Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Xiaochun Liu Date

Three Essays in Modeling and Forecasting Economic Dynamics

By

Xiaochun Liu
Doctor of Philosophy

Economics

Richard Luger, Ph.D.
Co-Advisor

Tao Zha, Ph.D.
Co-Advisor

Esfandiar Maasoumi, Ph.D.
Committee Member

David Jacho-Chavez, Ph.D.
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

Three Essays in Modeling and Forecasting Economic Dynamics

By

Xiaochun Liu
M.A., Emory University, 2013
M.S., Marquette University, 2008

Advisors: Richard Luger, Ph.D.
Tao Zha, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Economics
2014

Abstract

Three Essays in Modeling and Forecasting Economic Dynamics

By Xiaochun Liu

An important problem of modern financial economics is to understand and quantify the interaction between macroeconomics and financial markets under hypothetically distress scenarios. Modeling economic dynamics given a hypothetically distress economic scenario requires knowledge and techniques for extreme values or events. In this dissertation, on the econometrics side, I propose new econometric approaches for modeling evolutionary processes in tails of a data distribution, and constructing decomposition models to forecast excess stock returns by considering the role of dynamic higher moments, such as time-varying skewness. On the financial economics side, my research analyzes the counter-cyclical risk pattern of stock markets, asymmetric dynamics in macroeconomic variables, and systemic risk measure of financial institutions subject to regime switching in tails. The first essay proposes a new time-series econometric model to estimate quantiles of a data distribution subject to regime shifts, so-called Markov-Switching Quantile Autoregression (MSQAR). The purpose of this new econometric model is to characterize nonstationary natures of different parts of a data distribution. This is achieved via the assumption that quantile error terms follow a three-parameter asymmetric Laplace distribution. To deal with the difficulty in model estimation, I adopt a "block-at-a-time" Metropolis-Hastings sampling. The second essay applies the proposed MSQAR approach to stress-testing the U.S largest commercial banks by measuring systemic risk of individual banks subject to economic regime shifts. The new systemic risk measures show that the benchmark model of CoVaR approach underestimates systemic risk contributions of individual banks by around 131 basis points of asset loss on average. In addition, Banking Systemic Risk Index is constructed by value-weighted individual contributions. The third essay proposes a new approach to modeling time-varying skewness, the model performance of which is evaluated in out-of-sample forecast of the U.S. excess stock returns in terms of both statistical significance and economic values. Interestingly, a forecast combination, more robust to structural instability than the individual forecasts, performs significantly better out-of-sample than the benchmarks. The skewness timing of the proposed time-varying dependence models yields an average gain in the returns around 195 basis points per year over the forecast sample period.

Three Essays in Modeling and Forecasting Economic Dynamics

By

Xiaochun Liu
M.A., Emory University, 2013
M.S., Marquette University, 2008

Advisors: Richard Luger, Ph.D.
Tao Zha, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Economics
2014

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to my advisor Dr. Richard Luger for his continuous support throughout my Ph.D. program and dissertation development, even after he left Emory University in 2010. Dr. Luger has not only trained me superbly well in econometrics knowledge and programming skills, but has educated my attitude in academic rigorousness. Richard is very strict in academic integrity and details in research development. Most importantly, Richard was always patient to listen and ready to give me advice. He has taught me how to write precise and concise research papers. Also, I would like to thank my advisor Dr. Tao Zha. Tao has been very knowledgeable in advising me Markov-Switching model development which is the key of the first two chapters of my dissertation. His advice helps me gain many insights and economics intuitions beyond econometric techniques. In addition, I would like to grate for my committee members, Dr. Esfandiar Maasoumi and Dr. David Jacho-Chavez for their valuable comments and support.

Contents

List of Tables

List of Figures

Review

An important problem of modern financial economics is understanding and quantifying the interaction between macroeconomics and financial markets. Since the recent financial crisis of 2007-2009, studies in this area have emphasized on round effects between macroeconomics and financial sectors. Specifically, dynamic interactions are tested under hypothetically distress scenarios, known as stress-testing. Modeling economic dynamics given a hypothetically distress economic scenario requires knowledge and techniques for extreme values or events, largely apart from traditional econometric models on conditional means. Extreme values deviate from means and fall into lower and upper tails. And, reactions of market participants to extreme events can be to some extent characterized by higher moments of data distributions. In this dissertation, on the econometrics side, I propose new econometric approaches for modeling evolutionary processes in tails of a data distribution, and constructing decomposition models to forecast excess stock returns by considering the role of dynamic higher moments, such as time-varying skewness. On the financial economics side, my research analyzes the countercyclical risk pattern of stock markets, asymmetric dynamics in macroeconomic variables, and systemic risk measure of financial institutions subject to regime switching in tails. Below I briefly elaborate on the three papers that constitute this dissertation.

The first essay proposes a new time-series econometric model to estimate quantiles of a data distribution subject to regime shifts. The purpose of this new

econometric model is to characterize nonstationary natures of different parts of a data distribution. This is achieved via the assumption that quantile error terms follow a three-parameter asymmetric Laplace distribution. Particularly, I develop location-scale quantile autoregressive models in which the location and scale parameters are subject to regime shifts. Regime changes within a quantile is determined by a latent, discrete-state Markov process. Crucial inference for filtering transition probabilities of switching regimes is made through asymmetric Laplace distribution. The proposed model is referred to as Markov-Switching Quantile Autoregression (MSQAR) which nests Quantile Autoregression (QAR) of Koenker and Xiao (2006) as a special case.

MSQAR models are non-linear and involve indicator functions, which introduce kinks and discontinuities into the sample likelihood function. In addition, less observations fall in more extreme quantiles, which leads to the potential small sample issue. These issues make classical methods such as MLE very difficult for model estimation. In this essay based on the findings in Chernozhukov and Hong (2003), I adopt a "block-at-a-time" Metropolis-Hastings sampling by grouping highly correlated parameters as one block to be simultaneously updated at each Metropolis-Hasting step conditional on the remaining blocks. Despite that this sampling approach has been applied in previous studies, i.e., Tierney (1994), Ausin and Lopes (2010), Geweke and Tanizaki (2001), among others, to the best of my knowledge, this paper is the first to apply this sampling approach in quantile regressions to mitigate the estimation difficulty due to potential local optima on likelihood function surface. To achieve mixing properties, I further employ the adaptive scheme of Gerlach et al. (2011) and Chen et al. (2012), which is modified to the "block-at-a-time" Metropolis-Hasting sampling, in order to avoid stucking in a local mode for a long time.

The second essay applies the proposed MSQAR approach to stress-testing the U.S largest commercial banks by measuring systemic risk of individual banks subject to economic regime shifts. Specifically, I characterize two risk states: a normal risk level implied by good economic periods and a high risk level associated with economic recessions, crises or extreme events. Different sets of parameters are thus obtained from MSQAR model estimation for each risk state in stress-testing. Apparently, Markov-Switching conditional value-at-risk (MSCoVaR) measure of

systemic risk naturally fits to the Supervisory Stress Scenario required by Federal Reserve Bank in Comprehensive Capital Analysis and Review (CCAR). In CCAR, a supervisory stress scenario is a hypothetical scenario to be used to assess the strength and resilience of bank holding companies (BHCs) capital in a severely adverse economic environment. It represents an outcome in which the U.S. economy experiences a significant recession and economic activity in other major economies also contracts significantly. It is reasonable to expect that economic representatives react differently to different economic conditions. Hence, the set of MSQAR parameters obtained from the high risk level associated with economic recessions and crises can be appropriately applied to stress-testing BHCs in Fed's severely adverse economic scenarios.

The new systemic risk measures show in comparison that the CoVaR approach of Adrian and Brunnermeier (2011) underestimates systemic risk contributions of individual banks by around 131 basis points of asset loss on average. The empirical results also present that the banking system is more sensitive to marginal changes of an individual bank during high risk episodes than during normal risk periods. In addition, Banking Systemic Risk Index, which is constructed in this essay by value-weighted individual contributions, appears to have a high relevance in tracing financial distress situations over the sample period.

The third essay studies the importance of dynamic higher moments in forecasting financial market behaviors. Anatolyev and Gospodinov (2010) exploit inherent nonlinearity in excess return dynamics by a return decomposition to the product of absolute returns and return signs. The joint distribution of absolute values and signs is characterized by a copula function. Their decomposition approach, however, is restrictive as the dependence between absolute returns and signs is constant over sample periods. The constant dependence also imposes a constant skewness on excess returns. The literature has recognized that returns may in fact be better characterized by a conditional distribution with time-varying asymmetry. The importance of time-varying skewness has also been found in asset pricing and allocation by recent studies, i.e., Harvey and Siddique (1999, 2000a&b), Leon et al. (2005), and Jondeau and Rockinger (2003), among others. Leon et al. (2005) estimate time-varying skewness and kurtosis using a Gram-Charlier series expansion of the normal density function for the error term. It is found that specifications

allowing for time-varying skewness and kurtosis outperform specifications with constant third and fourth moments. Jondeau and Rockinger (2003) use the generalized student-t distribution with an autoregressive specification of the parameters to demonstrate the importance of time-varying asymmetry parameters.

In this essay, I propose a new approach to modeling time-varying skewness, the model performance of which is evaluated in out-of-sample forecast of the U.S. excess stock returns in terms of both statistical significance and economic values. Specifically, I extend AG's constant decomposition model by characterizing the joint distribution as a time-varying copula function. The nonlinear temporary interdependence between absolute returns and signs, which governs dynamic skewness processes of returns, is thus estimated by the dynamic copula function simultaneously with marginals. Besides modeling time-varying skewness in out-of-sample forecast, this paper also differs from AG's work in several important ways: (1) The out-of-sample forecast period is extended to cover the recent financial crisis of 2007-2009 which has attracted tremendous research interests in both economic and finance literature. (2) The Fluctuation tests and the decomposition of forecast performance, proposed by Giacomini and Rossi (2010) and Rossi and Sekhposyan (2011) respectively, show strong statistical evidence of the instability of forecast performance over the sample time paths. The forecast performance is evaluated by loss functions of forecast errors. Interestingly, a forecast combination, more robust to structural instability than the individual forecasts, performs significantly better out-of-sample than the benchmarks. (3) The economic values of skewness timing present substantial benefits from modeling time-varying skewness. The skewness timing of the proposed time-varying dependence models yields an average gain in the returns around 195 basis points per year over the forecast sample period.

Chapter 1

Markov-Switching Quantile Autoregression

Abstract: This chapter considers the location-scale quantile autoregression in which the location and scale parameters are subject to regime shifts. The regime changes are determined by the outcome of a latent, discrete-state Markov process. The new method provides direct inference and estimate for different parts of a nonstationary time series distribution. Bayesian inference for switching regimes within a quantile, via a three-parameter asymmetric-Laplace distribution, is adapted and designed for parameter estimation. The simulation study shows reasonable accuracy and precision in model estimation. From a distribution point of view, rather than from a mean point of view, the potential of this new approach is illustrated in the empirical applications to reveal the countercyclical risk pattern of stock markets and the asymmetric persistence of real GDP growth rates and real trade-weighted exchange rates.

Keywords: Asymmetric-Laplace Distribution, Metropolis-Hastings, Block-at-a-Time, Asymmetric Dynamics, Transition Probability

JEL: C22, C58, C51, C11, G23

1.1. Introduction

Koenker and Xiao (2006) study quantile autoregression models in which the autoregressive coefficients may take distinct values over different quantiles of the innovation process. Their models can capture systematic influences of conditioning variables on the location, scale and shape of the conditional distribution. Let $\{U_t\}$ be a sequence of i.i.d. standard uniform random variables. Consider the mth-order autoregressive process

$$y_t = \theta_0(U_t) + \theta_1(U_t)\, y_{t-1} + ... + \theta_m(U_t) y_{t-m} \tag{1.1.1}$$

where y_t is the time series observation at time t, and θ's are unknown functions $[0,1] \to \mathbb{R}$ to be estimated. Provided that the right side of (1.1.1) is monotone increasing in U_t, it follows that the τth conditional quantile function of y_t can be obtained as

$$Q_{y_t}\left(\tau | \boldsymbol{y}_{t-1}\right) = \theta_0(\tau) + \theta_1(\tau) y_{t-1} + ... + \theta_m(\tau) y_{t-m} \tag{1.1.2}$$

where $\boldsymbol{y}_{t-1} = (y_{t-1}, ..., y_{t-m})'$. The transition from (1.1.1) to (1.1.2) is an immediate consequence of equivariance to monotone transformations.[1] In (1.1.2), the quantile autoregressive coefficients may be τ-dependent and thus can vary over the quantiles. The conditioning variables not only shift the location of the distribution of y_t , but also may alter the scale and shape of the conditional distribution. Koenker and Xiao (2006) also show that quantile autoregressive models exhibit a form of asymmetric persistence and temporarily explosive behavior.

[1] See the theorem of equivariance to monotone transformations in Koenker (2005), page 39.

However, the linear quantile autoregressive models cannot accommodate many stylized facts such as structural breaks and nonlinearities in macroeconomic and financial time series. The aim of this article is to extend the quantile autoregression of Koenker and Xiao (2006) by modeling nonstationary quantile dynamics. Particularly, I consider the location-scale quantile autoregression in which the location and scale parameters are subject to regime shifts within a quantile. Switching quantile regimes is determined by the outcome of an unobserved state indicator variable that follows a Markov process with unknown transition probabilities. The proposed Markov-Switching Quantile Autoregression (MSQAR) nests the quantile autoregression of Koenker and Xiao (2006) as a special case when conditional distributions are stationary.

MSQAR is a convenient approach built on the vast literature of Markov-switching time series models.[2] Nonetheless, simply combining quantile autoregressive models with Markov-switching techniques is econometrically infeasible. The challenge is that the objective function of quantile autoregression is a non-likelihood based function generally estimated by nonlinear least square. The non-likelihood based function does not allow make inference on the latent state variable for switching regimes. To solve this problem, I assume that quantile error terms follow a three-parameter asymmetric-Laplace distribution (Yu and Zhang (2005)). This chapter shows that maximizing this distribution is mathematically equivalent to minimizing quantile objective functions. Importantly, it also satisfies the restrictive conditions of quantile regression. With this distribution, the inference for switch-

[2] See e.g., Sims and Zha, 2006, Gray (1996), Cheung and Erlandsson (2005), Hamilton and Susmel (1994), Kim et al. (2008), among many others. Guidolin (2012) provides a recent review for the applications of Markov-switching models in empirical finance.

ing quantile regimes can be made through the standard Hamilton filter approach (Hamilton (1994)).

This chapter adopts Bayesian approach for model estimation. As discussed in Yu and Moyeed (2001), the use of an asymmetric Laplace distribution for error terms provides a natural way to deal with some serious computational challenges through Bayesian quantile regression. Also see Chernozhukov and Hong (2003). In the terminology of Chib and Greenberg (1995), this chapter adopts a "block-at-a-time" Metropolis-Hastings sampling to reduce computational cost. This algorithm groups highly correlated parameters as one block to be simultaneously updated at each Metropolis-Hasting step conditional on the remaining blocks, see e.g., Tierney (1994), Ausin and Lopes (2010), Geweke and Tanizaki (2001), among others. To further speed up convergence and to achieve desirable mixing properties in MCMC chains, I employ the adaptive scheme of Gerlach et al. (2011) and Chen et al. (2012), which combines a random walk and an independent kernel Metropolis-Hastings algorithm, each based on a mixture of multivariate normal distributions.

This chapter examines the new approach in a simulation study to show its accuracy and precision in model estimation. The empirical application to S&P 500 returns illustrates the usefulness of this new approach in risk management, i.e., for stress-testing financial institutions from the perspective of central banks. In this chapter, asymmetric dynamics have also been found for quarterly real GDP growth rates but not for quarterly real trade-weighted U.S. dollars. In addition, the asymmetric dynamics appear to be different across economic regimes. Notably, modeling the regime persistence in lower tails of real GDP growth rates improves the predictabilities of switching economic states and turning points.

The rest of this chapter is structured as follows. Section 2 introduces the connection of asymmetric-Laplace distributions to the solution of quantile regressions. Section 3 defines Markov-Switching quantile autoregression. Section 4 describes the Bayesian methods in this chapter for model estimation. Section 5 presents model simulations and results. Section 6 reports the results of empirical applications to stock markets, real GDP growth rates and real trade-weighted exchange rates. Section 7 concludes this chapter.

1.2. Asymmetric Laplace Distribution Connection

The QAR(m) model of (1.1.2) can be reformulated in a more conventional regression form as

$$y_t = \theta_0(\tau) + \sum_{l=1}^{m} \theta_l(\tau) y_{t-l} + \varepsilon_t(\tau) \tag{1.2.1}$$

where $\varepsilon_t(\tau)$ is quantile error terms which follow an asymmetric-Laplace (AL) distribution, denoted by $AL(0, \varsigma, \tau)$, with the density function as

$$f_{\varepsilon_t}(\varepsilon; 0, \varsigma, \tau) = \frac{\tau(1-\tau)}{\varsigma} exp\left\{ -\frac{\varepsilon\left(\tau - I(\varepsilon \leq 0)\right)}{\varsigma} \right\} \tag{1.2.2}$$

where $I(\cdot)$ is an indicator function. τ determines the skewness of the distribution, $\varsigma > 0$ is a scale parameter. $AL(0, \varsigma, \tau)$ with the location parameter being zero provides that the τth quantile of the distribution is zero as $Pr\left(\varepsilon_t \leq 0\right) = \tau$, which satisfies the quantile regression condition $\int_{-\infty}^{0} f_{\varepsilon}(s)ds = \tau$. The asymmetric-Laplace distribution with the density function of (1.2.2) has the mean and variance, $E(\varepsilon_t) = \varsigma(1-2\tau)/[(1-\tau)\tau]$ and $Var(\varepsilon_t) = \varsigma^2(1-2\tau+2\tau^2)/[(1-\tau)^2\tau^2]$, respectively. See Yu and Zhang (2005) for details. With the assumption of i.i.d. $\varepsilon_t(\tau)$, the sample

likelihood function is given by

$$\begin{aligned} L(\boldsymbol{\theta},\tau) &= [\tau(1-\tau)/\varsigma]^T \\ & exp\left\{-\sum_{t=1}^{T}\frac{y_t - Q_{y_t}\left(\tau|\boldsymbol{y}_{t-1}\right)}{\varsigma}\left[\tau - I\left(y_t \leq Q_{y_t}\left(\tau|\boldsymbol{y}_{t-1}\right)\right)\right]\right\} \end{aligned} \tag{1.2.3}$$

In the literature the error density is often left unspecified, see e.g., Koenker and Bassett (1978), Koenker (2005), and Koenker and Xiao (2006), etc. Quantile autoregression is the solution to the following minimization problem

$$\boldsymbol{\theta}(\tau) = arg\ \underset{\boldsymbol{\theta}}{min}\ E\left(\rho_\tau\left(y_t - Q_{y_t}\left(\tau|\boldsymbol{y}_{t-1};\boldsymbol{\theta}\right)\right)\right) \tag{1.2.4}$$

where $\theta\left(\tau\right) = (\theta_0(\tau), ..., \theta_m\left(\tau\right))$ is the parameter space to be estimated. The quantile criterion (check or loss) function $\rho_\tau(\cdot)$ is defined as $\rho_\tau(\varepsilon) = \varepsilon\left(\tau - I(\varepsilon < 0)\right)$ in Koenker and Bassett (1978). Solving the sample analog gives the estimator of θ

$$\hat{\boldsymbol{\theta}}(\tau) = arg\ \underset{\boldsymbol{\theta}}{min}\ \sum_{t=1}^{T}\rho_\tau\left(y_t - Q_{y_t}\left(\tau|\boldsymbol{y}_{t-1};\boldsymbol{\theta}\right)\right) \tag{1.2.5}$$

Recently, Yu and Moyeed (2001), Yu and Zhang (2005) and Gerlach et al. (2011), among others, have illustrated the link between the quantile estimation problem and asymmetric-Laplace distribution. Since the quantile loss function is contained in the exponent of the asymmetric-Laplace likelihood, maximizing the sample likelihood of (1.2.3) is mathematically equivalent to minimizing the quantile loss function of (1.2.5). It is important to emphasize that, in practice, the parameter τ is chosen by researchers as quantile levels of interest during parameter estimation and only a single quantile of the distribution of y_t is estimated. More importantly, the asymmetric Laplace distribution transforms the non-likelihood

based quantile regression of (1.2.5) to a likelihood based approach, so that the inference for the probability of switching regimes is possibly made through Hamilton filter.

1.3. Markov-Switching Quantile Autoregression

For the τth conditional quantile of y_t, let $\{s_t\}$ be an ergodic homogeneous Markov chain on a finite set $S = \{1, ..., k\}$, with a transition matrix P defined by the following transition probabilities

$$\{p_{ij} = Pr\left(s_t = j | s_{t-1} = i\right)\}$$

and the unconditional probabilities

$$\{\pi_j = Pr\left(s_t = j\right)\}$$

for $i, j \in S$ and assume s_t follow a first-order Markov chain. The transition probabilities satisfy $\sum_{j \in S} p_{ij} = 1$ and $\sum_{j \in S} \pi_j = 1$. The stochastic process for s_t is strictly stationary if p_{ij} is less than unity and does not take on the value of 0 simultaneously.

Using transition probabilities above, this chapter defines Markov-Switching quantile autoregressive models (MSQAR) as

$$\begin{aligned} y_t &= Q_{y_t}(\tau | \mathbf{y}_{t-1}; \boldsymbol{\theta}_{s_t}) + \varepsilon_t(\tau) \\ &= \theta_{s_t,0}(\tau) + \sum_{l=1}^{m} \theta_{s_t,l}(\tau) y_{t-l} + \varepsilon_t(\tau) \end{aligned} \tag{1.3.1}$$

Suppose that $\mathbf{y}_t$ can be observed directly but can only make an inference about the value of s_t based on the observations as of date t. This inference gives the filtering probability as

$$\begin{aligned}\xi_{j,t|t} &= Pr\left(s_t = j|\mathbf{y}_t;\boldsymbol{\Theta}\right) \\ &= \sum_{i\in S} Pr\left(s_t = j, s_{t-1} = i|\mathbf{y}_t;\boldsymbol{\Theta}\right)\end{aligned}$$

where $\sum_{j\in S}\xi_{j,t|t} = 1$ and $\boldsymbol{\Theta} = (P, \boldsymbol{\theta}_{s_t}(\tau))$ is a vector of the parameters with $s_t \in S$. The formulation of filtering probabilities is obtained by Bayes theorem as

$$\xi_{j,t|t} = \frac{\sum_{i\in S} p_{ij}\xi_{i,t-1|t-1}\eta_{j,t}}{f\left(y_t|\mathbf{y}_{t-1},\tau;\boldsymbol{\Theta}\right)} \tag{1.3.2}$$

where $\eta_{j,t}$ is conditional likelihood as

$$\begin{aligned}\eta_{j,t} &= f\left(y_t|s_t = j,\mathbf{y}_{t-1},\tau;\boldsymbol{\theta}\right) \qquad (1.3.3)\\ &= \frac{\tau(1-\tau)}{\varsigma} exp\left\{-\frac{\left(y_t - Q_{y_t}\left(\tau|\mathbf{y}_{t-1};\boldsymbol{\theta}_j\right)\right)}{\varsigma}\left[\tau - I\left(y_t < Q_{y_t}\left(\tau|\mathbf{y}_{t-1};\boldsymbol{\theta}_j\right)\right)\right]\right\}\end{aligned}$$

and

$$f(y_t|\mathbf{y}_{t-1},\tau;\boldsymbol{\Theta}) = \sum_{j\in S}\sum_{i\in S} p_{ij}\xi_{i,t-1|t-1}\eta_{j,t}$$

Thus, the relationship between the filtering and prediction probabilities is given by

$$\xi_{j,t+1|t} = Pr\left(s_{t+1} = j|\mathbf{y}_t;\boldsymbol{\Theta}\right) = \sum_{i\in S} p_{ij}\xi_{i,t|t} \tag{1.3.4}$$

The inference, similar to Hamilton's filter (Hamilton, 1994), is performed iteratively for $t = 1, ..., T$ with the initial values, $\xi_{j,0|0}$ for $j \in S$. The sample likelihood

for the τth conditional quantile of y_t is then given by

$$L(\boldsymbol{\Theta}) = \prod_{t=1}^{T} f(y_t|\mathbf{y}_{t-1}, \tau; \boldsymbol{\Theta}) \tag{1.3.5}$$

The connection to the solution of quantile regression can also be viewed as follows. Based on quantile loss functions, $\boldsymbol{\Theta}$ is solved for the following minimization problem

$$\min_{\boldsymbol{\Theta}} E\left(\sum_{j\in S} \rho_\tau \left(y_t - Q_{y_t}\left(\tau|s_t = j,\ \mathbf{y}_{t-1}; \boldsymbol{\Theta}\right)\right) I\left(s_t = j\right)\right) \tag{1.3.6}$$

where $\mathbf{y}_t = \{y_t, y_{t-1}, ..., y_1, y_0\}$. Apply the law of iterated expectation to rewritten (1.3.6) as

$$\min_{\boldsymbol{\Theta}} \sum_{j\in S} E\left[\rho_\tau \left(y_t - Q_{y_t}\left(\tau|s_{t,\tau} = j,\ \mathbf{y}_{t-1}; \boldsymbol{\Theta}\right)\right) Pr\left(s_{t,\tau} = j|\mathbf{y}_t; \boldsymbol{\Theta}\right)\right] \tag{1.3.7}$$

Provided that τ is chosen by researchers of interest, maximizing the likelihood of (1.3.5) is mathematically equivalent to the minimization of (1.3.6), since the likelihood function can be alternatively rewritten as $L(\boldsymbol{\Theta}) = \prod_{t=1}^{T} \sum_{j\in S} f(y_t|s_t = j, \mathbf{y}_{t-1}, \tau; \boldsymbol{\Theta}) Pr\left(s_t = j|\mathbf{y}_t; \boldsymbol{\Theta}\right)$ with $Pr(s_t = j|\mathbf{y}_t; \boldsymbol{\Theta}) = \sum_{i\in S} p_{ij}\xi_{i,t-1|t-1}$. However, $Pr\left(s_t = j|\mathbf{y}_t; \boldsymbol{\Theta}\right)$ cannot be filtered by using the nonlinear least square estimation of (1.3.7); therefore, the likelihood function of the asymmetric Laplace distribution is used to infer transition probabilities.

To estimate smoothing transition probabilities $Pr\left(s_t = i|\mathbf{y}_T; \theta\right)$, this chapter follows the approach of Kim (1994). Apply the Bayes theorem and the Markov property to yield

$$Pr\left(s_t = i|s_{t+1} = j,\ \mathbf{y}_T; \mathbf{\Theta}\right) = \frac{p_{ji} Pr\left(s_t = j|\mathbf{y}_t; \mathbf{\Theta}\right)}{Pr\left(s_{t+1} = i|\mathbf{y}_t; \mathbf{\Theta}\right)} \tag{1.3.8}$$

It is therefore the case that

$$Pr\left(s_t = j, s_{t+1} = i|\mathbf{y}_T; \mathbf{\Theta}\right) = Pr\left(s_{t+1} = i|\mathbf{y}_T; \mathbf{\Theta}\right) \frac{p_{ji} Pr\left(s_t = j|\mathbf{y}_t; \mathbf{\Theta}\right)}{Pr\left(s_{t+1} = i|\mathbf{y}_t; \mathbf{\Theta}\right)} \tag{1.3.9}$$

The smoothed inference for date t is the sum of (1.3.9) over $i \in S$

$$\begin{aligned} \xi_{j,t|T} &= Pr\left(s_t = j|\mathbf{y}_T; \mathbf{\Theta}\right) \\ &= \sum_{i \in S} Pr\left(s_{t+1} = i|\mathbf{y}_T; \mathbf{\Theta}\right) \frac{p_{ji} Pr\left(s_t = j|\mathbf{y}_t; \mathbf{\Theta}\right)}{Pr\left(s_{t+1} = i|\mathbf{y}_t; \mathbf{\Theta}\right)} \end{aligned} \tag{1.3.10}$$

The smoothed transition probabilities are thus obtained by iterating on (1.3.10) backward for $t = T-1, T-2, ..., 1$. This iteration is started with $\xi_{j,T|T}$ for $j \in S$ which is estimated from (1.3.2) for $t = T$. This algorithm is valid only when s_t follows a first-order Markov chain.

From the conditional density (1.3.4), it is straightforward to forecast the one-step-ahead τth quantile of y_{t+1} at time t conditional on knowing $s_{t+1,\tau}$,

$$Q_{y_{t+1}}\left(\tau|s_{t+1} = j, \mathbf{y}_t; \boldsymbol{\theta}_j\right) = \theta_{j,0}(\tau) + \sum_{l=0}^{m-1} \theta_{j,l+1}(\tau) y_{t-l} \tag{1.3.11}$$

Further, from (1.3.4), the forecast for $t+1$ conditional on time t obtained as

$$Q_{y_{t+1}}\left(\tau|\mathbf{y}_t; \boldsymbol{\theta}\right) = \sum_{j=1}^{k} Q_{y_{t+1}}\left(\tau|s_{t+1} = j, \mathbf{y}_t; \boldsymbol{\theta}_j\right) Pr\left(s_{t+1} = j|\mathbf{y}_t, \mathbf{\Theta}\right) \tag{1.3.12}$$

which is to multiply the appropriate forecast of the quantile in the jth regime

given by (1.3.11) with the probability that the process will be in that regime given by (1.3.4), and to sum those products for every regime together. Note that h-step-ahead forecasts for $h > 2$ require different approaches since it involves forecasts of y_{t+h-1} in (1.3.11) for $Q_{y_{t+h}}(\tau|s_{t+h-1,\tau} = j, \mathbf{y}_{t+h-1}; \boldsymbol{\theta}_j)$, as shown in Cai (2010).

In MSQAR model estimation, similar to other Markov-Switching time series models, one must use some identification restrictions to avoid the label switching issue. See Bauwens et al. (2010) and Hamilton et al. (2007) for a discussion. In this chapter, regimes are labeled by the restrictions on quantile intercepts, for example, $\theta_{1,0}(\tau) > ... > \theta_{k,0}(\tau)$. In addition, in empirical applications, the transition probabilities are allowed but not imposed dependent on τ. The intuition is that even though economic states are common across quantiles implying the same unconditional probabilities, no theories show that regime persistence should be the same across quantiles. To obtain some insights on this empirical question, regime persistence is allowed to be driven by data across quantiles.

1.4. Bayesian Inference

MSQAR models are non-linear and involve indicator functions, which introduce kinks and discontinuities into the sample likelihood function in (1.3.5). In addition, less observations fall in more extreme quantiles, which leads to the potential small sample issue. These issues make classical methods such as MLE very difficult for model estimation. In this chapter, I instead prefer to use Bayesian MCMC methods to learn about the model parameters.

Given the sample realizations, y_t for $t = 1, ..., T$, the posterior distribution of

$\boldsymbol{\Theta}$ takes the usual form: $p(\boldsymbol{\Theta}|\mathbf{y}_t) \propto L(\mathbf{y}_t|\boldsymbol{\Theta})\pi(\boldsymbol{\Theta})$, where $L(\mathbf{y}_t|\boldsymbol{\Theta})$ is the sample likelihood function and $\pi(\boldsymbol{\Theta})$ is the prior distribution. Yu and Moyeed (2001) and Cai and Stander (2008) prove that the posterior distribution is proper under the improper prior for general quantile regression models. In this chapter, the prior distribution is taken as uniform over Ξ, the admissible parameter space of $\boldsymbol{\Theta}$, i.e., satisfying the label switching restrictions. The prior for the scale parameter is $\pi(\varsigma) \propto \varsigma^{-1}$ also used in Gerlach et al. (2011).

Just like Vrontos et al. (2002) and Ausin and Lopes (2010), I also find that MCMC mixing can be improved and the computational cost reduced by using simultaneous updating of the highly correlated parameter groups at each Metropolis-Hastings (MH) step. In the terminology of Chib and Greenberg (1995), this approach is therefore based on a "block-at-a-time" MH sampler which is carried out by cycling repeatedly through draws of each parameter block conditional on the remaining parameter blocks. Let $\boldsymbol{\Theta} = (\boldsymbol{P}, \boldsymbol{\theta}_1(\tau), ..., \boldsymbol{\theta}_k(\tau))$ represent the blocks of the population parameters. $\boldsymbol{P} = (p_{ij})$ contains all transition probability parameters and $\boldsymbol{\theta}_{j,\tau}$ includes all parameters in the jth regime for $j = 1, ..., k$. Hence, the parameters in $\boldsymbol{\Theta}$ are grouped in $k+1$ blocks and the parameters of each block are simultaneously updated conditional on the remaining blocks.

This chapter implements the MH sampler according to the adaptive scheme of Gerlach et al. (2011) and Chen et al. (2012) which combines the random walk MH (RW-MH) and the independent kernel MH (IK-MH) algorithms, each based on a mixture of multivariate normal distributions. The random walk part of this scheme is designed to allow occasional large jumps, perhaps away from local modes, thereby improving the chances that the Markov chain will explore

the posterior distribution space. Hence, this adaptive scheme allows for further speeding convergence and achieving desirable mixing properties in MCMC chains.

To illustrate this adaptive algorithm in the block-at-a-time MH sampler, I rewrite the notation of the parameter blocks as $\boldsymbol{\Theta} = (\boldsymbol{\theta}_{1,\tau}, \boldsymbol{\theta}_{2,\tau}, ..., \boldsymbol{\theta}_{k+1,\tau})$, where $\boldsymbol{\theta}_{1,\tau} = \boldsymbol{P}$ and $\boldsymbol{\theta}_{j,\tau}$ denotes the parameters in the $(j-1)$th regime for $j = 2, ..., k+1$. And, let $\boldsymbol{\Theta}_{-j}$ denote the vector $\boldsymbol{\Theta}$ excluding the block $\boldsymbol{\theta}_{j,\tau}$. Starting at $g = 1$ with $\boldsymbol{\Theta}^{[1]} = \left(\boldsymbol{\theta}^{[1]}_{1,\tau}, ..., \boldsymbol{\theta}^{[1]}_{k+1,\tau}\right)$, the G_1 random walk MH iterations for $\boldsymbol{\Theta}$ proceed as follows:

Step 1. Increment g by 1 and set $\boldsymbol{\Theta}^{[g]}$ equal to $\boldsymbol{\Theta}^{[g-1]}$.

Step 2. For $i = 1, ..., k+1$ in turn, generate $\boldsymbol{\theta}^{c}_{i,\tau}$ as

$$\boldsymbol{\theta}^{c}_{i,\tau} = \boldsymbol{\theta}^{[g]}_{i,\tau} + \boldsymbol{\varepsilon}, \qquad \boldsymbol{\varepsilon} \sim \rho N\left(\mathbf{0}, diag\left\{b_i\right\}\right) + (1-\rho) N\left(\mathbf{0}, \omega diag\left\{b_i\right\}\right)$$

and replace $\boldsymbol{\theta}^{[g]}_{i,\tau}$ in $\boldsymbol{\Theta}^{[g]}$ by $\boldsymbol{\theta}^{c}_{i,\tau}$ with the probability $min\left(\zeta_i, 1\right)$, where

$$\zeta_i = \frac{L\left(\boldsymbol{y}_t | \boldsymbol{\theta}^{c}_{i,\tau}, \boldsymbol{\Theta}^{[g]}_{-i}\right) \pi\left(\boldsymbol{\theta}^{c}_{i,\tau}, \boldsymbol{\Theta}^{[g]}_{-i}\right)}{L\left(\boldsymbol{y}_t | \boldsymbol{\Theta}^{[g]}\right) \pi\left(\boldsymbol{\Theta}^{[g]}\right)}$$

Step 3. If $g < G_1$, go to Step 1. Upon completion, these first G_1 iterations yield the burn-in sample. Following Chen et al. (2012), I set $\rho = 0.95$, $\omega = 100$, and tune the positive number b_i so that the empirical acceptance rate lies in the range $(0.2, 0.45)$ for the ith block. Tuning is done every 100 iterations by increasing b_i when the acceptance rate in the last 100 iterations is higher than 0.45, or decreasing b_i when that rate is lower than 0.2.

At the end of the first G_1 iterations, the burn-in sample mean $\boldsymbol{\mu}_{i,\tau}$ and covariance matrix $\boldsymbol{\Sigma}_{i,\tau}$ of $\boldsymbol{\theta}_{i,\tau}$ with corresponding lower triangular Cholesky factor $\boldsymbol{\Sigma}^{1/2}_{i,\tau}$

are computed for $i = 1, ..., k+1$. The MCMC sampling scheme then continues for G_2 additional iterations according to the following independent kernel MH steps:

Step 4. Increment g by 1 and set $\boldsymbol{\Theta}^{[g]}$ equal to $\boldsymbol{\Theta}^{[g-1]}$.

Step 5. For $i = 1, ..., k+1$ in turn, generate $\boldsymbol{\theta}_{i,\tau}^{c}$ as

$$\boldsymbol{\theta}_{i,\tau}^{c} = \boldsymbol{\mu}_{i,\tau} + \boldsymbol{\Sigma}_{i,\tau}^{1/2}\boldsymbol{\varepsilon}, \qquad \boldsymbol{\varepsilon} \sim \rho N(\mathbf{0}, \boldsymbol{I}) + (1-\rho) N(\mathbf{0}, \omega \boldsymbol{I})$$

and replace $\boldsymbol{\theta}_{i,\tau}^{[g]}$ in $\boldsymbol{\Theta}^{[g]}$ by $\boldsymbol{\theta}_{i,\tau}^{c}$ with the probability $min(\zeta_i, 1)$, where

$$\zeta_i = \frac{L\left(\boldsymbol{y}_t | \boldsymbol{\theta}_{i,\tau}^{c}, \boldsymbol{\Theta}_{-i}^{[g]}\right) \pi\left(\boldsymbol{\theta}_{i,\tau}^{c}, \boldsymbol{\Theta}_{-i}^{[g]}\right) q\left(\boldsymbol{\theta}_{i,\tau}^{[g]}\right)}{L\left(\boldsymbol{y}_t | \boldsymbol{\Theta}^{[g]}\right) \pi\left(\boldsymbol{\Theta}^{[g]}\right) q\left(\boldsymbol{\theta}_{i,\tau}^{c}\right)}$$

$$\begin{aligned} q(\boldsymbol{\theta}_{i,\tau}) \quad \propto \quad & \rho exp\left\{-\frac{1}{2}\left(\boldsymbol{\theta}_{i,\tau} - \boldsymbol{\mu}_{i,\tau}\right)' \boldsymbol{\Sigma}_{i,\tau}^{-1}\left(\boldsymbol{\theta}_{i,\tau} - \boldsymbol{\mu}_{i,\tau}\right)\right\} \\ & + \frac{1-\rho}{\omega^{dim(\boldsymbol{\theta}_{i,\tau})/2}} exp\left\{-\frac{1}{2}\left(\boldsymbol{\theta}_{i,\tau} - \boldsymbol{\mu}_{i,\tau}\right)' \boldsymbol{\Sigma}_{i,\tau}^{-1}\left(\boldsymbol{\theta}_{i,\tau} - \boldsymbol{\mu}_{i,\tau}\right)\right\} \end{aligned}$$

Step 6. If $g < G_1 + G_2$, go to Step 4. Observe that the use of $\boldsymbol{\Sigma}_{i,\tau}$ in Step 5 accounts for the posterior correlation among the elements of $\boldsymbol{\theta}_{i,\tau}$, thereby improving the efficiency of the Markov chain. The parameter updates are sequentially repeated until the convergence of the Markov chain is achieved. The burn-in draws are discarded, and the steps are iterated a large number of times to generate draws from which the desired features (means, variances, quantiles, etc.) of the posterior distribution can be estimated consistently.

In this chapter, $G_1 = 50,000$ for the random walk MH sampler and $G_2 = 50,000$ with a thinning of 5 for the independent kernel MH sampler, resulting in posterior

samples comprising 10,000 draws. The convergence of the IK-MH Markov chains is assessed using the Geweke (1992) test. For each parameter, I also assess the accuracy of its posterior mean by computing the numerical standard error (NSE) according to the batch-means method (Ripley, 1987). In all simulated and real data examples of this chapter, it is observed that MCMC chains are well converged inside 50,000 iterations.

1.5. Simulation

This section carries on a simulation study. In MSQAR nonlinear settings where the number of parameters increases with the number of regimes, it is very convenient to choose parsimonious models that require a low number of parameters. For simplicity in the exposition, data are simulated from the true model with 2 regimes and autoregressive order 1 as

$$y_t = \begin{cases} 2.0 + 0.2y_{t-1} + 0.5\varepsilon_t, & s_t = 1 \\ -2.0 + 0.4y_{t-1} + \varepsilon_t, & s_t = 2 \end{cases}$$

The true parameter values are referenced based on empirical data estimations in next section. Three underlying distributions are considered for error terms, including a standard normal distribution ($N(0,1)$), a standardized student-t distribution with 3 degrees of freedom (t_3), and a mixed distribution between $N(0,1)$ when $s_t = 1$ and t_3 when $s_t = 2$.

The theoretical τth conditional quantile of y_t can be expressed in a MSQAR form as

$$Q_{y_t}(\tau|\mathbf{y}_{t-1};\theta_{s_t}) = \begin{cases} \theta_{10}(\tau) + \theta_{11}(\tau)y_{t-1}, & s_t = 1 \\ \theta_{20}(\tau) + \theta_{21}(\tau)y_{t-1}, & s_t = 2 \end{cases}$$

with the corresponding quantile parameters as $\theta_{10}(\tau) = 2.0 + 0.5Q_{\varepsilon_t}(\tau)$, $\theta_{11}(\tau) = 0.2$, $\theta_{20}(\tau) = -2.0 + Q_{\varepsilon_t}(\tau)$, and $\theta_{21}(\tau) = 0.4$. $Q_{\varepsilon_t}(\tau)$ is the theoretical τth quantile of a underlying distribution.

200 data replications are simulated for each underlying distribution. 50,00 observations are generated for each data replication but only the last 500 observations are kept for estimation in order to reduce initial effects. MSQAR models are examined in different sample sizes, $T = \{200, 500\}$ and quantile levels, $\tau = \{0.05, 0.25, 0.5, 0.75, 0.95\}$.

Table 1.1 reports the simulation results. This table includes the true quantile parameters (True), posterior means (PM), standard errors (Std), the root of mean squared errors (RMSE), and the mean absolute deviation (MAD). RMSE and MAD errors in Table 1.1 are small over different quantile levels and distributions. The small difference between the true and estimated parameters indicates the reasonable accuracy in model estimation. The small standard errors also show a favorable precision in model estimation. Furthermore, the accuracy and the precision of model estimations are improved with the increase in sample sizes considered due to the reduction in RMSEs, MADs and standard errors. As expected, the model estimation for the less extreme quantiles present smaller RMSE and MAD errors than extreme quantiles. The MSQAR model estimation also

shows reasonable performance for the data generated from mixtures of normal and student-t distributions.

Figure 1.1 plotting the posterior kernel densities of parameter estimates along with true parameters indicated by the vertical lines. Figure 1 shows that the posteriors well contain the true quantile parameters with a slightly better performance for $\tau = 0.5$. In many cases, the posteriors appear skewed but still with most of the density concentrated near the true parameter values. To save space, Figure 1 plots results for $\tau = 0.05, 0.5, 0.95$ and $N = 200$ from the normal distribution. Other results are similar and available upon request.

Following Guerin and Marcellino (2013), Table 1.2 reports the quadratic probability scores (QPS), absolute probability scores (APS) and log probability scores (LPS) for the quantile autoregressive models with Markov-switching features to check how well these models can estimate the true regimes. QPS, APS and LPS criteria evaluate the qualitative prediction abilities of MSQAR models, that is, to what extent the true quantile regimes are predicted. The predictability of regime 2 is computed for QPS, APS, and LPS as follows:

$$QPS = \frac{2}{T}\sum_{t=1}^{T}\left(\xi_{2,t|t} - I\left(\tilde{s}_t = 2\right)\right)^2 \tag{1.5.1}$$

$$APS = \frac{1}{T}\sum_{t=1}^{T}\left|\xi_{2,t|t} - I\left(\tilde{s}_t = 2\right)\right| \tag{1.5.2}$$

$$LPS = -\frac{1}{T}\sum_{t=1}^{T}\left(1 - I\left(\tilde{s}_t = 2\right)\right) log\left(1 - \xi_{2,t|t}\right) + I\left(\tilde{s}_t = 2\right) log\left(\xi_{2,t|t}\right) \tag{1.5.3}$$

where $\xi_{2,t|t}$ is obtained from (1.3.2) and $\tilde{s}_t$ is the simulated states. A score of 0 occurs when perfect predictions are made. Note that QPS is bounded between

0 and 2. The worst score is 2 for QPS and occurs if at each period probability predictions of 0 or 1 are made but turn out to be wrong each time. Note that correct predictions have individual scores between 0 and 0.5, whereas incorrect predictions have individual scores between 0.5 and 2.0 for QPS. Nonetheless, a few incorrect predictions can therefore dominate a majority of correct predictions in QPS scores. For this reason, a modified version of probability scores, absolute probability score (APS), is also considered. Like QPS, the best possible score for APS is 0. The worst score is 1. Here correct predictions have individual scores between 0 and 0.5, whereas incorrect forecasts carry scores between 0.5 and 1. The range for LPS is 0 to ∞. LPS penalizes large prediction errors more than QPS and APS. See also Christoffersen et al. (2007).

Table 1.2 shows that all QPS and APS scores are small and less than 0.5, which indicate the dramatic model predictability for switching regimes. The probability scores are slightly lower with the increase in sample sizes. The results also show that regime predictions for lower tails are better than for upper tails. In addition, the statistics of LPS are also smaller than 0.5 which imply that no prediction outliers are penalized.

1.6. Empirical Applications

Many studies have employed quantile autoregressive models to estimate risks of financial markets and assets. In macroeconomics literature, asymmetric dynamics have also been found for macroeconomic variables. This section applies the proposed MSQAR model to S&P 500 returns for market risk assessment and to real U.S. GDP growth rates (RGDP) and real exchange rates of U.S. dol-

lars (trade-weighted by major currencies, RTWER) for asymmetric persistence. Monthly and weekly S&P 500 index returns are taken from the Center for Research in Security Prices (CRSP). The quarterly RGDP and RTWER data are taken from Federal Reserve Bank of St. Louis as percent changes from year ago. The data summary in Table 1.3 show negative skewness for S&P 500 returns and real exchange rates. The skewness for real GDP growth is positive but small. S&P 500 returns appear to have excess kurtosis. Jarque-Bera tests reject the null of data normality for S&P 500 returns, whereas the tests do not reject the null for real exchange rates. The normality for real GDP is rejected at 10% level. Figure 1.2 plots the time series of the empirical data.

As discussed in section 5, for empirical illustration, this chapter estimates MSQAR of order 1 with 2 regimes to keep a parsimonious parameter space. This chapter defines that regime 2 represents more extreme outcomes than regime 1. For instance, at lower tails, quantiles of regime 2 should be more negative or farther into the left tail areas than those of regime 1, which is mostly associated with the periods of economic recessions and crises. In contrary, at upper tails, quantiles of regime 2 should be more positive or farther into the right tail areas than those of regime 1.

1.6.1. Stock Market Risk

Table 1.4 reports the estimation results for monthly and weekly S&P 500 returns. The entries are the posterior means of parameters with associated numerical standard errors in parentheses. In general, the values of the Geweke (1992) test statistic in square brackets indicate convergence of the Markov chain to stationarity. Table 1.5 shows that the numerical standard errors are small and the Markov

chain appears to be converged well as indicated by the generally insignificant values of the Geweke (1992) test statistic. Figure 1.3 also plots the estimated parameters over quantiles with the 5% and 95% intervals of posterior distributions. As seen, the quantile intercepts monotonically increase with the increase of quantile levels. The quantile autoregressive coefficients are close to zero around median, while they deviate from zero at lower and upper tails. The zero coefficients around median seem to suggest market efficiency for S&P 500 index. However, it appears to be less efficient at tails. Moreover, the autoregressive coefficients of regime 2 are larger in magnitude than those of regime 1. This result implies that markets are less efficient when extreme events occur or during economic recessions and crises. Interestingly, the positive autoregressive coefficients at lower tails suggest that risk expectation is positively impacted by past risks, while the negative autoregressive coefficients at upper tails indicate that during market good times investors is expecting higher risk in future. These results clearly show countercyclical behaviors in financial markets estimated by MSQAR models.

The results also show that the variation of transition probabilities across quantiles is much smaller in regime 1 than in regime 2. The transition probabilities of regime 1 are ranging from 0.85 to 0.985, compared to the range for regime 2 from 0.381 to 0.945. It seems that the more extreme the quantile level is, the lower the persistence of regime 2 (p_{22}) is.[3] Despite the large variation in regime persistence, the unconditional probabilities are very similar across quantiles, i.e., π_1 and π_2 are around 0.84 and 0.16 for each quantile level, respectively.[4] This result

[3] The regime persistence for regime 1 and 2can be computed as $1/(1-p_{11})$ and $1/(1-p_{22})$, respectively.

[4] Unconditional probabilities of π_1 and π_2 can be obtained as $(1-p_{22})/(2-p_{11}-p_{22})$ and $(1-p_{11})/(2-p_{11}-p_{22})$, respectively.

is reasonable in that economic conditions provide the common economic states to different parts of a data distribution. However, persistence is possibly varying across quantiles. This observation is further consolidated by Figure 1.4 plotting the smoothed transition probability $\xi_{s_t=2,t|T}$ for $\tau = 0.05, 0.5$. The shaded areas are NBER-dated business cycles. This figure shows that the fluctuation within each economic recession period is much larger in $\tau = 0.05$ than in $\tau = 0.5$. The responses of the 5% lower tail to the economic recessions are much stronger than those of median, by showing much higher probabilities of switching to regime 2.

Value-at-Risk is implicitly defined on quantiles as a one-to-one function of a quantile, over a given time interval, of a conditional return distribution (see Jorion (2000)). For assessing S&P 500 return risks, Figure 1.5 plots 5% Value-at-Risk (VaR) estimated from the dynamic quantile of $\tau = 0.05$ as $Q_{y_t}(\tau|\boldsymbol{y}_{t-1}, s_t; \hat{\Theta}) = \sum_{i \in S} Q_{y_t}\left(\tau|\boldsymbol{y}_{t-1}, s_t = i; \hat{\theta}_i\right) Pr\left(s_t = i|\boldsymbol{y}_t; \hat{\Theta}\right)$. The dark lines in Figure 1.5 are the estimated 5% VaR dynamics ($Q_{y_t}(\tau|\boldsymbol{y}_{t-1}, s_t; \hat{\Theta})$) and the top and bottom light lines are the estimated 5% VaR dynamics of regime 1 ($Q_{y_t}(\tau|\boldsymbol{y}_{t-1}, s_t = 1; \hat{\theta}_1)$) and regime 2 ($Q_{y_t}(\tau|\boldsymbol{y}_{t-1}, s_t = 2; \hat{\theta}_2)$), respectively. As seen, the dynamics in regime 2 is larger than in regime 1 due to the larger autoregressive coefficients. This result indicates that market efficiency is different across regimes.

The usefulness of the proposed MSQAR model can be immediately recognized from Figure 1.5. Value-at-Risk estimated from existing methods are undistinguished from different distributions associated with i.e., good times or economic recessions. Thus, the VaR values from those approaches are at best the results of averaging on different economic states. However, Figure 1.5 shows VaR values for both regime 1 implied by good economic periods and regime 2 associated with economic recessions. Risk states identified by the MSQAR model are particularly

beneficial for risk management, as a risk manager would care more about the most extreme scenarios or the worst possible outcomes. For example, to stress-testing a hypothetically stressed financial institution, one should use VaR values estimated from regime 2 ($Q_{y_t}(\tau|\boldsymbol{y}_{t-1}, s_t = 2; \hat{\theta}_2)$) as the worst scenario hypothetically occurring. This may be an appropriate approach to measure systemic risks for considering capital buffer requirement on financial institutions from the perspectives of central banks.

1.6.2. Asymmetric Persistence in Macroeconomic Dynamics

To study asymmetric dynamics of macroeconomic variables, this chapter estimates MSQAR models for percentiles. Table 1.5 reports the estimation results for real GDP growth rates and real trade-weighted exchange rates. The results of ADF, KPSS, Phillips-Perron and Zivot-Andrews tests (not reported here) reject the null hypothesis of unit roots for these macroeconomic variables.The entries are the posterior means of parameters with associated numerical standard errors in parentheses and the Geweke (1992) test statistic in square brackets. Table 1.5 shows that the numerical standard errors are small and the Markov chain appears to be converged well as indicated by the generally insignificant values of the Geweke (1992) test statistic. Figure 1.6 also plots the estimated parameters over quantile levels with the 5% and 95% intervals of posterior distributions.

Figure 1.6 shows that the quantile autoregressive coefficients of real GDP growth rates vary over different quantiles, displaying asymmetric dynamics. Upper tails appear to have higher dynamic persistence than lower tails. The quantile autoregressive coefficients of regime 2 has the range from 0.623 to 0.979, compared to the range of 0.779 to 0.874 for the coefficients of regime 1. This result indicates

that economic regimes demonstrate different asymmetric dynamics. By contrast, the evidence of asymmetric persistence in real trade-weighted exchange rates is weak due to much less variation across the quantile autoregressive coefficients. This result is consistent with Jarque-Bera test in Table 1.4 showing that the null of data normality is not rejected for real trade-weighted exchange rates, where it is rejected for real GDP growth rates at 10% confidence level.

In addition, Figure 1.6 also shows that transition probabilities slightly vary across quantiles in both regimes. It implies that regime persistence of macroeconomic variables is mainly driven by common economic conditions, and hence much less dependent on τ. This result is very different from the regime behaviors of financial markets in section 5, but consistent with the fact that macroeconomic variables are common economic states and factors in an economy.

Table 1.6 examines the regime predictions of real GDP growth rates. This table reports QPS, APS and LPS values for regime 2 by using NBER-dated business cycles as true regimes. The probability scores of QPS and APS are smaller than 0.5 across quantiles, which indicates a significant predictability of economic regimes based on real GDP growth. Interestingly, the predictability of regimes from lower tails is much stronger than from upper tails. In addition, LPS values are larger than one at upper tails than at lower tails. This result implies the issue of regime predictive outliers.

The different regime predictabilities across quantiles are also shown by Figure 1.7 plotting the smoothed transition probability $\xi_{s_t=2,t|T}$ for $\tau = 0.1, 0.5, 0.9$. The shaded areas are NBER-dated business cycles. As seen, the predicated regimes from $\tau = 0.05$ and $\tau = 0.5$ seem closely to trace NBER dated business cycles, whereas the predicated regimes from $\tau = 0.9$ appear to be lagged. In addition, the

responses of the 10% quantile to the economic recessions are much stronger than those of median, by showing much higher probabilities (close to 1) of switching to regime 2. These results suggest that lower tails of real GDP growth rates reveal more information of economic states than upper tails. This might be due to the economic behaviors of risk aversion and also reflect the effects of macroeconomic policies.

1.6.3. Quantile Monotonicity

It is important to evaluate the model by the monotonicity requirement on the conditional quantile functions. If the monotonicity is satisfied, there should be no crossings over quantiles. Severe crossing problems violate the theorem of equivariance to monotone transformation from (1.1.1) to (1.1.2). Figure 1.8 plots the estimated quantiles of each single regime. The straight lines are $Q_{y_t}(\tau|\boldsymbol{y}_{t-1}, s_t = 1; \hat{\theta}_1)$ and $Q_{y_t}(\tau|\boldsymbol{y}_{t-1}, s_t = 2; \hat{\theta}_2)$ for regime 1 and 2. The dots are the scatter plots with y_t as y-axis and y_{t-1} as x-axis. Despite that the MSQAR model is nonlinear, it takes a linear form within a single regime. Quantiles within a regime are not parallel due to its location-scale quantile autoregressive model, unlike location-shift quantile functions. Quantiles in regime 2 have no crossing issues, while crossing problems occur in regime 1 between $\tau = 0.4, 0.5, 0.6$. Nonetheless, the proportion of violations of the monotonicity in regime 1 is below 2% between $\tau = 0.4, 0.5, 0.6$, except around 10% for the quantiles of regime 1 of real GDP growth rates crossing between $\tau = 0.5$ and $\tau = 0.6$. Overall, Figure 1.8 does not show severe crossing issues for the data considered in this chapter.

1.7. Conclusion

This chapter proposes a new location-scale quantile autoregression, so-called Markov- switching quantile autoregression, to characterize behaviors of different parts of a nonstationary time series distribution. The new method directly inferences and estimates dynamic quantiles by allowing the location and scale parameters subject to regime shifts. Unobservable economic regimes are inferred by standard Hamilton filter approach in which quantile error terms follow a three parameter asymmetric Laplace distribution. Bayesian estimation is adopted to deal with some serious computational challenges in this nonlinear model which has differentiable likelihood functions. The empirical application to S&P 500 returns is able to show countercyclical risk accumulations in financial markets. It also illustrates that the dynamic quantiles associated with economic recessions should be an appropriate extreme scenario for stress-testing hypothetically stressed financial institutions from the perspective of central banks. Furthermore, the estimation results for macroeconomic variables show evidence of asymmetric dynamics for quarterly real GDP growth rates but not for quarterly real trade-weighted U.S. dollars. The transition probabilities are similar across quantiles within a single regime for macroeconomic variables, whereas they vary in financial markets. In addition, this chapter has found that the lower tails of real GDP growth provide more valuable information than the upper tails for predicting economic regimes.

Table 1.1: Simulation Results

	(1) Normal errors								
	True	$N = 200$				$N = 500$			
		PM	Std	RMSE	MAD	PM	Std	RMSE	MAD
$\tau = 0.05$									
p_{11}	0.900	0.883	0.034	0.042	0.031	0.895	0.019	0.022	0.017
p_{22}	0.900	0.890	0.034	0.039	0.030	0.901	0.022	0.025	0.019
$\theta_{10}(\tau)$	1.178	1.262	0.113	0.120	0.099	1.249	0.081	0.092	0.075
$\theta_{11}(\tau)$	0.200	0.181	0.038	0.210	0.168	0.184	0.032	0.178	0.145
$\theta_{20}(\tau)$	-3.645	-3.584	0.295	0.082	0.065	-3.606	0.193	0.054	0.043
$\theta_{21}(\tau)$	0.400	0.401	0.097	0.172	0.200	0.400	0.059	0.147	0.116
$\tau = 0.25$									
p_{11}	0.900	0.887	0.034	0.040	0.029	0.898	0.019	0.021	0.016
p_{22}	0.900	0.889	0.032	0.038	0.029	0.899	0.021	0.024	0.018
$\theta_{10}(\tau)$	1.663	1.671	0.088	0.053	0.041	1.665	0.060	0.036	0.029
$\theta_{11}(\tau)$	0.200	0.194	0.030	0.151	0.122	0.195	0.024	0.120	0.099
$\theta_{20}(\tau)$	-2.674	-2.677	0.221	0.083	0.067	-2.660	0.123	0.046	0.037
$\theta_{21}(\tau)$	0.400	0.394	0.067	0.168	0.133	0.402	0.039	0.098	0.081
$\tau = 0.5$									
p_{11}	0.900	0.889	0.034	0.040	0.029	0.900	0.019	0.021	0.015
p_{22}	0.900	0.888	0.032	0.038	0.029	0.897	0.021	0.023	0.018
$\theta_{10}(\tau)$	2.000	1.997	0.085	0.043	0.035	1.988	0.055	0.028	0.022
$\theta_{11}(\tau)$	0.200	0.196	0.028	0.142	0.112	0.198	0.021	0.106	0.085
$\theta_{20}(\tau)$	-2.000	-2.057	0.209	0.108	0.089	-2.041	0.119	0.063	0.052
$\theta_{21}(\tau)$	0.400	0.382	0.064	0.165	0.139	0.388	0.037	0.097	0.080
$\tau = 0.75$									
p_{11}	0.900	0.891	0.035	0.040	0.030	0.903	0.019	0.022	0.018
p_{22}	0.900	0.884	0.032	0.040	0.031	0.894	0.021	0.024	0.019
$\theta_{10}(\tau)$	2.337	2.323	0.087	0.038	0.031	2.311	0.057	0.027	0.021
$\theta_{11}(\tau)$	0.200	0.203	0.029	0.144	0.116	0.206	0.021	0.107	0.054
$\theta_{20}(\tau)$	-1.326	-1.499	0.231	0.118	0.172	-1.469	0.148	0.105	0.125
$\theta_{21}(\tau)$	0.400	0.358	0.075	0.177	0.173	0.365	0.047	0.140	0.115
$\tau = 0.95$									
p_{11}	0.900	0.884	0.040	0.048	0.035	0.896	0.024	0.027	0.021
p_{22}	0.900	0.863	0.036	0.057	0.046	0.874	0.025	0.040	0.033
$\theta_{10}(\tau)$	2.822	2.779	0.123	0.046	0.037	2.772	0.080	0.034	0.027
$\theta_{11}(\tau)$	0.200	0.217	0.042	0.225	0.188	0.215	0.033	0.181	0.148
$\theta_{20}(\tau)$	-0.355	-0.453	0.124	0.144	0.380	-0.451	0.103	0.123	0.322
$\theta_{21}(\tau)$	0.400	0.337	0.086	0.187	0.215	0.328	0.067	0.145	0.204

RMSE and MAD are the root of mean squared errors and the mean absolute deviation errors.

	(2) t_3 errors							
	True	$N = 200$				$N = 500$		
		PM	Std	RMSE	MAD	PM	Std	RMSE	MAD
$\tau = 0.05$								
p_{11}	0.900	0.879	0.031	0.042	0.033	0.885	0.021	0.029	0.022
p_{22}	0.900	0.880	0.041	0.050	0.037	0.892	0.024	0.028	0.022
$\theta_{10}(\tau)$	1.321	1.428	0.101	0.111	0.094	1.441	0.077	0.108	0.094
$\theta_{11}(\tau)$	0.200	0.191	0.035	0.180	0.148	0.193	0.028	0.142	0.117
$\theta_{20}(\tau)$	-3.359	-3.319	0.398	0.119	0.093	-3.319	0.246	0.074	0.059
$\theta_{21}(\tau)$	0.400	0.416	0.118	0.197	0.240	0.407	0.080	0.121	0.161
$\tau = 0.25$									
p_{11}	0.900	0.885	0.031	0.038	0.029	0.891	0.021	0.026	0.020
p_{22}	0.900	0.879	0.035	0.045	0.033	0.889	0.021	0.027	0.021
$\theta_{10}(\tau)$	1.779	1.770	0.062	0.035	0.028	1.777	0.043	0.024	0.018
$\theta_{11}(\tau)$	0.200	0.202	0.024	0.118	0.095	0.202	0.016	0.078	0.060
$\theta_{20}(\tau)$	-2.442	-2.465	0.156	0.064	0.049	-2.443	0.091	0.037	0.031
$\theta_{21}(\tau)$	0.400	0.393	0.048	0.120	0.093	0.400	0.029	0.073	0.058
$\tau = 0.5$									
p_{11}	0.900	0.886	0.031	0.037	0.028	0.892	0.021	0.025	0.020
p_{22}	0.900	0.876	0.034	0.046	0.034	0.886	0.021	0.028	0.022
$\theta_{10}(\tau)$	2.000	1.994	0.058	0.029	0.022	1.993	0.034	0.017	0.014
$\theta_{11}(\tau)$	0.200	0.199	0.023	0.115	0.091	0.201	0.013	0.067	0.051
$\theta_{20}(\tau)$	-2.000	-2.025	0.117	0.060	0.048	-2.025	0.074	0.039	0.028
$\theta_{21}(\tau)$	0.400	0.394	0.037	0.095	0.077	0.396	0.024	0.062	0.049
$\tau = 0.75$									
p_{11}	0.900	0.886	0.032	0.039	0.029	0.892	0.022	0.026	0.020
p_{22}	0.900	0.870	0.034	0.050	0.039	0.881	0.022	0.033	0.026
$\theta_{10}(\tau)$	2.221	2.218	0.075	0.034	00.025	2.212	0.043	0.020	0.016
$\theta_{11}(\tau)$	0.200	0.199	0.027	0.133	0.101	0.202	0.016	0.080	0.063
$\theta_{20}(\tau)$	-1.558	-1.643	0.137	0.103	0.083	-1.647	0.085	0.079	0.065
$\theta_{21}(\tau)$	0.400	0.383	0.042	0.113	0.092	0.385	0.027	0.076	0.062
$\tau = 0.95$									
p_{11}	0.900	0.881	0.039	0.048	0.036	0.888	0.026	0.032	0.024
p_{22}	0.900	0.853	0.037	0.067	0.055	0.861	0.025	0.051	0.044
$\theta_{10}(\tau)$	2.679	2.690	0.186	0.069	0.053	2.675	0.128	0.048	0.038
$\theta_{11}(\tau)$	0.200	0.197	0.066	0.153	0.265	0.202	0.047	0.137	0.186
$\theta_{20}(\tau)$	-0.641	-0.596	0.110	0.185	0.141	-0.596	0.073	0.134	0.106
$\theta_{21}(\tau)$	0.400	0.354	0.066	0.201	0.167	0.358	0.048	0.158	0.130

RMSE and MAD are the root of mean squared errors and the mean absolute deviation errors.

	(3) Mixed errors								
	True	$N = 200$				$N = 500$			
		PM	Std	RMSE	MAD	PM	Std	RMSE	MAD
$\tau = 0.05$									
p_{11}	0.900	0.885	0.032	0.042	0.030	0.896	0.019	0.025	0.017
p_{22}	0.900	0.893	0.032	0.039	0.028	0.905	0.020	0.023	0.018
$\theta_{10}(\tau)$	1.178	1.271	0.118	0.128	0.105	1.252	0.085	0.096	0.078
$\theta_{11}(\tau)$	0.200	0.182	0.038	0.212	0.167	0.186	0.029	0.160	0.126
$\theta_{20}(\tau)$	-3.359	-3.406	0.226	0.127	0.096	-3.348	0.155	0.076	0.059
$\theta_{21}(\tau)$	0.400	0.416	0.128	0.163	0.163	0.409	0.085	0.114	0.169
$\tau = 0.25$									
p_{11}	0.900	0.886	0.034	0.041	0.030	0.895	0.021	0.023	0.016
p_{22}	0.900	0.887	0.031	0.038	0.027	0.898	0.019	0.022	0.016
$\theta_{10}(\tau)$	1.663	1.671	0.082	0.049	0.040	1.668	0.059	0.035	0.029
$\theta_{11}(\tau)$	0.200	0.197	0.033	0.166	0.13	0.199	0.022	0.108	0.086
$\theta_{20}(\tau)$	-2.442	-2.462	0.153	0.063	0.049	-2.446	0.094	0.039	0.030
$\theta_{21}(\tau)$	0.400	0.400	0.046	0.115	0.090	0.400	0.029	0.072	0.057
$\tau = 0.5$									
p_{11}	0.900	0.886	0.035	0.039	0.029	0.894	0.021	0.022	0.015
p_{22}	0.900	0.883	0.030	0.037	0.026	0.893	0.019	0.021	0.016
$\theta_{10}(\tau)$	2.000	1.998	0.080	0.040	0.033	1.990	0.053	0.027	0.022
$\theta_{11}(\tau)$	0.200	0.200	0.031	0.153	0.125	0.202	0.022	0.111	0.088
$\theta_{20}(\tau)$	-2.000	-2.037	0.119	0.062	0.047	-2.033	0.076	0.041	0.033
$\theta_{21}(\tau)$	0.400	0.393	0.037	0.093	0.073	0.393	0.024	0.062	0.049
$\tau = 0.75$									
p_{11}	0.900	0.886	0.035	0.042	0.032	0.894	0.023	0.026	0.020
p_{22}	0.900	0.878	0.030	0.041	0.031	0.888	0.019	0.025	0.019
$\theta_{10}(\tau)$	2.337	2.329	0.084	0.036	0.028	2.315	0.061	0.028	0.022
$\theta_{11}(\tau)$	0.200	0.202	0.033	0.167	0.134	0.208	0.025	0.129	0.105
$\theta_{20}(\tau)$	-1.558	-1.648	0.130	0.104	0.081	-1.651	0.087	0.082	0.070
$\theta_{21}(\tau)$	0.400	0.382	0.040	0.110	0.088	0.382	0.026	0.079	0.064
$\tau = 0.95$									
p_{11}	0.900	0.877	0.041	0.052	0.039	0.884	0.027	0.035	0.027
p_{22}	0.900	0.856	0.034	0.062	0.050	0.866	0.022	0.045	0.039
$\theta_{10}(\tau)$	2.822	2.813	0.153	0.054	0.042	2.798	0.095	0.035	0.028
$\theta_{11}(\tau)$	0.200	0.204	0.058	0.218	0.167	0.208	0.035	0.181	0.145
$\theta_{20}(\tau)$	-0.641	-0.600	0.103	0.173	0.135	-0.601	0.087	0.128	0.098
$\theta_{21}(\tau)$	0.400	0.355	0.062	0.192	0.160	0.354	0.044	0.158	0.133

RMSE and MAD are the root of mean squared errors and the mean absolute deviation errors.

Table 1.2: Summary Statistics for the Predictability of Simulated Regimes.

	Normal			t_3			Mixed		
	QPS	APS	LPS	QPS	APS	LPS	QPS	APS	LPS
$N = 200$									
$\tau = 0.05$	0.019	0.025	0.046	0.049	0.042	0.251	0.043	0.042	0.206
$\tau = 0.25$	0.013	0.014	0.025	0.018	0.015	0.040	0.013	0.011	0.031
$\tau = 0.5$	0.006	0.011	0.014	0.014	0.011	0.038	0.010	0.010	0.029
$\tau = 0.75$	0.033	0.032	0.065	0.034	0.026	0.101	0.030	0.025	0.088
$\tau = 0.95$	0.055	0.052	0.138	0.056	0.050	0.305	0.048	0.044	0.286
$N = 500$									
$\tau = 0.05$	0.019	0.025	0.046	0.045	0.038	0.196	0.038	0.038	0.136
$\tau = 0.25$	0.012	0.013	0.024	0.016	0.012	0.038	0.012	0.011	0.030
$\tau = 0.5$	0.006	0.011	0.013	0.014	0.011	0.040	0.010	0.010	0.025
$\tau = 0.75$	0.031	0.031	0.063	0.033	0.025	0.102	0.030	0.024	0.086
$\tau = 0.95$	0.052	0.050	0.140	0.054	0.048	0.297	0.046	0.042	0.274

QPS, APS and LPS represent quadratic, absolute and log probability scores, respectively.

Table 1.3: Data Summary Statistics

	Monthly S&P 500	Weekly S&P 500	Real GDP	Real TWER
Sample Periods	1926:01-2013:02	01/09/1950-02/25/2013	1948Q1-2013Q2	1974Q1-2013Q2
# of obs.	1047	3294	263	159
Mean	0.461	0.136	3.263	-0.132
Median	0.907	0.282	3.200	0.201
Std. dev.	5.505	2.091	2.675	7.087
Skewness	-0.525	-0.567	0.004	-0.256
Kurtosis	10.75	8.744	0.692	-0.044
Jarque-Bera	$<$0.001	$<$0.001	0.058	0.412

Note: the $p-$values are reported for Jarque-Bera statistics test of data normality.

Table 1.4: Estimation Results for S&P 500 Index Returns

	Monthly S&P 500					Weekly S&P 500				
	$\tau = 0.05$	$\tau = 0.25$	$\tau = 0.5$	$\tau = 0.75$	$\tau = 0.95$	$\tau = 0.05$	$\tau = 0.25$	$\tau = 0.5$	$\tau = 0.75$	$\tau = 0.95$
p_{11}	0.902	0.925	0.985	0.987	0.842	0.850	0.868	0.985	0.947	0.787
	(0.018)	(0.055)	(0.010)	(0.009)	(0.016)	(0.017)	(0.034)	(0.011)	(0.102)	(0.019)
	[1.440]	[-0.447]	[-0.405]	[0.062]	[-0.274]	[0.631]	[0.065]	[1.273]	[0.685]	[-1.832]
p_{22}	0.572	0.625	0.914	0.917	0.381	0.389	0.641	0.945	0.880	0.415
	(0.045)	(0.063)	(0.043)	(0.039)	(0.066)	(0.046)	(0.037)	(0.027)	(0.146)	(0.044)
	[-0.685]	[-0.570]	[-2.405]	[-0.483]	[-0.244]	[-0.276]	[0.446]	[-0.803]	[0.568]	[-0.509]
$\theta_{10}(\tau)$	-4.121	-0.114	1.045	3.522	4.971	-1.619	-0.039	0.346	1.066	1.647
	(0.089)	(0.051)	(0.064)	(0.062)	(0.063)	(0.039)	(0.041)	(0.040)	(0.132)	(0.035)
	[-2.444]	[-0.464]	[0.648]	[-0.108]	[0.836]	[-0.322]	[-0.747]	[-1.678]	[0.723]	[-1.007]
$\theta_{11}(\tau)$	0.084	0.035	0.026	-0.087	-0.082	0.085	0.015	0.035	-0.025	-0.032
	(0.042)	(0.026)	(0.027)	(0.031)	(0.042)	(0.036)	(0.026)	(0.022)	(0.081)	(0.020)
	[1.835]	[-0.307]	[-0.518]	[0.098]	[0.064]	[-0.328]	[1.246]	[-0.788]	[-0.616]	[-0.126]
$\theta_{20}(\tau)$	-15.77	-3.978	-2.124	5.784	12.05	-5.865	-1.952	-0.094	2.422	4.851
	(0.147)	(0.047)	(0.318)	(0.174)	(0.200)	(0.094)	(0.040)	(0.147)	(0.147)	(0.080)
	[-0.478]	[0.101]	[-1.615]	[-0.212]	[0.169]	[-0.296]	[1.404]	[0.965]	[-0.835]	[-0.665]
$\theta_{21}(\tau)$	0.378	0.115	0.085	0.042	0.032	0.207	0.010	0.016	-0.196	-0.251
	(0.039)	(0.045)	(0.048)	(0.054)	(0.046)	(0.038)	(0.017)	(0.056)	(0.069)	(0.018)
	[-0.224]	[-0.342]	[0.889]	[0.395]	[0.621]	[0.317]	[0.348]	[0.124]	[0.617]	[0.538]

Values in parentheses are numerical standard errors and the Geweke (1992) test statistic in square brackets.
The test distribution for Geweke (1992) statistic is standard normal distribution.

Table 1.5: Estimation Results for Macroeconomic Variables

	Real GDP Growth Rates						Real Trade-weighted Exchange Rates					
	p_{11}	p_{22}	$\theta_{10}(\tau)$	$\theta_{11}(\tau)$	$\theta_{20}(\tau)$	$\theta_{21}(\tau)$	p_{11}	p_{22}	$\theta_{10}(\tau)$	$\theta_{11}(\tau)$	$\theta_{20}(\tau)$	$\theta_{21}(\tau)$
$\tau = 0.1$	0.927	0.647	-0.178	0.781	-2.310	0.623	0.789	0.500	-2.318	0.774	-7.949	0.761
	(0.023)	(0.046)	(0.087)	(0.032)	(0.162)	(0.099)	(0.056)	(0.093)	(0.291)	(0.055)	(0.264)	(0.079)
	[-1.085]	[1.193]	[0.699]	[-1.382]	[0.737]	[1.090]	[-1.110]	[0.409]	[1.434]	[-1.067]	[1.083]	[-1.205]
$\tau = 0.2$	0.926	0.667	0.196	0.779	-1.734	0.668	0.811	0.573	-1.149	0.777	-6.199	0.788
	(0.021)	(0.047)	(0.085)	(0.037)	(0.089)	(0.047)	(0.066)	(0.053)	(0.209)	(0.047)	(0.327)	(0.047)
	[0.375]	[-1.077]	[-2.107]	[0.067]	[-1.015]	[-0.535]	[1.368]	[-0.837]	[-1.439]	[1.294]	[-1.110]	[-0.585]
$\tau = 0.3$	0.930	0.638	0.310	0.801	-1.561	0.716	0.832	0.582	-0.620	0.784	-5.156	0.781
	(0.043)	(0.063)	(0.048)	(0.031)	(0.143)	(0.076)	(0.054)	(0.041)	(0.103)	(0.051)	(0.211)	(0.057)
	[0.559]	[0.085]	[0.227]	[-0.407]	[-1.193]	[-1.127]	[-0.392]	[-0.794]	[0.700]	[0.911]	[-0.083]	[0.489]
$\tau = 0.4$	0.943	0.687	0.630	0.784	-1.125	0.758	0.849	0.554	-0.323	0.799	-3.488	0.832
	(0.054)	(0.134)	(0.182)	(0.055)	(0.276)	(0.208)	(0.046)	(0.066)	(0.110)	(0.048)	(0.193)	(0.058)
	[-0.125]	[0.076]	[0.113]	[0.292]	[0.047]	[0.175]	[1.295]	[-1.579]	[-0.804]	[0.771]	[-0.858]	[-0.453]
$\tau = 0.5$	0.939	0.706	0.890	0.779	-0.629	0.883	0.926	0.489	-0.196	0.820	-2.307	0.770
	(0.015)	(0.039)	(0.068)	(0.034)	(0.114)	(0.054)	(0.033)	(0.056)	(0.092)	(0.025)	(0.222)	(0.052)
	[-0.925]	[0.229]	[0.043]	[-0.178]	[1.245]	[-0.825]	[0.457]	[-1.276]	[-1.231]	[0.357]	[-0.514]	[0.460]
$\tau = 0.6$	0.965	0.609	0.917	0.790	3.302	0.938	0.808	0.486	0.485	0.829	2.893	0.831
	(0.025)	(0.043)	(0.032)	(0.016)	(0.222)	(0.043)	(0.066)	(0.065)	(0.130)	(0.058)	(0.299)	(0.126)
	[-1.041]	[-0.643]	[0.874]	[-0.637]	[-0.242]	[-1.456]	[-1.150]	[0.160]	[-1.222]	[-0.654]	[-0.241]	[-0.100]
$\tau = 0.7$	0.955	0.609	0.974	0.809	3.514	0.958	0.792	0.465	0.995	0.810	4.660	0.878
	(0.018)	(0.039)	(0.036)	(0.020)	(0.057)	(0.013)	(0.094)	(0.089)	(0.177)	(0.053)	(0.419)	(0.135)
	[-1.217]	[0.589]	[0.051]	[-0.745]	[0.441]	[-0.345]	[0.666]	[-2.952]	[-0.132]	[-0.109]	[1.301]	[1.115]
$\tau = 0.8$	0.940	0.592	1.113	0.834	3.626	0.971	0.779	0.440	1.676	0.791	5.963	0.900
	(0.023)	(0.054)	(0.042)	(0.020)	(0.074)	(0.015)	(0.073)	(0.088)	(0.230)	(0.074)	(0.303)	(0.096)
	[1.427]	[-0.311]	[0.032]	[-0.465]	[0.421]	[-0.301]	[0.370]	[1.522]	[-0.580]	[1.120]	[0.095]	[-0.253]
$\tau = 0.9$	0.924	0.579	1.375	0.874	3.762	0.979	0.808	0.412	2.799	0.754	8.335	0.820
	(0.029)	(0.058)	(0.065)	(0.027)	(0.175)	(0.013)	(0.086)	(0.069)	(0.226)	(0.033)	(0.506)	(0.102)
	[0.437]	[-0.124]	[-0.288]	[1.182]	[0.423]	[1.072]	[-1.057]	[-1.162]	[-0.114]	[0.029]	[-1.054]	[0.991]

Table 1.6: Real GDP Growth Rates: Predictability of Regime 2

	QPS	APS	LPS
$\tau = 0.1$	0.135	0.111	0.281
$\tau = 0.2$	0.131	0.117	0.235
$\tau = 0.3$	0.134	0.119	0.223
$\tau = 0.4$	0.132	0.132	0.219
$\tau = 0.5$	0.151	0.179	0.259
$\tau = 0.6$	0.403	0.227	1.039
$\tau = 0.7$	0.439	0.249	1.075
$\tau = 0.8$	0.481	0.277	1.134
$\tau = 0.9$	0.516	0.298	1.370

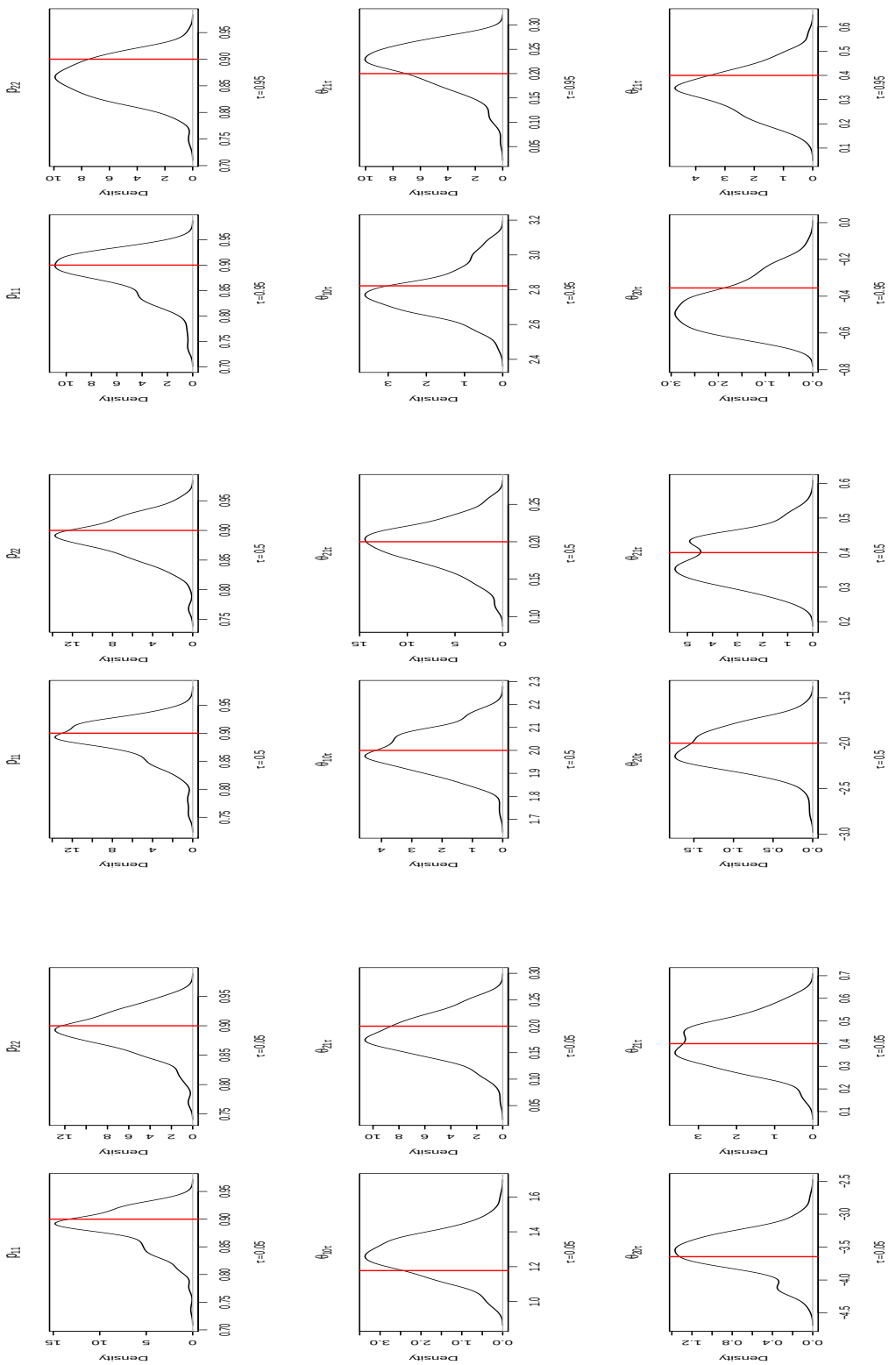


Figure 1.1: Posteriors of the Parameter Estimates

(1) $\tau = 0.05$ (2) $\tau = 0.5$ (3) $\tau = 0.95$

Note: the true parameters are indicated by the vertical lines for $\tau = 0.05, 0.5, 0.95$

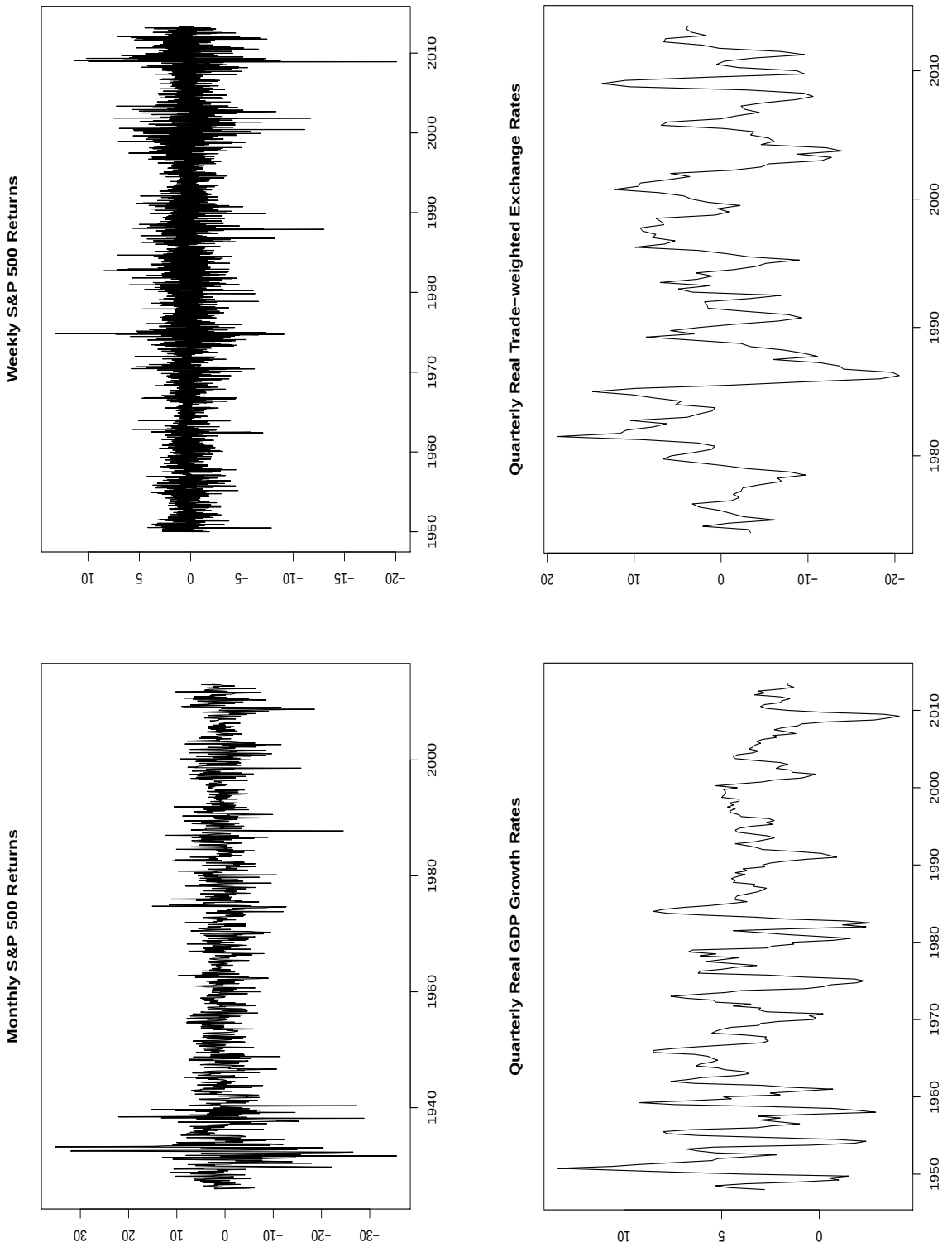


Figure 1.2: Time Series Data Plots

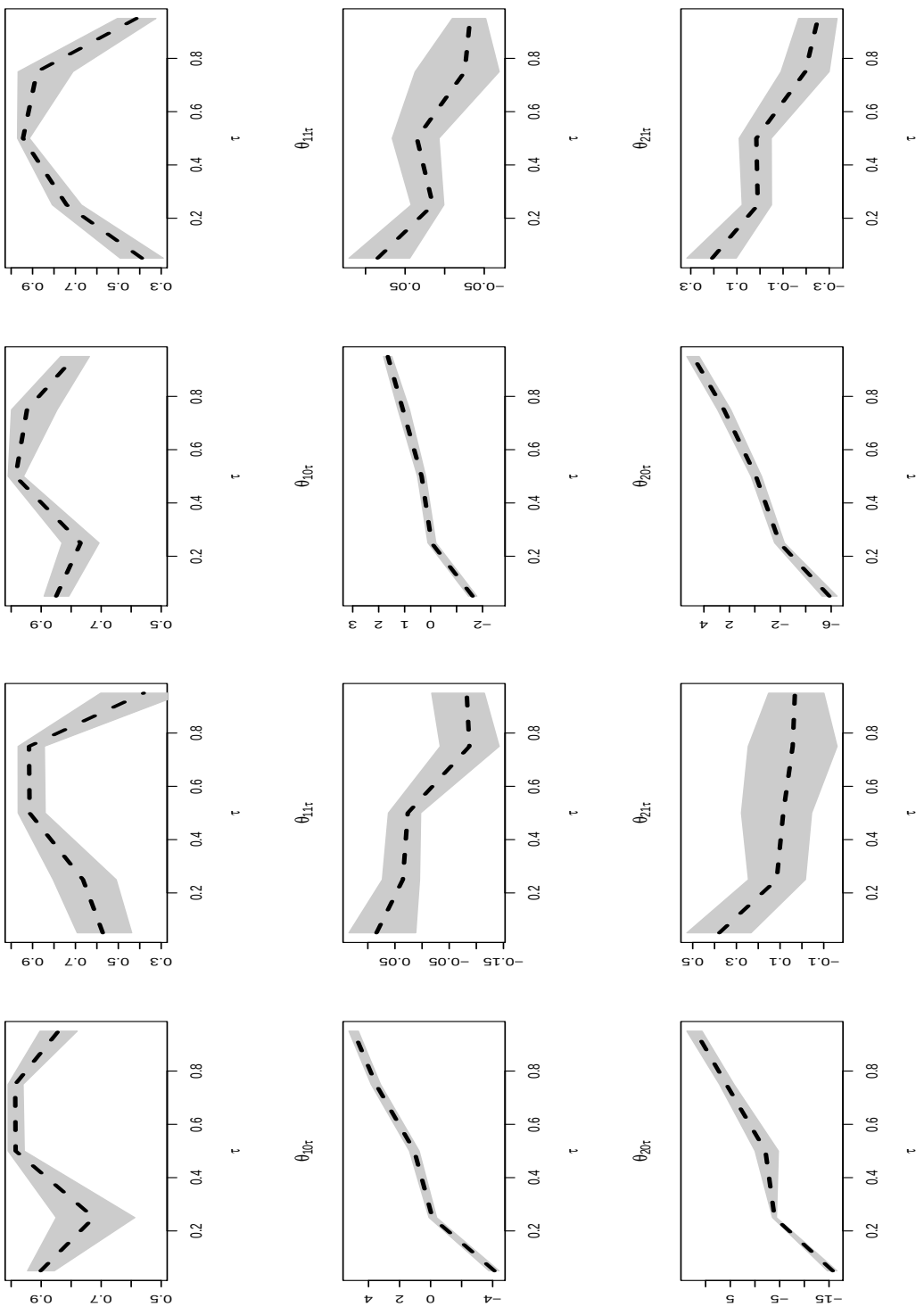
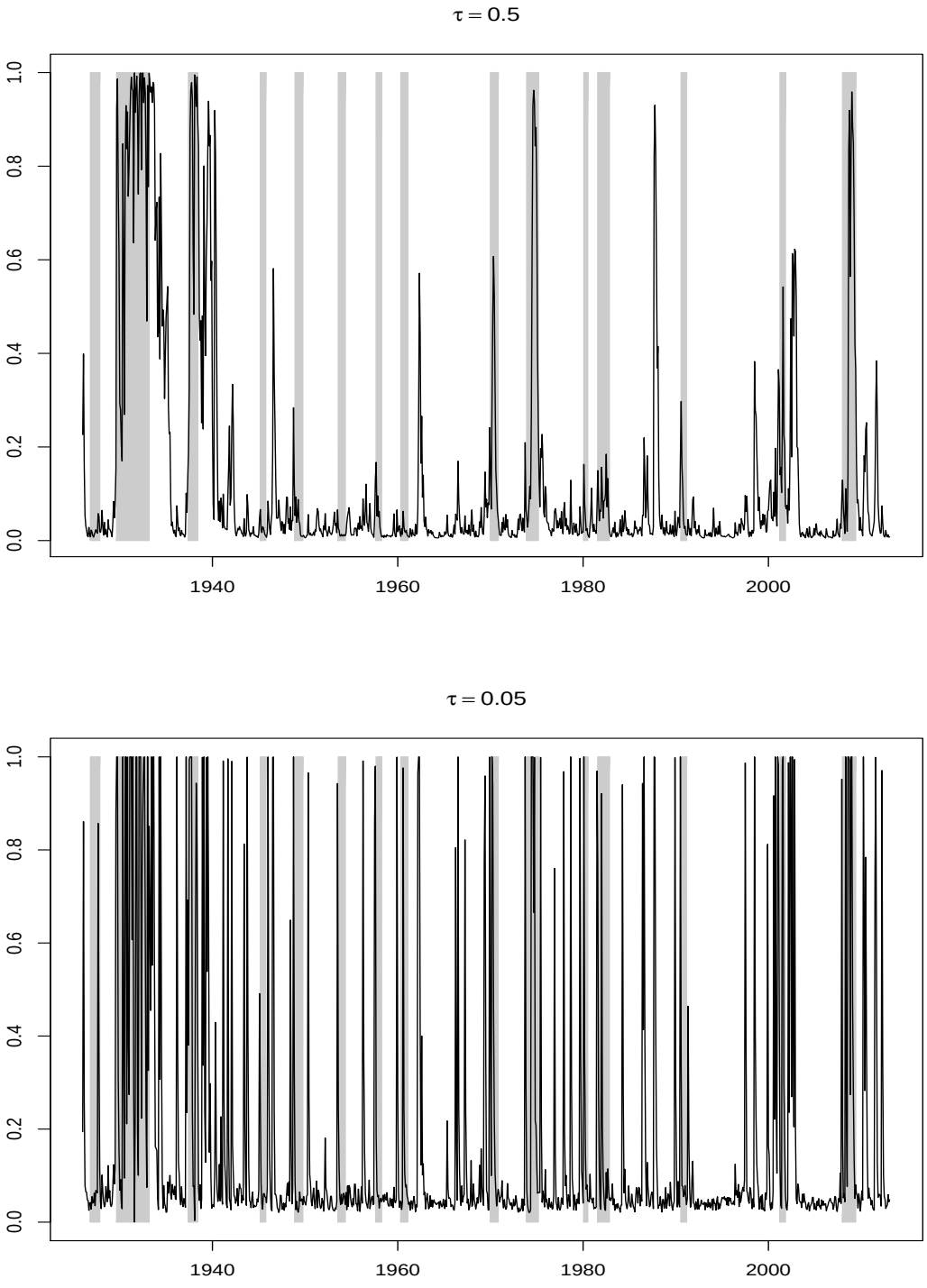


Figure 1.3: Quantile Parameter Estimation: S&P 500

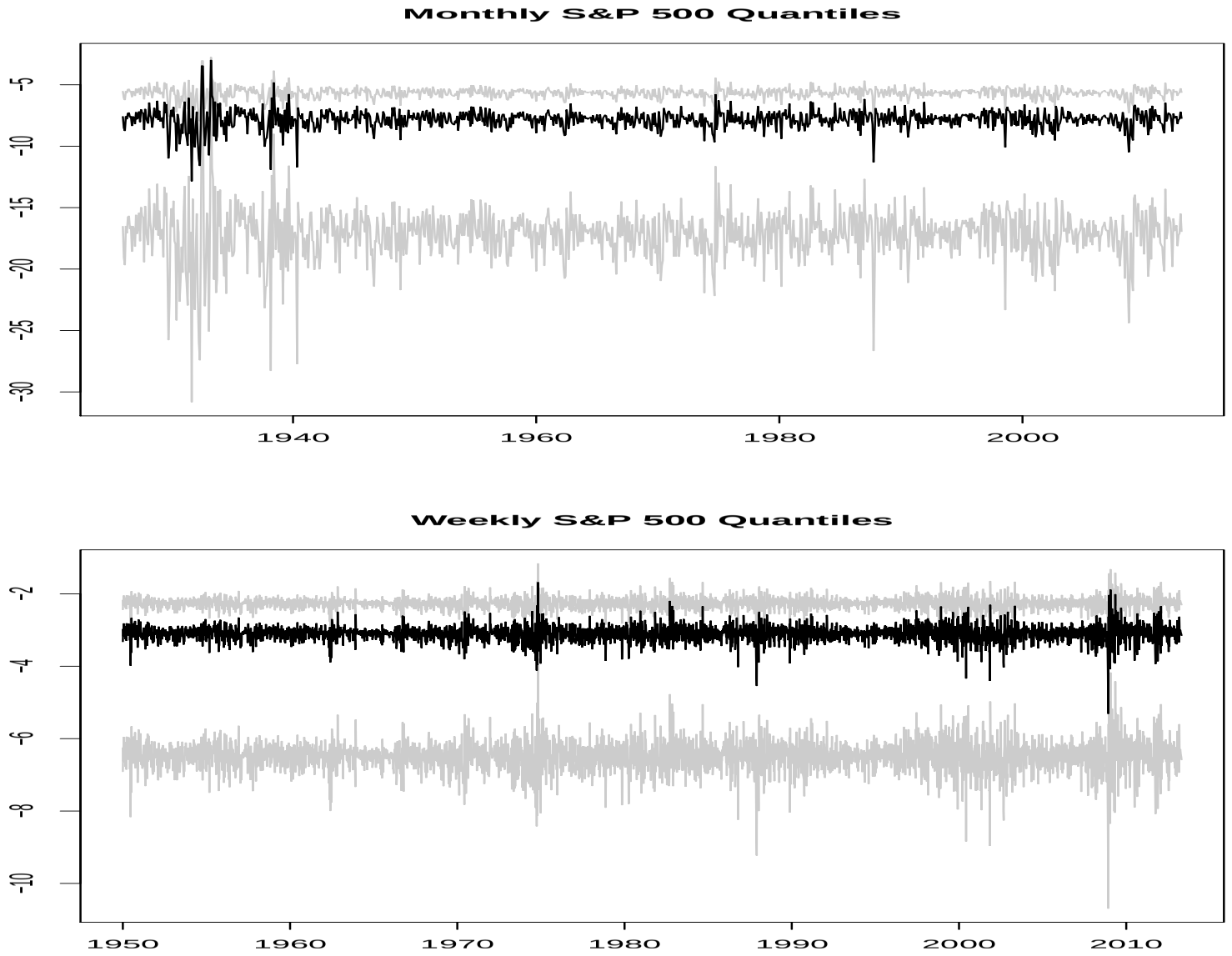
(1) Monthly S&P 500 Returns (2) Weekly S&P 500 Returns

Figure 1.4: Smoothed Transition Probability



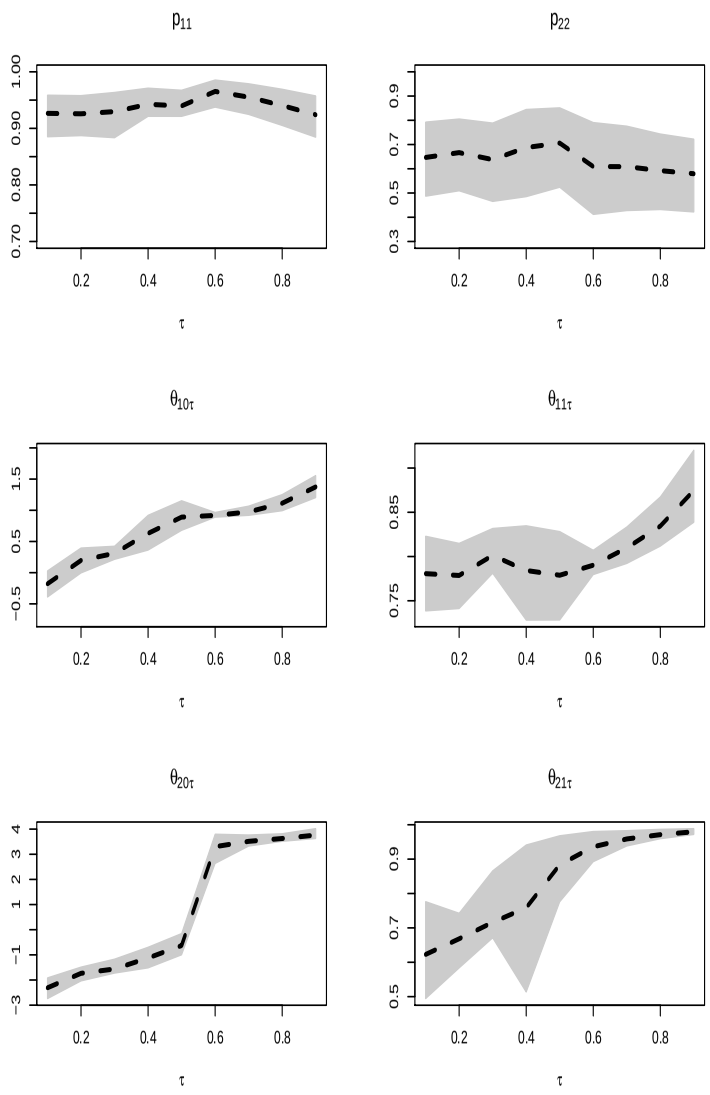
Note: the shaded areas are NBER-dated business cycles

Figure 1.5: The Estimated Quantiles.

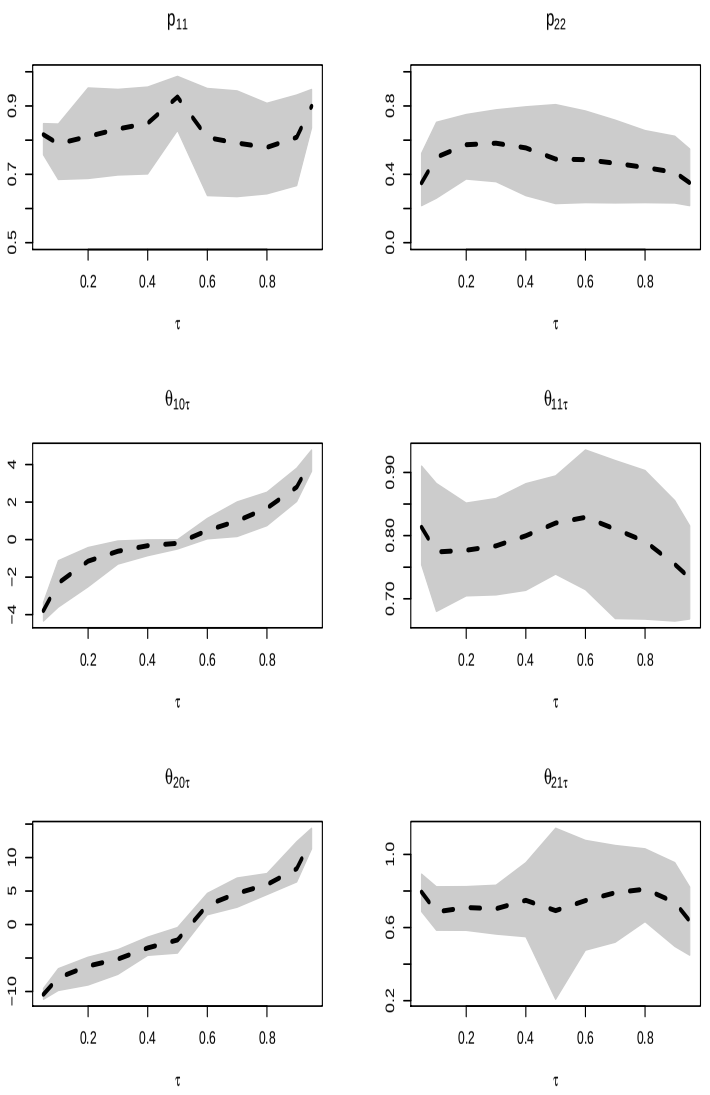


Note: the quantiles are estimated for $Q_{y_t}(\tau = 0.05|s_t = 1)$ (top light lines), $Q_{y_t}(\tau = 0.05|s_t)$ (dark lines), $Q_{y_t}(\tau = 0.05|s_t = 2)$ (bottom light lines)

Figure 1.6: Quantile Parameter Estimations for Macroeconomic Variables

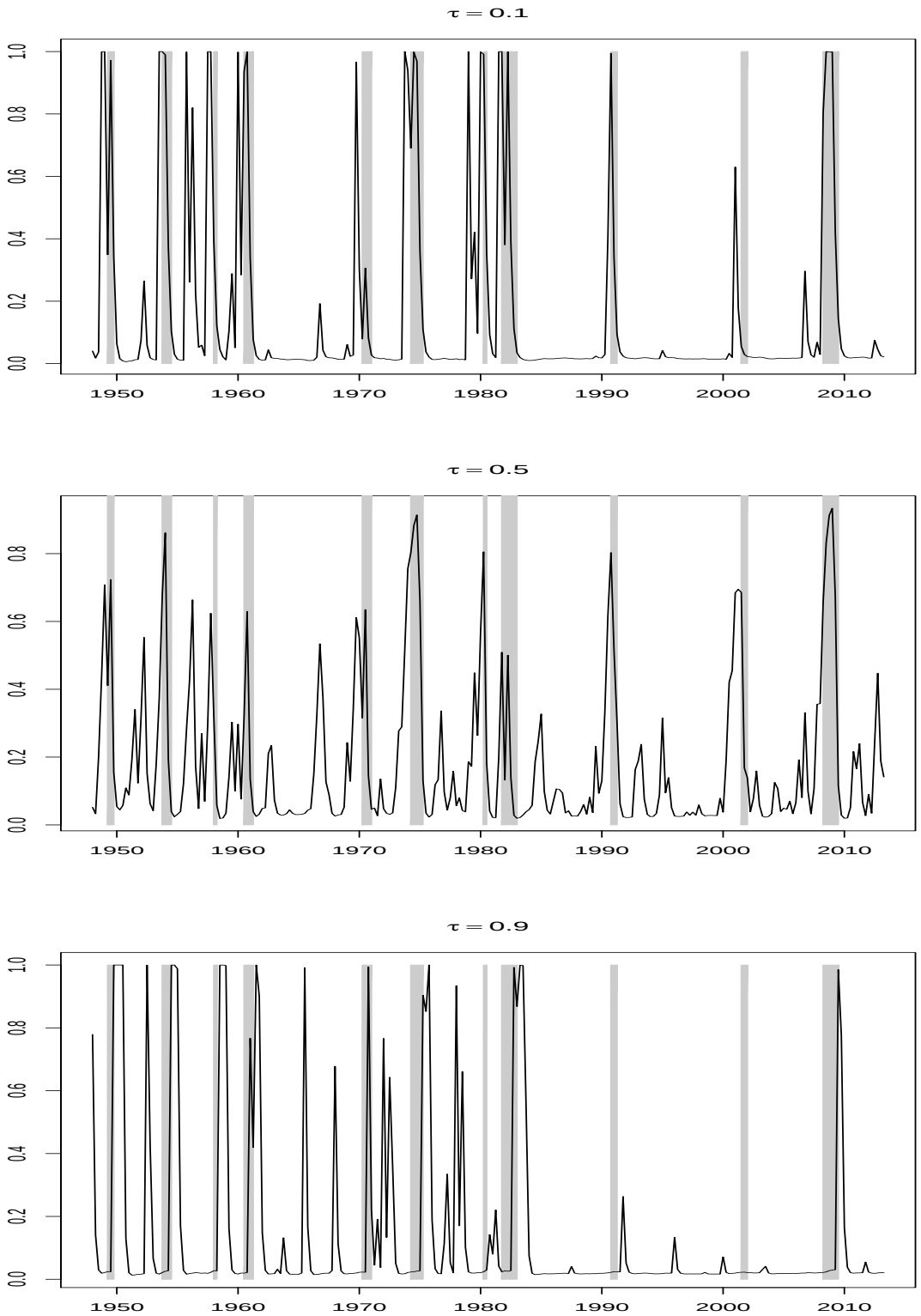


(1) Real GDP Growth Rates



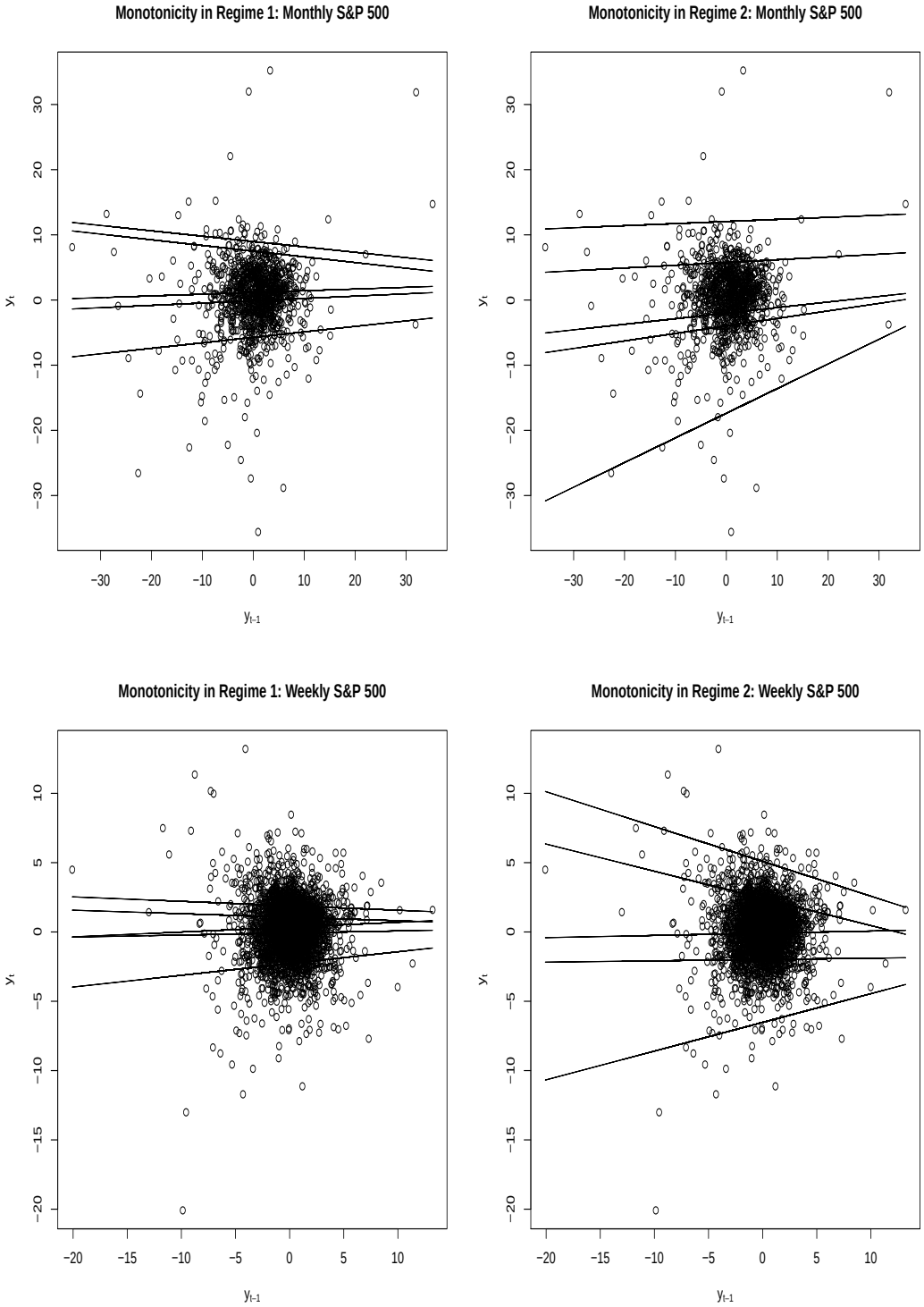
(2) Real Trade-weighted Exchange Rates

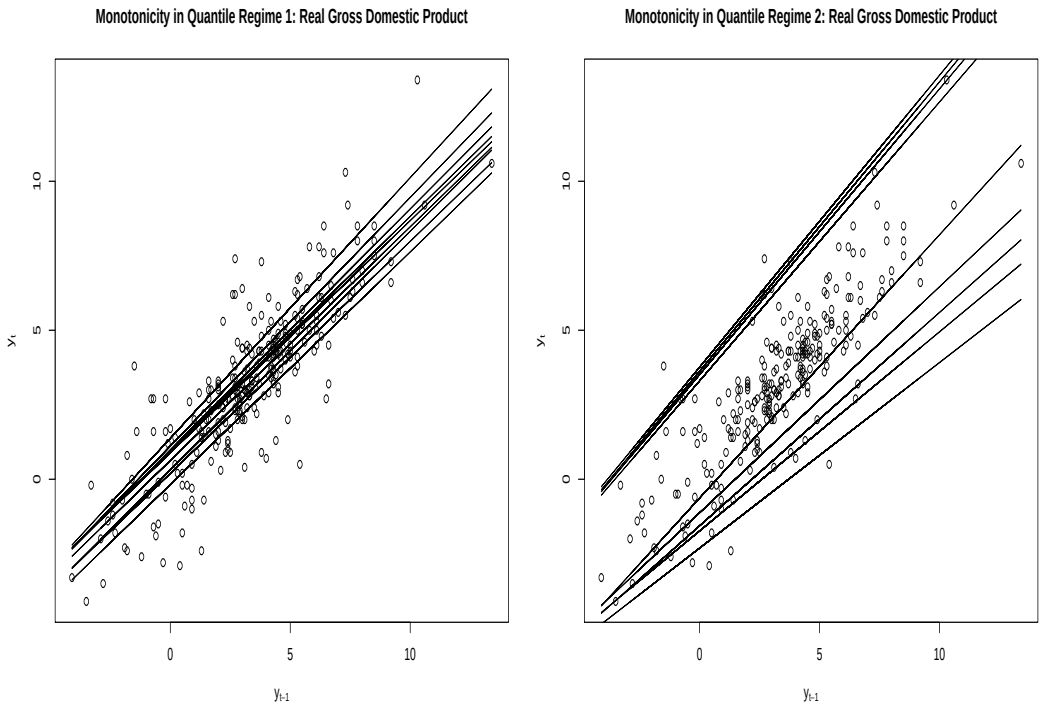
Figure 1.7: Smoothed Transition Probability for Real GDP.

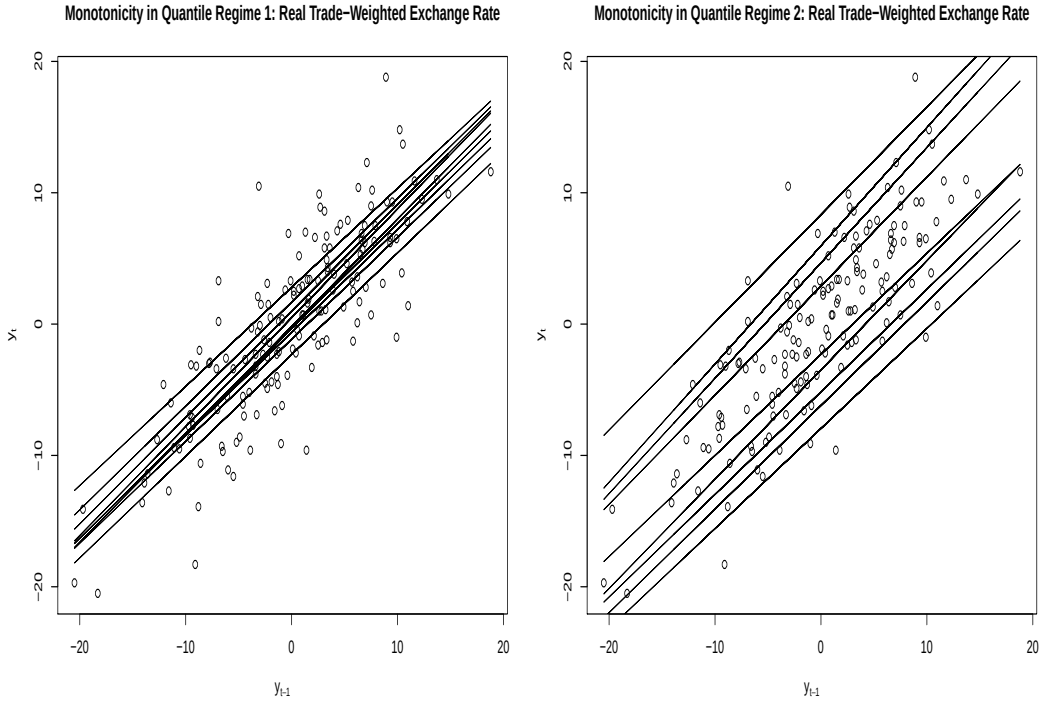


Note: The shaded areas are NBER-dated business cycles

Figure 1.8: Quantile Monotonicity for each regime.




Monotonicity in Quantile Regime 1: Real Gross Domestic Product
Monotonicity in Quantile Regime 2: Real Gross Domestic Product
y_t
y_{t-1}
0
5
10


Monotonicity in Quantile Regime 1: Real Trade−Weighted Exchange Rate
Monotonicity in Quantile Regime 2: Real Trade−Weighted Exchange Rate
y_t
y_{t-1}
−20
−10
0
10
20

Bibliography

[1] Ausin, M.C. and H.F. Lopes (2010) Time-varying joint distribution through copulas. Computational Statistics and Data Analysis 54: 2383-2399

[2] Bauwens, L., A. Preminger and J.V.K. Rombouts (2010) Theory and inference for a Markov-Switching GARCH model. The Econometrics Journal 13: 218-244

[3] Cai, Y. and J. Stander (2008) Quantile Self-Exciting Threshold Autoregressive Time Series Models. Journal of Time Series Analysis 29(1): 186-202

[4] Cai, Y. (2010) Forecasting for quantile self-exciting threshold autoregressive time series models. Biometrika 97(1): 199-208

[5] Chen, Q., R. Gerlach, and Z. Lu (2012) Bayesian Value-at-Risk and expected shortfall forecasting via the asymmetric Laplace distribution. Computational Statistics and Data Analysis 56(11): 3498-3516

[6] Chernozhukov, V. and H. Hong (2003) An MCMC approach to classical estimation. Journal of Econometrics 115: 293-346

[7] Cheung, Y. and U.G. Erlandsson (2005) Exchange rates and Markov-Switching dynamics. Journal of Business & Economic Statistics 23(3): 314-320

[8] Chib, S. and E. Greenberg (1995) Understanding the Metropolis-Hastings algorithm. American Statistician 49: 327-335

[9] Christoffersen, P.F., F.X. Diebold, R.S. Mariano, A.S. Tay and Y.K. Tse (2007) Direction-of-change forecasts based on conditional variance, skewness and kurtosis dynamics: international evidence. Journal of Financial Forecasting 1(2): 1-22

[10] Gerlach, R., C.W.S. Chen and N.Y.C. Chan (2011) Bayesian Time-Varying Quantile Forecasting for Value-at-Risk in Financial Markets. Journal of Business & Economic Statistics 29(4): 481-492

[11] Geweke, J. and H. Tanizaki (2001) Bayesian estimation of state-space models using the Metropolis-Hastings algorithm within Gibbs sampling. Computational Statistics & Data Analysis 37: 151-170

[12] Geweke, J. (1992) Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In: Bernardo, J., Berger, J., David, A., Smith, A. (Eds.), Bayesian Statistics. Vol. 4. Oxford University Press, Oxford, pp. 169{193.

[13] Gray, S.F. (1996) Modeling the conditional distribution of interest rates as a regime-switching process. Journal of Financial Economics 42: 27-62

[14] Guerin, P. and M. Marcellino (2013) Markov-Switching MIDAS Models. Journal of Business & Economic Statistics 31(1): 45-56

[15] Guidolin, M. (2012) Markov Switching Models in Empirical Finance. Advances in Econometrics, ISBN: 978-1-78052-526-6

[16] Hamilton, J.D. (1994) Time Series Analysis. Princeton University Press

[17] Hamilton, J., D. Waggoner and T. Zha (2007) Normalization in econometrics. Econometric Reviews 26: 221-252

[18] Hamilton, J.D. and R. Susmel (1994) Autoregressive conditional heteroskedasticity and changes in regime. Journal of Econometrics 64: 307-333

[19] Jorion, P. (2000) Value-at-Risk: The New Benchmark for Managing Financial Risk. McGraw-Hill. ISBN-13: 978-0071355025

[20] Kim, C.J. (1994) Dynamic linear models with Markov-switching. Journal of Econometrics 60: 1-22

[21] Kim, C.J., J. Piger, and R. Startz (2008) Estimation of Markov regime-switching regression models with endogenous switching. Journal of Econometrics 143: 263-273

[22] Koenker, R. (2005) Quantile Regression. Cambridge University Press.

[23] Koenker, R. and G. Bassett (1978) Regression quantile. Econometrica 46: 33-50

[24] Koenker, R. and Z. Xiao (2006) Quantile Autoregression. Journal of the American Statistical Association 101(475): 980-990

[25] Ripley, B. (1987) Stochastic Simulation. John Wiley, New York

[26] Sims, C.A. and T. Zha (2006) Were There Regime Switching in US Monetary Policy. American Economic Review 96(1): 54-81

[27] Tierney, L. (1994) Markov Chains for Exploring Posterior Distributions. Ann. Statist. 22: 1701-1728

[28] Vrontos, I., P. Dellaportas, D. Politis (2002) Full Bayesian inference for GARCH and EGARCH models. Journal of Business and Economic Statistics 18: 187-198

[29] Yu, K. and J. Zhang (2005) A Three-Parameter Asymmetric Laplace Distribution and Its Extension. Communications in Statistics- Theory and Methods 34: 1867-1879

[30] Yu, K. and R.A. Moyeed (2001) Bayesian quantile regression. Statistics & Probability Letters 54: 437-447

Chapter 2

Systemic Risk of Commercial Banks with Regime Switching in Tails

Abstract: This chapter extends the Conditional Value-at-Risk approach of Adrian and Brunnermeier (2011) by allowing systemic risk structures subject to economic regime shifts, which are governed by a discrete, latent Markov process. This proposed Markov-Switching Conditional Value-at-Risk is more suitable to Supervisory Stress Scenario required by Federal Reserve Bank in conducting Comprehensive Capital Analysis and Review, since it is capable of identifying the risk states in which the estimated risk levels are characterized. Applying MSCoVaR to stress-testing the U.S. largest commercial banks, this chapter finds that the CoVaR approach underestimates systemic risk contributions of individual banks by around 131 basis points of asset loss on average. In addition, this chapter constructs Banking Systemic Risk Index by value-weighted individual risk contributions for specifically monitoring the systemic risk of the banking system as a whole.

Keywords: Markov-Switching Conditional Value-at-Risk, Conditional Expected Shortfall, Bayesian Quantile Inference, Stress-testing, Value-at-Risk, Commercial Banks, Banking Systemic Risk Index

JEL: C22, C58, C51, C11, G23

2.1. Introduction

Recently, Adrian and Brunnermeier (2011) propose to measure systemic risk via the conditional value-at-risk (CoVaR) of the financial system, conditional on institutions being in a state of distress. In their work, an institution's contribution to systemic risk is defined as the difference between CoVaR conditional on the institution being in distress and CoVaR in the median ("normal") state of the institution. Hence, it characterizes the marginal contribution of a particular institution (in a non-causal sense) to the overall systemic risk.

The CoVaR approach is particularly appealing in that it outlines a method to construct a countercyclical, forward-looking systemic risk measure by predicting future systemic risk using current institutional characteristics. This is a time-varying systemic risk measure which does not rely on contemporaneous price movements and thus can be used to anticipate systemic risk. This method relates systemic risk measure to macroeconomic variables and the balance sheet deleveraging and characteristics of individual institutions. This is essentially a main regulatory concern of central banks.

A number of recent studies have extended and estimated the CoVaR measure of systemic risk for a variety of financial systems.[1] Adams et al. (2011) estimate a system of quantile regressions for four sets of major financial institutions (commercial banks, investment banks, hedge funds and insurance companies). Wong and Fong (2010) estimate CoVaR for the CDS of Asia-Pacific banks. Brunnermeier et al. (2012) use the CoVaR approach to examine the contribution of non-interest income

[1] See e.g., Brunnermeier et al. (2012), Lopez-Espinosa et al. (2012), Rodriguez-Moreno and Pena (2012), Arias et al. (2010), Girardi and Ergun (2012), Roengiptya and Rungcharoenkitkul (2011), and Van Oordt and Zhou (2010), etc. Bisias et al. (2012) and Brunnermeier and Oehmke (2012) provide comprehensive reviews on systemic risk analytics.

to systemic bank risk. They find that banks with a higher non-interest income to interest income ratio have a higher contribution to systemic risk and their contributions appear to be countercyclical to systemic risk build-up. Lopez-Espinosa et al. (2012) use the CoVaR approach to identify the main factors behind systemic risk in a set of large international banks. They find that short-term wholesale funding is a key determinant in triggering systemic risk episodes.

However, Bisias et al. (2012) raise the important econometric issue of nonstationarity which is particularly relevant to systemic risk measurement. Virtually the existing methods of systemic risk estimation and inference rely on the assumption of stationarity. In other words, the joint distribution of the relevant variables is stable over time. Nonetheless, the literature has recognized the stylized fact of structural breaks in macroeconomic and financial time series, so that the distribution structures of a time series might, driven by economic states, evolve over time. Hence, the very nature of systemic risk implies a certain degree of nonstationarity that may not always be consistent with the econometric framework in which risk measures are typically estimated.

Brunnermeier and Oehmke (2012) also concern that the CoVaR approach is vulnerable to regime changes based on historical data. The estimated CoVaR value is undistinguished from the distributions associated with i.e., a good economic state or an economic downturn. In this regard, without informing its associated risk states, the CoVaR measure is at best an averaging across different economic regimes and hence less advisable to or even misleading market participants and regulators in managing risks with ambiguous targets. Evidently, Adams et al. (2011) have shown the sensitivity of systemic risk to tranquil, normal and volatile economic states, while Lopez-Espinosa et al. (2012) have found that asymmetries

based on the sign of bank returns play an important role in capturing sensitivity of system-wide risk to individual bank returns. These concerns highlight the need for new systemic risk methods that are able to address nonstationarity in a more sophisticated way.

This chapter specifically considers the systemic risk measure subject to regime shifts. I extend the CoVaR measure of systemic risk to a nonlinear dynamic structure, namely Markov-switching CoVaR (MSCoVaR), in which an institution's contribution to systemic risk is measured by allowing the joint distribution evolving over time. Switching regimes is determined by the outcome of a latent, discrete Markov process, so that the conditional value-at-risk can be obtained with the filtered probabilities of risk states.

This chapter characterizes two risk states: a normal risk level implied by good economic periods and a high risk level associated with economic recessions, crises or extreme events. MSCoVaR is thus obtained for each risk state in stress-testing. Particularly, this chapter obtains MSCoVaR by estimating Markov-switching quantile autoregressive models (MSQAR) recently developed by Liu (2014). MSQAR is the location-scale quantile autoregression in which the location and scale parameters are permitted to evolve over time.

The MSCoVaR measure of systemic risk appears to have the advantage of naturally fitting to the Supervisory Stress Scenario required by Federal Reserve Bank in Comprehensive Capital Analysis and Review (CCAR).[2] In CCAR, a supervisory stress scenario is a hypothetical scenario to be used to assess the strength and re-

[2] See Comprehensive Capital Analysis and Review 2012 : Methodology and Results for Stress Scenario Projections. Board of Governors of the Federal Reserve System: March13 , 201 2; and Comprehensive Capital Analysis and Review 2013: Assessment Framework and Results. Board of Governors of the Federal Reserve System: March 2013

silience of BHC capital in a severely adverse economic environment. It represents an outcome in which the U.S. economy experiences a significant recession and economic activity in other major economies also contracts significantly, i.e., a deep recession in the United States, significant declines in asset prices and increases in risk premia, and a slowdown in global economic activity, etc. Therefore, the MSCoVaR result from a high risk episode is well-defined for the stress-testing in Fed's supervisory stress scenario since it estimates a separate set of parameters for high risk episodes.

In addition, the MSCoVaR measure of systemic risk provides various ways to test different stress scenarios. For instance, if an institution is systematically important, its hypothetically distressed scenario should also cause a distress in financial system. The systemic risk of a systemically important institution can thus be measured by the high risk episodes of both financial system and the institution. By contrast, as a non-systemically important institution, its hypothetical stress scenario, unless leading to a herding effect, does not cause a distress in financial system. Hence, its systemic risk can be measured by using the high risk episode of the institution and the normal risk period of financial system.

Importantly, the assumption in Liu (2014) that quantile error terms follow a three-parameter asymmetric Laplace distribution (ADL) for filtering transition probabilities of regimes can also be used to simulate the Markov-switching conditional expected shortfall (MSCoES) from the MSQAR results. This provides a natural solution to the theoretical issue that CoVaR is not a coherent risk measure due to its nonsubadditive nature.[3] Note that MSCoES takes distributional aspects within the tail into account. To this end, a banking systemic risk index

[3] See Adrian and Brunnermeier (2011) and Artzner et al. (1999).

by value-weighted individual contributions is constructed for monitoring systemic risk specific to the banking system as a whole.

This chapter estimates MSCoVaR and MSCoES as risk contributions of the largest U.S. commercial banks. The empirical results show strong evidence that financial institutions and the banking system as a whole experience regime shifts in their lower tails. The new systemic risk measure shows that the CoVaR approach of Adrian and Brunnermeier (2011) underestimates systemic risk contributions of individual banks by around 131 basis points of asset loss on average. The empirical results also show that the banking system is more sensitive to marginal changes of an individual bank during high risk episodes than during normal risk periods. In addition, Banking Systemic Risk Index presents the high relevance of tracing financial distress situations over the sample period.

The rest of this chapter is structured as follows. Section 2 defines the Markov-switching systemic risks measured by MSCoVaR and MSCoES. Markov-Switching Quantile Autoregression of Liu (2014) for estimating MSCoVaR and MSCoES are described in *Appendix A*. Section 3 applies MSCoVaR and MSCoES methods to stress-testing the U.S. largest commercial banks. In this section, the banking systemic risk index is also constructed. Section 4 concludes this chapter.

2.2. Systemic Risk Measure

This section briefs the CoVaR measure of systemic risk and then extends it to define the Markov-Switching CoVaR to identify risk states for a potential nonstationary time series. It is followed by a discussion of simulating the Markov-switching conditional expected shortfall as a coherent risk measure.

2.2.1. CoVaR

Recall that the value-at-risk of institution n given the probability of τ is

$$Pr\left(X_t^n \leq VaR_t^n\right) = \tau \tag{2.2.1}$$

where X_t^n denotes the asset return value of institution n at time t. The VaR of the financial system return (X_t^w) conditional on the event $\left\{\mathbb{C}(X_t^n) : X_t^n = VaR_{t,\tau}^n\right\}$, i.e., institution n's asset-return attains its VaR value, is denoted by $CoVaR_{t,\tau}^{w|n}$, such that

$$Pr\left(X_t^w \leq CoVaR_t^w | \mathbb{C}\left(X_t^n\right)\right) = \tau$$

Institution n's contribution to the system risk is thus defined as

$$\triangle CoVaR_{t,\tau}^{w|n} = CoVaR_{t,\tau}^{w|n} - CoVaR_{t,\tau}^{w|n,50\%} \tag{2.2.2}$$

where $CoVaR_{t,\tau}^{w|n,50\%}$ denotes the VaR of the financial system when the institution n's returns are at their median ("normal") state as $Pr(X_t^w \leq CoVaR_t^w | X_t^n = VaR_{t,50\%}^n) = \tau$. For simplicity, this chapter suppresses the superscript w. Hence, $\triangle CoVaR_{t,\tau}^n$ denotes the difference between the VaR of the financial system conditional on the distress of a particular financial institution n and the VaR of the financial system conditional on the median state of the institution n. Thus, $\triangle CoVaR_{t,\tau}^n$ quantifies how much an institution n adds to overall systemic risk. It captures the amount of additional risk that an institution inflicts upon financial system when the institution attains its VaR value.

Adrian and Brunnermeier (2011) apply quantile autoregressive models (QAR) of Koenker and Xiao (2006) to estimate CoVaR in two steps as follows

$$X_t^n = \alpha_\tau^n + \rho_\tau^n X_{t-1}^n + \gamma_\tau^{'n} Z_{t-1} + \varepsilon_{t,\tau}^n \tag{2.2.3}$$

$$X_t^w = \alpha_\tau^{w|n} + \rho_\tau^{w|n} X_{t-1}^w + \beta_\tau^{w|n} X_t^n + \gamma_\tau^{'w|n} Z_{t-1} + \varepsilon_{t,\tau}^{w|n} \tag{2.2.4}$$

where ε_t is quantile error terms and Z_t is the predictive variables. From (2.2.2), the risk contribution of an institution n to financial system is then given by

$$\triangle CoVaR_{t,\tau}^n = \beta_\tau^{w|n} \left(VaR_{t,\tau}^n - VaR_{t,50\%}^n \right) \tag{2.2.5}$$

where $VaR_{t,\tau}^n = \alpha_\tau^n + \rho_\tau^n X_{t-1}^n + \gamma_\tau^{'n} Z_{t-1}$ is estimated from (2.2.3) and $\beta_\tau^{w|n}$ is estimated from (2.2.4). In this framework, the existence of risk spillovers is captured through the parameter $\beta_\tau^{w|n}$: for non-zero values of this parameter, the left tail of the system distribution can be predicted by observing the predetermined distribution of an institution's returns.

2.2.2. Markov-Switching CoVaR

To address the vulnerability of CoVaR to regime shifts and the requirement of stress-testing of an institution in a hypothetically stressed scenario, i.e., a deep economic recession or asset price downturn, this section defines the Markov-switching CoVaR measure of systemic risk to identify distinct risk states as CoVaR subject to regime changes.

Let $\{s_t\}$ be an ergodic homogeneous Markov chain on a finite set $K = \{1, ..., k\}$ with a transition matrix P defined by the following transition probabilities

$$\{p_{ij} = Pr\left(s_t = j | s_{t-1} = i\right)\}$$

for $i, j \in K$ and assume s_t follow a first-order Markov chain. Transition probabilities satisfy $\sum_{j \in S} p_{ij} = 1$. In this chapter, I define two distinct risk regimes, $K = \{1, 2\}$. Regime 1 ($s_t = 1$) represents a normal risk level which is implied by a good economic state and regime 2 ($s_t = 2$) represents a high risk episode most likely associated with an economic recession or financial crisis. The risk structures are determined by data distributions of each regime over time. Note that economic states, s_t, are unobservable so that switching in s_t is inferred by transition probabilities which are estimated from data.

Suppose that X_t can be observed directly but can only make an inference about the value of s_t based on the observations as of date t. From (2.2.1), Markov-switching VaR (MSVaR) of an institution n can be defined as

$$Pr\left(X_t^n \leq VaR_t^n | s_t^n = j\right) = \tau$$

and denoted by $MSVaR_{s_t,\tau}^n$ which represents the value-at-risk level of an institution n in its risk regime j. Accordingly, the VaR of the financial systemic returns conditional on the event $\left\{\mathbb{C}(X_t^n) : X_t^n = MSVaR_{s_t,\tau}^n\right\}$, denoted by $MSCoVaR_{s_t,\tau}^n$, is given by

$$Pr\left(X_t^w \leq CoVaR_t^w | \mathbb{C}\left(X_t^n | s_t^n = i\right), s_t^w = j\right) = \tau$$

Note that the risk states of an institution and the financial system are not neces-

sarily coincided, i.e., $i \neq j$. For instance, a non-systemically important institution being distressed does not cause the same high risk episode to the whole financial system. However, a distressed financial system may indeed cause a high risk episode for a non-systemically important institution.

Apply the definition in (2.2.2) to obtain an institution n's contribution to systemic risk as

$$\triangle MSCoVaR^n_{s_t,\tau} = MSCoVaR^n_{s_t,\tau} - MSCoVaR^{n,50\%}_{s_t,\tau}$$

In this chapter, MSCoVaR is estimated by Markov-Switching quantile autoregressive models (MSQAR), specified as

$$X^n_t = \alpha^n_{s_t,\tau} + \rho^n_{s_t,\tau} X^n_{t-1} + \gamma'^n_{s_t,\tau} Z_{t-1} + \varepsilon^n_{t,\tau} \tag{2.2.6}$$

$$X^w_t = \alpha^{w|n}_{s_t,\tau} + \rho^{w|n}_{s_t,\tau} X^w_{t-1} + \beta^{w|n}_{s_t,\tau} X^n_t + \gamma'^{w|n}_{s_t,\tau} Z_{t-1} + \varepsilon^{w|n}_{t,\tau} \tag{2.2.7}$$

such that $MSVaR^n_{s_t,\tau} = \alpha^n_{s_t,\tau} + \rho^n_{s_t,\tau} X^n_{t-1} + \gamma'^n_{s_t,\tau} Z_{t-1}$ is estimated from (2.2.6) and then

$$MSCoVaR^n_{s_t,\tau} = \beta^{w|n}_{s_t,\tau} MSVaR^n_{s_t,\tau}$$

can also be computed based on the estimation results of (2.2.7). See *Appendix A* for details of the MSQAR model estimation for (2.2.6) and (2.2.7).

In this MSCoVaR measure, $\beta^{w|n}_{s_t,\tau}$ depends on risk states. The response of financial system to a negative shock to an institution's balance sheet during a high risk episode ($\beta^{s|i}_{s_t=2,\tau}$), hence, allows to be different from a normal risk period ($\beta^{s|i}_{s_t=1,\tau}$). The set of coefficients estimated from high risk episodes describes the distributional

structures of data in economic recessions, crises or extreme events. Therefore, it is suitable to be applied to stress-testing financial institutions in supervisory stress scenario required by Federal Reserve Bank. Note that if no risk regime-switching presents, MSCoVaR is equivalent to CoVaR. In this sense, the CoVaR approach is a special case of the MSCoVaR measure when there is no structural breaks. In this chapter, I assume the presence of distinct economic regimes based on the findings in literature. However, an appropriate approach of testing the number of regimes should be considered in future research.

The new framework of the MSCoVaR approach indeed provides flexibility to test different stress scenarios. For instance,

Scenario(1) An extreme scenario is that the financial system depends on the regimes of systemically important banks. This scenario describes the recent financial crisis as: the financial system is distressed once a systemically important bank is distressed, while the financial system is away from distress only if none of systemically important banks are distressed. Hence, systemic risk contribution might be measured by

$$\triangle MSCoVaR^n_{t,\tau} = \beta^{w|n}_{s_t=2,\tau}\left(MSVaR^n_{s_t=2,\tau} - MSVaR^{n,50\%}_{s_t=1,\tau}\right) \tag{2.2.8}$$

The first product in the right side of (2.2.8) is the value-at-risk of financial system conditional on hypothetically assuming both the financial system and the institution n in their high risk episodes. The second product in (2.2.8) is the value-at-risk of financial system conditional on normal states of that institution.

Scenario(2) In comparison, assuming current financial system in regime 1, a distressed institution n contributes systemic risk to financial system given by

$$\triangle MSCoVaR^{n}_{t,\tau} = \beta^{w|n}_{s_t=1,\tau}\left(MSVaR^{n}_{s_t=2,\tau} - MSVaR^{n,50\%}_{s_t=1,\tau}\right) \tag{2.2.9}$$

This scenario implies that the institution n is assumed to be not systemically important. Its high risk state does not cause a distressed financial system. However, it might still accumulate and contribute systemic risk to financial system, especially when herding effects occurring.

Scenario(3) Even during a normal time if a systemically important financial institution reaches its VaR level, it also likely shocks financial system into its high risk episode. Hence, the systemic risk contribution of a distressed institution i can also be measured by

$$\triangle MSCoVaR^{n}_{t,\tau} = \beta^{w|n}_{s_t=2,\tau}\left(MSVaR^{n}_{s_t=1,\tau} - MSVaR^{n,50\%}_{s_t=1,\tau}\right) \tag{2.2.10}$$

For instance, an institution reaching its risk level during a normal period might be caused by short-term maturity mismatch, while an institution reaching its high risk episode might be caused by the large number of defaults on loans like the recent subprime crisis. Despite that risk during a normal time is less severe than during a high risk period, the highly interconnected banking system, herding effects, and market panic might contagiously amplify these negative impacts on financial system and hence lead to crises by i.e., fire-sales and domino effects, etc.

2.2.3. Markov-Switching CoES

VaR is not a coherent risk measure due to its nonsubadditivity and does not take distributional aspects within the tail into account. This theoretical issue to some extent makes the CoVaR and MSCoVaR measures of systemic risk invalid. However, the asymmetric Laplace distribution assumption in the MSQAR framework of Liu (2014) provides a convenient solution by obtaining expected shortfall through Monte Carlo simulation based on model estimation results. Expected shortfall computed as conditional tail expectation is a coherent risk measure and considers risks beyond the point of a VaR value. See Artzner et al. (1999).

Using the simulation method in *Appendix A* and the estimation results from (2.2.6), Markov-switching expected shortfall ($MSES$) for an institution n can be obtained and denoted by $MSES^n_{s_t,\tau}$. Then, conditional on the event $\{\mathbb{C}(X^n_t) : X^n_t = MSES^n_{s_t,\tau}\}$, an institution n's contribution to systemic risk is given by

$$\triangle MSCoES^n_{s_t,\tau} = MSCoES^n_{s_t,\tau} - MSCoES^{n,50\%}_{s_t,\tau}$$

with $MSCoES^n_{s_t,\tau} = \beta^{w|n}_{s_t,\tau} MSES^n_{s_t,\tau}$. Expected shortfall can also be applied to the three scenarios of measuring systemic risk discussed previously:

$$(1) \quad \triangle MSCoES^n_{t,\tau} = \beta^{w|n}_{s_t=2,\tau} \left(MSES^n_{s_t=2,\tau} - MSES^{n,50\%}_{s_t=1,\tau} \right) \tag{2.2.11}$$

$$(2) \quad \triangle MSCoES^n_{t,\tau} = \beta^{w|n}_{s_t=1,\tau} \left(MSES^n_{s_t=2,\tau} - MSES^{n,50\%}_{s_t=1,\tau} \right) \tag{2.2.12}$$

$$(3) \quad \triangle MSCoES^n_{t,\tau} = \beta^{w|n}_{s_t=2,\tau} \left(MSES^n_{s_t=1,\tau} - MSES^{n,50\%}_{s_t=1,\tau} \right) \tag{2.2.13}$$

2.3. Stress-testing Commercial Banks

In this section, the MSCoVaR and MSCoES measures of systemic risk are estimated for stress-testing the largest U.S. commercial banks, using the CoVaR measure of systemic risk as the benchmark model. In addition, given the sub-additivity property, the Markov-switching expected shortfall is used to construct a banking systemic risk index (BSRI) via value-weighted individual systemic risk contributions for monitoring dynamic systemic risk of the financial system.

2.3.1. Data

Daily market equity data were taken from The Center for Research in Security Prices (CRSP). The universe of bank holding companies (BHCs) are the stocks corresponding to CRSP SIC codes 6000-6199 and 6712. Daily market data is used to form weekly returns on market-valued total assets of individual banks. Following Adrian and Brunnermeier (2011), a bank market-valued total asset is transformed from book-valued total assets into market-valued total assets by applying market-to-book equity ratios. Then, the financial system return is computed as a value-weighted average on the returns of the universe of banks.[4]

This chapter considers the largest U.S. commercial banks since they are the targets of current regulatory efforts and would likely be considered too-big-to-fail by central banks. Table 2.1 provides a bank list considered for stress-testing in this chapter. The ultimate criterion to configure the sample of potentially systemically important banks is the availability of comparable data over a long enough period of time. This sifting criterion rules out some large banks, i.e., HSBC, etc. The

[4] See details in Adrian and Brunnermeier (2011) and Lopez-Espinosa et al. (2012).

resulting sample is formed by a total of the 27 largest BHCs sampled from June 1993 to June 2012 with 1000 weekly observations. Note that this chapter estimates the systemic risk contributions of the 27 commercial banks to the financial system, while the financial system is constructed by the universe of financial institutions with the SIC code of 6000-6199 and 6712. Hence, the financial system defined in this chapter is equivalently referred to as the banking system hereafter.

The identification of risk regimes is enhanced by using a set of macro-financial predictive variables that are acknowledged to capture the expected return in financial markets. I choose a small set of predictive variables to avoid over-fitting the data. The predictive variables (Z_t) used in this chapter include: (1) the change in the credit spread ($\triangle cs$) between the 10-year Moody's seasoned Baa corporate bond and the 10-year U.S. Treasury bond; (2) The change in the U.S. Treasury bill secondary market 3-month rate ($\triangle 3mtb$); (3) the change in the slope of the yield curve ($\triangle ys$), measured by the yield spread between the U.S. Treasury benchmark 10-year bonds and the U.S. 3-month T-bill rate; (4) liquidity spread (ls), defined as the difference between the 3-month U.S. repo rate and the 3-month T-bill rate; (5) the S&P500 Composite Index return (sp); (6) the volatility Index (vix) of the Chicago Board Options Exchange (CBOE). All these variables are sampled weekly and obtained from CBOE, the Federal Reserve Board's H.15 Release and the Datastream database, respectively.

2.3.2. Empirical Results

Table 2.2 reports the results of the MSQAR model estimation with $\tau = 5\%$. Panel A presents the results estimated from (2.2.6) for individual banks (X_t^n) conditional on predictive variables (Z_{t-1}), and Panel B estimated from (2.2.7) for

the banking system (X_t^w) conditional on a individual bank n (X_t^n) and predictive variables (Z_{t-1}). This table displays the medians of the coefficient estimates, the numerical standard errors in square brackets, and the posterior credible intervals (PCI) in parentheses, across banks.[5]

In Table 2.2, the quantile intercepts ($\alpha_{s_t,\tau}$) of both individual banks and the banking system appear to have the non-overlapped PCIs between regimes ($s_t = 1$ and $s_t = 2$). This indicates an effective identification of risk regimes by the label switching restriction. The regime identification is further enhanced by predictive variables: S&P 500 returns, the changes in T-bill rates, market volatility for individual banks; and contemporaneous returns of individual banks, the lagged banking system returns, S&P 500 returns, the change in yield curve, and market volatility for the banking system. These predictive variables have non-overlapped PCIs between regimes.

In addition, the transition probabilities, which have the non-overlapped PCIs between regimes, present a much higher level of the regime persistence during regime 1 than during regime 2. The transition probability of regime 2 at 5% VaR is around 50%, which is much lower than that at median levels around 92%.[6] The explanation to this result is that, compared to a deviation from the median or a normal risk period, whenever an individual bank attains its 5% VaR (tail risk) in a high risk episode, the bank more likely takes measures to resolve the risky situation immediately, i.e., adjusting capital structure to reduce debt levels, implementing

[5] The detail estimation results of each bank are not reported here to save space, but available upon request. Numerical standard errors are obtained using batch mean method, e.g., Ripley (1987). The posterior credible intervals are computed using the highest posterior probability regions with the 95% credible level.

[6] The estimation results for $\tau = 50\%$ are not reported in this chapter to preserve space, but available upon request.

more conservative loan policies, etc. Those measures affect the persistence of a high risk episode. Similarly, when the banking system is stressed in a high risk episode, regulators also likely intervene markets by monetary and/or fiscal policies. The scale parameters (ς_{s_t}) imply much higher standard deviations (around 20.85 and 5.988 for individual banks and the banking system, respectively) during regime 2 than those during regime 1 (around 4.447 and 1.532 for individual banks and the banking system, respectively).[7] This result is highly consistent with the findings in literature that financial returns are more volatile during economic recessions and crises than economic good times.

Panel A of Table 2.2 shows that the predictive variables, including S&P 500, changes in T-bill rates, changes in yield curves and market volatility, which have their PCIs excluded zero values, show the predictability for the VaRs of individual banks. By contrast in Panel B of Table 2.2, the predictors, including contemporaneous returns of individual banks, the lagged banking system returns, S&P 500 returns, the change in yield curve, and market volatility, which have their PCIs excluded zero values, present the predictability for the VaRs of the banking system. For instance, among these predictors, a widening of yield spreads and spikes in market volatility are generally associated with a larger one-period ahead VaR value, and hence could be used to anticipate higher levels of downside risk. As a result, the conditioning variables considered in the analysis have shown the predictability for financial systemic risk.

Interestingly, S&P 500 returns appear to countercyclically contribute to the systemic risk of the banking system: the negative coefficient of S&P 500 return in regime 1 implies that a stock market boom accumulates tail risks in the banking

[7] The implied variance is computed based on the formula provided in *Appendix A*.

system, while its positive coefficient in regime 2 provides that the increase in stock market prices recovers tail risks of the banking system. Additionally, the contemporaneous returns of individual banks appear to have a strong positive relationship with system risk. This contemporaneous effect exacerbates the downside risk level of the banking system due to the drop of a bank return. The small numerical standard errors in Table 2.2 indicate reasonable model estimation accuracy.

Table 2.3 reports the VaR and MSVaR values of individual banks (X_t^n) estimated from (2.2.3) and (2.2.6) conditional on predictive variables (Z_{t-1}), respectively. MSES values are simulated based on the model estimation results using the approaches in *Appendix A*. Table 2.3 shows that given 5% probability, the worst possible outcome is MS estimated from $VaR_{t,\tau}$ (around 1,017 basis points) and STT (around 5,695 basis points) estimated from $MSVaR_{s_t=2,\tau}$. On average, $MSVaR_{s_t=2,5\%}$ values are about 800 basis points more riskier than $VaR_{t,5\%}$ results and about 1,200 basis points more riskier than $MSVaR_{s_t=1,5\%}$ results. From the coherent risk measure, $MSES_{s_t=2,5\%}$ and $MSES_{s_t=1,5\%}$ results have about 110 and 120 basis points on average more riskier than $MSVaR_{s_t=2,5\%}$ and $MSVaR_{s_t=1,5\%}$, respectively.

Note that these estimated values are used in (2.2.5), (2.2.8)-(2.2.10), and (2.2.11)-(2.2.13) to compute $\triangle CoVaR$, $\triangle MSCoVaR$, and $\triangle MSCoES$ for measuring systemic risk contributions of individual banks. Due to the clear difference between VaR and $MSVaR$ values in Table 2.3, this evidence shows that existing VaR methods, which provide the results averaging across different economic regimes, do not well reflect extreme risk scenarios for stress-testing purposes. In contrary, the risk levels obtained from high risk episodes (regime 2) are more

suitable for measuring hypothetically distressed contributions under supervisory stress scenarios.

The disparity between regimes can also be observed in Figure 2.1, which plots the $MSES^n_{s_t,\tau}$ values for the six largest U.S. commercial banks. The solid dark lines are the $MSES^n_{s_t,\tau}$ estimates from regime 1 and the dashed light lines from regime 2. Generally speaking, high risk episodes show higher dynamics and larger volatilities than normal risk periods. The difference between regimes exists over time, and the gap is dramatically enlarged during recessions and financial crises. For instance, the risk level during the recent financial crisis of 2008-2009 is well reflected in regime 2 by showing a deep drop into far left tails.

Table 2.4 reports the systemic risk sensitivities of the banking system as a whole conditional on individual banks. The banks in this table are ranked based on risk sensitivity coefficients ($\beta^{w|n}_{s_t=2,\tau}$). The risk sensitivity coefficients are the important elements for computing systemic risk contributions in (2.2.5), (2.2.8)-(2.2.10), and (2.2.11)-(2.2.13). For comparison, this table also includes the estimation results of the QAR model as the benchmark for 1-regime using (2.2.4).

The systemic risk sensitivity coefficients in Table 2.4 show that many individual banks tend to impact the banking system heavier during high risk episodes than during normal risk periods, whereas for some other banks the opposite is true. For instance, the marginal impact of BK on the banking system is 0.414 during high risk episodes much larger than 0.169 during normal risk periods. Different sensitivities across regimes show asymmetric effects of individual banks on the banking system. Generally, it is observed that the systemic risk sensitivity coefficients of $\beta^{w|n}_{s_t=2,\tau}$ are also largely different from the sensitivity results of 1-regime estimations ($\beta^{w|n}_{t,\tau}$). The higher value of a sensitivity coefficient represents the larger response of

the banking system to individual bank's shocks. The negative coefficients of BBT and CMA banks imply that during high risk episodes these banks do not worsen the systemic risk of the banking system, despite that their negative coefficients are small in magnitudes.

Table 2.5 reports the systemic risk contributions of individual banks to the banking system as a whole. $\triangle MSCoVaR_1$, $\triangle MSCoVaR_2$, $\triangle MSCoVaR_3$, $\triangle MS CoES_1$, $\triangle MSCoES_2$, and $\triangle MSCoES_3$, are computed in each scenario of (2.2.8)-(2.2.10) and (2.2.11)-(2.2.13), respectively. For comparison, the systemic risk contributions without switching regimes are also computed from (2.2.5) as benchmarks. The ingredients for computing systemic risk contributions are the systemic risk coefficients ($\beta^{w|n}_{s_t,\tau}$) and individual bank's $MSVaR^n_{s_t,\tau}$ and $MSES^n_{s_t,\tau}$ values. The banks in each scenario are ordered by their values of the systemic risk contributions.

On average across banks, the systemic risk contribution from scenario (1) is around 131 basis points higher than that measured by the CoVaR approach. In addition, scenario (2) generates the systemic risk contribution to the banking system about 72 basis points on average higher than that measured by the CoVaR approach. These results clearly show empirical evidence of the underestimated systemic risk contributions by the CoVaR approach.

The orders of individual bank's systemic risk contributions are very different between $\triangle MSCoVaR_1$ and $\triangle CoVaR$ measures as well. For instance, the systemic risk contribution of STT is the highest in the $\triangle MSCoVaR_1$ measure, while the highest systemic risk contribution in the $\triangle CoVaR$ measure is the AXP bank. The difference between their contributions is as large as about 827 basis points. A strong negative relationship between systemic risk contributions and bank sizes

has also been found through a OLS regression (not reported here). This result indicates that the bigger the bank asset sizes are, the larger the banks impact on the banking system. This result provides quantitative evidence for the recent debate of "too big to fail" of banks.

Apparently, the $\triangle MSCoVaR_1$ measure of systemic risk provides the most extreme stressed outcomes among the 3 scenarios considered. Even in the case that a bank is not systemically important but distressed during high risk episodes (scenario (2)), the average systemic risk contribution is around 169 basis points which cannot be neglected. The orders of systemic risk contributions also vary across the 3 scenarios.

In addition, Table 2.5 reports the simulated results of $MSCoES_{s_t,\tau}$. As seen, the systemic risks are similar between $\triangle MSCoES_1$ and $\triangle MSCoVaR_1$, and between $\triangle MSCoES_2$ and $\triangle MSCoVaR_2$, while the results from scenario (3) are very different. However, this chapter suggests to adopt the systemic risk measurement results of $\triangle MSCoES_{s_t,\tau}$ since $\triangle MSCoVaR$ is not a coherent risk measure.

Figure 2.2 plots the dynamics of systemic risk contributions measured by $\triangle MSCo\ VaR_1$ and $\triangle CoVaR$ approaches along with the correlation.[8] The results show that the $\triangle MSCoVaR_1$ measure of systemic risk contributions are more dynamic than the $\triangle CoVaR$ measure. Some banks, i.e., JPMorgan Chase, Citi Financial Group and Morgan Stanley, etc., appear to have high correlations (about 83%-95%) between $\triangle MSCoVaR_1$ and $\triangle CoVaR$, while other banks, i.e., Bank of America, Well Fargo, etc., have correlations below 50%. Furthermore, $\triangle MSCoVaR_1$ and $\triangle CoVaR$ are negatively correlated for the bank of USB. These

[8] Instead of $\triangle MSCoES_1$ and $\triangle CoES$, this chapter makes the comparison between $\triangle MSCoVaR_1$ and $\triangle CoVaR$, because Adrian and Brunnermier (2011) approach cannot be used to compute expected shortfall.

results show that systemic risk contributions measured by $\triangle MSCoVaR_1$ and $\triangle CoVaR$ are not only different in magnitudes, but also in the dynamics over sample periods.

Table 2.6 reports the correlation matrix for banks' systemic risk contributions measured by $\triangle MSCoVaR_1$. The correlation matrix shows that banks are highly interconnected. For instance, Bank of America is positively correlated with other banks ranging from 75%-95%. Bank of America has the highest correlation of 96% with JPMorgan Chase bank. Among all the banks sampled, BBT, CMA and SCHW are the only banks negatively correlated with other banks. Table 2.6 shows that the potential contagious channels of a crisis are hidden behind the high interconnections between banks.

2.3.3. Banking Systemic Risk Index

Figure 2.3 plots the quarterly systemic risk index of the banking sector (BSRI). The solid line is the quarterly Financial Stress Index constructed by Federal Reserve Bank of St. Louis (STLFSI). The dashed line is quarterly BSRI constructed by the value-weighted $\triangle MSCoES_1$ on individual banks as

$$BSRI_t = -\sum_{n=1}^{N} w_t^n \triangle MSCoES_{t,\tau}^n$$

where weekly $\triangle MSCoVaR_{t,\tau}^n$ is aggregated to quarterly frequency and w_t^n is the bank n's weight based on its market capitalization at time t. The shaded areas are NBER-dated business cycle phases. Figure 2.3 shows that the constructed systemic risk index for the banking sector is capable of reproducing the recent economic recession. The quarterly BSRI reaches the highest risk during the recent

financial crisis of 2007-2009. The BSRI also shows a milder risk increase than STLFSI for the economic recession during the IT Bubble Bust period since it is not a recession highly related to the banking sector. Figure 2.3 presents a positive 61.5% comovement between the BSRI and the STLFSI. Furthermore, a simple linear regression shows that the BSRI is able to significantly explain the dynamics of the Financial Stress Index by 37.83% (R^2). Hence, the constructed BSRI index is supplementary to monitoring financial market risks by very specific to the risk nature of the banking sector.

2.4. Conclusion

This chapter has defined a Markov-switching conditional Value-at-Risk (MSCoVaR) approach to measure systemic risk of commercial banks. Applying the Markov-Switching Quantile Autoregression framework of Liu (2014), systemic risks are estimated subject to regime shifts within tails. The new method presents the advantage and flexibility in supervisory stress scenarios required by Federal Reserve Bank. I estimated systemic risk contributions of the U.S. largest commercial banks and found around 131 basis points of the underestimated asset loss by the existing CoVaR measure of systemic risk. The banking system is more sensitive to marginal changes of an individual bank during high risk episodes than during normal risk periods. In addition, systemic risk contributions of individual banks are highly interconnected. Furthermore, Banking Systemic Risk Index, constructed in this chapter by value-weighted individual systemic risk contributions, presents not only a high relevance to trace financial distress situations, but also very specific to the risk nature of the banking industry.

Table 2.1: The Sample
List of the U.S. Largest Commercial Banks as of 06/30/2012 Ranked in Total Assets

Institution Name	Ticker	Total Assets in thousand dollars as of 06/30/2012
JPMORGAN CHASE & CO.	JPM	$2,290,146,000
BANK OF AMERICA CORPORATION	BAC	$2,162,083,396
CITIGROUP INC.	C	$1,916,451,000
WELLS FARGO & COMPANY	WFC	$1,336,204,000
MORGAN STANLEY	MS	$748,517,000
U.S. BANCORP	USB	$353,136,000
BANK OF NEW YORK MELLON CORPORATION, THE	BK	$330,490,000
PNC FINANCIAL SERVICES GROUP, INC., THE	PNC	$299,712,018
STATE STREET CORPORATION	STT	$200,368,976
BB&T CORPORATION	BBT	$178,560,000
SUNTRUST BANKS, INC.	STI	$178,307,292
AMERICAN EXPRESS COMPANY	AXP	$146,890,000
REGIONS FINANCIAL CORPORATION	RF	$122,344,664
FIFTH THIRD BANCORP	FITB	$117,542,579
CHARLES SCHWAB CORPORATION	SCHW	$111,816,000
NORTHERN TRUST CORPORATION	NTRS	$94,455,895
KEYCORP	KEY	$86,741,424
M&T BANK CORPORATION	MTB	$80,807,578
BBVA USA BANCSHARES, INC.	BBVA	$66,013,042
COMERICA INCORPORATED	CMA	$62,756,597
HUNTINGTON BANCSHARES INCORPORATED	HBAN	$56,622,959
ZIONS BANCORPORATION	ZION	$53,418,819
POPULAR, INC.	BPOP	$36,612,000
PEOPLE'S UNITED FINANCIAL, INC.	PBCT	$28,134,752
SYNOVUS FINANCIAL CORP.	SNV	$26,294,110
BOK FINANCIAL CORPORATION	BOKF	$25,561,731
FIRST HORIZON NATIONAL CORPORATION	FHN	$25,493,925

Note: The composition of the banks is based on consolidated assets, lagged by one quarter.

Table 2.2: MSQAR model estimation results

Panel A	p_{ii}^n	$\alpha_{s_t,\tau}$	X_{t-1}^n	sp	$\triangle 3mtb$	$\triangle ys$	$\triangle cs$	ls	vix	$\varsigma_{s_t}^n$
$s_{t,\tau}=1$	0.862	-2.418	-0.105	0.441	-1.740	-0.168	0.008	-0.893	-0.040	0.222
	[0.028]	[0.143]	[0.044]	[0.057]	[0.301]	[0.086]	[0.056]	[0.184]	[0.029]	[0.026]
	(0.823,0.900)	(-2.927,-1.944)	(-0.156,-0.055)	(0.350,0.531)	(-2.523,-0.029)	(-0.355,0.016)	(-0.071,0.096)	(-1.657,-0.076)	(-0.064,-0.018)	(0.203,0.242)
$s_{t,\tau}=2$	0.499	-10.70	-0.170	0.991	-5.404	-1.340	0.179	-1.723	-0.229	1.041
	[0.050]	[0.336]	[0.086]	[0.144]	[0.384]	[0.238]	[0.143]	[0.347]	[0.067]	[0.072]
	(0.370,0.627)	(-13.04,-8.427)	(-0.368,-0.035)	(0.517,1.582)	(-8.245,-2.461)	(-2.620,-0.146)	(-0.347,0.708)	(-4.119,0.630)	(-3.58,-0.113)	(0.870,1.222)

Panel B	$p_{ii}^{w\|n}$	$\alpha_{s_t,\tau}^{w\|n}$	X_t^n	X_{t-1}^w	sp	$\triangle 3mtb$	$\triangle ys$	$\triangle cs$	ls	vix	$\varsigma_{s_t}^{w\|n}$
$s_{t,\tau}=1$	0.852	-0.695	0.125	0.085	-0.109	0.238	-0.115	-0.006	-0.191	-0.026	0.077
	[0.028]	[0.080]	[0.023]	[0.043]	[0.037]	[0.172]	[0.051]	[0.033]	[0.099]	[0.017]	[0.015]
	(0.813,0.889)	(-0.850,-0.537)	(0.110,0.140)	(0.037,0.131)	(-0.145,-0.031)	(-0.324,0.827)	(-0.179,-0.050)	(-0.036,0.023)	(-0.424,0.031)	(-0.034,-0.019)	(0.070,0.083)
$s_{t,\tau}=2$	0.469	-2.824	0.164	-0.073	0.201	1.153	-0.669	0.026	0.033	-0.177	0.299
	[0.048]	[0.187]	[0.043]	[0.072]	[0.081]	[0.236]	[0.142]	[0.067]	[0.195]	[0.049]	[0.040]
	(0.350,0.589)	(-3.515,-2.129)	(0.136,0.212)	(-0.204,-0.023)	(0.042,0.361)	(0.048,2.173)	(-1.102,-0.249)	(-0.085,0.139)	(-0.710,0.762)	(-0.238,-0.123)	(0.254,0.347)

Note: Panel A reports the results estimated from (2.2.6)for individual banks (X_t^n) conditional on predictive variables. Panel B reports the results estimated from (2.2.7) for financial system (X_t^w) conditional on individual banks (X_t^n) and predictive variables. This table displays the medians of the coefficient estimates, the medians of the numerical standard errors in square brackets for evaluating the estimation accuracy, and the medians of the posterior credible intervals in parenthesis. Numerical standard errors are obtained using batch mean method, e.g., Refly (1987). The posterior credible intervals are computed using the highest posterior probability regions with the 95% confidence level. The detail estimation results for each bank are not reported here to save space, but available upon request. The results this table is for $\tau = 5\%$.

Table 2.3: VaR, MSVaR and MSES estimates of individual banks

	$VaR_{t,5\%}$	$MSVaR$		$MSES$	
		$MSVaR_{s_t=1,5\%}$	$MSVaR_{s_t=2,5\%}$	$MSES_{s_t=1,5\%}$	$MSES_{s_t=2,5\%}$
JPM	-6.286	-2.642	-9.204	-2.835	-9.760
BAC	-7.324	-2.919	-12.21	-3.141	-13.14
C	-9.446	-4.711	-35.62	-5.038	-37.56
WFC	-5.813	-2.096	-9.047	-2.269	-9.699
MS	-10.17	-5.384	-18.45	-5.750	-20.12
USB	-6.670	-4.141	-22.91	-4.419	-25.01
BK	-6.069	-2.482	-9.572	-2.687	-10.10
PNC	-6.172	-3.105	-10.07	-3.325	-10.79
STT	-7.273	-5.403	-56.95	-5.761	-59.75
BBT	-6.207	-2.497	-10.16	-2.686	-11.00
STI	-6.804	-2.530	-10.89	-2.712	-11.72
AXP	-5.664	-2.892	-9.574	-3.105	-10.10
RF	-9.327	-3.521	-16.27	-3.757	-17.40
FITB	-7.176	-3.360	-15.21	-3.598	-16.19
SCHW	-9.652	-5.767	-21.68	-6.158	-23.06
NTRS	-5.382	-1.600	-7.555	-1.759	-7.979
KEY	-6.730	-3.047	-12.03	-3.256	-12.83
MTB	-5.212	-2.711	-11.91	-2.914	-14.17
BBVA	-8.094	-3.991	-15.38	-4.277	-17.28
CMA	-7.454	-2.990	-11.64	-3.220	-12.44
HBAN	-7.482	-2.784	-11.81	-2.990	-12.82
ZION	-7.620	-2.709	-12.49	-2.914	-13.45
BPOP	-8.345	-2.543	-12.57	-2.729	-13.42
PBCT	-5.087	-2.865	-9.344	-3.084	-9.881
SNV	-8.161	-2.821	-12.937	-3.039	-13.80
BOKF	-5.360	-2.532	-9.309	-2.731	-9.876
FHN	-7.372	-2.421	-12.18	-2.623	-13.32

The entries are VaR and MSVaR values of individual banks (X_t^n) estimated from (2.2.3) and (2.2.6) conditional on predictive variables (Z_{t-1}), respectively. $MSES$ values are simulated based on the model estimation results using the approaches in Appendix A. The values are ordered by banks' total asset values.

Table 2.4: Systemic risk sensitivities

Banks	$\beta^{w\|n}_{5\%}$	$\beta^{w\|n}_{s_t=1,5\%}$	$\beta^{w\|n}_{s_t=2,5\%}$
BK	0.227	0.169	0.414
NTRS	0.208	0.178	0.374
AXP	0.272	0.213	0.359
BOKF	0.111	0.087	0.279
WFC	0.222	0.207	0.256
USB	0.185	0.116	0.238
PNC	0.209	0.196	0.231
PBCT	0.178	0.080	0.231
JPM	0.195	0.221	0.228
KEY	0.193	0.146	0.224
BBVA	0.096	0.039	0.214
STI	0.158	0.222	0.194
SNV	0.107	0.163	0.184
MTB	0.170	0.119	0.179
STT	0.152	0.071	0.173
BAC	0.143	0.197	0.147
RF	0.136	0.106	0.119
MS	0.090	0.047	0.118
ZION	0.096	0.111	0.113
BPOP	0.016	0.009	0.097
HBAN	0.068	0.162	0.058
FITB	0.112	0.123	0.055
FHN	0.115	0.059	0.049
C	0.040	0.085	0.023
SCHW	0.072	0.019	0.011
BBT	0.136	0.122	-0.048
CMA	0.129	0.110	-0.075

β_τ and $\beta_{s_t,\tau}$ are estimated from QAR and MSQAR models on (2.2.4) and (2.2.7), respectively. The banks in this table are ranked based on the risk sensitivity coefficients ($\beta^{w|n}_{s_t=2,5\%}$).

Table 2.5: Systemic Risk Contribution of each bank to financial system

1-regime		2-regime											
Banks	$\triangle CoVaR$	Banks	$\triangle MSCoVaR_1$	Banks	$\triangle MSCoVaR_2$	Banks	$\triangle MSCoVaR_3$	Banks	$\triangle MSCoES_1$	Banks	$\triangle MSCoES_2$	Banks	$\triangle MSCoES_3$
AXP	-1.617	STT	-9.886	STT	-4.064	AXP	-1.134	STT	-9.857	STT	-4.052	USB	-0.636
BK	-1.413	USB	-5.542	C	-3.037	BK	-1.096	USB	-5.528	C	-2.960	STT	-0.537
KEY	-1.339	BK	-4.029	USB	-2.713	USB	-1.084	BBVA	-3.241	USB	-2.705	BBVA	-0.461
WFC	-1.336	AXP	-3.533	BAC	-2.492	STT	-0.987	BK	-3.174	STI	-2.288	AXP	-0.435
PNC	-1.330	BBVA	-3.384	STI	-2.491	BBVA	-0.951	AXP	-2.947	BAC	-2.215	KEY	-0.327
RF	-1.284	NTRS	-3.116	SNV	-2.192	PNC	-0.941	KEY	-2.470	SNV	-1.974	MS	-0.316
USB	-1.272	KEY	-2.752	PNC	-2.162	NTRS	-0.887	BOKF	-2.308	JPM	-1.765	BOKF	-0.313
JPM	-1.248	BOKF	-2.691	JPM	-2.110	PBCT	-0.829	MTB	-2.252	HBAN	-1.748	PNC	-0.264
NTRS	-1.166	PNC	-2.550	AXP	-2.094	BOKF	-0.800	NTRS	-2.237	AXP	-1.747	SNV	-0.254
STT	-1.118	SNV	-2.476	HBAN	-1.933	KEY	-0.741	SNV	-2.230	FITB	-1.727	STI	-0.252
STI	-1.092	WFC	-2.373	WFC	-1.921	JPM	-0.684	WFC	-2.035	PNC	-1.686	RF	-0.246
BAC	-1.071	PBCT	-2.323	FITB	-1.898	MS	-0.664	MS	-1.996	RF	-1.664	JPM	-0.245
CMA	-0.985	MS	-2.202	KEY	-1.794	SNV	-0.619	STI	-1.996	WFC	-1.647	MTB	-0.243
PBCT	-0.948	JPM	-2.178	RF	-1.763	WFC	-0.595	PNC	-1.989	KEY	-1.611	BAC	-0.180
SNV	-0.914	MTB	-2.174	BK	-1.642	STI	-0.554	RF	-1.865	MTB	-1.497	ZION	-0.138
MS	-0.905	STI	-2.173	NTRS	-1.480	MTB	-0.532	JPM	-1.822	ZION	-1.309	WFC	-0.135
MTB	-0.904	RF	-1.976	MTB	-1.445	BAC	-0.491	PBCT	-1.670	BK	-1.293	BPOP	-0.118
BBT	-0.872	BAC	-1.857	ZION	-1.425	RF	-0.462	BAC	-1.651	BBT	-1.163	BK	-0.107
FHN	-0.857	ZION	-1.443	CMA	-1.327	ZION	-0.342	ZION	-1.325	CMA	-1.161	PBCT	-0.103
FITB	-0.819	BPOP	-1.253	BBT	-1.306	BPOP	-0.284	BPOP	-1.151	NTRS	-1.062	FITB	-0.080
BBVA	-0.805	FITB	-0.853	MS	-0.873	FITB	-0.198	C	-0.805	MS	-0.792	HBAN	-0.056
ZION	-0.747	C	-0.826	BOKF	-0.836	HBAN	-0.169	FITB	-0.777	BOKF	-0.717	C	-0.055
SCHW	-0.681	HBAN	-0.691	PBCT	-0.808	FHN	-0.137	HBAN	-0.625	FHN	-0.666	FHN	-0.031
BOKF	-0.621	FHN	-0.616	FHN	-0.739	C	-0.113	FHN	-0.555	BBVA	-0.593	SCHW	-0.026
HBAN	-0.507	SCHW	-0.235	BBVA	-0.619	SCHW	-0.062	SCHW	-0.210	PBCT	-0.581	BBT	0.058
C	-0.367	BBT	0.515	SCHW	-0.407	BBT	0.145	BBT	0.459	SCHW	-0.364	NTRS	0.092
BPOP	-0.134	CMA	0.906	BPOP	-0.120	CMA	0.256	CMA	0.792	BPOP	-0.110	CMA	0.100

Note: the systemic risk contributions are measured for 2-regimes using $MSCoVaR$ approach and for 1-regime using $CoVaR$ method. The banks are ranked based on each systemic risk measure.

Table 2.6: Correlation matrix of banks' systemic risk contributions measured by $\triangle MSCoVaR_1$

	BAC	BBT	BBVA	BK	BOKF	BPOP	C	CMA	FHN	FITB	HBAN	JPM	KEY	MS	MTB	NTRS	PBCT	PNC	RF	SCHW	SNV	STI	STT	USB	WFC	ZION
AXP	0.55	-0.25	0.42	0.06	0.66	0.19	0.31	-0.16	0.35	0.23	0.29	0.49	0.39	0.17	0.07	0.33	0.27	0.30	0.12	0.13	0.21	0.29	0.12	0.17	0.23	0.47
BAC	1.00	-0.66	0.57	0.74	0.74	0.72	0.83	-0.76	0.76	0.89	0.90	0.96	0.83	0.85	0.43	0.89	0.55	0.88	0.63	-0.34	0.67	0.90	-0.09	0.28	0.87	0.72
BBT		1.00	-0.84	-0.60	-0.63	-0.78	-0.37	0.87	-0.65	-0.68	-0.83	-0.57	-0.92	-0.66	-0.83	-0.78	-0.28	-0.87	-0.88	0.50	-0.81	-0.73	-0.46	-0.59	-0.68	-0.65
BBVA			1.00	0.27	0.75	0.46	0.35	0.74	0.58	0.55	0.65	0.43	0.79	0.41	0.76	0.60	0.25	0.70	0.64	-0.36	0.60	0.57	0.49	0.56	0.45	0.52
BK				1.00	0.46	0.84	0.70	-0.70	0.58	0.87	0.87	0.81	0.75	0.88	0.49	0.87	0.57	0.82	0.80	-0.49	0.78	0.91	-0.08	0.31	0.88	0.66
BOKF					1.00	0.55	0.58	-0.57	0.49	0.65	0.68	0.68	0.78	0.43	0.50	0.67	0.54	0.72	0.59	-0.29	0.68	0.72	0.34	0.56	0.59	0.78
BPOP						1.00	0.47	-0.77	0.54	0.73	0.88	0.74	0.89	0.76	0.55	0.81	0.39	0.90	0.90	-0.50	0.93	0.85	0.31	0.58	0.89	0.84
C							1.00	-0.61	0.62	0.89	0.75	0.85	0.57	0.77	0.29	0.75	0.62	0.70	0.42	0.35	0.46	0.83	-0.38	0.07	0.75	0.46
CMA								1.00	-0.62	-0.85	-0.9	-0.66	-0.89	-0.76	-0.75	-0.79	-0.45	-0.93	-0.85	0.68	-0.78	-0.85	-0.27	-0.56	-0.84	-0.65
FHN									1.00	0.71	0.78	0.77	0.69	0.85	0.49	0.88	0.17	0.69	0.56	-0.03	0.47	0.69	-0.23	0.02	0.58	0.31
FITB										1.00	0.94	0.88	0.81	0.89	0.58	0.91	0.64	0.92	0.73	-0.54	0.72	0.97	-0.12	0.31	0.91	0.63
HBAN											1.00	0.89	0.94	0.91	0.64	0.96	0.48	0.98	0.84	-0.49	0.85	0.97	0.07	0.42	0.93	0.73
JPM												1.00	0.77	0.88	0.33	0.91	0.52	0.83	0.59	-0.25	0.66	0.91	-0.21	0.17	0.87	0.69
KEY													1.00	0.75	0.73	0.88	0.44	0.97	0.91	-0.51	0.91	0.89	0.38	0.63	0.85	0.84
MS														1.00	0.50	0.94	0.40	0.84	0.69	-0.33	0.63	0.88	-0.28	0.10	0.85	0.48
MTB															1.00	0.62	0.32	0.70	0.79	-0.54	0.62	0.58	0.43	0.53	0.46	0.43
NTRS																1.00	0.47	0.92	0.80	-0.33	0.76	0.93	-0.06	0.29	0.84	0.63
PBCT																	1.00	0.53	0.43	-0.50	0.40	0.61	-0.01	0.31	0.53	0.53
PNC																		1.00	0.89	-0.58	0.89	0.96	0.22	0.55	0.93	0.80
RF																			1.00	-0.59	0.89	0.80	0.44	0.67	0.76	0.74
SCHW																				1.00	-0.55	-0.53	-0.35	-0.55	-0.57	-0.49
SNV																					1.00	0.84	0.47	0.71	0.84	0.86
STI																						1.00	0.02	0.42	0.95	0.76
STT																							1.00	0.81	0.05	0.49
USB																								1.00	0.43	0.71
WFC																									1.00	0.79

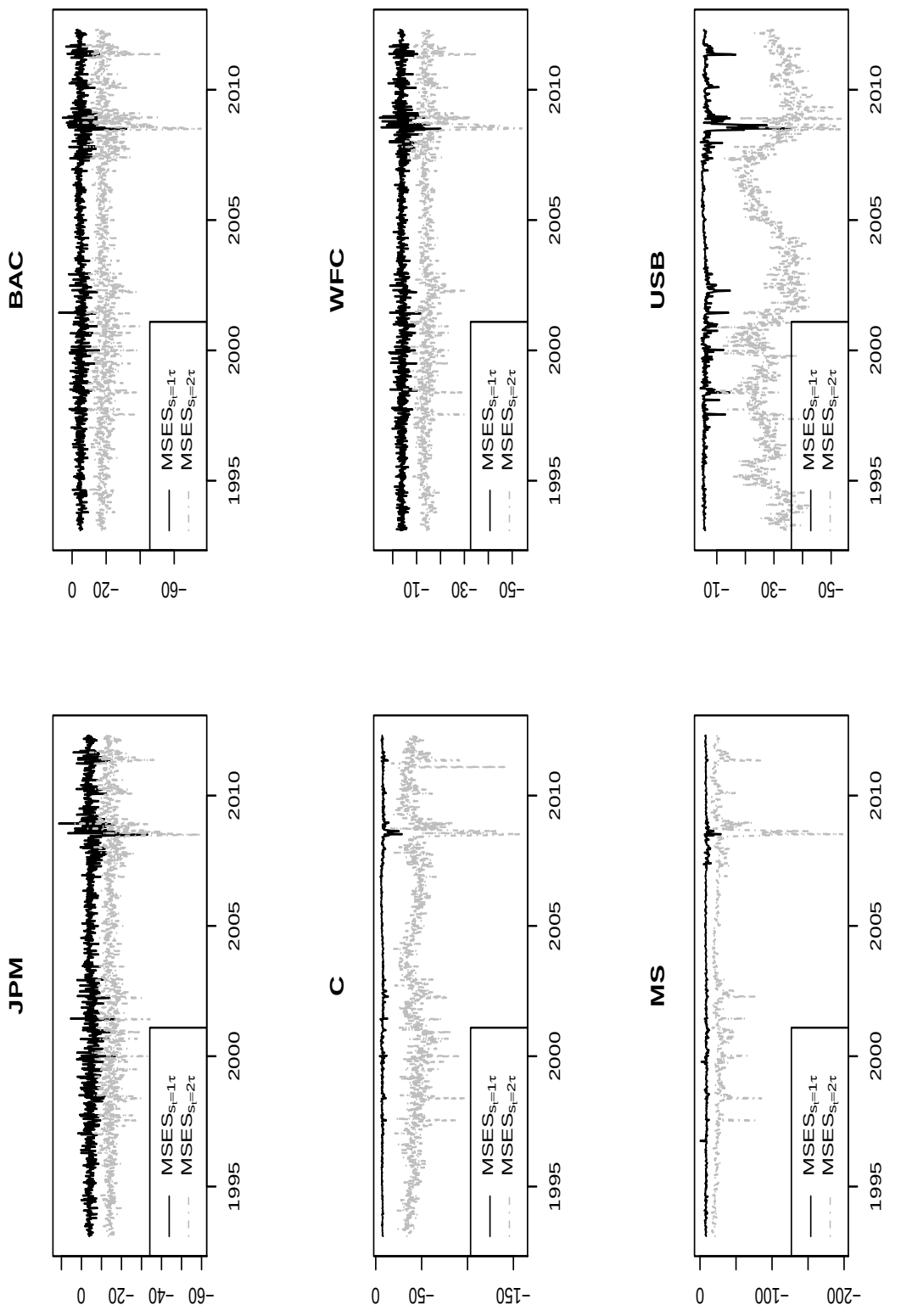


Figure 2.1: $MSES^n_{s_t,\tau}$ for a subset of the sample banks.

Note: The solid dark lines are the estimated $MSES^n_{s_t,\tau}$ values for $s_t = 1$, and the dash light lines for $s_t = 2$, respectively.

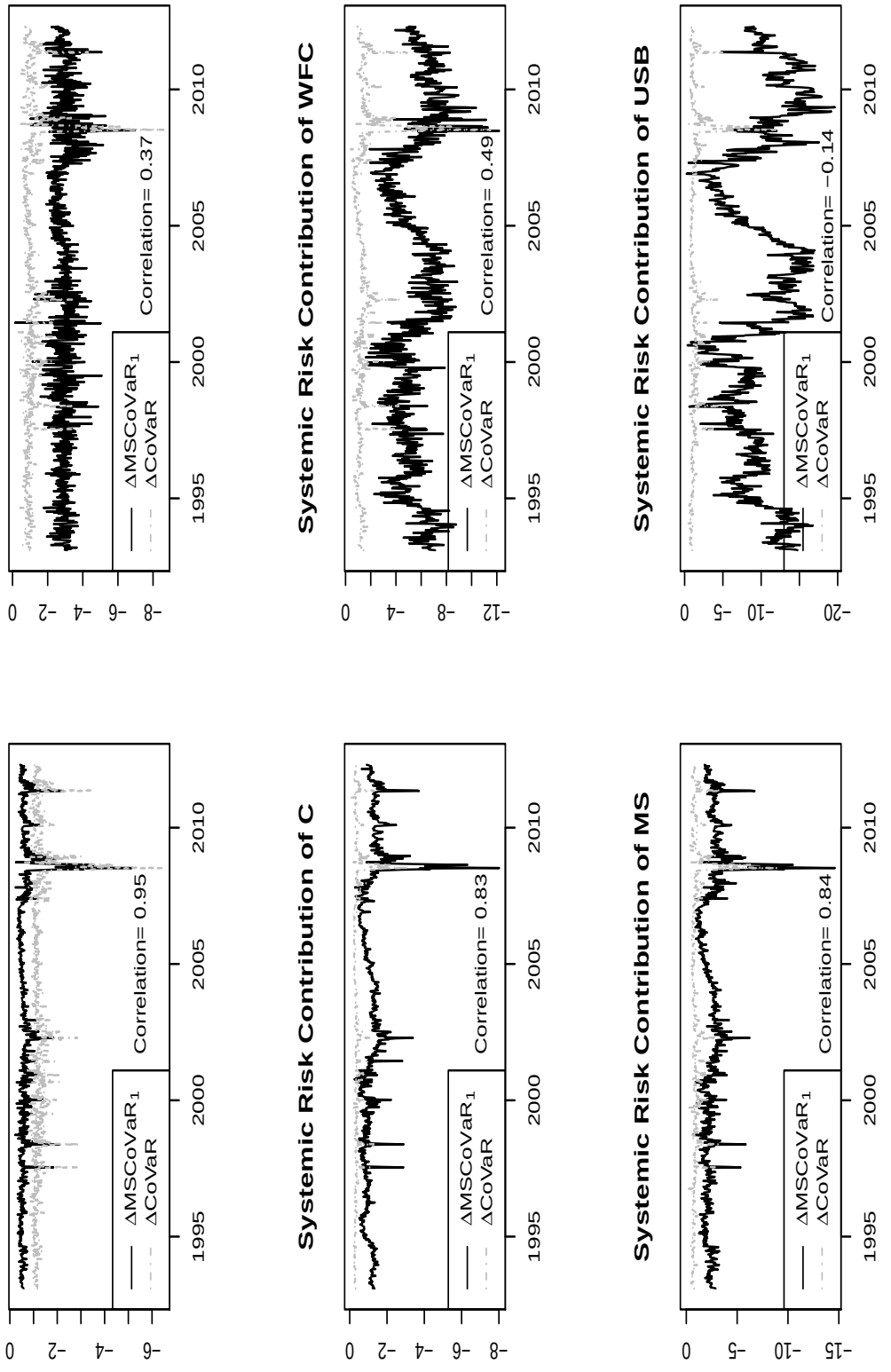
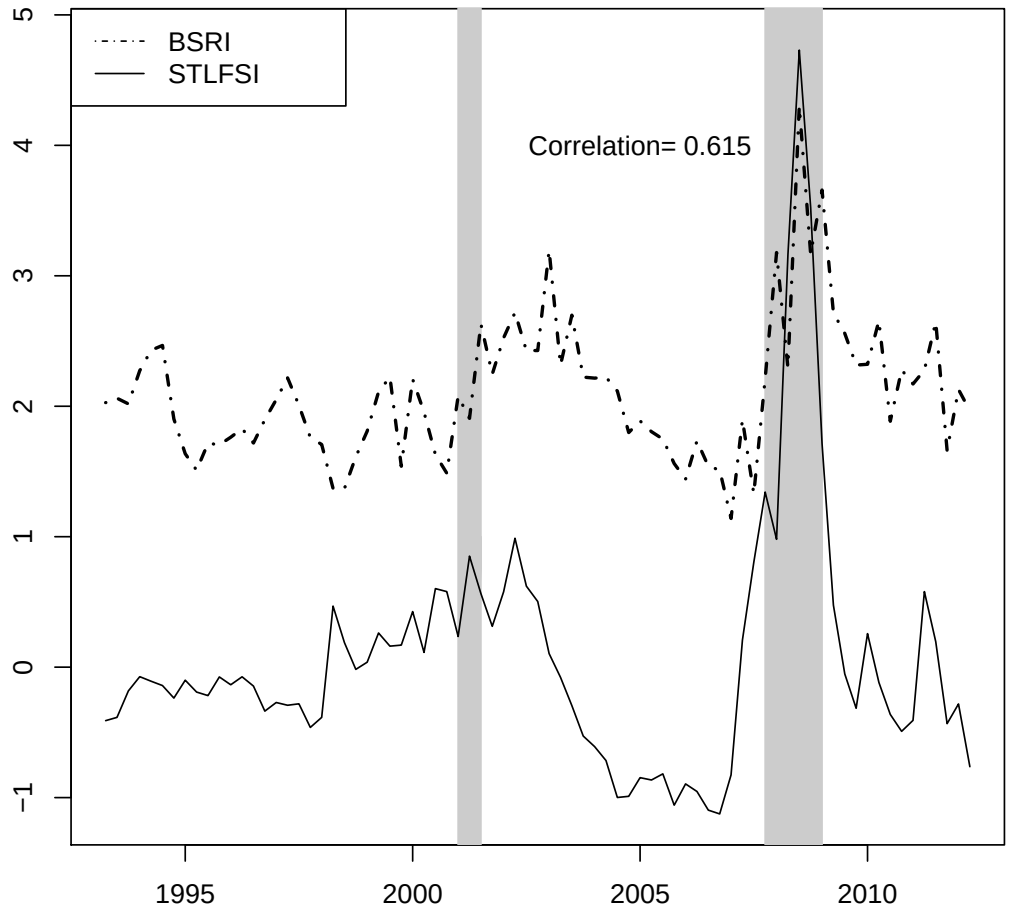


Figure 2.2: Systemic risk contributions for a subset of the sample banks: $\triangle MSCoVaR_1$ vs $\triangle CoVaR$

Figure 2.3: Banking Systemic Risk Index (BSRI).



Note: The solid line is the financial stress index constructed by Federal Reserve Bank of St. Louis and the dashed line is BSRI constructed by the value-weighted $\triangle MSCoES_1$ on individual banks.

Bibliography

[1] Adams, Z., R. Fuss and R. Gropp (2011) Spillover effects among financial institutions: a state-dependent sensitivity Value-at-Risk (SDSVaR) Approach. Working Paper

[2] Adrian, T. and M.K. Brunnermeier (2011) CoVaR. NBER Working Paper No. 17454

[3] Arias, M., J.C. Mendoza and D. Perez-Reyna (2010) Applying CoVaR to measure systemic market risk: the Colombian case. Working paper

[4] Artzner, P., F. Delbaen, J. Eber and D. Heath (1999) Coherent Measures of Risk. Mathematical Finance 9(3): 203-228

[5] Bisias, D., M. Flood, A. W. Lo and S. Valavanis (2012) A survey of systemic risk analytics. The Annual Review of Financial Economics 4: 255-96

[6] Brunnermeier, M. and M. Oehmke (2012) Bubbles, Financial Crises and Systemic Risk. NBER Working Paper No. 18398

[7] Brunnermeier, M.K., G. Dong and D. Palia (2012) Banks' Non-Interest Income and Systemic risk. Working Paper

[8] Girardi, G. and A.T. Ergun (2012) Systemic risk measurement: Multivariate GARCH estimation of CoVaR. Working paper

[9] Hamilton, J.D. (1994) Time Series Analysis. Princeton University Press

[10] Koenker, R. and Z. Xiao (2006) Quantile Autoregression. Journal of the American Statistical Association 101(475): 980-990

[11] Liu, X. (2014) Markov-Switching Quantile Autoregression. Working Paper

[12] Lopez-Espinosa, G., A. Moreno, A. Rubia, and L. Valderrama (2012) Short-term wholesale

funding and systemic risk: A global CoVaR approach. Journal of Banking & Finance 36: 3150-3162

[13] Ripley, B. (1987) Stochastic Simulation. John Wiley, New York

[14] Rodriguez-Moreno, M. and J. I. Pena (2012) Systemic risk measures: The simpler the better? Journal of Banking & Finance, forthcoming

[15] Roengiptya, R. and P. Rungcharoenkitkul (2011) Measuring Systemic Risk and Financial Linkages in the Thai Banking System. Discussion paper 02/2010. Bank of Thailand.

[16] Van Oordt, M. and C. Zhou (2010) Systematic Risk under Adverse Market Conditions. Working paper. De Nederlandsche Bank.

[17] Wong, A. and T. Fong (2010) Analysing Interconnectivity among Economies. Hong Kong Monetary Authority Working Paper 03/2010

[18] Yu, K. and J. Zhang (2005) A Three-Parameter Asymmetric Laplace Distribution and Its Extension. Communications in Statistics- Theory and Methods 34: 1867-1879

Chapter 3

Modeling Time-Varying Skewness via Decomposition for Out-of-Sample Forecast

Abstract: This chapter models time-varying skewness via a return decomposition framework which splits a return to the product of absolute return and sign. Particularly, the nonlinear dependence between absolute returns and signs is characterized by a dynamic copula function which governs a dynamic skewness process of financial returns. The importance of modeling time-varying skewness is evaluated in out-of-sample forecast for the U.S. excess stock returns in terms of both statistical significance and economic relevance. I find that the skewness timing of the proposed time-varying dependence models yields an average gain in the returns around 195 basis points per year over the forecast sample period. Statistically, the Fluctuation test shows strong evidence that the forecasting performance of the decomposition models is unstable over the sample time path. In this regard, a forecast combination, more robust to structural instability than the individual forecasts, performs significantly better out-of-sample than the benchmarks.

Keywords: Nonlinear dependence, Copula constancy tests, Dynamic tail dependence and asymmetry, Fluctuation tests, Skewness timing, Volatility timing, Forecast combination

3.1. Introduction

Recently, Anatolyev and Gospodinov (2010) (reference as AG herein) take an alternative approach to predicting excess stock returns: instead of trying to identify better predictors, they look for better ways of using predictors. They accomplish this by modeling individual multiplicative components of excess stock returns and combining information in the components to recover the conditional expectation of the original variable of interest. Let r_t denote the excess stock return at time t. Specifically, AG's approach utilizes a return decomposition given by

$$r_t = |r_t| sign(r_t) \tag{3.1.1}$$

which is also called "an intriguing decomposition" in Christoffersen and Diebold (2006). The joint distribution of the multiplicative components in (3.1.1) is obtained by combining a multiplicative error model for absolute returns, a dynamic binary choice model for signs, and a copula function for their interaction.

AG's decomposition models are able to incorporate important nonlinearities in excess return dynamics that cannot be captured in the standard predictive regression setup. Their approach, however, is restrictive as the dependence between absolute returns and signs is constant over sample periods. The constant dependence also imposes a constant skewness on excess returns. The literature has recognized that returns may in fact be better characterized by a conditional distribution with time-varying asymmetry. Some results from AG constant decomposition models are difficult to be interpreted given that the distribution of excess returns might be time-varying. For instance, their constant copula structure may give an averaging on symmetric and asymmetric distributions of excess returns over a sample period, so that the high degree of return asymmetry in some subperiods cannot be well distinguished from small asymmetry or symmetry in other subperiods.[1] This may result

[1] If the excess return at time t is symmetrically distributed around zero, then absolute returns and signs are independent and expected sign equals zero (Randles and Wolfe, 1979, Lemma 2.4.2). If the distribution of the excess return at time t has a small degree of asymmetry, then absolute returns and signs might be

in a conditionally weak dependence between multiplicative components found by AG. In this sense, a time-invariant dependency structure may miss important distribution timing of excess returns during i.e., economic recessions and financial crises, etc.

In addition, time-varying return skewness can be naturally characterized by dynamic dependence between absolute returns and signs. If the excess return at time t is distributed around zero, its predictability might be statistically small; nonetheless, it can be improved if time-varying skewness is present and modeled.[2] As a pre-testing for the U.S. excess stock returns, the test statistics of Busetti and Harvey (2011) in section 3 provide significant evidence of dynamic tail dependence and asymmetry. A time-varying process of dependence between absolute returns and signs is thus desirable for improving the predictability of excess returns.

The importance of time-varying skewness has also been found in asset pricing and allocation by recent studies. Harvey and Siddique (1999, 2000a&b) show that the inclusion of autoregressive conditional skewness affects the persistence of variance and helps explain the time-variation of the ex ante market risk premiums and the cross-sectional variation of expected returns across assets. Leon et al. (2005) estimate time-varying skewness and kurtosis using a Gram-Charlier series expansion of the normal density function for the error term. It is found that specifications allowing for time-varying skewness and kurtosis outperform specifications with constant third and fourth moments. Jondeau and Rockinger (2003) use the generalized student-t distribution with an autoregressive specification of the parameters to demonstrate the importance of time-varying asymmetry parameters.

However, these studies are mainly concerned with in-sample fit of time-varying skewness. An exception is Jondeau and Rockinger (2012) who study the importance of time-varying higher moments in out-of-sample asset allocation. Their results show that an investor might receive a sizable benefit from distribution timing comparable to volatility timing.

---

weakly dependent. By contrast, if the excess return at time t is asymmetrically distributed, then absolute returns and signs are dependent. In the second and third cases, expected sign is nonzero.

[2] Christoffersen et al. (2007) have documented that even if expected returns are zero and regardless of whether volatility dynamics are present, sign predictability arises as long as conditional skewness dynamics are present. This property remains intact in conditionally non-Gaussian environments. See online supplement of Christoffersen and Diebold (2006).

In this chapter, I propose a new approach to modeling time-varying skewness, the model performance of which is evaluated in out-of-sample forecast of the U.S. excess stock returns in terms of both statistical significance and economic values. Specifically, I extend AG's constant decomposition model by characterizing the joint distribution as a time-varying copula function.[3] The nonlinear temporary interdependence between absolute returns and signs, which governs dynamic skewness processes of returns, is thus estimated by the dynamic copula function simultaneously with marginals. Importantly, this approach provides a flexible way to estimate time-varying skewness in that the joint distribution is specified in three components, namely a copula function with two marginals, whereas conventional approaches assume a single distribution for returns. Hence, the proposed dynamic decomposition model is expected to capture both important hidden nonlinearities and the time-varying distributional natures of excess returns.

Besides modeling time-varying skewness in out-of-sample forecast, this chapter also differs from AG's work in several important ways: (1) The out-of-sample forecast period is extended to cover the recent financial crisis of 2007-2009 which has drawn tremendous research interests in both economic and finance literature. (2) The Fluctuation tests and the decomposition of forecast performance, proposed by Giacomini and Rossi (2010) and Rossi and Sekhposyan (2011) respectively, show strong statistical evidence of the instability of forecast performance over the sample time paths. This finding reconciles the insignificant results of average forecast performance from AG's constant decomposition models. Interestingly, a forecast combination, more robust to structural instability than the individual forecasts, performs significantly better out-of-sample than the benchmarks. (3) The economic values of skewness timing present substantial benefits from modeling time-varying skewness. The skewness timing of the proposed time-varying dependence models yields an average gain in the returns around 195 basis points per year over the forecast sample period. By comparison, the forecast results further show that an investor is willing to pay an extra

[3] See Manner and Reznikova (2012) for a recent survey of time-varying copulas.

442 basis points of the returns per year beyond volatility timing to acquire skewness timing information for his/her portfolio during the recent financial crisis of 2007-2009.

The rest of the chapter is organized as follows. Section 2 presents the proposed methodology of modeling time-varying skewness and is then followed by a discussion of forecasting and simulation methods for conditional mean forecasts. Section 3 describes data construction. This section also presents the pretesting results of the U.S. excess stock returns for some preliminary evidence of potential nonlinearity and time-varying tail dependence and asymmetry between absolute returns and signs. Section 4 reports empirical results of both statistical significance and economic values of forecast performance. Section 5 concludes this chapter.

3.2. The Model

The return decomposition in (3.1.1) can be rewritten as

$$r_t = |r_t| \, (2s_t - 1) \tag{3.2.1}$$

where $s_t = \mathbb{I}(r_t > 0)$ and $\mathbb{I}(\cdot)$ is an indicator function. The decomposition in (3.2.1) implies that the conditional expected return is given by

$$E_{t-1} r_t = 2E_{t-1}|r_t|s_t - E_{t-1}|r_t| \tag{3.2.2}$$

where $E_{t-1}|r_t|s_t$ is the expected cross-product of $|r_t|$ and s_t and the expectations are conditional on I_{t-1}, the information set available at time $t-1$ suppressed in subscript for simplicity.

The key to predicting returns is to model the joint distribution for the product of absolute returns and signs. Following AG's work, I obtain the joint distribution by combining

a multiplicative error model for absolute returns, a dynamic binary choice model for signs, and a copula function for the interaction between absolute returns and signs.

However, this chapter parameterizes copula functions with time-varying dependency parameters for estimating time-varying return skewness. To allow the complete emphasis in the forecasting performance of modeling time-varying skewness, I keep marginals and information variables the same as in AG, and only introduce time-varying copula dependency parameters to isolate changes irrelevant to skewness modeling which may possibly affect the forecasting performance. The proposed dynamic decomposition model, by keeping everything else the same, allows a direct examination of the role of time-varying skewness in return predictability, compared to AG's constant decomposition model. The next subsections describe the proposed dynamic decomposition structures. In addition, Appendix A summarizes the marginals from AG's chapter for convenient reference.

3.2.1. Joint Distribution

In order to construct the bivariate distribution of $Y_t = (|r_t|,\ s_t)^{'}$, I appeal to the theory of copulas as the joint distribution of a continuous variable ($|r_t|$) and a discrete binary variable (s_t). Specifically, it is well known that a conditional meta-distribution can be created as[4]

$$F_{Y_t}(u, v|I_{t-1}) = C\left(F_{|r_t|}(u|I_{t-1}),\ F_{s_t}(v|I_{t-1})\,|I_{t-1}\right)$$

where $F_{|r_t|}(u|I_{t-1})$ and $F_{s_t}(v|I_{t-1})$ are the conditional distribution functions (see Appendix A) of $|r_t|$ and s_t, respectively. $C(\cdot,\cdot|I_{t-1})$ is a bivariate conditional copula distribution function. From AG's original work, the joint conditional density/mass function of $|r_t|$ and s_t is given by

$$f_{Y_t}(u, v|I_{t-1}) = f_{|r_t|}(u|I_{t-1})\,\varrho_t\left(F_{|r_t|}(u|I_{t-1})\right)^v\left(1-\varrho_t\left(F_{|r_t|}(u|I_{t-1})\right)\right)^{1-v} \tag{3.2.3}$$

[4] see e.g., Trivedi and Zimmer (2005), Nelsen (2006), Cherubini et al. (2004) and Patton (2012) for recent surveys.

where $\varrho_t(z) = 1 - \partial C(z, 1-p_t|I_{t-1})/\partial\omega_1$ with $\omega_1 = F_{|r_t|}(u|I_{t-1})$ and $p_t = E_{t-1}(r_t > 0)$. $f_{|r_t|}(u|I_{t-1})$ is the marginal PDF function of absolute returns (see Appendix A). Observe that (3.2.3) is a product of the marginal density of $|r_t|$ and a "deformed" Bernoulli mass of s_t whose success probability is given by $\varrho_t\left(F_{|r_t|}(u|I_{t-1})\right)$.

This chapter considers a Clayton copula for the empirical application given by

$$C(\omega_1, \omega_2|I_{t-1}) = (\omega_1^{-\alpha_t} + \omega_2^{-\alpha_t} - 1)^{-1/\alpha_t} \tag{3.2.4}$$

and

$$\varrho_t(z) = \begin{cases} 1 - \left(1 + \frac{(1-p_t)^{-\alpha_t} - 1}{z^{-\alpha_t}}\right)^{-1/\alpha_t - 1}, & for\ \alpha_t \neq 0 \\ p_t, & for\ \alpha_t = 0 \end{cases} \tag{3.2.5}$$

where $\alpha_t > 0\ \forall t$ is the dependency parameter for Clayton copula and $\omega_2 = F_{s_t}(v|I_{t-1})$. By contrast, AG's constant decomposition models restrict the copula dependency parameters to be constant as α in (3.2.4) and (3.2.5).

Specifically, the copula parameter measures the dependence between $|r_t|$ and s_t. If $|r_t|$ and s_t are independent in the case of $\alpha_t \to 0$, it implies $\varrho_t \to p_t$. Also, a conditional independence between $|r_t|$ and s_t might practically occur if r_t is symmetrically distributed with a small conditional mean at time t. Note that a Clayton copula permits only positive dependence between the marginals. AG's results show that this is not restrictive for the application of monthly returns in that: (i) both positive and negative skewness values can be obtained from different dependency parameter values conditional on available information sets; (ii) the constant estimators of different copula functions from AG's original chapter are all positive; (iii) expected monthly returns are generally nonzero. Particularly, this chapter stays with the choice of a Clayton copula as empirically supported by goodness-of-fit tests in section 3, but leaves the flexibility of dependency directions open to future research.

In addition, the tail dependence between $|r_t|$ and s_t can also be captured by Clayton copula dependency parameter as $2^{-1/\alpha_t}$, while symmetric copulas, i.e., Gaussian copula,

Student-t copula, Frank copula, etc., nonetheless, have zero tail dependence.[5] In this sense, the dependency structure of Clayton copula reflects the degree of tail dependence between random variables.

3.2.2. Dynamic Dependence

To capture the potentially time-varying conditional return skewness, I parameterize the copula parameter in "observation-driven" dynamic processes. The suitable restrictions for each specification are explicitly imposed to ensure the admissible conditions that the Clayton copula has the positive support and its dynamic process is stationary.

Dynamic Tail Dependence Patton (2006) proposes the observation driven copula models for which the time-varying dependence parameter of a copula is a parametric function of the lagged data and an autoregressive term. Akin to a restricted ARMA$(1, m)$ process, the dependence parameter (or tail dependence) is specified as

$$log\left(\alpha_t\right) = \omega + \beta log\left(\alpha_{t-1}\right) + \phi \frac{1}{m} \sum_{i=1}^{m} \left|w_{1,t-i} - w_{2,t-i}\right| \tag{3.2.6}$$

where $|\beta| < 1$ and $\omega_{1,t}$ and $\omega_{2,t}$ are ω_1 and ω_2 at time t, respectively. Following Patton (2006), this chapter sets $m = 10$. Note that the expectation of this distance measure is inversely related to the concordance ordering of copulas. In the empirical section, (3.2.6) is referred to as Patton.

Exponential-weighted (ExpWeight) Dynamic Dependence As pointed out in Patton (2006), the difficulty in specifying observation-driven copula parameters lies in defining the forcing variables of evolution equations. (3.2.6) is a martingale process depending on the past path of fixed length m. Particularly, the following specification allows (3.2.6) to depend on the whole past path using exponential weights as

$$log\left(\alpha_t\right) = \omega + \beta log(\alpha_{t-1}) + \phi \frac{1-\lambda}{1-\lambda^{t-1}} \sum_{i=1}^{t-1} \lambda^{i-1} \left|w_{1,t-i} - w_{2,t-i}\right| \tag{3.2.7}$$

[5] See Cherubini et al. (2004), pp. 127 and Nelsen (2006), pp. 215.

where $0 < \lambda < 1$ and $\frac{1-\lambda}{1-\lambda^{t-1}} \sum_{i=1}^{t-1} \lambda^{i-1} = 1$. This weighting framework assigns more weights to recent observations, whereas it gives less weights to past observations. The intuition of this evolution equation comes from the general fact that the latest events exert a larger influence on current and near future dependence evolution than events far past.

GARCH-type (Gtype) Dynamic Dependence In contrast to (3.2.6) and (3.2.7), I take a first-order martingale process depending on only one past path, analog to a GARCH(1,1) process, as

$$\alpha_t = \omega + \beta \alpha_{t-1} + \phi \left| w_{1,t-1} - w_{2,t-1} \right| \tag{3.2.8}$$

with that the stability of the dynamics is assumed, for example, $\omega \geq 0, 0 \leq \beta < 1$ and $\phi \geq 0$. (3.2.8) is similar to the model (3) in Jondeau and Rockinger (2003), who present various possible specifications for the dynamics of Skewed Student-t distribution parameters.

Integrated GARCH-type (IGtype) Dynamic Dependence The integrated-GARCH type dynamics is a variant of (3.2.8) by restricting $\omega = 0$ and $\beta + \phi = 1$, such that

$$\alpha_t = \beta \alpha_{t-1} + \phi \left| w_{1,t-1} - w_{2,t-1} \right| \tag{3.2.9}$$

where $0 \leq \beta, \phi < 1$ as the strict stationary restrictions.

One-sided Asymmetric (OSA) Dynamic Dependence Dynamic dependency parameters reflect the nonlinear relations between $|r_t|$ and s_t; it can be described that α_t measures the interdependence of absolute returns and signs as the uncentered product, $\vartheta_t^+ = |r_t| \, s_t$ in 3.2.1. Hence, a dynamic dependence can also be specified as

$$\alpha_t = \omega + \beta \alpha_{t-1} + \phi \vartheta_{t-1}^+ \tag{3.2.10}$$

Despite the fact that ϑ_t^+ reveals the useful information for modeling α_t, their expectations are not equal. For the strict stationarity, following the proof in Zakoian (1994) and Nelsen

(1990, 1991), I have the conditions, $\omega \geq 0$, $0 \leq \beta < 1$, and $\phi \geq 0$. In the empirical section, (3.2.10) is estimated via a logarithmic specification for the purpose of easy convergence as

$$log\left(\alpha_t\right) = \omega + \beta log\left(\alpha_{t-1}\right) + \phi\vartheta_{t-1}^{+}$$

with $|\beta| < 1$.

Two-sided Asymmetric (TSA) Dynamic Dependence An extension to (3.2.10) is a two-sided asymmetric tail dependence as

$$log\left(\alpha_t\right) = \omega + \beta log\left(\alpha_{t-1}\right) + \phi\vartheta_{t-1}^{+} + \lambda\vartheta_{t-1}^{-} \tag{3.2.11}$$

where $\vartheta_t^{-} = |r_t|\left(1 - s_t\right)$. Glosten et al. (1993), Zakoian (1994) and Jondeau and Rockinger (2003) have also suggested the similar specifications for asymmetric effects.

TVC Dynamic Dependence The Clayton parameter is related to Kendall's τ via $\tau_t = \frac{\alpha_t}{\alpha_t+2}$, which implies that $0 < \tau_t < 1$, due to $\alpha_t > 0$. The time variation in α_t can be modeled as $\alpha_t = \frac{2\tau_t}{1-\tau_t}$ with τ_t itself governed by the TVC-type equation of Tse and Tsui (2002) as

$$\tau_t = \omega + \beta\tau_{t-1} + \phi\tilde{\tau}_{t-1} \tag{3.2.12}$$

where $\omega \geq 0$, $0 \leq \beta, \phi < 1$ and $\beta + \phi < 1$. $\tilde{\tau}_{t-1}$ is the non-negative estimators transformed from sample Kendall's $\hat{\tau}$ between periods $t - m$ and $t - 1$. For each of the $m(m-1)/2$ possible pairs, a sample estimator of Kendall's $\hat{\tau}$ between times $t - m$ and $t - 1$ is first computed as

$$\hat{\tau}_{t-1} = \frac{2}{m(m-1)} \sum_{t-m \leq t_1 < t_2 \leq t-1} sign\left[\left(w_{1,t_1} - w_{1,t_2}\right)\left(w_{2,t_1} - w_{2,t_2}\right)\right]$$

and then $\hat{\tau}_t \in [-1,\, 1]$ is converted to ensure $\tilde{\tau}_{t-1} \in [0, 1]$ by keeping the original movement of sample Kendall's $\hat{\tau}$ intact:

$$\tilde{\tau}_{t-1} = \frac{exp(\hat{\tau}_{t-1})}{1 + exp(\hat{\tau}_{t-1})}$$

(3.2.12) implicitly assumes a martingale process to capture the variation of sample Kendall's $\hat{\tau}_t$. A similar TVC approach is also specified in Jondeau and Rockinger (2006) for a time-varying Student-t copula.

This chapter applies a variance-targeting-like method as in Engle (2009) for TVC dependence estimation,

$$\tau_t = (1 - \beta - \phi)\bar{\tau} + \beta\tau_{t-1} + \phi\tilde{\tau}_{t-1}$$

where $0 \leq \beta, \phi < 1$, and $\beta + \phi < 1$. $\bar{\tau} = \sum_{i=t-R}^{t-1} \tilde{\tau}_i / R$ where R is the rolling window of fixed length specified in the next sections.

Integrated TVC (ITVC) Dynamic Dependence The integrated-TVC is simply a special case of TVC specification by restricting $\omega = 0$ and $\beta + \phi = 1$, such that

$$\tau_t = \beta\tau_{t-1} + \phi\tilde{\tau}_{t-1} \tag{3.2.13}$$

where $0 \leq \beta, \phi < 1$.

3.2.3. Likelihood Function

Given the joint distribution ((3.2.4) and (3.2.5)) with the dynamic dependence structures ((3.2.6)-(3.2.13)) and the marginals ((A.0.3), (A.0.4) and (A.0.5)), the sample log-likelihood function can be computed from (3.2.3) as:

$$\begin{aligned}
\mathcal{L}(\Phi) \quad = \quad & \sum_{t=1}^{T} s_t ln\varrho_t \left(F_{|r_t|}\left(u|I_{t-1}\right)\right) \\
& + (1 - s_t)\, ln\left(1 - \varrho_t\left(F_{|r_t|}\left(u|I_{t-1}\right)\right)\right) \\
& + lnf_{|r_t|}(u|I_{t-1})
\end{aligned}$$

With all the specified ingredients, the set of parameters to be estimated by maximum likelihood estimation is $\Phi = (\omega_v,\ \beta_v,\ \gamma_v,\ \rho_v,\ \delta_v,\ \kappa,\ \omega_d,\ \phi_d,\ \delta_d,\ \Theta)$, where Θ contains the dynamic dependence parameters, $(\omega,\ \beta,\ \phi,\ \lambda)$ from (3.2.6)-(3.2.13). Note that in this chapter, all parameters in the set Φ are simultaneously estimated by maximizing the sample log-likelihood of the full decomposition models.

3.2.4. Forecasting Methods

From (3.2.2), the forecast of excess returns for time $t+1$ is given by

$$\hat{r}_{t+1} = 2\hat{\xi}_{t+1} - \hat{\psi}_{t+1} \tag{3.2.14}$$

where $\psi_{t+1} = E_t|r_{t+1}|$ and $\xi_{t+1} = E_t|r_{t+1}|s_{t+1}$ is given by AG as

$$\hat{\xi}_{t+1} = \int_0^\infty u f_{|r_t|}\left(u|\hat{\psi}_{t+1}\right) \varrho_{t+1}\left(F_{|r_t|}\left(u|\hat{\psi}_{t+1}\right)|\hat{p}_{t+1}\right) du$$

which must be evaluated numerically. Upon the transformation of variable $z = F_{|r_t|}(u|\psi_t)$, this integral can be rewritten as

$$\hat{\xi}_{t+1} = \int_0^1 Q_{t+1}(z)\varrho_{t+1}(z)dz$$

where $Q_{t+1}(z) = F^{-1}_{|r_t|}\left(z|\hat{\psi}_{t+1}\right)$ is the quantile function of $F_{|r_t|}(u|\psi_t)$. From (A.0.3) and (A.0.4), I have that

$$F^{-1}_{|r_t|}\left(z|\hat{\psi}_{t+1}\right) = \hat{\psi}_{t+1}\Gamma^{-1}(1+\frac{1}{\hat{\kappa}_{[t]}})\left[-ln\left(1-z\right)\right]^{\frac{1}{\hat{\kappa}_{[t]}}} \tag{3.2.15}$$

where $\hat{\kappa}_{[t]}$ is the estimator of Weibull distribution using the data sample from $t-R+1$ to t. In this chapter a moving window of fixed length $R = 360$ is employed for one-step-ahead forecast. Simulation steps are described as follows. At time t,

1. Estimate the decomposition model to obtain the estimators for $\hat{\Phi}_{[t]}$ using the sample period of $t-R+1$ to t.

2. Compute $\hat{\psi}_{t+1}$ and $\hat{p}_{t+1}$ from (A.0.2) and (A.0.5) respectively by updating predictors from I_{t-1} to I_t and using $\hat{\Phi}_{[t]}$.
3. Randomly draw a vector $z = \{z_i\}_{i=1}^N$ from a uniform distribution $U(0,1)$. In this chapter $N = 20000$.
4. Compute $Q_{t+1}(z)$ from (3.2.15) using $\hat{\psi}_{t+1}$ and $\hat{\kappa}_{[t]}$
5. Compute $\hat{\varrho}_{t+1}(z)$ from (3.2.5) using $\hat{p}_{t+1}$ and $\hat{\alpha}_{t+1}$, where $\hat{\alpha}_{t+1}$ is computed from (3.2.6)-(3.2.13).
6. Compute $\hat{\xi}_{t+1} = \frac{1}{N}\sum_{i=1}^N Q_{t+1}(z_i)\varrho_{t+1}(z_i)$
7. Obtain the forecast $\hat{r}_{t+1} = 2\hat{\xi}_{t+1} - \hat{\psi}_{t+1}$

Repeat (1)-(7) to obtain a set of out-of-sample forecasts $\hat{r}_{t+1}$ for $t = R, ..., T-1$.

Recall from previous sections that if at time $t+1$, a return is approximately symmetrically distributed with a small conditional mean, its absolute return and sign might be conditionally independent. Anatolyev and Gospodinov (2010) have found some empirical evidence of weak conditional dependence. In this case, one may ignore conditional dependence to simplify the forecasting approach as

$$\hat{r}_{t+1} = \hat{\psi}_{t+1}\,(2\hat{p}_{t+1} - 1) \tag{3.2.16}$$

If the conditional dependence is weak, the feasible forecasts (3.2.16) may well dominate the feasible optimal forecasts (3.2.15) by screening noises possibly during some periods when financial data is approximately symmetrically distributed or less skewed.

In the empirical section, this chapter simultaneously estimates the parameters from full decomposition models. However, the forecast exercises are proceeded in two scenarios: conditional independence (ignoring dependence using (3.2.16)) and conditional dependence (exploiting dependence using (3.2.14) through numerical simulation).

3.3. Data

The sample period of the monthly U.S. excess returns is from 1952:01 to 2010:12, which extends the data period in AG's work to cover the recent financial crisis of 2007-2009. The data for the sample period from 1952:01 to 2002:12 is available at AG's website. For the extended period, the value-weighted excess return is taken from the Center for Research in Security Prices (CRSP); the earnings-price ratio (ep) and dividend-price ratio (dp) data in logs are constructed using the dataset provided by Shiller (2005); the 3-month T-bill rate ($ir3$) and Moody's Aaa and Baa corporate bond yield data are taken from Federal Reserve Bank of St. Louis. The yield spread (irs) is computed as the difference between Moody's Aaa and Baa yields. Daily data on NYSE/AMEX value-weighted index from CRSP is used to construct realized volatility and higher moments.

The U.S. excess stock returns of the full sample period were pretested for the existence of dependence and copula constancy between absolute returns and signs. A goodness-of-fit is also tested for the empirical choice of copula functions. Panel A in Table 3.1 reports the independence test (Genest and Rémillard (2004, 2006, 2007)). The test result rejects the null hypothesis of independence to support potential nonlinear dependence between absolute returns and signs. Panel B reports goodness-of-fit tests (Genest et al. (2009)) for bivariate copula functions. The test results favor a Clayton copula and reject a Gumbel copula among other elliptical copulas. The choice of a Clayton copula with lower tail dependence provides strong evidence for asymmetric distributions of the U.S. stock returns, consistent with the stylized fact in financial data.[6]

Panel C reports the results of the copula constancy tests (Busetti and Harvey (2011)). This approach associated with different quantile levels is flexible and useful in pointing to changes in the different parts of a copula distribution. τ represents the chosen quantile confidence level. The test results reject both constant lower and upper tail dependences.

[6] Independence test on the difference between the empirical copula distribution function and the product of the marginal empirical distributions is based on asymptotically independent Cramér-von Mises statistics derived from a Mobius decomposition of the empirical copula process. Parametric bootstrap is used to obtain critical values. In addition, for the properties of the copulas considered in Table 3.1, see Nelsen (2006) and Cherubini et al. (2004).

The evidence of time-varying lower tail dependence seems stronger than time-varying upper tail dependence, due to p-value rejecting the null at 10% confidence level for upper tail dependence with τ=75% but rejecting the null at 5% confidence level for lower tail dependence with τ=25%.[7] Also, the test results reject constant asymmetry and overall copula constancy across quantile levels, except the copula constancy test for τ=0.5. As a result, the preliminary investigations in Table 3.1 demonstrate some empirical evidences for: the existence of significant nonlinear dependence which cannot be estimated by linear regressions; and time-varying tail dependence and asymmetry between the U.S. absolute returns and signs, which are not considered in AG's constant decomposition models.

The model naming convention in empirical section is introduced as follows. In this chapter, CDM and DDM denote the constant and dynamic decomposition models, respectively. CI and CD represent the forecast scenarios for ignoring and exploiting dependence, respectively. For instance, the dynamic decomposition model of IGtype, if its forecasts by exploiting dependence structure, is denoted as IGtype-CD; otherwise as IGtype-CI. A constant decomposition model with the presumption of conditional independence is denoted as CDM-CI.

3.4. Empirical Results

Table 3.2 summarizes the in-sample estimation results. p-values are reported in parentheses. The subsample estimation from 1952:01 to 2002:12 is made comparable to Anatolyev and Gospodinov (2010). The estimation results are statistically significant at conventional confidence levels. p-values for Wald joint significance tests and the likelihood ratio tests reject the null hypothesis of dynamic dependence parameters being jointly equal to zero across sample periods and models. These results provide in-sample evidence for significant nonlinear and time-varying dependence between absolute returns and signs, and are also consistent with the pretesting results. In addition, Table 3.2 shows that among the dy-

[7] Busetti and Harvey (2011) (pp. 115) show that the lower quadrant test is more powerful for τ<0.5, that is in the lower tail of the distribution, while the upper quadrant test is more powerful in the upper tail (τ>0.5).

namic decomposition models, OSA and TSA obtain the highest persistence levels, estimated through the parameter β's and comparable to the persistence level of volatility processes commonly documented in GARCH literature. Nonetheless, the integrated models (IGtype and ITVC) have the persistence of around 0.75, which is lower than 0.94 the value calibrated for the integrated GARCH model by J.P. Morgan RiskMetrics.

3.4.1. Density Forecasts of Copula Specifications

Diks et al. (2010) propose a statistical test for comparing the predictive accuracies of competing copula specifications in multivariate density forecasts, using the Kullback-Leibler information criterion (KLIC). The test method is valid under general conditions on the competing copulas including the density forecasts from copulas with time-varying dependence parameters. The test statistic of equal KLIC scores, a heteroskedasticity and auto-covariance consistent (HAC) estimator, asymptotically follows the standard normal distribution by applying Theorem 4 of Giacomini and White (2006).

Table 3.3 presents the test statistic of Diks et al. (2010). The entries in this table are the mean values of KLIC score difference between the benchmark (constant decomposition) model and competing (dynamic decomposition) models, scaled by 100. p-values are reported in parentheses for the null hypothesis of equal predictive accuracy in density forecasts. In the case of rejection, a negative mean of KLIC score difference provides the statistical evidence for the density forecastability of a competing model over the benchmark model and vice versa. As seen in Table 3.3, the null hypothesis is statistically rejected by IGtype, OSA and TSA models. Further, in the rejections, the mean values of KLIC score difference are negative, which indicate their better predictive accuracy in terms of density forecasts than the benchmark model, as a higher average copula score is preferred. By contrast, the mean score differences of ExpWeight and Gtype dynamic models are negative but statistically insignificant.

To preserve space, the rest of this chapter reports the empirical results from the dynamic copula specifications of IGtype and OSA, the density forecasts of which are significantly

better than the constant decomposition models.[8] Besides, interest is also given to the average forecasting performance on all proposed dynamic decomposition models.

3.4.2. Statistical Significance of Out-of-Sample Forecasts

To evaluate out-of-sample forecast performance, this chapter uses the out-of-sample coefficient suggested by Campbell and Thompson (2008) as

$$OS = 1 - \frac{\sum_{t=R+1}^{T} L(r_t - \hat{r}_t)}{\sum_{t=R+1}^{T} L(r_t - \bar{r}_t)}$$

where $L(r_t - \hat{r}_t)$ is a loss function based on forecast errors and R is the length of the rolling window in forecasting. The OS statistic measures the reduction in forecast errors of a competing model ($\hat{r}_t$) relative to the historical average ($\bar{r}_t$) model. Particularly, a positive OS value implies that a competing model performs better out-of-sample than the benchmark model and vice versa.[9] This chapter considers both squared and absolute forecast errors in loss functions.

Table 3.4 presents the OS statistic results scaled by 100. Bold values indicate the highest forecast gains in terms of the OS statistic among competing models within a given sample period. The OS results from the full forecast period of 1982:01 to 2010:12 show that the decomposition models have positive forecast gains, despite that the positive gain is marginally small for the CDM-CI model. Importantly, the dynamic decomposition models of exploiting time-varying return skewness have the average relative forecast performance of 2.62% higher than other models. Among these models considered, the IGtype-CD model obtains the highest relative forecast gain of 3.59% followed by the OSA-CD model with the relative gain of 3.57%.

[8] OSA is chosen mainly due to its forecasting performance similar to TSA but with a relatively simpler copula specification.

[9] Following finance literature, this chapter also uses the historical average model as benchmark for OS statistic. Given a sample period, the forecast errors of historical average model in the denominator of OS remain the same across different competing models. In this section, all forecast performance comparisons among competing models are made relative to the historical average model, hereafter referred to as relative forecast performance. Note that converting relative performance to direct performance comparisons among pairwise competing models does not alter the conclusions in this section.

However, by excluding the recent financial crisis to be comparable with the forecast period in Anatolyev and Gospodinov (2010), DDM-CI models obtain the average relative forecast gain of 3.66% higher than other models. Among these, the IGtype-CI model appears to have the highest relative forecast performance of 4.62%. Based on this interesting observation, this chapter further excludes the forecast period of 1998:01 to 2012:12 which is the economic recession period including the Hi-tech Bubble Bust and Asian financial crisis. During this relatively more tranquil period after the exclusion of this economic recession period, DDM-CI models continuously perform better out-of-sample with the average relative forecast gain of 4.64%, and among these IGtype-CI model obtains the highest relative gain of 5.05%.

In comparison to the tranquil periods, this chapter also evaluates the forecast performance during economic recessions. Conversely, the results show that during both the Hi-tech Bubble Bust recession and the recent financial crisis, exploiting the time-varying skewness provides the average relative forecast performance of 8.53% and 8.31%, respectively, dramatically higher than other decomposition models. Also, among these, the IGtype models of exploiting dynamic dependence obtain the highest relative forecast gains of 9.36% and 12.31%, respectively.

Intuitively, the forecast performance of exploiting conditional dependence during a turmoil period reflects that volatility (absolute returns) is higher, asymmetric effects (signs) are stronger, and temporary interdependence between volatility and sign increases during a market downturn. This intuition is further consolidated by Figure 3.1 which plots the out-of-sample estimation of nonlinear copula dependence structures. The estimated time-varying dependence structures are remarkably higher and more volatile during turmoil times, while the degree of dependence from the constant decomposition model is relatively stable over time. Specifically, the OSA dynamic decomposition model obtains the mean dependence of 0.275 which is 3.4 times higher than the constant decomposition model. The averages of dependence on all proposed dynamic decomposition specifications are also consistently above the constant decomposition model over the sample period.

As a result, the empirical evidence above show the presence of possible instabilities of forecast performance which depend on economic conditions. The instability of forecast performance implies that a competing model may not dominate a benchmark model over the whole forecast period: better in some subperiods and no difference in other subperiods. This forecast instability feature might suggest that a risk-averse investor should consider the dynamic decomposition models with the presumption of conditional independence during tranquil periods, whereas during turmoil periods he/she should exploit temporary interdependence between absolute returns and signs. The similar conclusion can be also drawn from the results of absolute forecast errors with the highest relative forecast gains from OSA-CD models during turmoil periods. The next subsections formally test the statistical significance of predictive accuracy and the instability of the forecast performance, and also provide more detailed insights on the sources of forecast performance.

Statistical Significance Test

Table 3.5 reports the test results of statistical significance. Entries in this table are p-values of Giacomini and White (2006) test statistic for the null hypothesis of equal conditional predictive ability.[10] Values in square brackets are relative performance that indicates the proportion of times over the sample period that a competing model in the row heading dominates a benchmark model in the column heading. Despite that the relative performance suggests that the dynamic decomposition models dominate both the historical average and the constant decomposition models, p-values show no clear evidence that the predictive accuracies of the dynamic decomposition models are statistically significantly different from the benchmarks. A few exceptions are that the IGtype-CD and OSA-CD models reject the null hypothesis to show better predictive abilities than the benchmarks. However, the

[10] Let $\triangle L_{t+1}$ denote the difference of the loss functions (squared or absolute losses) of two models at time $t+1$. Then, the null of equal predictive ability of two models can be expressed as $H_0 : E(h_t \triangle L_{t+1}) = 0$ where h_t is a $q \times 1$ vector that belongs to the information set at time t. Following Anatolyev and Gospodinov (2010) and Giacomini and White (2006), this chapter sets the conditional variable $h_t = (1, \triangle L_t)'$. The test statistic of equal predictive ability is χ^2_q-distributed under the null. For details, see Giacomini and White (2006).

average predictive abilities on the DDM-CD models are insignificantly different from the benchmarks.

The insignificant test results in Table 3.5, also found in AG's work, might be due to the presence of forecast instabilities as pointed out in subsection 4.2. Giacomini and Rossi (2010) show that in unstable environments the forecasting performance of models may itself change over time. See also Stock and Watson (2003). Hence, tests of overall predictive ability selecting the model that forecasts best on average, e.g., Diebold and Mariano (1995), Clark and West (2006), and Giacomini and White (2006), etc., may result in a loss of information and possibly lead to incorrect selection decisions. To address this issue, the next subsection investigates the stability of forecasting performance over time by means of statistical tests.

Fluctuation Test

Figure 3.2 plots the fluctuation test statistics (Giacomini and Rossi (2010)).[11] Dashed lines indicate critical values at 5% confidence level. A positive test statistic implies the smaller forecast loss of a competing model. A one-sided test statistic (solid line) above the critical value represents that the competing model statistically significantly performs better than the benchmark at the time point when the test statistic is evaluated. The naming convention in Figure 3.2 is as a benchmark model vs. a competing model. For instance, the notation of CDM-CI vs. OSA-CI means that CDM-CI is used as the benchmark model and OSA-CI as the competing model for computing test statistics of Giacomini and Rossi (2010).

[11] Giacomini and Rossi (2010) propose the Fluctuation test for comparing the out-of-sample forecasting performance of two competing models in the presence of possible instabilities. The main idea is to develop a measure of the relative local forecasting performance on the entire time path, which may contain useful information that is lost when looking for the model that forecasts best on average. Particularly, this chapter sets $\mu = 0.2$ for the centered rolling windows of size m as in Giacomini and Rossi (2010). The corresponding critical values are provided by Table 1 of Giacomini and Rossi (2010). I would like to thank an anonymous referee who suggests this Fluctuation test.

Note that the forecast instability addressed by the Fluctuation test is not the instability in the Giacomini and White (2006) conditional predictive ability test in that the Fluctuation test traces the Giacomini and White (2006) unconditional predictive ability test over time. For instance, the results of the Giacomini and White (2006) conditional predictive ability test cannot be obtained by averaging the resuls of the Fluctuation tests over time. For details see their original works and also Giacomini (2011).

The fluctuation test results show that forecast performance statistically varies over time. For instance, the OSA-CI forecasts have better predictive ability than the historical average during 1986-1989, while their difference is insignificant during 2002-2009. By contrast, the OSA-CD forecasts perform significantly better during 2007-2010, but are insignificantly different from the historical average during 1986-1989. In addition, both forecasts appear statistically insignificantly different from the historical average during the housing market boom period of 2004-2006.

Turning to the comparisons between constant and dynamic decomposition models, both the forecasts of the OSA-CI and OSA-CD models are significantly better than the CDM-CI and CDM-CD models, respectively, except during 1986-1989 for the OSA-CI forecasts and during 2003-2006 for the OSA-CD forecasts. Further looking at conditional independence vs. conditional dependence, both the forecasts of the CDM-CD and OSA-CD models perform significantly better than the CDM-CI and OSA-CI models, respectively, except for the period of 1986-1989. It is also clearly observed that the CDM-CD and OSA-CD forecastabilities exceed the CDM-CI and OSA-CI models respectively during the turmoil periods (after 1997), whereas they do not appear to be significantly better during the relative tranquil periods (prior to 1997).

Table 3.6 presents more detailed test results from the decomposition of forecasting performance (Rossi and Sekhposyan (2011)).[12] Entries in this table are p-values. The models in the first column and the first row are benchmarks and competing models respectively. The tests for the first component exclusively reject the null hypothesis of no time variation in the expected relative forecasting performance, in line with the fluctuation test results. The tests for the second component show that, compared to historical average, in-sample losses of the decomposition models provide better explanation for out-of-sample losses, while the

[12] Rossi and Sekhposyan (2011) decompose the forecasting performance into 3 components, namely time variation in the forecasting performance, predictive content and over-fitting. The first component measures the presence of time variation in the models' performance relative to their average performance. Predictive content measures the models' out-of-sample relative forecasting ability reflected in the in-sample relative performance, for instance, whether in-sample losses have predictive content for out-of-sample losses. Over-fitting measures models' in-sample fits not reflected in the out-of-sample forecast. Particularly, this chapter sets $\mu = 0.2$ for the centered rolling windows of size m in Rossi and Sekhposyan (2011). The corresponding critical values are provided by Table 1 of their original work.

predictive contents are similar between constant and dynamic decomposition models. The tests for the third component do not show evidence for difference in the significance of the over-fitting component across the models. This implies that the proportion of out-of-sample losses which cannot be explained by in-sample losses is statistically similar between benchmarks and competing models.

The statistical test results above have revealed that one model performs better in certain periods and the competing model is more accurate in other periods. A natural question to ask, however, is what a forecaster should do if the tests find instability in the relative performance of competing models. In this case, a forecast combination may be more robust to structural instability than either of the individual forecasts. Table 3.7 reports the results from simple forecast combination experiments. Following Rapach et al. (2010), the combination weights are determined by using the discount mean square prediction error (DMSPE) with a discount factor of 0.9.[13] The first row represents the combined models. Compared to Table 3.4, combining forecasts obtains higher positive forecast gains. More importantly, in contrast to the insignificant test results from Table 3.5, p-values in Table 3.7 show that the forecast combinations perform significantly better forecastabilities than the historical average. However, the fluctuation tests (not reported here) also show that instabilities of forecast performance do not completely disappear from forecast combinations. In-depth analysis of forecast combinations in this context is open to future research.

3.4.3. Economic Values of Out-of-Sample Forecasts

Next, this chapter assesses the economic values of out-of-sample forecasts in terms of portfolio profits from market timing, volatility timing and skewness timing. A relatively naive market timing strategy is employed to actively allocate portfolio between stocks and

[13] Different discount factor values are also considered (not reported here but available upon request to the author) and do not alter the conclusion. In addition, the purpose of using the weighting scheme from Rapach et al. (2010) is to show empirical evidence for forecast combination robust to instability. Hence, the choice of it is not based on whether it is better than other candidate weighting schemes such as simple averaging, Bayesian model averaging, information criterion weighting, etc. For example, I also conduct a simple averaging which does not alter the conclusion of forecast combination (not reported here). However, given the evidence of the robustness of forecasting combination found in this chapter, the comparisons and choices between various weighting schemes might be interesting to pursue in future research.

bonds as in Guo (2006) and Anatolyev and Gospodinov (2010), among others: investing in stocks if the predicted excess return is positive or in bonds if the predicted excess return is negative. The profits of market timing are computed from actual stock returns and risk-free rates and compared to the benchmark buy&hold strategy. Economic values of volatility timing and skewness timing are then evaluated through utility functions.

Economic Values of Market Timing

Table 3.8 reports the market timing results for buy&hold and decomposition-based trading rules. Bold values indicate the best annualized performance. Average portfolio returns in this table show strong evidence for the economic relevance of the proposed dynamic decomposition approach. During the full sample period of 1982:01-2010:12, the DDM-CD portfolios produce the average return of 11.96% slightly higher than the buy&hold and the constant decomposition portfolios. Among these, the OSA-CD portfolio has the highest return of 12.48%. Compared to buy&hold portfolio, all decomposition-based portfolios are accomplished with a large reduction in standard deviation, for instance, 15.09% of the buy&hold vs. 13.75% of the DDM-CD portfolios on average. As a result, among portfolios the OSA-CD model has the highest Sharpe ratio of 0.559 (versus 0.473 from buy&hold portfolio).

In sharp contrast, considering only the 1998-2002 and 2007-2010 periods leads to a significant deterioration of the statistics for buy&hold portfolios, while the decomposition-based strategies of exploiting time-varying skewness are still practically very profitable during these two economic recessions. Specifically, the DDM-CD portfolios have presented very impressive performance even during the recent financial crisis period. For instance, the OSA-CD portfolio achieves around 710 and 523 basis points of the returns per year higher than buy&hold and CDM-CI portfolios, respectively, during the crisis. In addition, the IGtype-CI portfolio has the highest average returns and Sharpe ratios during relative tranquil periods, whereas the OSA-CD portfolio obtains the best profits during turmoil times. This result also suggests the robust trading strategies that a risk-averse investor should consider the dynamic decomposition models by ignoring dependence during tranquil periods

but exploiting dependence during turmoil times. Alternatively, portfolios based on forecast combinations can be considered as well.

Economic Values of Volatility and Skewness Timings

To evaluate economic values of volatility and skewness timings, I employ the performance fee measure as in Fleming et al. (2001), Guo (2006), and Jondeau and Rockinger (2012), among others. A positive management fee represents that an investor is willing to pay to switch from a benchmark strategy to a competing strategy.

Consider a power utility function $U(W_{t+1}) = W_{t+1}^{1-\gamma}/(1-\gamma)$, where $\gamma > 0$ ($\gamma \neq 1$) measures the investor's constant relative risk aversion. The performance fees for volatility timing and skewness timing (denoted by $\triangle$ and $\triangle'$, respectively) are estimated by equating the average utilities of the second- and third-order Taylor series expansions between competing and benchmark strategies, respectively,

$$\sum_{t=R}^{T-1}\left[(r_{p,t+1}-\triangle)-\frac{\gamma}{2}(r_{p,t+1}-\triangle)^2\right]=\sum_{t=R}^{T-1}\left(r_{b,t+1}-\frac{\gamma}{2}r_{b,t+1}^2\right) \tag{3.4.1}$$

$$\sum_{t=R}^{T-1}\left[\left(r_{p,t+1}-\triangle'\right)-\frac{\gamma}{2}\left(r_{p,t+1}-\triangle'\right)^2+\frac{\gamma(\gamma+1)}{3!}\left(r_{p,t+1}-\triangle'\right)^3\right]=\sum_{t=R}^{T-1}\left(r_{b,t+1}-\frac{\gamma}{2}r_{b,t+1}^2+\frac{\gamma(\gamma+1)}{3!}r_{b,t+1}^3\right) \tag{3.4.2}$$

where $r_{b,t+1}$ and $r_{p,t+1}$ represent the buy&hold and the decomposition-based portfolio returns, respectively. (3.4.1) and (3.4.2) are solved numerically for $\triangle$ and $\triangle'$.[14] In this set-up, the performance fee means how much an investor is willing to pay for acquiring the information from a decomposition-based portfolio versus the buy&hold portfolio.

Table 3.9 reports the economic values of volatility and skewness timings.[15] The results clearly show that the OSA portfolios provide the highest performance fees for an investor acquiring skewness information, while the IGtype portfolios obtain the highest performance fees for volatility timing. This result demonstrates the economic importance of modeling

[14] I assume that $\gamma = 5$. The results are not sensitive to reasonable variations in γ. See also Rapach et al. (2010), Guo (2006) and Jondeau and Rockinger (2012), among others.

[15] Particularly, volatility timing is evaluated by (3.4.1) using the decomposition portfolios with the presumption of conditional independence (ignoring dependence), while skewness timing is evaluated by (3.4.2) using the decomposition portfolios of exploiting dependence.

time-varying skewness, which generates the performance fees of 55-174 basis points of the returns per year more than the constant decomposition portfolios across different sample periods.

The comparisons between volatility timing and skewness timing also show large variations over time depending on economic conditions.[16] In the full forecast sample period of 1982-2010, an investor is willing to pay 68 basis points of the returns per year on average more than volatility timing to acquire skewness timing for his/her portfolio. Specifically, skewness timing from the OSA trading strategy receives a management fee of around 131 basis points of the returns per year more than volatility timing from an investor.

The economic values of skewness timing are much higher during the turmoil times of 1998-2002 and 2007-2010 than the relatively tranquil periods of 1982-2002 and 1982-1997. Particularly, the incorporation of skewness timing into an investor's portfolio during the recent financial crisis receives 442 basis points of the returns per year on average more than volatility timing. By contrast, during the relatively tranquil periods, performance fees of volatility timing are 35-249 basis points of the returns per year higher than those of skewness timing.

An explanation of this performance difference might be explicated as follows. During tranquil periods, return distributions might be more symmetric such that the relationship between absolute returns and signs is either conditional independence or weakly conditionally dependent. Nonetheless, during market downturns returns are more likely asymmetrically distributed such that a strong temporary interdependence exists between absolute returns and signs. Apparently, the increase in return asymmetry during market downturns raises the importance of modeling time-varying skewness.

[16] The difference between $\triangle'$ and $\triangle$ represents the extra performance fee for which an investor is willing to pay beyond volatility timing to obtaining benefits from skewness timing.

3.5. Conclusion

This chapter estimates time-varying return skewness by specifying dynamic dependence in a return decomposition framework which splits a return to the product of absolute return and sign. The importance of time-varying skewness is evaluated in out-of-sample forecasts of excess stock returns in terms of both statistical significance and economic relevance.

The empirical results show strong statistical evidence for the instabilities of forecast performance: one model performs better in certain periods and the competing model is more accurate in other periods. Particularly, the results of this chapter suggest that an investor might consider the proposed dynamic decomposition model with the presumption of conditional independence during relatively tranquil periods, while by exploiting conditional dependence between absolute return and sign during turmoil times. Alternatively, a forecast combination, more robust to structural instability than either of the individual forecasts, performs statistically significantly better out-of-sample than the benchmarks.

The economic values of the proposed dynamic decomposition models also show clear economic relevance. This chapter finds that the skewness timing of dynamic decomposition models yields an average gain in the returns around 195 basis points per year over the forecast sample period. Specifically, modeling time-varying skewness during the recent financial crisis generates a substantial average gain of 818 basis points in the returns per year. It has also been found that an investor is willing to pay an extra 68-442 basis points of the returns per year beyond volatility timing to acquire skewness timing information for his/her portfolio.

Additionally, the following interests might be considered for future research: (1) an in-depth analysis of forecast combinations robust to structural instability of forecasting performance; (2) the comparisons of the proposed dynamic decomposition models to the approaches of Harvey and Siddique (1999) and Leon et al. (2005); (3) the applications of other copula functions which allow negative dependence structures between absolute returns and signs.

Table 3.1: Pretesting the U.S. Excess Stock Returns

Panel A						
Independence Tests	0.041					
Panel B	Gaussian Copula	t Copula	Frank	FGM	Gumbel	Clayton
Goodness-of-fit Tests	0.040	0.043	0.040	0.035	<0.001	0.154
Panel C: Copula Constancy Tests	τ					
	0.1	0.25	0.5	0.75	0.9	
Constant Lower Tail Dependence	0.012	0.031	0.019	0.007	0.013	
Constant Upper Tail Dependence	<0.001	0.002	0.006	0.076	0.019	
Constant Asymmetry	0.041	0.006	<0.001	<0.001	0.022	
Overall Copula Constancy	<0.001	0.026	0.378	0.074	0.014	

Entries in this table are p-values. In Panel C, the tests for constant lower and upper tail dependences are based on (4) and (5) in Busetti and Harvey (2011); the test for constant asymmetry is described on pp. 114 of Busetti and Harvey (2011) and overall copula constancy is based on the quadrant association test.

Table 3.2: In-Sample Estimation Results

	Constant	Dynamic Decomposition Models							
	Decomposition Model	Patton	ExpWeight	IGtype	Gtype	OSA	TSA	TVC	ITVC
1952:01-2010:12									
ω		-2.940	-2.222		0.127	-0.221	-0.314		
		(<0.001)	(0.032)		(0.044)	(0.048)	(0.166)		
β		0.218	0.512	0.767	0.689	0.932	0.901	0.645	0.732
		(0.316)	(0.098)	(<0.001)	(0.023)	(<0.001)	(<0.001)	(0.036)	(0.021)
ϕ		4.815	3.909		0.256	4.865	3.785	0.087	
		(<0.001)	(<0.001)		(0.083)	(<0.001)	(0.013)	(0.218)	
λ			0.702				1.848		
			(0.007)				(0.218)		
α	0.112								
	(0.016)								
Likelihood Ratio Test		<0.001	<0.001	<0.001	0.004	0.04	0.005	0.002	0.036
Wald Test		<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.041
1952:01-2002:12									
ω		-2.664	-2.236		0.092	-0.235	-0.247		
		(0.017)	(0.046)		(0.041)	(0.064)	(0.108)		
β		0.436	0.492	0.768	0.636	0.933	0.927	0.608	0.753
		(0.127)	(0.118)	(<0.001)	(0.051)	(<0.001)	(<0.001)	(0.221)	(0.063)
ϕ		4.749	3.802		0.192	5.301	4.031	0.046	
		(0.046)	(0.056)		(0.062)	(0.011)	(0.006)	(0.382)	
λ			0.677				1.016		
			(0.010)				(0.296)		
α	0.093								
	(0.047)								
Likelihood Ratio Test		<0.001	<0.001	<0.001	0.003	0.041	0.006	0.010	0.030
Wald Test		0.002	0.012	<0.001	0.007	<0.001	<0.001	0.095	0.120

p-values are reported in parentheses, and also for likelihood ratio tests and Wald joint significance tests.

Table 3.3: Density Forecast Comparisons of Copula Specifications

	Patton	ExpWeight	IGtype	Gtype	OSA	TSA	TVC	ITVC
1982:01-2010:12	0.236	0.052	-0.909	-0.521	-0.243	-0.516	0.175	0.052
	(<0.001)	(0.478)	(<0.001)	(0.146)	(<0.001)	(0.007)	(0.238)	(0.423)
1982:01-2002:12	0.484	-1.020	-1.272	-0.204	-0.277	-0.423	0.181	0.247
	(0.217)	(0.159)	(<0.001)	(0.107)	(0.023)	(<0.001)	(<0.001)	(0.127)

Entries in this table are the mean values of KLIC score difference between the benchmark (constant decomposition) model and competing (dynamic decomposition)) models, scaled by 100. The tests are based on Diks et al. (2010) with the null hypothesis of equal density forecast accuracy. p-values are reported in parentheses. In the case of rejection, a negative value provides the statistical evidence for the density forecastability of a competing model over the benchmark model.

Table 3.4: Out-of-Sample Forecast Statistics

	Constant Decomposition Models		Dynamic Decomposition Models					
	Ignoring dependence	Exploiting dependence	Ignoring dependence			Exploiting dependence		
			IGtype	OSA	Average	IGtype	OSA	Average
(a) Squared forecast errors								
1982:01-2010:12	0.18	2.16	1.69	1.57	1.41	**3.59**	3.57	2.62
1982:01-2002:12	3.23	2.79	**4.62**	3.34	3.66	3.44	2.88	3.01
1982:01-1997:12	4.13	0.79	**5.05**	4.56	4.64	1.66	1.25	1.38
1998:01-2002:12	1.12	5.18	3.62	0.49	1.35	**9.36**	8.77	8.53
2007:07-2010:12	-3.57	6.72	1.79	1.57	1.06	**12.31**	11.21	8.31
(b) Absolute forecast errors								
1982:01-2010:12	0.70	2.23	1.59	1.67	1.40	2.79	**3.12**	2.59
1982:01-2002:12	2.45	2.30	**3.39**	1.76	2.79	2.83	2.53	2.92
1982:01-1997:12	3.61	1.87	**4.20**	4.07	4.17	2.31	1.79	2.39
1998:01-2002:12	-0.50	3.42	1.33	-0.58	-0.26	4.16	**4.43**	4.25
2007:07-2010:12	2.06	6.55	6.15	5.73	4.57	10.99	**11.35**	8.00

Entries in this table are OS statistics scaled by 100. The bold values indicate the best OS statistic among models within a given sample period. "Ignoring dependence" means that the decomposition model is estimated but predictions are constructed using (3.2.16) under the presumption of conditional independence between signs and absolute returns. "Exploiting dependence" means that the decomposition model is estimated and fully used in constructing predictions by (3.2.14) through numerical simulation. Averages on all proposed dynamic models are also reported.

Table 3.5: Tests of Conditional Predictive Ability

	Constant Decomposition Models		Dynamic Decomposition Models					
	Ignoring	Exploiting	Ignoring Dependence			Exploiting Dependence		
	Dependence	Dependence	IGtype	OSA	Average	IGtype	OSA	Average
(a) Squared forecast errors								
Historical Average	0.357	0.114	0.215	0.199	0.244	0.094	0.064	0.099
	[54.9%]	[89.9%]	[75.0%]	[75.3%]	[73.9%]	[91.7%]	[89.7%]	[90.7%]
CDM-CI		0.181	0.129	0.060	0.198	0.083	0.008	0.041
		[78.2%]	[93.9%]	[95.1%]	[90.7%]	[82.2%]	[82.2%]	[82.7%]
CDM-CD						0.242	0.492	0.211
						[82.5%]	[96.3%]	[71.8%]
(b) Absolute forecast errors								
Historical Average	0.814	0.097	0.660	0.781	0.793	0.012	0.009	0.167
	[74.4%]	[77.8%]	[87.9%]	[92.8%]	[90.8%]	[74.4%]	[79.3%]	[80.1%]
CDM-CI		0.031	0.020	0.181	0.234	0.654	0.012	0.265
		[79.0%]	[82.5%]	[99.1%]	[78.6%]	[95.4%]	[81.3%]	[85.6%]
CDM-CD						0.020	0.176	0.201
						[81.5%]	[97.7%]	[78.7%]

Entries in this table are p-values of the test statistics of Giacomini and White (2006) for 1982:01-2010:12. Values in square brackets are relative performance which indicates the percentage of times that a competing model in the first row dominates the benchmark model in the first column. Averages on all proposed dynamic decomposition models are also reported for ignoring and exploiting dependence. CDM denotes constant decomposition model. CI and CD represent the forecast scenarios of ignoring and exploiting dependence, respectively.

Table 3.6: Sources of Forecasting Performance

	Constant Decomposition Models		Dynamic Decomposition Models					
	Ignoring	Exploiting	Ignoring Dependence			Exploiting Dependence		
	Dependence	Dependence	IGtype	OSA	Average	IGtype	OSA	Average
Components (1): Time Variation of Forecasting Performance								
(a) Squared forecast errors								
Historical Average	0.057	0.006	0.007	0.010	0.034	<0.001	<0.001	0.045
CDM-CI		0.008	0.003	<0.001	<0.001	0.007	<0.001	0.011
CDM-CD			0.005	<0.001	<0.001	0.025	<0.001	0.096
(b) Absolute forecast errors								
Historical Average	0.009	0.006	0.007	0.010	0.008	0.005	<0.001	0.039
CDM-CI		0.083	0.006	<0.001	0.040	0.009	0.006	0.064
CDM-CD			0.007	0.008	0.008	0.070	<0.001	0.108
Components (2): Predictive Content								
(a) Squared forecast errors								
Historical Average	0.003	0.003	0.010	0.001	0.005	0.001	0.005	0.016
CDM-CI		0.183	0.092	0.158	0.174	0.033	0.111	0.088
CDM-CD			0.214	0.183	0.126	0.026	0.345	0.136
(b) Absolute forecast errors								
Historical Average	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
CDM-CI		0.121	0.011	0.258	0.098	0.002	0.081	0.016
CDM-CD			0.179	0.474	0.192	0.004	0.466	0.108
Components (3): Over-Fitting								
(a) Squared forecast errors								
Historical Average	0.063	0.425	0.218	0.039	0.151	0.456	0.435	0.444
CDM-CI		0.319	0.147	0.154	0.153	0.173	0.287	0.259
CDM-CD			0.485	0.445	0.449	0.127	0.105	0.226
(b) Absolute forecast errors								
Historical Average	0.020	0.359	0.197	0.048	0.104	0.492	0.444	0.452
CDM-CI		0.214	0.036	0.065	0.098	0.103	0.110	0.132
CDM-CD			0.405	0.383	0.345	0.127	0.127	0.196

Entries in this table are p-values of the test statistics of Rossi and Sekhposyan (2011).

Table 3.7: Forecast Combination

	CDM-CI & CDM-CD	IGtype-CI & IGtype-CD	OSA-CI & OSA-CD	Average	IGtype-CI & OSA-CD	OSA-CI & IGtype-CD	IGtype-CD & OSA-CD	All-DDMs
(a) Squared Forecast Errors	1.56	3.63	3.82	2.98	3.92	4.53	6.54	3.96
	(0.043)	(0.003)	(0.006)	(0.031)	(<0.001)	(0.006)	(0.018)	(0.015)
(b) Absolute Forecast Errors	1.23	2.22	2.63	2.08	2.78	2.64	3.64	2.48
	(0.265)	(0.008)	(0.017)	(0.076)	(0.013)	(0.053)	(0.018)	(0.058)

This table reports the forecast combination results for the forecast period of 1982:01-2010:12. The first row represents the combined models. Average is the average on the pairwise forecast combinations between conditional independence and conditional dependence of dynamic decomposition models. All-DDMs is a forecast combination of all DDMs including both CI and CD forecasts, which includes a total of 16 DDM models. Entries are OS statistics scaled by 100. Values in parentheses are p-values for the predictive accuracy test of Giacomini and White (2006).

Table 3.8: Economic Values of Market Timing

		Constant Decomposition Models		Dynamic Decomposition Models					
	Buy	Ignoring	Exploiting	Ignoring Dependence			Exploiting Dependence		
	& Hold	Dependence	Dependence	IGtype	OSA	Average	IGtype	OSA	Average
1982:01-2010:12									
Average returns	11.85%	10.81%	11.80%	11.15%	10.73%	10.93%	12.16%	**12.48%**	11.96%
Std. Dev.	15.09%	13.89%	13.81%	**13.66%**	13.93%	13.77%	13.80%	13.87%	13.75%
Sharpe Ratio	0.473	0.440	0.513	0.472	0.436	0.453	0.540	**0.559**	0.527
1982:01-2002:12									
Average returns	11.95%	13.03%	12.49%	**13.59%**	13.21%	13.00%	13.09%	12.50%	12.79%
Std. Dev.	14.81%	13.65%	13.72%	**13.47%**	13.53%	13.55%	13.72%	13.77%	13.71%
Sharpe Ratio	0.421	0.535	0.492	**0.567**	0.556	0.531	0.537	0.501	0.522
1982:01-1997:12									
Average returns	15.55%	15.88%	14.73%	**16.33%**	15.69%	15.84%	15.12%	15.77%	15.14%
Std. Dev.	14.10%	13.21%	13.51%	**13.05%**	13.15%	13.11%	13.63%	13.83%	13.58%
Sharpe Ratio	0.658	0.727	0.626	**0.771**	0.716	0.731	0.650	0.688	0.655
1998:01-2002:12									
Average returns	0.45%	3.91%	5.31%	3.22%	2.32%	3.03%	6.60%	**6.62%**	6.15%
Std. Dev.	16.59%	14.77%	14.23%	14.46%	14.39%	14.64%	13.95%	**13.94%**	14.02%
Sharpe Ratio	-0.221	-0.013	0.085	-0.061	-0.124	-0.073	0.179	**0.181**	0.148
2007:07-2010:12									
Average returns	2.89%	4.76%	8.88%	7.03%	4.76%	5.57%	9.78%	**9.99%**	9.06%
Std. Dev.	21.91%	21.35%	20.10%	20.87%	21.34%	20.92%	20.05%	20.04%	**19.90%**
Sharpe Ratio	0.086	0.176	0.343	0.288	0.175	0.219	0.444	**0.446**	0.403

Bold values indicate the best annualized performance. Portfolio returns are computed from actual stock returns and risk-free rates. Sharpe ratios are computed as the average of portfolio returns minus risk free rate divided by its standard deviation.

Table 3.9: Economic Values of Volatility and Skewness Timings

	Constant Decomposition Models	Dynamic Decomposition Models		
		IGtype	OSA	Average
1982:01-2010:12				
Volatility timing ($\triangle$)	1.05%	**1.74%**	1.00%	1.27%
Skewness timing ($\triangle'$)	1.76%	1.99%	**2.31%**	1.95%
$\triangle' - \triangle$	0.71%	0.25%	**1.31%**	0.68%
1982:01-2002:12				
Volatility timing ($\triangle$)	3.12%	**3.63%**	2.67%	2.99%
Skewness timing ($\triangle'$)	2.15%	2.54%	**3.02%**	2.64%
$\triangle' - \triangle$	-0.97%	-1.09%	**0.35%**	-0.35%
1982:01-1997:12				
Volatility timing ($\triangle$)	2.09%	**2.85%**	1.92%	2.18%
Skewness timing ($\triangle'$)	0.31%	0.36%	**0.96%**	0.62%
$\triangle' - \triangle$	-1.78%	-2.49%	**-0.96%**	-1.56%
1998:01-2002:12				
Volatility timing ($\triangle$)	5.60%	**6.22%**	4.94%	5.48%
Skewness timing ($\triangle'$)	7.77%	9.21%	**9.31%**	8.82%
$\triangle' - \triangle$	2.17%	2.99%	**4.37%**	3.34%
2007:07-2010:12				
Volatility timing ($\triangle$)	2.49%	**5.25%**	2.49%	3.76%
Skewness timing ($\triangle'$)	7.21%	8.91%	**8.95%**	8.18%
$\triangle' - \triangle$	4.72%	3.66%	**6.46%**	4.42%

Volatility timing is evaluated by the trading strategy of ignoring dependence and skewness timing by the trading strategy of exploiting dependence. Bold values indicate the highest performance fees among trading strategies.

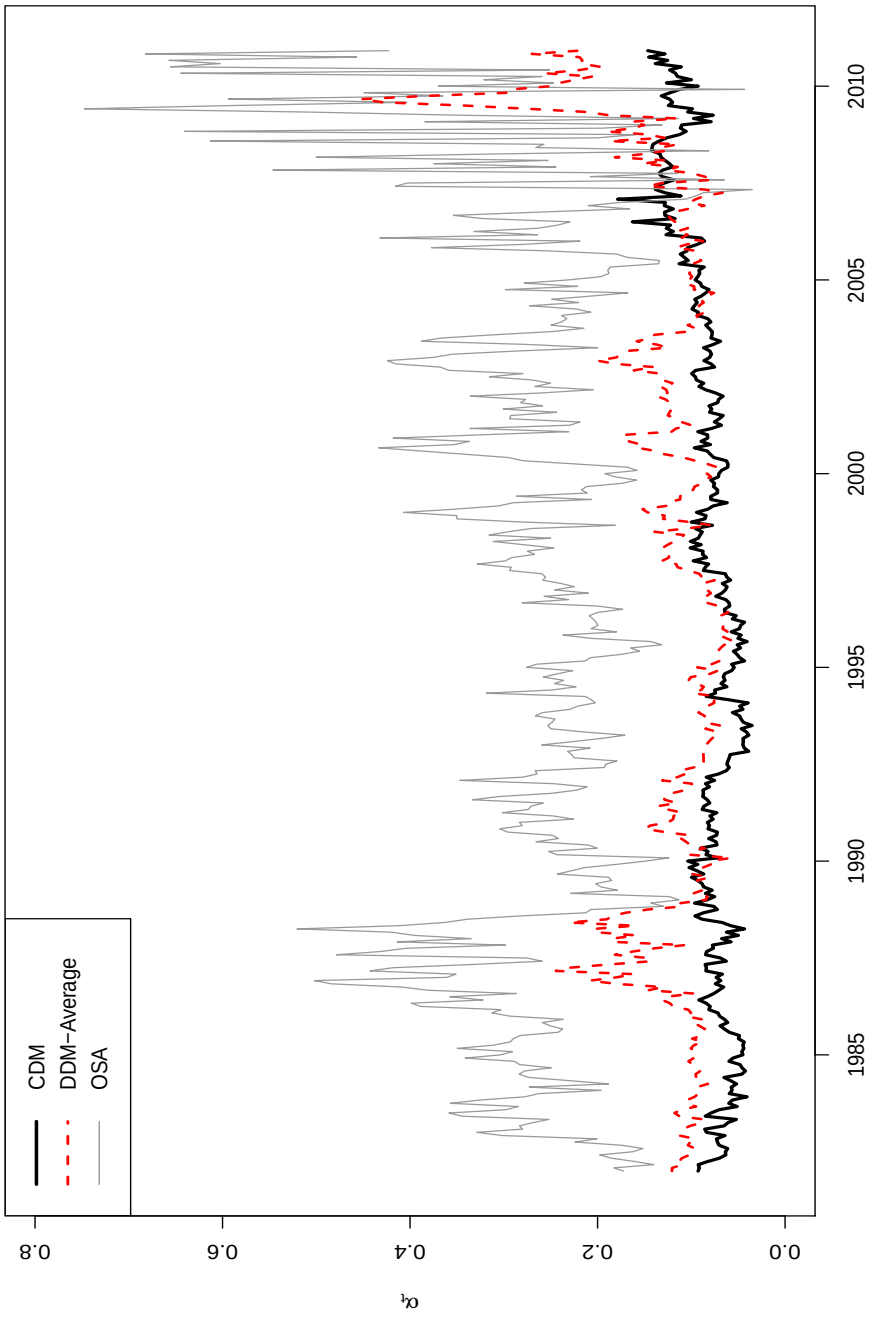


Figure 3.1: Out-of-Sample Estimators of Dependency Structures

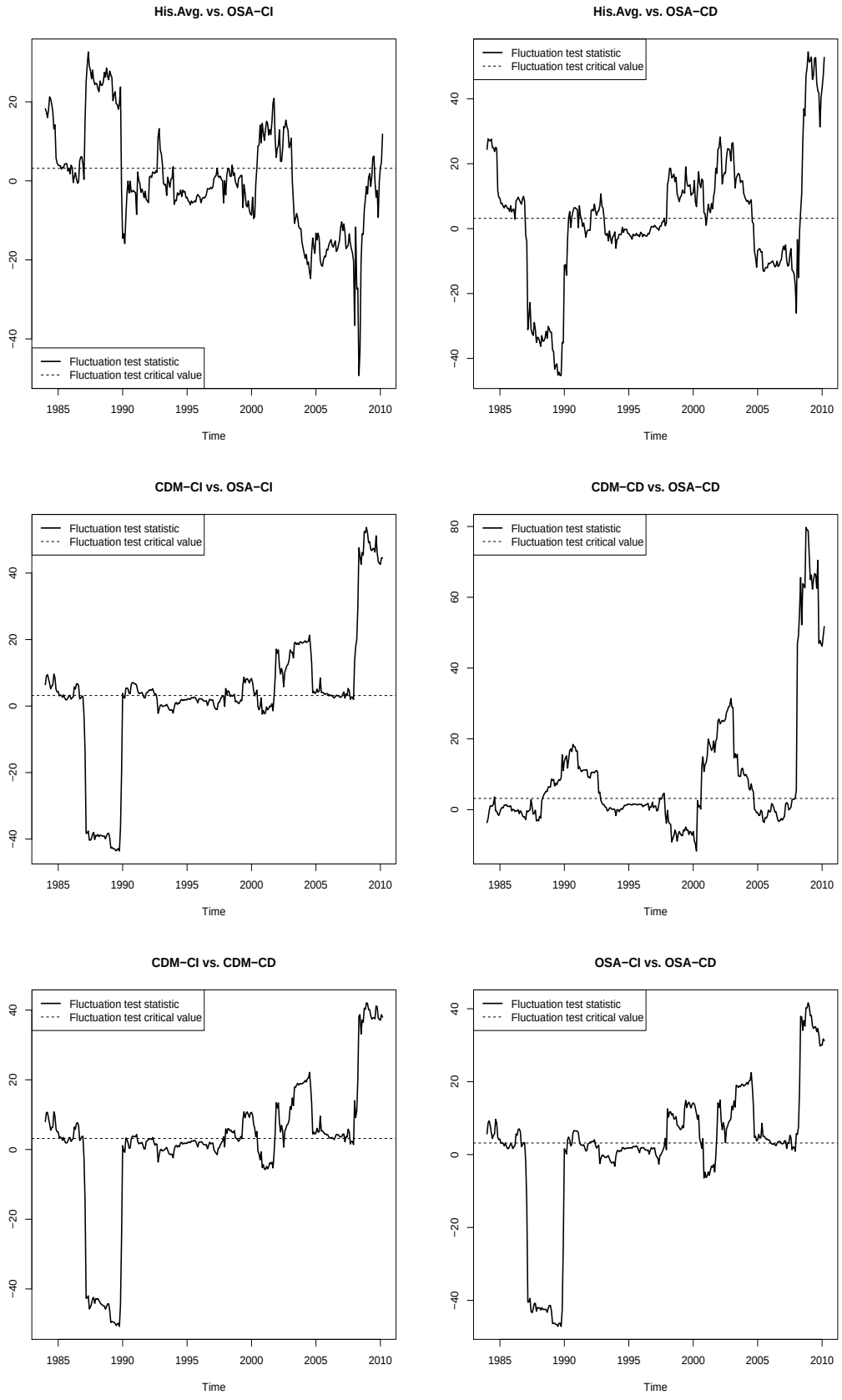


Figure 3.2: Fluctuation Tests

Bibliography

[1] Amisano, G., and R. Giacomini (2007) Comparing Density Forecasts via Weighted Likelihood Ratio Tests. Journal of Business & Economic Statistics 25(2): 177-190

[2] Anatolyev, S., and N. Gospodinov (2010) Modeling Financial Return Dynamics via Decomposition. Journal of Business and Economic Statistics 28(2): 232-245

[3] Busetti, F., and A. Harvey (2011) When is a Copula Constant? A Test for Changing Relationships. Journal of Financial Econometrics 9(1): 106-131

[4] Campbell, J. Y., and S. B. Thompson (2008) Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average? Review of Financial Studies 21: 1509–1531

[5] Cherubini, U., E. Luciano, and W. Vecchiato (2004) Copula Methods in Finance. The Wiley Finance Series, Wiley Press

[6] Christoffersen, P. F., and F. X. Diebold (2006) Financial Asset Returns, Direction-of-Change Forecasting, and Volatility Dynamics. Management Science 52(8): 1273–1288

[7] Christoffersen, P.F., F.X. Diebold, R.S. Mariano, A. S. Tay and Y.K. Tse (2007) Direction-of-change forecasts based on conditional variance, skewness and kurtosis dynamics: international evidence. Journal of Financial Forecasting 1(2): 1-22

[8] Clark, T.E., and K.D. West (2006) Using out-of-sample mean squared prediction errors to test the martingale difference hypothesis. Journal of Econometrics 135 (1–2): 155–186

[9] Diebold, F.X., and R.S. Mariano (1995) Comparing predictive accuracy. Journal of Business and Economic Statistics 13 (3): 253–263

[10] Diks, C., V. Panchenko and D. van Dijk (2010) Out-of-sample comparison of copula spec-

ifications in multivariate density forecasts. Journal of Economic Dynamics & Control 34: 1596-1609

[11] Engle, R.F. (2009) Anticipating Correlations: A New Paradigm for Risk Management. Princeton University Press, Princeton and Oxford

[12] Fleming, J., C. Kirby and B. Ostdiek (2001) The Economic Value of Volatility Timing. The Journal of Finance 56(1): 329-352

[13] Genest, C., and B. Rémillard (2004) Tests of independence and randomness based on the empirical copula process. Test 13: 335–369

[14] Genest, C., and B. Rémillard (2006) Local efficiency of a Cramér-von Mises test of independence. Journal of Multivariate Analysis 97(1): 274–294.

[15] Genest, C., and B. Rémillard (2007) Asymptotic local efficiency of Cramér-von Mises tests for multivariate independence. The Annals of Statistics 35(1): 166–191

[16] Genest, C. , B. Remillard and D. Beaudoin (2009) Goodness-of-fit tests for copulas: A review and a power study. Insurance: Mathematics and Economics 44: 199-214.

[17] Giacomini, R. (2011) Testing conditional predictive ability. In Oxford Handbook of Economic Forecasting. ed. M. Clements and D. Hendry, Oxford University Press.

[18] Giacomini, R. and B. Rossi (2010) Forecast comparisons in unstable environments. Journal of Applied Econometrics 25: 595-620

[19] Giacomini, R., and H. White (2006) Tests of Conditional Predictive Ability. Econometrica 74: 1545–1578

[20] Glosten, R.T., R. Jagannathan, and D. Runkle (1993) On the relation between the expected value and the volatility of the nominal excess return on stocks. Journal of Finance 48(5): 1779-1801

[21] Guo, H. (2006) On the out-of-sample predictability of stock market returns. The Journal of Business 79(2): 645-670

[22] Harvey, C.R. and A. Siddique (1999) Autoregressive Conditional Skewness. Journal of Financial and Quantitative Analysis 34(4): 465-487

[23] Harvey, C.R. and A. Siddique (2000a) Time-varying conditional skewness and the market risk premium. Research in Banking and Finance 1: 25-58

[24] Harvey, C.R. and A. Siddique (2000b) Conditional Skewness in Asset Pricing Tests. The Journal of Finance 55(3): 1263-1295

[25] Jondeau, E., and M. Rockinger (2003) Conditional volatility, skewness, and kurtosis: existence, persistence, and comovement. Journal of Economic Dynamics & Control 27: 1699-1737

[26] Jondeau, E., and M. Rockinger (2006) The copula-garch model of conditional dependencies: An international stock market application. Journal of International Money and Finance 25: 827-853

[27] Jondeau, E., and M. Rockinger (2012) On the importance of time variability in higher moments for asset allocation. Journal of Financial Econometrics 10(1): 84-123

[28] Leon, A., G. Rubio and G. Serna (2005) Autoregressive conditional volatility, Skewness and Kurtosis. The Quarterly Review of Economics and Finance 45: 599-618

[29] Manner, H., and O. Reznikova (2012) A survey on time-varying copulas: specification, simulations, and application. Econometric Reviews 31: 654-687

[30] Nelson, D. B. (1990) Stationarity and Persistence in the GARCH(1,1) Model. Econometric Theory 6(3): 318-334

[31] Nelson, D. B. (1991) Conditional Heteroscedasticity in Asset Returns: A New Approach. Econometrica 59(2): 347-370

[32] Nelsen, R.B. (2006) An introduction to Copulas. Springer Series in Statistics, Springer Press

[33] Patton, A. (2006) Modelling asymmetric exchange rate dependence. International Economic Review 47: 527-556

[34] Patton, A. (2012) A review of copula models for economic time series. Journal of Multivariate Analysis 110: 4-18

[35] Randles, R. and D. Wolfe (1979) Introduction to the Theory of Nonparametric Statistics. Wiley, New York

[36] Rapach, D.E., J.K. Strauss, and G. Zhou (2010) Out-of-Sample Equity Premium Prediction: Combination Forecasts and Links to the Real Economy. The Review of Financial Studies 223(2):821-862

[37] Rossi, B., and T. Sekhposyan (2011) Understanding model's forecasting performance. Journal of Econometrics 164(1): 158-172

[38] Shiller, R. J. (2005) Irrational Exuberance. 2nd edition, Princeton University Press

[39] Stock, J.H. and M.W. Watson (2003) Forecasting output and inflation: the role of asset prices. Journal of Economic Literature 41: 788-829

[40] Trivedi, P. and D. Zimmer (2005) Copula modeling: an introduction for practitioners. Foundations and Trends in Econometrics 1: 1-111

[41] Tse, Y.K., and A.K.C. Tsui (2002) A Multivariate Generalized Autoregressive Conditional Heteroscedasticity Model with Time-Varying Correlations. Journal of Business & Economic Statistics 20(3): 351-362

[42] Zakoian, J. (1994) Threshold heteroskedastic models. Journal of Economic Dynamics and Control 18: 931-955

Appendix A

Marginals

Following AG's work, the dynamics for the positively valued variable $|r_t|$ are specified in a multiplicative error model as

$$|r_t| = \psi_t \eta_t \tag{A.0.1}$$

where $\psi_t = E_{t-1}|r_t|$ and η_t is a positive multiplicative error with $E(\eta_t) = 1$. A logarithmic conditional autoregressive model is used for ψ_t as

$$log\psi_t = \omega_v + \beta_v log\psi_{t-1} + \gamma_v log|r_{t-1}| + \rho_v \mathbb{I}(r_{t-1} > 0) + x'_{t-1}\delta_v \tag{A.0.2}$$

where x_t are economic predictors. The persistence of the process is governed by the parameter $|\beta_v + \gamma_v| < 1$.

Assume that η_t follows a Weibull distribution with the cumulative distribution function

$$F_{\eta_t}(x;\, \kappa, \varsigma) = 1 - e^{-\left(\frac{x}{\varsigma}\right)^{\kappa}}$$

and the probability density function

$$f_{\eta_t}(x;\, \kappa,\, \varsigma) = \begin{cases} \frac{\kappa}{\varsigma}\left(\frac{x}{\varsigma}\right)^{\kappa-1} e^{-\left(\frac{x}{\varsigma}\right)^{\kappa}} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

where $\kappa, \varsigma > 0$ are shape and scale parameters, respectively. Weibull distributions have the mean of $E(\eta_t) = \varsigma\Gamma\left(1 + 1/\kappa\right)$, such that the restriction $E(\eta_t) = 1$ implies $\varsigma = 1/\Gamma\left(1 + 1/\kappa\right)$. $\Gamma(\cdot)$ is a gamma function. Hence, in this case the Weibull distribution can be reparameterized in terms of κ. Based on the Jacobian transformation matrix $|J| = \psi_t^{-1}$ from (A.0.1),

$|r_t|$ has the CDF and PDF functions respectively as

$$F_{|r_t|}(u|I_{t-1}) = F_{\eta_t}(u/\psi_t|I_{t-1}) \tag{A.0.3}$$

and

$$f_{|r_t|}(u|I_{t-1}) = \psi_t^{-1} f_{\eta_t}(u/\psi_t|I_{t-1}) \tag{A.0.4}$$

Marginal distribution for the sign dynamics ($s_t = I(r_t > 0)$) is a Bernoulli distribution with the marginal CDF and PDF functions respectively as

$$F_{s_t}(v|I_{t-1}) = 1 - p_t(1 - v)$$

and

$$f_{s_t}(v|I_{t-1}) = p_t^v \left(1 - p_t\right)^{1-v}$$

with $p_t = E_{t-1}s_t$ modeled by a dynamic logit-linked model

$$p_t = \frac{exp(\theta_t)}{1 + exp(\theta_t)} \tag{A.0.5}$$

and

$$\theta_t = \omega_d + \phi_d \mathbb{I}(r_{t-1} > 0) + y_{t-1}^{'}\delta_d \tag{A.0.6}$$

where y_t includes the set of predictors such as macroeconomic variables.

Note that the predictive variables of x_t and y_t are not necessarily identical so as to allow the dynamics of absolute returns and signs driven by different information sets. As discussed in section 2, this chapter uses the same marginals and information variables as defined in AG's work and only introduces time-varying copula dependency parameters. Hence, the comparison to AG's constant decomposition models can be easily made. To preserve space, I would refer readers to AG's original work for detail data description and construction on x_t and y_t.