, 280–310.

LUNDBERGH, S., AND T. TERÄSVIRTA (1998): "Modelling Economic High-Frequency Time Series with STAR-STGARCH Models," Working Paper Series in Economics and Finance 291, Stockholm School of Economics.

LUUKKONEN, R., P. SAIKKONEN, AND T. TERÄSVIRTA (1988): "Testing Linearity Against Smooth Transition Autoregressive Models," *Biometrika*, 75, 491–499.

MEDEIROS, M., E. MENDES, AND L. OXLEY (2009): "Cointegrating Smooth Transition Regressions with a Stationary Transition Variable," Discussion paper, Department of Economics, Pontifical Catholic University of Rio de Janeiro.

MEDEIROS, M. C., AND A. VEIGA (2003): "Diagnostic Checking in a Flexible Nonlinear Time Series Model," *Journal of Time Series Analysis*, 24, 461–482.

——— (2005): "A Flexible Coefficient Smooth Transition Time Series Model," *IEEE Transactions on Neural Networks*, 16, 97–113.

MENG, C., AND A. SENGUPTA (2011): "CDO Tranche Sensitivities in the Gaussian Copula Model," *International Journal of Theoretical and Applied Finance*, Upcoming.

MIRA, S., AND A. ESCRIBANO (2000): "Nonlinear Time Series Models: Consistency and Asymptotic Normality of NLS Under New Conditions," in *Nonlinear Econometric Modeling in Time Series Analysis*, ed. by W. A. Barnett, D. Hendry, S. Hylleberg, T. Teräsvirta, D. Tjøsthein, and A. Würtz, pp. 119–164. Cambridge University Press.

MURPHY, D. (2008): "A Preliminary Enquiry into the Causes of the Credit Crunch," *Quantitative Finance*, 8(5), 435–451.

NEWEY, W., AND K. WEST (1987): "A Simple, Positive Semi-Definite, Heteroscedastic and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.

PETERSEN, M. (2009): "Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches," *Review of Financial Studies*, 22(1), 435–480.

REINHART, C. M., AND K. S. ROGOFF (2008): "Is the US Sub-Prime Financial Crisis So Different? An International Historical Comparison," *American Economic Review*, 98(2), 339–344.

——— (2009): "The Aftermath of Financial Crises," *American Economic Review*, 99(2), 466–472.

ROGERS, W. (1993): "Regression Standard Errors in Clustered Samples," *Stata Technical Bulletin Reprints*, 3, 88–94.

SAIKKONEN, P., AND I. CHOI (2004): "Cointegrating Smooth Transition Regressions," *Econometric Theory*, 20, 301–340.

SCHÖNBUCHER, P. (2003): *Credit Derivatives Pricing Models: Models, Pricing and Implementation*. Wiley, Chichester, UK.

SCHÖNBUCHER, P., AND D. SCHUBERT (2001): "Copula Dependent Default Risk in Intensity Models," Working paper, Bonn University.

SERVIGNY, A., AND O. RENAULT (2002): "Default Correlation: Empirical Evidence," Working paper, Standard and Poor's.

SHIRYAEV, A. (1996): *Probability*. Springer, New York, second edn.

SIDENIUS, J., V. PITERBARG, AND L. ANDERSEN (2008): "A New Framework for Dynamic Credit Portfolio Loss Modelling," *International Journal of Theoretical and Applied Finance*, 11(2), 163–197.

SKLAR, A. (1959): "Fonctions de Repartition a *n* Dimensions et Leurs Marges," *Publ. Inst. Stat. Univ. Paris*, 8, 229–231.

STRIKHOLM, B. (2006): "Determining the number of breaks in a piecewise linear regression model," Working Paper Series in Economics and Finance 648, Stockholm School of Economics.

SUAREZ-FARIÑAS, M., C. E. PEDREIRA, AND M. C. MEDEIROS (2004): "Local Global Neural Networks: A New Approach for Nonlinear Time Series Modeling," *Journal of the American Statistical Association*, 99, 1092–1107.

TERÄSVIRTA, T. (1994): "Specification, Estimation, and Evaluation of Smooth Transition Autoregressive Models," *Journal of the American Statistical Association*, 89, 208–218.

——— (1998): "Modelling Economic Relationships with Smooth Transition Regressions," in *Handbook of Applied Economic Statistics*, ed. by A. Ullah, and D. E. A. Giles, pp. 507–552. Dekker.

VAN DIJK, D., T. TERÄSVIRTA, AND P. H. FRANSES (2002): "Smooth Transition Autoregressive Models - A Survey of Recent Developments," *Econometric Reviews*, 21, 1–47.

WELCH, I., AND A. GOYAL (2008): "A Comprehensive Look at the Empirical Performance of Equity Premium Prediction," *Review of Financial Studies*, 21(4).

Appendix: Proofs for Chapter 5

.1 Proof of Consistency

Proof of Theorem 3. We establish the conditions for consistency according to Theorem 4.1.1 of Amemiya (1985). We have $\widehat{\boldsymbol{\theta}}_T \xrightarrow{p} \boldsymbol{\theta}_0$ if the following conditions hold:

1. $\boldsymbol{\Theta}$ is a compact parameter set.
2. $\mathcal{L}_T(\boldsymbol{\theta}, \boldsymbol{\varepsilon}_t)$ is continuous in $\boldsymbol{\theta}$ and measurable in $\boldsymbol{\varepsilon}_t$.
3. $\mathcal{L}_T(\boldsymbol{\theta})$ converges to a deterministic function $\mathcal{L}(\boldsymbol{\theta})$ in probability uniformly on $\boldsymbol{\Theta}$ as $T \to \infty$.
4. $\mathcal{L}(\boldsymbol{\theta})$ attains a unique global maximum at $\boldsymbol{\theta}_0$.

Item (1) is given by Assumption 1. Item (2) holds by definition of the quasi-maximum likelihood estimator (5.73) from the definition of the normal density. For item (3) we refer to Theorem 4.2.1 of Amemiya (1985): This holds for i.i.d. data if $\mathbb{E}\sup_{\boldsymbol{\theta}\in\Theta} |\ell_t(\boldsymbol{\theta})| < \infty$ and $\ell_t(\boldsymbol{\theta})$ is continuous in $\boldsymbol{\theta}$ for each ε_t. The extension to stationary and ergodic data using the same set of assumptions is achieved in Ling and McAleer (2003, Theorem 3.1). We have $\mathbb{E}\sup_{\boldsymbol{\theta}\in\Theta} |\ell_t(\boldsymbol{\theta})| < \infty$ by Jensen's inequality and $\mathbb{E}\sup |\phi(\boldsymbol{\varepsilon}_t, \boldsymbol{\theta})| < \infty$, where ϕ denotes the normal density function. The finiteness of the last expression follows from the assumption that $\sigma_\varepsilon^2 > c > 0$ for some constant c. The log density $\log \phi(\varepsilon_t, \boldsymbol{\theta})$ is continuous in $\boldsymbol{\theta}$ for every ε_t.

Consider Item (4). By the Ergodic Theorem, $\mathbb{E}\ell_t(\boldsymbol{\theta}) = \mathcal{L}(\boldsymbol{\theta})$. Rewrite the maximization problem as

$$\max_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \mathbb{E}\left[\ell_t\left(\boldsymbol{\theta}\right) - \ell_t(\boldsymbol{\theta}_0)\right].$$

Now, for a given number σ_ε^2,

$$\begin{aligned}
\mathbb{E}\left[\ell_t\left(\boldsymbol{\theta}\right) - \ell_t\left(\boldsymbol{\theta}_0\right)\right] &= \mathbb{E}\log\left[\frac{\phi(\boldsymbol{\varepsilon}_t, \boldsymbol{\theta})}{\phi(\boldsymbol{\varepsilon}_t, \boldsymbol{\theta}_0)}\right], \\
&= \mathbb{E}\left[-\frac{1}{2}\log\frac{\sigma_\varepsilon^2}{\sigma_{\varepsilon,0}^2} - \frac{1}{2}\left(\frac{\varepsilon_t^2}{\sigma_\varepsilon^2} - \frac{\varepsilon_{t,0}^2}{\sigma_{\varepsilon,0}^2}\right)\right], \\
&= -\frac{1}{2}\log\frac{\sigma_\varepsilon^2}{\sigma_{\varepsilon,0}^2} - \frac{1}{2}\left[\mathbb{E}(\varepsilon_t^2\sigma_\varepsilon^{-2}) - 1\right].
\end{aligned} \tag{75}$$

We show that $\mathbb{E}\varepsilon_t^2(\boldsymbol{\theta}) \geq \mathbb{E}\varepsilon_{t,0}^2 = \sigma_{\varepsilon,0}^2$ and that (75) attains an upper bound at $\boldsymbol{\theta} = \boldsymbol{\theta}_0$ uniquely. Consider

$$\mathbb{E}\varepsilon_t^2(\boldsymbol{\theta}) = \mathbb{E}\left[y_t - g(\boldsymbol{x}_t; \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{c})\right]^2.$$

Substituting for $y_t = g(\boldsymbol{x}_t; \boldsymbol{\beta}_0, \boldsymbol{\gamma}_0, \boldsymbol{c}_0) + \varepsilon_{t,0}$ and rearranging, we obtain

$$\begin{aligned}\mathbb{E}\varepsilon_t^2(\boldsymbol{\theta}) &= \mathbb{E}\left[g(\boldsymbol{x}_t; \boldsymbol{\beta}_0, \boldsymbol{\gamma}_0, \boldsymbol{c}_0) + \varepsilon_{t,0} - g(\boldsymbol{x}_t; \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{c})\right]^2, \\ &\geq \mathbb{E}\varepsilon_{t,0}^2 = \sigma_{\varepsilon,0}^2.\end{aligned}$$

The inequality holds from Assumption 2 (3). We have established that for any given σ_ε^2, the objective function (75) attains its maximum of

$$-\frac{1}{2}\left(\log\frac{\sigma_\varepsilon^2}{\sigma_{\varepsilon,0}^2} + \frac{\sigma_{\varepsilon,0}^2}{\sigma_\varepsilon^2} - 1\right)$$

at $\boldsymbol{\beta} = \boldsymbol{\beta}_0$, $\boldsymbol{\gamma} = \boldsymbol{\gamma}_0$, $\boldsymbol{c} = \boldsymbol{c}_0$. Define $x = \sigma_\varepsilon^2/\sigma_{\varepsilon,0}^2$, then

$$f(x) = -\frac{1}{2}\left(\log x + \frac{1}{x} - 1\right)$$

attains its maximum of 0 at $x = 1$, therefore the maximizer is $\sigma_\varepsilon^2 = \sigma_{\varepsilon,0}^2$. This shows that $\mathbb{E}(\ell_t(\boldsymbol{\theta}) - \ell_t(\boldsymbol{\theta}_0))$ is uniquely maximized at $\boldsymbol{\theta} = \boldsymbol{\theta}_0$. $\square$

.2 Proof of Asymptotic Normality

REMARK 1.

1. *In this proof, terms will sometimes involve expectations of cross-products of the type $\mathbb{E}(XY)$, where X and Y are correlated random variables. Note that by the Cauchy-Schwarz inequality, we have*
$$\mathbb{E}XY \leq \left(\mathbb{E}X^2\right)^{\frac{1}{2}} \left(\mathbb{E}Y^2\right)^{\frac{1}{2}},$$
and thus in order to show that the cross-product has finite expectation, it suffices to show that both random variables have finite second moments.

2. *By the same token, if both X and Y have finite second moments,*
$$\begin{aligned}\mathbb{E}(X+Y)^2 &\leq \mathbb{E}X^2 + \mathbb{E}Y^2 + 2\left(\mathbb{E}X^2\right)^{\frac{1}{2}} \left(\mathbb{E}Y^2\right)^{\frac{1}{2}}, \\ &\leq K(\mathbb{E}X^2 + \mathbb{E}Y^2),\end{aligned}$$
for some $K < \infty$.

In the outline of the proof we follow Theorem 4.1.3 of Amemiya (1985). Therefore we have to establish the conditions

1. $\frac{\partial^2 \ell_t}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}$ exists and is continuous in an open neighborhood of $\boldsymbol{\theta}_0$.

2. $\boldsymbol{A}_T(\boldsymbol{\theta}_T^*) \xrightarrow{p} \boldsymbol{A}(\boldsymbol{\theta}_0)$ for all sequences $\boldsymbol{\theta}_T^* \xrightarrow{p} \boldsymbol{\theta}_0$.

3. $$\boldsymbol{B}(\boldsymbol{\theta}_0)^{-\frac{1}{2}}\frac{1}{\sqrt{T}}\sum_{t=1}^{[rT]}\left.\frac{\partial \ell_t}{\partial\boldsymbol{\theta}}\right|_{\boldsymbol{\theta}_0} \xrightarrow{d} W(s),\ s \in [0,1],$$
   where W is standard Brownian motion on the unit interval.

Item (1) is shown in Lemma 3. Item (2) needs consistency of $\widehat{\boldsymbol{\theta}}_T$ for $\boldsymbol{\theta}_0$, which we established in Theorem 3. It further needs uniform convergence of $\boldsymbol{A}_T$ to $\boldsymbol{A}$, i.e.

$$\sup_{\boldsymbol{\theta}\in\Theta} |\boldsymbol{A}_T(\boldsymbol{\theta}) - \boldsymbol{A}(\boldsymbol{\theta})| \xrightarrow{p} 0.$$

We use Ling and McAleer (2003, Theorem 3.1) to establish this, which achieved invocation of the Ergodic Theorem without having to show finiteness of third order derivative information. We show the uniform convergence in Lemma 4.

Item (3) uses Billingsley (1999, Theorem 18.3) and needs (a) that $\{\partial\ell_t/\partial\boldsymbol{\theta}|\boldsymbol{\theta}_0, \mathcal{F}_t\}$ is a stationary martingale difference sequence and (b) that $\boldsymbol{B}(\boldsymbol{\theta}_0)$ exists. Both with be proved in Lemma 3. The first two lemmas show a few technical properties of $g(\boldsymbol{x}_t; \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{c})$ that are needed in the following.

LEMMA 1. *The transition function given by Equation* (5.71) *is bounded, and so are its first and second derivatives with respect to* γ_m *and* c_m, $\forall m = 1, 2, \dots M$.

Proof. We will use shorthand notation f for $f\left[\gamma_m\left(\frac{T_0}{T}t - c_m\right)\right]$ below unless otherwise stated. Defining $a_m(t) := \frac{T_0}{T}t - c_m, t = 1, 2\dots, T$, it is easy to verify that $-\infty < -c_m < a_m(t) \leq T_0 - c_m < \infty$. Since the transition function has the range $(0, 1)$, it is clearly bounded. For the first derivative of f with respect to γ_m, $\forall m = 1, 2, \dots M$,

$$\left|\frac{\partial f}{\partial \gamma_m}\right| = \left|\frac{a_m(t)e^{-\gamma_m a_m(t)}}{(1 + e^{-\gamma_m a_m(t)})^2}\right| \leq |a_m(t)f| < \infty.$$

The first inequality follows from the fact that $1+e^{-\gamma_m a_m(t)} > e^{-\gamma_m a_m(t)} > 0$. The second inequality holds because both $a_m(t)$ and f are bounded. For the second derivative of f with respect to c_m, $\forall m = 1, 2, \dots M$,

$$\begin{aligned}
\left|\frac{\partial^2 f}{\partial \gamma_m^2}\right| &= \left|\frac{2a_m(t)^2 e^{-2\gamma_m a_m(t)}}{\left(1 + e^{-\gamma_m a_m(t)}\right)^3} + \frac{a_m(t)^2 e^{-\gamma_m a_m(t)}}{\left(1 + e^{-\gamma_m a_m(t)}\right)^2}\right|, \\
&\leq \left|\frac{2a_m(t)^2 e^{-2\gamma_m a_m(t)}}{\left(1 + e^{-\gamma_m a_m(t)}\right)^3}\right| + \left|\frac{a_m(t)^2 e^{-\gamma_m a_m(t)}}{\left(1 + e^{-\gamma_m a_m(t)}\right)^2}\right|, \\
&\leq \left|\frac{2a_m(t)^2}{1 + e^{-\gamma_m a_m(t)}}\right| + \left|\frac{a_m(t)^2}{1 + e^{-\gamma_m a_m(t)}}\right|, \\
&= \left|3a_m(t)^2 f\right| < \infty.
\end{aligned}$$

The second inequality follows from the fact that $1 + e^{-\gamma_m a_m(t)} > e^{-\gamma_m a_m(t)} > 0$, the last inequality holds because both $a_m(t)$ and f are bounded. The proof of the boundedness of the first and second derivatives of f with respect to c_m is almost identical to the one above and is omitted for brevity. □

LEMMA 2.
Let $\boldsymbol{\xi} := (\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{c})$, *then*

1. $\mathbb{E}\left\|\frac{\partial}{\partial\boldsymbol{\xi}} g(\boldsymbol{x}_t; \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{c})\right\|^2 < \infty.$

2. $\mathbb{E}\left\|\frac{\partial^2}{\partial\boldsymbol{\xi}\partial\boldsymbol{\xi}'}g(\boldsymbol{x}_t;\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{c})\right\|^2<\infty$, *where $\|\cdot\|$ denotes the standard vector and matrix norms.*

Proof. We will prove the statements element by element. For statement (1),

$$\mathbb{E}\left\|\frac{\partial}{\partial\boldsymbol{\beta}_0}g(\boldsymbol{x}_t;\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{c})\right\|^2=\mathbb{E}\left\|\boldsymbol{x}_t\right\|^2<\infty$$

by Assumption 3 (2).

$$\mathbb{E}\left\|\frac{\partial}{\partial\boldsymbol{\beta}_m}g\left(\boldsymbol{x}_t;\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{c}\right)\right\|^2=\mathbb{E}\left\|\boldsymbol{x}_t f\right\|^2\leq\mathbb{E}\left\|\boldsymbol{x}_t\right\|^2<\infty,$$

by the fact that $|f|<1$.

$$\begin{aligned}\mathbb{E}\left\|\frac{\partial}{\partial\gamma_m}g(\boldsymbol{x}_t;\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{c})\right\|^2&=\mathbb{E}\left\|\boldsymbol{x}_t'\boldsymbol{\beta}_m\frac{\partial f}{\partial\gamma_m}\right\|^2,\\&\leq\mathbb{E}\left\|\boldsymbol{x}_t\right\|^2\left\|\boldsymbol{\beta}_m\right\|^2\left|\frac{\partial f}{\partial\gamma_m}\right|^2<\infty\end{aligned}$$

by Lemma 1, Assumption 1, and Assumption 3 (2). Similarly,

$$\begin{aligned}\mathbb{E}\left\|\frac{\partial}{\partial c_m}g(\boldsymbol{x}_t;\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{c})\right\|^2&=\mathbb{E}\left\|\boldsymbol{x}_t'\boldsymbol{\beta}_m\frac{\partial f}{\partial c_m}\right\|^2,\\&\leq\mathbb{E}\left\|\boldsymbol{x}_t\right\|^2\left\|\boldsymbol{\beta}_m\right\|^2\left|\frac{\partial f}{\partial c_m}\right|^2<\infty.\end{aligned}$$

For statement (2),

$$\mathbb{E}\left\|\frac{\partial^2}{\partial\boldsymbol{\beta}_0\partial\boldsymbol{\beta}_0'}g(\boldsymbol{x}_t;\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{c})\right\|^2=\boldsymbol{0},$$

$$\mathbb{E}\left\|\frac{\partial^2}{\partial\boldsymbol{\beta}_m\partial\boldsymbol{\beta}_m'}g(\boldsymbol{x}_t;\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{c})\right\|^2=\boldsymbol{0},$$

$$\begin{aligned}\mathbb{E}\left\|\frac{\partial^2}{\partial\gamma_m^2}g(\boldsymbol{x}_t;\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{c})\right\|^2&=\mathbb{E}\left\|\boldsymbol{x}_t'\boldsymbol{\beta}_m\frac{\partial^2 f}{\partial^2\gamma_m^2}\right\|^2,\\&\leq\mathbb{E}\left\|\boldsymbol{x}_t\right\|^2\left\|\boldsymbol{\beta}_m\right\|^2\left|\frac{\partial^2 f}{\partial\gamma_m^2}\right|^2<\infty.\end{aligned}$$

For the second inequality, we use the fact that $\left|\frac{\partial^2 f}{\partial\gamma_m^2}\right|$ is bounded from Lemma 1.

Similarly,

$$\begin{aligned}\mathbb{E}\left\|\frac{\partial^2}{\partial c_m^2}g(\boldsymbol{x}_t;\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{c})\right\|^2&=\mathbb{E}\left\|\boldsymbol{x}_t'\boldsymbol{\beta}_m\frac{\partial f}{\partial^2 c_m^2}\right\|^2,\\&\leq\mathbb{E}\left\|\boldsymbol{x}_t\right\|^2\left\|\boldsymbol{\beta}_m\right\|^2\left|\frac{\partial^2 f}{\partial c_m^2}\right|^2<\infty.\end{aligned}$$

□

LEMMA 3.

1. *The sequence* $\left\{ \frac{\partial \ell_t}{\partial \boldsymbol{\theta}}\Big|_{\boldsymbol{\theta}_0}, \mathcal{F}_t \right\}_{t=1,\ldots,T}$ *is a stationary martingale difference sequence.* $\mathcal{F}_t$ *is the sigma-algebra given by all information up to time* t.

2. $$\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \mathbb{E}\left\| \frac{\partial \ell_t}{\partial \boldsymbol{\theta}} \right\| < \infty,$$

3. $$\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \mathbb{E}\left\| \frac{\partial \ell_t}{\partial \boldsymbol{\theta}} \frac{\partial \ell_t}{\partial \boldsymbol{\theta}'} \right\| < \infty.$$

Proof. For part (1) of the proof, all derivatives are evaluated at $\boldsymbol{\theta} = \boldsymbol{\theta}_0$. The nought-subscript is suppressed to reduce notational clutter. Let $\boldsymbol{\xi} = (\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{c})$, as before.

$$\mathbb{E}\left(\frac{\partial \ell_t}{\partial \boldsymbol{\xi}} \middle| \mathcal{F}_{t-1} \right) = \mathbb{E}\left(-\frac{\varepsilon_t}{\sigma_\varepsilon^2} \frac{\partial \varepsilon_t}{\partial \boldsymbol{\xi}} \middle| \mathcal{F}_{t-1} \right) = \mathbb{E}\left(\frac{\varepsilon_t}{\sigma_\varepsilon^2} \frac{\partial}{\partial \boldsymbol{\xi}} g(\boldsymbol{x}_t; \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{c}) \middle| \mathcal{F}_{t-1} \right) = 0,$$

since $g(\boldsymbol{x}_t; \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{c})$ is independent of ε_t and its derivatives are bounded (Lemma 2).

$$\mathbb{E}\left(\frac{\partial \ell_t}{\partial \sigma_\varepsilon^2} \middle| \mathcal{F}_{t-1} \right) = \mathbb{E}\left(-\frac{1}{2\sigma_\varepsilon^2} + \frac{1}{2} \frac{\varepsilon_t^2}{\sigma_\varepsilon^4} \middle| \mathcal{F}_{t-1} \right) = 0,$$

since ε_t has mean zero and variance σ_ε^2.

For part (2) and (3) of the proof, the expressions are evaluated at any $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ if not otherwise stated. The data-generating parameters will be explicitly denoted by a nought-subscript. The process y_t is *data* and thus evaluated at $\boldsymbol{\theta}_0$ throughout.

We first consider the gradient vectors of $\boldsymbol{\xi}$,

$$\begin{aligned}
\mathbb{E}\left\| \frac{\partial \ell_t}{\partial \boldsymbol{\xi}} \right\| &= \mathbb{E}\left\| \frac{\varepsilon_t}{\sigma_\varepsilon^2} \frac{\partial}{\partial \boldsymbol{\xi}} g(\boldsymbol{x}_t; \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{c}) \right\|, \\
&\leq \left(\mathbb{E}\left| \frac{\varepsilon_t}{\sigma_\varepsilon^2} \right|^2 \right)^{\frac{1}{2}} \left(\mathbb{E}\left\| \frac{\partial}{\partial \boldsymbol{\xi}} g(\boldsymbol{x}_t; \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{c}) \right\|^2 \right)^{\frac{1}{2}}, \\
&\leq \left(\frac{\mathbb{E}\varepsilon_t^2}{c} \right)^{\frac{1}{2}} \left(\mathbb{E}\left\| \frac{\partial}{\partial \boldsymbol{\xi}} g(\boldsymbol{x}_t; \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{c}) \right\|^2 \right)^{\frac{1}{2}} < \infty.
\end{aligned}$$

The finiteness of the second factor follows from Lemma 2 (1). For the first factor, note that

$$\begin{aligned}
\varepsilon_t^2 &= \left(y_t - \boldsymbol{x}_t' \boldsymbol{\beta}_0 - \sum_{m=1}^{M} \boldsymbol{x}_t' \boldsymbol{\beta}_m f[\gamma_m (t - c_m)] \right)^2, \\
&= \left(\boldsymbol{x}_t' (\boldsymbol{\beta}_{0,0} - \boldsymbol{\beta}_0) + \sum_{m=1}^{M} \boldsymbol{x}_t' \left[\boldsymbol{\beta}_{m,0} f\left(\gamma_{m,0}(t - c_{m,0})\right) - \boldsymbol{\beta}_m f\left(\gamma_m (t - c_m)\right) \right] \right)^2.
\end{aligned}$$

Therefore, there exists $K \in \mathbb{N}$ such that

$$
\begin{aligned}
\varepsilon_t^2 &\leq K \left| \boldsymbol{x}_t'(\boldsymbol{\beta}_{0,0} - \boldsymbol{\beta}_0) \right|^2 + K \sum_{m=1}^{M} \left| \boldsymbol{x}_t' \left(\boldsymbol{\beta}_{m,0} f[\gamma_{m,0}(t - c_{m,0})] - \boldsymbol{\beta}_m f[\gamma_m(t - c_m)] \right) \right|^2, \\
&\leq KL \left\| \boldsymbol{x}_t \right\|^2 + KL \sum_{m=1}^{M} \left\| \boldsymbol{x}_t \right\|^2, \\
&= KL(M+1) \left\| \boldsymbol{x}_t \right\|^2,
\end{aligned}
$$

where L is some positive constant. The existence of such L is guaranteed by the compactness of the parameter space and the fact that f is bounded. Using Assumption 3 (2), it is clear that $\mathbb{E}\varepsilon_t^2$ is bounded.

For σ_ε^2,

$$
\begin{aligned}
\mathbb{E} \left| \frac{\partial \ell_t}{\partial \sigma_\varepsilon^2} \right| &= \mathbb{E} \left| \frac{1}{2\sigma_\varepsilon^2} - \frac{1}{2} \frac{\varepsilon_t^2}{\sigma_\varepsilon^4} \right|, \\
&\leq \frac{1}{2\sigma_\varepsilon^2} + \frac{1}{2} \mathbb{E} \left| \frac{\varepsilon_t^2}{\sigma_\varepsilon^4} \right|, \\
&= \frac{1}{\sigma_\varepsilon^2} < \infty.
\end{aligned}
$$

This shows statement (2) of Lemma 3. Statement (3) use similar techniques in the proof. We will only show the case of γ_m, which requires most work. The rest of the proof will be omitted for brevity.

$$
\begin{aligned}
\mathbb{E} \left| \frac{\partial \ell_t}{\partial \gamma_m} \frac{\partial \ell_t}{\partial \gamma_m'} \right| &= \mathbb{E} \left| \frac{\varepsilon_t^2}{\sigma_\varepsilon^4} \left(\frac{\partial f}{\partial \gamma_m} \right)^2 \boldsymbol{x}_t' \boldsymbol{\beta}_m \boldsymbol{\beta}_m' \boldsymbol{x}_t \right|, \\
&\leq \left(\mathbb{E} \left| \frac{\varepsilon_t^2}{\sigma_\varepsilon^4} \right|^2 \right)^{\frac{1}{2}} \left(\mathbb{E} \left| \boldsymbol{x}_t' \boldsymbol{\beta}_m \boldsymbol{\beta}_m' \boldsymbol{x}_t \right|^2 \right)^{\frac{1}{2}} \left| \frac{\partial f}{\partial \gamma_m} \right|^2, \\
&\leq \left(\frac{\mathbb{E}\varepsilon_t^4}{c^3} \right)^{\frac{1}{2}} \left(\mathbb{E} \left\| \boldsymbol{x}_t \right\|^4 \left\| \boldsymbol{\beta}_m \right\|^4 \right)^{\frac{1}{2}} \left| \frac{\partial f}{\partial \gamma_m} \right|^2 < \infty.
\end{aligned}
$$

The finiteness of $\mathbb{E} \left\| \boldsymbol{x}_t \right\|^4$ follows from Assumption 3 (2). $\left\| \boldsymbol{\beta}_m \right\|^4$ is finite due to Assumption 1. Lemma 1 ensures that the last factor is bounded. To see the finiteness of the first factor, recall in part (2) we have shown that

$$
\varepsilon_t^2 \leq KL(M+1) \left\| \boldsymbol{x}_t \right\|^2.
$$

It follows that

$$
\varepsilon_t^4 \leq (KL)^2 (M+1)^2 \left\| \boldsymbol{x}_t \right\|^4.
$$

Therefore,

$$
\mathbb{E}\varepsilon_t^4 \leq (KL)^2 (M+1)^2 \mathbb{E} \left\| \boldsymbol{x}_t \right\|^4 < \infty
$$

by Assumption 3 □

LEMMA 4. *The function*

$$g_t(\boldsymbol{\theta}) := -\frac{\partial^2 \ell_t}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} - \boldsymbol{A}(\boldsymbol{\theta})$$

where

$$\boldsymbol{A}(\boldsymbol{\theta}) = -\mathbb{E}\frac{\partial^2 \ell_t}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}$$

is absolutely uniformly integrable:

$$\mathbb{E} \sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \|g_t(\boldsymbol{\theta})\| < \infty;$$

it is continuous in $\boldsymbol{\theta}$ and has zero mean: $\mathbb{E} g_t(\boldsymbol{\theta}) = \mathbf{0}$.

Proof. ¿From the triangular inequality,

$$\mathbb{E} \sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \|g_t(\boldsymbol{\theta})\| \leq \mathbb{E} \sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \left\| \frac{\partial^2 \ell_t}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \right\| + \mathbb{E} \sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \|\boldsymbol{A}(\boldsymbol{\theta})\|.$$

If $\mathbb{E} \sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \|\partial^2 \ell_t / \partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'\| < \infty$, $\boldsymbol{A}(\boldsymbol{\theta})$ exists and by the Ergodic Theorem, there is pointwise convergence. Thus showing absolute uniform integrability reduces to showing that

$$\mathbb{E} \sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \left\| \frac{\partial^2 \ell_t}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \right\| < \infty.$$

Proving finiteness of the expected value of the supremum consists of repeated application of the Lebesgue Dominated Convergence Theorem (Shiryaev (1996, p. 187), Ling and McAleer (2003), Lemmas 5.3 and 5.4). We will show the statement for second derivatives element by element, starting with $\boldsymbol{\beta_0}$,

$$\frac{\partial^2 \ell_t}{\partial \boldsymbol{\beta}_0 \partial \boldsymbol{\beta}_0'} = -\frac{\boldsymbol{x}_t \boldsymbol{x}_t'}{\sigma_\varepsilon^2}.$$

According to Assumption 2 (1) there exists a constant c such that $\sigma_\varepsilon^2 > c > 0$, therefore

$$\sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \left\| \frac{\partial^2 \ell_t}{\partial \boldsymbol{\beta}_0 \partial \boldsymbol{\beta}_0'} \right\| \leq \left\| \frac{\boldsymbol{x}_t \boldsymbol{x}_t'}{c} \right\|.$$

By Assumption 3 (3),

$$\mathbb{E} \sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \left\| \frac{\partial^2 \ell_t}{\partial \boldsymbol{\beta}_0 \partial \boldsymbol{\beta}_0'} \right\| \leq \mathbb{E} \left\| \frac{\boldsymbol{x}_t \boldsymbol{x}_t'}{c} \right\| < \infty.$$

For $\boldsymbol{\beta}_m, m = 1, 2, \ldots, M$,

$$\sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \left\| \frac{\partial^2 \ell_t}{\partial \boldsymbol{\beta}_m \partial \boldsymbol{\beta}_m'} \right\| = \sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \left\| \frac{\boldsymbol{x}_t \boldsymbol{x}_t' f^2}{\sigma_\varepsilon^2} \right\| \leq \sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \left\| \frac{\boldsymbol{x}_t \boldsymbol{x}_t' f^2}{c} \right\| \leq \left\| \frac{\boldsymbol{x}_t \boldsymbol{x}_t'}{c} \right\|.$$

The last inequality follows from the fact that $|f| \leq 1$. Therefore,

$$\mathbb{E} \sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \left\| \frac{\partial^2 \ell_t}{\partial \boldsymbol{\beta}_m \partial \boldsymbol{\beta}_m'} \right\| \leq \frac{\mathbb{E} \|\boldsymbol{x}_t \boldsymbol{x}_t'\|}{c} < \infty.$$

We next examine the second derivatives of the log likelihood with respect to σ_ε^2,

$$\left|\frac{\partial^2 \ell_t}{\partial(\sigma_\varepsilon^2)^2}\right| = \left|\frac{1}{2\sigma_\varepsilon^4} - \frac{\varepsilon_t^2}{\sigma_\varepsilon^6}\right| \leq \left|\frac{1}{2\sigma_\varepsilon^4}\right| + \left|\frac{\varepsilon_t^2}{\sigma_\varepsilon^6}\right|,$$

$$\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \left|\frac{\partial^2 \ell_t}{\partial(\sigma_\varepsilon^2)^2}\right| \leq \frac{1}{2c^2} + \frac{1}{c^3} \sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \varepsilon_t^2.$$

In order to show $\mathbb{E}\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \left|\frac{\partial^2 \ell_t}{\partial(\sigma_\varepsilon^2)^2}\right| < \infty$, it is sufficient to show that $\mathbb{E}\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}(\varepsilon_t^2) < \infty$. Recall we have already proved in Lemma 3 (2) that

$$\varepsilon_t^2 \leq KL(M+1)\left\|\boldsymbol{x}_t\right\|^2.$$

It follows that

$$\mathbb{E}\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}(\varepsilon_t^2) \leq KL(M+1)\mathbb{E}\left\|\boldsymbol{x}_t\right\|^2 < \infty.$$

To show that $\mathbb{E}\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \left|\frac{\partial^2 \ell_t}{\partial\gamma_i^2}\right| < \infty$, consider

$$\begin{aligned}
\left|\frac{\partial^2 \ell_t}{\partial\gamma_m^2}\right| &= \left|\frac{-\left(\boldsymbol{x}_t'\boldsymbol{\beta}_m \frac{\partial f}{\partial\gamma_m}\right)^2 + \varepsilon_t\left(\boldsymbol{x}_t'\boldsymbol{\beta}_m \frac{\partial^2 f}{\partial\gamma_m^2}\right)}{\sigma_\varepsilon^2}\right|, \\
&\leq \frac{1}{c}\left(\frac{\partial f}{\partial\gamma_m}\right)^2 \left|\boldsymbol{x}_t'\boldsymbol{\beta}_m\right|^2 + \frac{1}{c}\left|\frac{\partial^2 f}{\partial\gamma_m^2}\right| \left|\varepsilon_t\right| \left|\boldsymbol{x}_t'\boldsymbol{\beta}_m\right|, \\
&\leq \frac{L}{c}\left(\frac{\partial f}{\partial\gamma_m}\right)^2 \left\|\boldsymbol{x}_t\right\|^2 + \frac{1}{c}\left|\frac{\partial^2 f}{\partial\gamma_m^2}\right| \left|\varepsilon_t\right| \left|\boldsymbol{x}_t'\boldsymbol{\beta}_m\right|,
\end{aligned}$$

where L is some positive constant. The second term on the right side can be written as

$$\begin{aligned}
\frac{1}{c}\left|\frac{\partial^2 f}{\partial\gamma_m^2}\right| \left|\varepsilon_t\right| \left|(\boldsymbol{x}_t'\boldsymbol{\beta}_m)\right| &= \frac{1}{c}\left|\frac{\partial^2 f}{\partial\gamma_m^2}\right| \left|\boldsymbol{x}_t'(\boldsymbol{\beta}_{0,0} - \boldsymbol{\beta}_0) + \sum_{m=1}^{M} \boldsymbol{x}_t'(\boldsymbol{\beta}_{m,0} f_{m,0} - \boldsymbol{\beta}_m f_m)\right| \left|\boldsymbol{x}_t'\boldsymbol{\beta}_m\right|, \\
&= \frac{1}{c}\left|\frac{\partial^2 f}{\partial\gamma_m^2}\right| \left|\boldsymbol{x}_t'(\boldsymbol{\beta}_{0,0} - \boldsymbol{\beta}_0)\right| \left|\boldsymbol{x}_t'\boldsymbol{\beta}_m\right| + \left|\sum_{m=1}^{M} \boldsymbol{x}_t'(\boldsymbol{\beta}_{m,0} f_{m,0} - \boldsymbol{\beta}_m f_m)\right| \left|\boldsymbol{x}_t'\boldsymbol{\beta}_m\right|, \\
&\leq \frac{1}{c}\left|\frac{\partial^2 f}{\partial\gamma_m^2}\right| K \left\|\boldsymbol{x}_t\right\|^2,
\end{aligned}$$

where K is some positive constant. Again, the compactness of the parameter space, boundedness of f, and stationarity of $\boldsymbol{x}_t$ ensures the existence of K and L. It follows that

$$\left|\frac{\partial^2 \ell_t}{\partial\gamma_i^2}\right| \leq \left(\frac{L}{c}\left(\frac{\partial f}{\partial\gamma_i}\right)^2 + \frac{1}{c}\left|\frac{\partial^2 f}{\partial\gamma_i^2}\right| K\right) \left\|\boldsymbol{x}_t\right\|^2.$$

The finiteness of the derivatives of f was shown in Lemma 1. Thus,

$$\mathbb{E}\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \left|\frac{\partial^2 \ell_t}{\partial\gamma_i^2}\right| \leq \left(\frac{L}{c}\left(\frac{\partial f}{\partial\gamma_i}\right)^2 + \frac{1}{c}\left|\frac{\partial^2 f}{\partial\gamma_i^2}\right| K\right) \mathbb{E}\left\|\boldsymbol{x}_t\right\|^2 < \infty.$$

The proof that $\mathbb{E}\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}\left|\frac{\partial^2 \ell_t}{\partial c_i^2}\right| < \infty$ closely resembles the proof above and is omitted for brevity. $\square$

Proof of Theorem 4. The proof establishes the conditions of Theorem 4.1.3 of Amemiya (1985) with a generalization due to Ling and McAleer (2003, Theorem 3.1). We need consistency of $\widehat{\boldsymbol{\theta}}_T$ for $\boldsymbol{\theta}_0$, which was shown in Theorem 3. Then we show

$$\boldsymbol{B}(\boldsymbol{\theta}_0)^{-\frac{1}{2}}\frac{1}{\sqrt{T}}\sum_{t=1}^{[rT]}\left.\frac{\partial \ell_t}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}_0} \xrightarrow{d} W(s),\ s\in[0,1],$$

where $W(r)$ is N-dimensional standard Brownian motion on the unit interval. This is condition (C) in Theorem 4.1.3 of Amemiya (1985). The convergence follows from Theorem 18.3 in Billingsley (1999) if (a) $\left\{\left.\frac{\partial \ell_t}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}_0}, \mathcal{F}_t\right\}$ is a stationary martingale difference, and (b) $\boldsymbol{B}(\boldsymbol{\theta}_0)$ exists. Both conditions were shown in Lemma 3.

To satisfy condition (B) of Theorem 4.1.3 of Amemiya (1985), we have to establish

$$\boldsymbol{A}_T(\boldsymbol{\theta}_T^*) \xrightarrow{p} \boldsymbol{A}(\boldsymbol{\theta}_0)$$

for any sequence $\boldsymbol{\theta}_T^* \xrightarrow{p} \boldsymbol{\theta}_0$,

$$\boldsymbol{A}_T(\boldsymbol{\theta}_T^*) = -\frac{1}{T}\sum_{t=1}^{T}\left.\frac{\partial^2 \ell_t}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'}\right|_{\boldsymbol{\theta}_T^*},$$

and

$$\boldsymbol{A}(\boldsymbol{\theta}_0) = -\mathbb{E}\left.\frac{\partial^2 \ell_t}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'}\right|_{\boldsymbol{\theta}_0}$$

is non-singular. Conditions for the double stochastic convergence can be found in Theorem 21.6 of Davidson (1994). We need to show

1. consistency of $\widehat{\boldsymbol{\theta}}_T$ for $\boldsymbol{\theta}_0$ (Theorem 3), and
2. uniform convergence of $\boldsymbol{A}_T$ to $\boldsymbol{A}$ in probability, i.e.

$$\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}|\boldsymbol{A}_T(\boldsymbol{\theta}) - \boldsymbol{A}(\boldsymbol{\theta})| \xrightarrow{p} 0.$$

We prove uniform convergence of $\boldsymbol{A}_T$ using Theorem 3.1 of Ling and McAleer (2003), who generalize Theorem 4.2.1 of Amemiya (1985) from i.i.d. data to stationary and ergodic data. This allows the immediate invocation of the Ergodic Theorem without having to check finiteness of third derivatives of ℓ_t as in Andrews (1992, Theorem 2). To apply Theorem 3.1 of Ling and McAleer (2003) we need that

$$g_t(\boldsymbol{\theta}) = -\frac{\partial^2 \ell_t}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'} - \boldsymbol{A}(\boldsymbol{\theta})$$

is continuous in $\boldsymbol{\theta}$ (this also establishes condition (A) of Theorem 4.1.3. of Amemiya (1985) along the way), has expected value $\mathbb{E}g_t(\boldsymbol{\theta}) = 0$ and is absolutely uniformly integrable:

$$\mathbb{E}\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}|g_t(\boldsymbol{\theta})| < \infty.$$

This was shown in Lemma 4. Thus, we have established all conditions for asymptotic normality according to Theorem 4.1.3 of Amemiya (1985). $\square$

Proof of Proposition 2. The proof of uniform convergence in probability of $\boldsymbol{A}_T$ to $\boldsymbol{A}$ is given in Lemma 4 and Theorem 4. We need to show uniform convergence of $\boldsymbol{B}_T$ to $\boldsymbol{B}$. We employ Theorem 3.1 of Ling and McAleer (2003) again and show that

$$h_t(\boldsymbol{\theta}) := \frac{\partial \ell_t}{\partial \boldsymbol{\theta}} \frac{\partial \ell_t}{\partial \boldsymbol{\theta}'} - \boldsymbol{B}(\boldsymbol{\theta}),$$

is absolutely uniformly integrable, continuous in $\boldsymbol{\theta}$, and has expected value $\mathbb{E}h_t(\boldsymbol{\theta}) = 0$. The detailed proof is in complete analogy to Lemma 4 and is omitted for brevity. $\square$

Vita

Junyue Xu was born in February 1982, in Shanghai, China. He finished his undergraduate studies at Fudan University in 2004 with a major in finance. He earned a master of science degree in economics from Louisiana State University in 2008. While in the Economic Department at Louisiana State University, he conducted many researches in the area of time series econometrics and financial economics under the guidance of Dr. Eric Hillebrand and Dr. Ambar Sengupta. All of these researches involved substantial amount of scientific programming in Matlab and C++. He also served as an instructor to teach Principle of Microeconomics during this period.