Journal of Applied Finance & Banking, vol. 6, no. 6, 2016, 123-156
ISSN: 1792-6580 (print version), 1792-6599 (online)
Scienpress Ltd, 2016

Stress Testing and a Comparison of Alternative Methodologies for Scenario Generation

Michael Jacobs, Jr., Ph.D., CFA[1]

Abstract

A critical question that banking supervisors are trying to answer is what is the amount of capital or liquidity resources required by an institution in order to support the risks taken in the course of business. The financial crises of the last several years have revealed that traditional approaches such as regulatory capital ratios to be inadequate, giving rise to supervisory stress testing as a primary tool. A critical input into this process are macroeconomic scenarios that are provided by the prudential supervisors to institutions for exercises such as the Federal Reserve's *Comprehensive Capital Analysis and Review* ("CCAR") program. Additionally, supervisors are requiring that banks develop their own macroeconomic scenarios. A common approach is to combine management judgment with a statistical model, such as a *Vector Autoregression* ("VAR"), to exploit the dependency structure between both macroeconomic drivers, as well between modeling segments. However, it is well-known that linear models such as VAR are unable to explain the phenomenon of fat-tailed distributions that deviate from normality, an empirical fact that has been well documented in the empirical finance literature. We propose a challenger approach, widely used in the academic literature, but not commonly employed in practice, the *Markov Switching VAR* ("MS-VAR") model. We empirically test these models using Federal Reserve Y-9 filing and macroeconomic data, gathered and released by the regulators for CCAR purposes, respectively. We find the MS-VAR model to be more conservative than the VAR model, and also to exhibit greater accuracy in model testing, as the latter model can better capture extreme events observed in history.

---

[1]Principal Director, Accenture Consulting, Finance and Risk Services Advisory / Models, Methodologies & Analytics, 1345 Avenue of the Americas, New York, N.Y., 10105, 917-324-2098.

Article Info: *Received* : August 3, 2016. *Revised* : August 31, 2016.
Published online : November 1, 2016

JEL classification numbers: C31, C53, E27, E47, E58, G01, G17, C54, G21, G28, G38
Keywords: Stress Testing, CCAR, DFAST, Credit Risk, Financial Crisis, Model Risk, Vector Autoregression, Markov Switching Model, Scenario Generation.

1 Introduction

In the aftermath of the financial crisis, regulators have utilized stress testing as a means to which to evaluate the soundness of financial institutions' risk management procedures ([1], [2]). The primary means of risk management, particularly in the field of credit risk ([3]), is through advanced mathematical, statistical and quantitative techniques and models, which leads to model risk. Model risk can be defined as the potential that a model does not sufficiently capture the risks it is used to assess, and the danger that it may underestimate potential risks in the future [4]. *Stress testing* ("ST") has been used by supervisors to assess the reliability of credit risk models, as can be seen in the revised Basel framework ([6], [7], [8], [9], [10, [11]) and the Federal Reserve's *Comprehensive Capital Analysis and Review* ("CCAR") program.

ST may be defined, in a general sense, as a form of deliberately intense or thorough testing used to determine the stability of a given system or entity. This involves testing beyond normal operational capacity, often to a breaking point, in order to observe the results. In the financial risk management context, this involves scrutinizing the viability of an institution in its response to various adverse configurations of macroeconomic and financial market events, which may include simulated financial crises. ST is closely related to the concept and practice of *scenario analysis* ("SC"), which in economics and finance is the attempt to forecast several possible scenarios for the economy (e.g. growth levels) or an attempt to forecast financial market returns (e.g., for bonds, stocks and cash) in each of those scenarios. This might involve sub-sets of each of the possibilities and even further seek to determine correlations and assign probabilities to the scenarios.

Current risk models consider both capital adequacy and liquidity concerns, which regulators use to assess the relative health of banks in adverse potential scenarios. The assessment process can be further segmented into a consideration of capital versus liquidity resources, corresponding to right and left sides of the balance sheet (i.e., net worth versus the share of "liquid" assets), respectively. In the best case scenario, not only do supervisory and bank models result in similar outputs, but also both do not produce outputs that far exceed the regulatory floor.

Prior to the-financial crisis, most of the most prominent financial institutions to fail (e.g., Lehman, Bear Stearns, Washington Mutual, Freddie Mac and Fannie Mae) were considered to be well-capitalized according to the standards across a wide span of regulators Another commonality among the large failed firms included a general exposure to residential real estate, either directly or through securitization. Further, it is widely believed that the internal risk models of these institutions were not wildly out of line with those of the regulators [13]. We learned through these unanticipated failures that the answer to the question of how much capital an institution needs to avoid failure was not satisfactory. While capital models accept a non-zero probability of default according to the risk aversion of the institution or the supervisor, the utter failure of these constructs to even come close to projecting the perils that these institutions faced was a great motivator for considering alternative tools to assess capital adequacy, such as the ST discipline.

Bank Holding Companies (BHCs) face a number of considerations in modeling losses for wholesale and retail lending portfolios. CCAR participants face some particular challenges in estimating losses based on scenarios and their associated risk drivers. The selection of modeling methodology must satisfy a number of criteria, such as suitability for portfolio type, materiality, data availability as well as alignment with chosen risk drivers.

This paper shall proceed as follows. Section 2 reviews the available literature on ST and scenario generation. Section 3 presents the competing econometric methodologies for generating scenarios, a time series *Vector Autoregressive* ("VAR") and *Markov Switching VAR* ("MS-VAR") models. Section 4 presents the empirical implementation, the data description, a discussion of the estimation results and their implications. Section 5 concludes the study and provides directions for future avenues of research.

2 Review of the Literature

Regulators first introduced ST within the Basel I According, with the 1995 Market Risk Amendment (Basel Committee on Banking Supervision 1988, 1996). Around the same time, the publication of RiskMetrics™ in 1994 (J.P. Morgan, 1994) marked risk management as a separate technical discipline, and therein all of the above mentioned types of ST are referenced. The seminal handbook on *Value-at-Risk* ("VaR"), also had a part devoted to the topic of ST (Jorion, 1996), while other authors (Kupiec (1999), Berkowitz and Jeremy(1999)) provided detailed discussions of VaR-based stress tests as found largely in the trading and treasury functions. The *Committee on Global Financial Systems* ("CGFS") conducted a survey on stress testing in 2000 that had similar findings [20]. Another study highlighted that the majority of the stress testing exercises performed to date were shocks to market observables based upon historical events, which have the advantage of being well-defined and easy to understand, especially when dealing with the trading book constituted of marketable asset classes [21].

However, in the case of the banking book (e.g., corporate / C&I or consumer loans), this approach of asset class shocks does not carry over as well, as to the extent these are less marketable there are more idiosyncracies to account for. Therefore, stress testing with respect to credit risk has evolved later and as a separate discipline in the domain of credit portfolio modeling. However, even in the seminal examples of CreditMetrics™ ([22]) and CreditRisk+™ ([23]), stress testing was not a component of such models. The commonality of all such credit portfolio models was subsequently demonstrated ([24]), as well as the correspondence between the state of the economy and the credit loss distribution, and therefore that this framework is naturally amenable to stress testing. In this spirit, a class of models was built upon the CreditMetrics™ ([22]) framework through macroeconomic stress testing on credit portfolios using credit migration matrices ([25]).

ST supervisory requirements with respect to the banking book were rather undeveloped prior to the crisis, although it was rather prescriptive in other domains, examples including the joint policy statement on interest rate risk ([26]), guidance on counterparty credit risk ([27]), as well as country risk management ([28]).

Following the financial crisis of the last decade, we find an expansion in the literature on stress testing, starting with a survey of the then extant literature on stress testing for credit risk ([29]). As part of a field of literature addressing various modeling approaches to stress testing, we find various papers addressing alternative issues in stress testing and stressed capital, including the aggregation of risk types of capital models ([30]) and also with respect to validation of these models (Jacobs, Jr., 2010). Various papers have laid out the reasons why ST has become such a dominant tool for regulators, including rationales for its utility, outlines for its execution, as well as guidelines and opinions on disseminating the output under various conditions ([13]). This includes a survey of practices and supervisory expectations for stress tests in a credit risk framework, and presentation of simple examples of a ratings migration based approach, using the CreditMetrics™ framework ([32]). Another set of papers argues for a Bayesian approach to stress testing, having the capability to cohesively incorporate expert knowledge model design, proposing a methodology for coherently incorporating expert opinion into the stress test modeling process [33]. Finally, yet another recent study features the application of a Bayesian regression model for credit loss implemented using Fed Y9 data, wherein regulated financial institutions report their stress test losses in conjunction with Federal Reserve scenarios, which can formally incorporate exogenous factors such as such supervisory scenarios, and also quantify the uncertainty in model output that results from stochastic model inputs [34].

One of the previously mentioned stress test surveys highlights the 2009 U.S. stress testing exercise, the *Supervisory Capital Assessment Program* ("SCAP") as an informative model (Schuermann, 2014). In that period there was incredible concern amongst investors over the viability of the U.S. financial system, given the looming and credible threat of massive equity dilution stemming from government action, such as bailouts mandated by regulators. The concept underlying the application of a macro-prudential stress test was that a bright line, delineating failure or survival under a credibly severe systematic scenario, would convince investors that failure of one or more financial institutions was unlikely, thus making the likelihood of capital injections remote. The SCAP exercise covered 19 banks in the U.S., having book value of assets greater than $100 billion (comprising approximately two-thirds the total in the system) as of the year-end 2008. The SCAP resulted in 10 of those banks having to raise a total of $75 billion in capital ($77 billion in Tier 1 common equity) in a six month period.

On study notes that CCAR was initially planned in 2010 and rolled out in 2011. It initially covered the 19 banks covered under SCAP, but as they document, a rule in November 2011 required all banks above $50 billion in assets to adhere to the CCAR regime [35]. The CCAR regime includes *Dodd-Frank Act Stress Tests* ("DFAST"), with the sole difference between CCAR and DFAST being that DFAST uses a homogenous set of capital actions on the part of the banks, while CCAR takes banks' planning distribution of capital into account when calculating capital ratios. The authors further document that the total increase in capital in this exercise, as measured by Tier 1 common equity, was about $400 Billion. Finally, the authors highlight that ST is a regime that allows regulators to not only set a quantitative hurdle for capital that banks must reach, but also to make qualitative assessments of key inputs into the stress test process, such as data integrity, governance, and reliability of the models.

The outcome of the SCAP was rather different from the *Committee of European Bank Supervisors* ("CEBS") stress tests conducted in 2010 and 2011, which coincided with the sovereign debt crisis that hit the periphery of the Euro-zone. In 2010, the ECBS stressed a total of 91 banks, as with the SCAP covering about two-thirds of assets and one-half of banks per participating juris-diction. There are several differences between the CEBS stress tests and SCAP worth noting. First, the CEBS exercise stressed the values of sovereign bonds held in trading books, but neglected to address that banking books where in fact the majority of the exposures in sovereign bonds were present, resulting in a mild requirement of just under $5B in additional capital. Second, in contrast to the SCAP, the CEBS stress testing level of disclosure was far less granular, with loss rates reported for only two broad segments (retail vs. corporate) as opposed to major asset classes (e.g., first-lien mortgages, credit cards, commercial real estate, etc.) The 2011 *European Banker's Association* ("EBA") exercise, covering 90 institutions in 21 jurisdictions, bore many similarities to the 2011 EBA tests, with only 8 banks required to raise about as much capital in dollar terms as the previous exercise. However, a key difference was the more granular disclosure requirements, such as a breakdowns of loss rates by not only major asset class but also by geography, as well availability of results to the public in a user-friendly form that admit-ted the application of analysts' assumptions. Similarly to the 2010 CEBS exercise, in which the CEBS test did not ameliorate nervousness about the Irish banks, the 2011 EBA version failed to ease concerns about the Spanish banking system, as while 5 of 25 passed there was no additional capital required [35].

The available public literature on scenario generation is rather limited to date. A San Francisco Federal Reserve Bank study argues that while in recent years ST has become an important component of financial and macroprudential regulation, nevertheless the techniques of stress testing are still being honed and debated [36]. The authors claim to contribute to the debate in proposing the use of robust forecasting analysis to identify and construct adverse scenarios that are naturally interpretable as stress tests. Their scenarios emerge from a particular pessimistic twist to a benchmark forecasting model, referred to as a “worst case distribution”, which they argue offers regulators a method of identifying vulnerabilities, while at the same time acknowledging that their models are mis-specified in possibly unknown ways. An Atlanta Federal Reserve Bank study presents a case study of a failed U.S. experience in tying stress test results to capital requirements was a spectacular failure due to issues associated with the specification of stress scenarios, namely the *Office of Federal Housing Enterprise Oversight's* (“OFHEO”) risk-based capital stress test for Fannie Mae and Freddie Mac [37]. The authors study a key component of OFHEOs model, the 30-year fixed-rate mortgage performance, and identify two key problems. They point out that OFHEO had left the model specification and associated parameters static for the entire time the rule was in force, and furthermore that the house price stress scenario was insufficiently dire, resulting in a significant underprediction of mortgage credit losses and associated capital needs at Fannie Mae and Freddie Mac during the housing bust.

3 Time Series VAR Methodologies for Scenario Generation

In macroeconomic forecasting, there are 4 basic tasks that we set out to do: characterize macroeconomic time series, conduct forecasts of macroeconomic or related data, make inferences about the structure of the economy, and finally advise policy-makers [38]. In the ST application, we are mainly concerned with the forecasting and policy advisory functions, as stressed loss projections help banker risk manager and bank supervisors make decisions about the potential viability of their institutions during periods of extreme economic turmoil. Going back a few decades, these functions were accomplished by a variety of means, ranging from large-scale models featuring the interactions of many variables, to simple univariate relationships motivated by stylized and parsimonious theories (e.g., Okun’s Law or the Phillips Curve). However, following the economic crises of the 1970s, most established economic relationships started to break down and these methods proved themselves to be unreliable. In the early 1980s, a new macroeconometric paradigm started to take hold, VAR, a simple yet flexible way to model and forecast macroeconomic relationships [39]. In contrast to the univariate autoregressive model ([40], [41], [42]), a VAR model is a multi-equation linear model in which variables can be explained by their own lags, as well as lags of other variables. As in the CCAR / ST application we are interested in modeling the relationship and forecasting multiple macroeconomic variables, the VAR methodology is rather suitable to this end.

Let $\mathbf{Y}_t = \left(Y_{1t},...,Y_{kt}\right)^T$ be a k-dimensional vector valued time series, the output variables of interest, in our application with the entries representing some loss measure in a particular segment, that may be influenced by a set of observable *input variables* denoted by $\mathbf{X}_t = \left(X_{1t},...,X_{rt}\right)^T$, an r-dimensional vector valued time series also referred as *exogenous variables*, and in our context representing a set of macroeconomic factors. We say that that $\mathbf{Y}_t$ follows a *multiple transfer function process* if it can be written in the following form:

$$\mathbf{Y}_t = \sum_{j=0}^{\infty} \mathbf{\Psi}^*_j \mathbf{X}_{t-j} + \mathbf{N} \tag{1}$$

Where $\mathbf{\Psi}^*_j$ are a sequence of $k \times r$ dimensional matrices and $\mathbf{N}_t$ is a k-dimensional vector of noise terms which follow an stationary *vector autoregressive-moving average process*, denoted by $VARMA\left(p,q,s\right)$:

$$\mathbf{\Phi}\left(B\right)\mathbf{N}_t = \mathbf{\Theta}\left(B\right)\boldsymbol{\varepsilon}_t \tag{2}$$

Where $\mathbf{\Phi}\left(B\right) = \mathbf{I}_r - \sum_{j=1}^{p} \mathbf{\Phi}_j B^j$ and $\mathbf{\Theta}\left(B\right) = \mathbf{I}_r - \sum_{j=1}^{q} \mathbf{\Theta}_j B^j$ are *autoregressive lag polynomials* of respective orders p and q, and B is the *back-shift operator* that satisfies $B^i X_t = X_{t-i}$ for any process $\left\{X_t\right\}$**.** It is common to assume that the input process $\left\{\mathbf{X}_t\right\}$ is generated independently of the noise process $\left\{\mathbf{N}_t\right\}$. In fact, the exogenous variables $\left\{\mathbf{X}_t\right\}$ can represent both stochastic and non-stochastic (deterministic) variables, examples being sinusoidal seasonal (periodic) functions of time, used to represent the seasonal fluctuations in the output process $\left\{\mathbf{Y}_t\right\}$, or intervention analysis modeling in which a simple step (or pulse indicator) function taking the values of 0 or 1 to indicate the effect of output due to unusual intervention events in the system.

Now let us assume that the transfer function operator can be represented by a *rational factorization* of the form $\mathbf{\Psi}^*\left(B\right) = \sum_{j=0}^{\infty} \mathbf{\Psi}^*_j B^j = \mathbf{\Phi}^{-1}\left(B\right)\mathbf{\Theta}^*\left(B\right)$, where $\mathbf{\Theta}^*\left(B\right) = \sum_{j=0}^{s} \mathbf{\Theta}^*_j B^j$ is of order s and $\mathbf{\Theta}^*_j$ are $k \times r$ matrices. This implies that $\mathbf{Y}_t$ follows a *vector autoregressive-moving average process with exogenous variables*, denoted by $VARMAX\left(p,q,s\right)$, which is motivated by assuming that the transfer function operator in (3.1) can be represented as a rational factorization of the form:

$$\mathbf{\Psi}^*(B) = \sum_{j=0}^{s} \mathbf{\Psi}_j^* B^j = \mathbf{\Phi}^{-1}(B)\mathbf{\Theta}^*(B) = \left(\mathbf{I}_r - \sum_{j=1}^{p} \mathbf{\Phi}_j B^j\right)^{-1} \left(\sum_{j=1}^{s} \mathbf{\Theta}_j^* B^j\right) \quad (3)$$

Where $\mathbf{\Theta}^*(B)$ is of order s and $\mathbf{\Theta}_j^* \in R^{k\times r}$ are $k\times r$ matrices. Without loss of generality, we assume that the factor $\mathbf{\Phi}(B) = \mathbf{I}_r - \sum_{j=1}^{p} \mathbf{\Phi}_j B^j$ is the same as the AR factor in the model for the noise process $\mathbf{N}_t$. This gives rise to the *VARMAX* (p,q,s) representation, where X stands for the sequence of exogenous (or input) vectors:

$$\mathbf{Y}_t - \sum_{j=1}^{p} \mathbf{\Phi}_j \mathbf{Y}_{t-j} = \sum_{j=1}^{s} \mathbf{\Theta}_j \mathbf{X}_{t-j} + \varepsilon_t - \sum_{j=1}^{q} \mathbf{\Theta}_j^* \varepsilon_{t-j} \quad (4)$$

Note that the VARMAX model (4) could be written in various equivalent forms, involving a lower triangular coefficient matrix for $\mathbf{Y}_t$ at lag zero, or a leading coefficient matrix for ε_t at lag zero, or even a more general form that contains a leading (non-singular) coefficient matrix for $\mathbf{Y}_t$ at lag zero that reflects instantaneous links amongst the output variables that are motivated by theoretical considerations (provided that the proper identifiability conditions are satisfied ([43], [44]). In the econometrics setting, such a model form is usually referred to as a *dynamic simultaneous equations model* or a *dynamic structural equation model*. The related model in the form of equation (4), obtained by multiplying the dynamic simultaneous equations model form by the inverse of the lag 0 coefficient matrix, is referred to as the *reduced form model*. In addition, (4) has the state space representation of the form [45]:

$$\begin{aligned} \mathbf{Z}_t &= \mathbf{\Phi}\mathbf{Z}_{t-1} + \mathbf{B}\mathbf{X}_{t-1} + \mathbf{a}_t \\ \mathbf{Y}_t &= \mathbf{H}\mathbf{Z}_t + \mathbf{F}\mathbf{X}_t + \mathbf{N}_t \end{aligned} \quad (5)$$

The ARMAX model (3)-(4) is said to be stable if the roots of $\det\{\mathbf{\Phi}(B)\} = 0$ all possess an absolute value greater than unity. In that case, if both the input $\{\mathbf{X}_t\}$ and the noise $\{\mathbf{N}_t\}$ processes are stationary, then so is the output process $\{\mathbf{Y}_t\}$ having the following convergent representation:

$$\mathbf{Y}_t = \sum_{j=0}^{\infty} \mathbf{\Psi}_j^* \mathbf{X}_{t-j} + \sum_{j=0}^{\infty} \mathbf{\Psi}_j \boldsymbol{\varepsilon}_{t-j} \quad (6)$$

Where $\mathbf{\Psi}(B) = \sum_{i=0}^{\infty} \mathbf{\Psi}_i B^i = \mathbf{\Phi}^{-1}(B)\mathbf{\Theta}(B)$ and $\mathbf{\Psi}^*(B) = \sum_{j=0}^{\infty} \mathbf{\Psi}_j^* B^j = \mathbf{\Phi}^{-1}(B)\mathbf{\Theta}^*(B)$.

The transition matrices $\mathbf{\Psi}_j^*$ of the transfer function $\mathbf{\Psi}^*(B)$ represent the partial effects that changes in the exogenous (or input variables; macroeconomic variables or scenarios in our application) variables have on the output variables $\mathbf{Y}_t$ at various time lags, and are sometimes called *response matrices*. The *long-run effects* or *total gains* of the dynamic system (3.6) is given by the elements of the matrix:

$$\mathbf{G} = \mathbf{\Psi}^*(1) = \sum_{j=0}^{\infty} \mathbf{\Psi}_j^* \tag{7}$$

And the entry $\mathbf{G}^{i,j}$ represents the long-run (or equilibrium) change in the i$^{\text{th}}$ output variable that occurs when a unit change in the j$^{\text{th}}$ exogenous variable occurs and is held fixed at some starting point in time, with all other exogenous variables held constant. In econometric terms, the elements of the matrices $\mathbf{\Psi}_j^*$ are referred to as *dynamic multipliers at lag* j, and the elements of $\mathbf{G}$ are referred to as *total multipliers*. In this study we consider a *vector autoregressive model with exogenous variables* (“VARX”), denoted by $VARX(p,s)$, which restricts the *Moving Averag*e (“MA”) terms beyond lag zero to be zero, or $\mathbf{\Theta}_j^* = \mathbf{0}_{k\times k} \quad j>0$:

$$\mathbf{Y}_t - \sum_{j=1}^{p} \mathbf{\Phi}_j \mathbf{Y}_{t-j} = \sum_{j=1}^{s} \mathbf{\Theta}_j \mathbf{X}_{t-j} + \varepsilon_t \tag{8}$$

The rationale for this restriction is three-fold. First, in MA terms were in no cases significant in the model estimations, so that the data simply does not support a VARMA representation. Second, the VARX model avails us of the very convenient DSE package in R, which has computational and analytical advantages [46]. Finally, the VARX framework is more practical and intuitive than the more elaborate VARMAX model, and allows for superior communication of results to practitioners.

We now consider the MS-VAR (or more generally MS-VARMAX) generalization of the VAR (or more generally (ARMAX) methodology with changes in regime, where the parameters of the VARMAX system $\mathbf{B} \square \left(\mathbf{\Phi}^T, \mathbf{\Theta}^T, \mathbf{\Theta}^{*T}\right)^T \in R^{p+q+s}$ will be time-varying. However, the process might be time-invariant conditional on an unobservable regime variable $s_t \in (1,...,M)$, denoting the state at time t out of M feasible states. In that case, then the conditional probability density of the observed time series $\mathbf{Y}_t$ is given by:

$$p\left(\mathbf{y}_t \middle| \mathbf{\Psi}_t, s_t\right) = \begin{cases} f\left(\mathbf{y}_t \middle| \mathbf{\Psi}_{t-1}, \mathbf{B}_1\right) & if\ s_t = 1 \\ \vdots & \\ f\left(\mathbf{y}_t \middle| \mathbf{\Psi}_{t-1}, \mathbf{B}_M\right) & if\ s_t = M, \end{cases} \tag{9}$$

Where $\mathbf{B}_m$ is the VAR parameter in regime $m \in (1,...,M)$ and $\mathbf{\Psi}_{t-1}$ are the observations $\left\{\mathbf{y}_{t-j}\right\}_{j=1}^{\infty}$. Therefore, given a regime s_t, the conditional $VARX\left(p,s\middle|s_t\right)$ system in expectation form can be written as:

$$E\left[\mathbf{y}_t\middle|\mathbf{\Psi}_{t-1},s_t\right]=\sum_{j=1}^{p}\mathbf{\Phi}_j\mathbf{Y}_{t-j}\left(s_t\right)+\sum_{j=1}^{s}\mathbf{\Theta}_j\left(s_t\right)\mathbf{X}_{t-j} \tag{10}$$

We define the innovation term as:

$$\varepsilon_t=\mathbf{y}_t-E\left[\mathbf{y}_t\middle|\mathbf{\Psi}_t,s_t\right] \tag{11}$$

The innovation process ε_t is a Gaussian, zero-mean white noise process having variance-covariance matrix $\mathbf{\Sigma}\left(s_t\right)$:

$$\varepsilon_t\sim NID\left(\mathbf{0},\mathbf{\Sigma}\left(s_t\right)\right) \tag{12}$$

If the $VARX\left(p,s\middle|s_t\right)$ process is defined conditionally upon an unobservable regime s_t as in equation (9), the description of the process generating mechanism should be made complete by specifying the stochastic assumption of the MS-VAR model. In this construct, the s_t follows a discrete state homogenous Markov chain:

$$\Pr\left(s_t\middle|\left\{\mathbf{y}_{t-j}\right\}_{j=1}^{\infty},\left\{s_{t-j}\right\}_{j=1}^{\infty}\right)=\Pr\left(\left\{s_{t-j}\right\}_{j=1}^{\infty}\middle|\boldsymbol{\rho}\right) \tag{13}$$

Where $\boldsymbol{\rho}$ denotes the parameter vector of the regime generating process. We estimate the MS-VAR model using MSBVAR the package in R [46].

The MS-VAR paradigm is based for the most upon three schools of thought. The first of these traditions is the linear time-invariant VAR model, as introduced and discussed at the beginning of this section. This framework analyzes the relationships of random variables in a dynamic system, the dynamic propagation of innovations the system, and the effects of regime change. The second foundation is the statistics of *probabilistic functions of Markov chai*ns ([47], [48]). Furthermore, the MS-VAR model also encompasses the even older traditions of *mixtures of normal distributions* ([49]) and the hidden Markov-chain ([50], [51]). Finally, another root can be found in the construction of rather basic Markov-chain regression models in econometrics [52]. The first holistic approach to the statistical analysis to the statistical analysis of the Markov-switching model can be found in the literature ([53], [54]). Finally, the treatment of the MS-VAR model as a Gaussian autoregressive process conditioned on an exogenous regime generating process is closely related to the theory of a doubly stochastic processes [55].

4 Empirical Implementation

As part of the Federal Reserve's CCAR stress testing exercise, U.S. domiciled top-tier BHCs are required to submit comprehensive capital plans, including pro forma capital analyses, based on at least one BHC defined adverse scenario. The adverse scenario is described by quarterly trajectories for key *macroeconomic variables* (“MVs”) over the next nine quarters or for thirteen months to estimate loss allowances. In addition, the Federal Reserve generates its own supervisory stress scenarios, so that firms are expected to apply both BHC and supervisory stress scenarios to all exposures, in order to estimate potential losses under stressed operating conditions. Firms engaged in significant trading activities (e.g., Goldman Sachs or Morgan Stanley) are asked to estimate a one-time trading-related market and counterparty credit loss shock under their own BHC scenarios, and a market risk stress scenario provided by the supervisors. Large custodian banks are asked to estimate a potential default of their largest counterparty. In the case of the supervisory stress scenarios, the Federal Reserve provides firms with global market shock components that are one-time, hypothetical shocks to a large set of risk factors. During the last two CCAR exercises, these shocks involved large and sudden changes in asset prices, rates, and CDS spreads that mirrored the severe market conditions in the second half of 2008.

Since CCAR is a comprehensive assessment of a firm's capital plan, the BHCs are asked to conduct an assessment of the expected uses and sources of capital over a planning horizon. In the 2009 SCAP, firms were asked to submit stress losses over the next two years, on a yearly basis. Since then, the planning horizon has changed to nine quarters. For the last three CCAR exercises, BHCs are asked to submit their pro forma, post-stress capital projections in their capital plan beginning with data as of September 30, spanning the nine-quarter planning horizon. The projections begin in the fourth quarter of the current year and conclude at the end of the fourth quarter two years forward. Hence, for defining BHC stress scenarios, firms are asked to project the movements of key MVs over the planning horizon of nine quarters. As for determining the severity of the global market shock components for trading and counterparty credit losses, it will not be discussed in this paper, because it is a one-time shock and the evaluation will be on the movements of the market risk factors rather the MVs. In the 2011 CCAR, the Federal Reserve defined the stress supervisory scenario using nine MVs:

- Real GDP (“RGDP”)
- Consumer Price Index (“CPI”)
- Real Disposable Personal Income (“RDPI”)
- Unemployment Rate (“UNEMP”)
- Three-month Treasury Bill Rate (“3MTBR”)
- Ten-year Treasury Bond Rate (“10YTBR”)

- BBB Corporate Rate (“BBBCR”)
- Dow Jones Index (“DJI”)
- National House Price Index (“HPI”)

In CCAR 2012, the number of MVs that defined the supervisory stress scenario increased to 14. In addition to the original nine variables, the added variables were:

- Real GDP Growth (“RGDPG”)
- Nominal Disposable Income Growth (“NDPIG”)
- Mortgage Rate (“MR”)
- CBOE’s Market Volatility Index (“VIX”)
- Commercial Real Estate Price Index (“CREPI”)

Additionally, there is another set of 12 international macroeconomic variables, three macroeconomic variables and four countries / country blocks, included in the supervisory stress scenario. For CCAR 2013, the Federal Reserve System used the same set of variables to define the supervisory adverse scenario as in 2012. For the purposes of this research, let us consider the supervisory base and severely adverse scenario in 2015, focusing on 5 of the 9 most commonly used national Fed CCAR MVs, and one prevalent non-Fed MV:

- High Yield Spread (“HYS”)
- Real GDP Growth (“RGDPG”)
- Unemployment Rate (“UNEMP”)
- BBB Corporate Credit Spread (“BBBCS”)
- Commercial Real Estate Price Index (“CREPI”)
- CBOE’s Market Volatility Index (“VIX”)

This historical data, 87 quarterly observations from 2Q94 to 4Q15, are summarized in Table 1 of this section, and in Figures 7 through 18 of Appendix 1. In this study we focus on 6 macroeconomic variables: HYS, RDIG, UNEMP, CREPI, BBBCS and VIX. The summary statistics of HYS are shown in the 1st column of Table 1, levels in the top panel and percent changes in the bottom panel, with the corresponding time series plots and histograms shown in Figures 7 and 8. The level of the HYS over the historical period averages 5.20%, with a median of 4.69, displaying significant departures from normality in terms of right skewness (2.09) and fat-tails (kurtosis of 5.5). The series is rather volatile, displaying a coefficient of variation of 47.6%, and ranging from 2.39% to 16.27% over the sample period. On the other hand, the percent of the HYS over the historical period averages 2.17%, with a median of -0.81%, displaying significant departures from normality in terms of right skewness (2.23) and fat-tails (kurtosis of 8.11), in excess of the raw variable. The percent change of the series is also extremely volatile, displaying a coefficient of variation of 881.1%, and ranging from -26.37% to 101.9% over the sample period.

The summary statistics of RDIG are shown in the 2nd column of Table 4.1, levels in the top panel and percent changes in the bottom panel, with the corresponding time series plots and histograms shown in Figures 9 and 10. The level of the RDIG over the historical period averages 2.80%, with a median of 2.90, displaying significant departures from normality in terms of left skewness (-1.61) and fat-tails (kurtosis of 7.45). The series is extremely volatile, displaying a coefficient of variation of 133.6%, and ranging from -15.9% to 10.9% over the sample period. On the other hand, the percent of the HYS over the historical period averages -0.92%, with a median of 0.00%, displaying significant departures from normality in terms of left skewness (0.87) and fat-tails (kurtosis of 5.93), in excess of the raw variable. The percent change of the series is also extremely volatile, displaying a coefficient of variation of 64,617.5%, and ranging from -2,860.0% to 1,860.0% over the sample period.
The summary statistics of UR are shown in the 3rd column of Table 1, levels in the top panel and percent changes in the bottom panel, with the corresponding time series plots and histograms shown in Figures 11 and 12. The level of the RDIG over the historical period averages 5.97%, with a median of 5.50, displaying some departures from normality in terms of right skewness (1.05) and fat-tails (kurtosis of 0.02). The series is only somewhat volatile, displaying a coefficient of variation of 27.66%, and ranging from 3.90% to 9.90% over the sample period. On the other hand, the percent change of the UR over the historical period averages -0.13%, with a median of -1.75%, displaying significant departures from normality in terms of right skewness (1.92) and fat-tails (kurtosis of 4.49), in excess of the raw variable. The percent change of the series is also extremely volatile, displaying a coefficient of variation of -3,630.3%, and ranging from -7.46% to 20.29% over the sample period.
The summary statistics of CREPI are shown in the 4th column of Table 1, levels in the top panel and percent changes in the bottom panel, with the corresponding time series plots and histograms shown in Figures 13 and 14. The level of the RDIG over the historical period averages 166.7, with a median of 160.0, displaying rather mild departures from normality in terms of right skewness (0.23) and thin-tails (kurtosis of -0.99). The series is only somewhat volatile, displaying a coefficient of variation of 31.42%, and ranging from 88.0 to 273.4 over the sample period. On the other hand, the percent of the CREPI over the historical period averages 1.40%, with a median of 1.54%, displaying some departures from normality in terms of left skewness (-1.0) and fat-tails (kurtosis of 4.13), in excess of the raw variable. The percent change of the series is also extremely volatile, displaying a coefficient of variation of 292.9%, and ranging from -14.83% to 12.09% over the sample period.
The summary statistics of BBBCS are shown in the 5th column of Table 1, levels in the top panel and percent changes in the bottom panel, with the corresponding time series plots and histograms shown in Figures 15 and 16. The level of the BBBCS over the historical period averages 6.35%, with a median of 6.50%, displaying very mild departures from normality in terms of left skewness (-0.04),

and only moderate deviations from a Gaussian distribution with respect to thin-tails (kurtosis of -0.81). The series is only somewhat volatile, displaying a coefficient of variation of 21.31%, and ranging from 3.9% to 9.4% over the sample period. On the other hand, the percent of the BBBCS over the historical period averages -0.45%, with a median of -1.37%, displaying significant departures from normality in terms of right skewness (-1.32) and fat-tails (kurtosis of 5.21), far in excess of the raw variable. The percent change of the series is also extremely volatile, displaying a coefficient of variation of -1460.3%, and ranging from -17.07% to 30.56% over the sample period.

The summary statistics of VIX are shown in the 6th column of Table 1, levels in the top panel and percent changes in the bottom panel, with the corresponding time series plots and histograms shown in Figures 17 and 18. The level of the VIX over the historical period averages 27.26, with a median of 23.60, displaying significant

res from normality in terms of right skewness (1.73), as well as material deviations from a Gaussian distribution with respect to fat-tails (kurtosis of 4.92). The series is only somewhat volatile, displaying a coefficient of variation of 41.49%, and ranging from 12.7 to 80.9 over the sample period. On the other hand, the percent of the VIX over the historical period averages 4.50%, with a median of -1.42%, displaying significant departures from normality in terms of right skewness (1.37) and fat-tails (kurtosis of 1.82), far in excess of the raw variable. The percent change of the series is also extremely volatile, displaying a coefficient of variation of 741.3%, and ranging from -49.45% to 115.34% over the sample period.

In Table 2 we display the correlation matrices of the macroeconomic variables. First, we note that there is a diversity of magnitudes in with respect to the correlations amongst the pairs, although for the most part they are sizable. In looking at the levels, we see high correlations (i.e., >50%) in the HYS / VIX (78.8%) and CREPI / BBBCS (-62.2%) pairs, and low correlations (i.e., <10%) in the RDIG / BBBCS (9.6%) and RDIG / VIX (-9.5%) pairs. In looking at the percent changes, we see high correlations in the HYS / VIX (54.7%) and CREPI / BBBCS (-62.2%) pairs, and low correlations in the RDIG / BBBCS (1.2%) and RDIG / VIX (-0.3%) pairs. In general, we also note that correlations are lower amongst the variables in percent rather than in level form. However, to that point we further observe that in some case the signs of the correlations are not always intuitive (highlighted in red in Table 4.2), and this is more of a problem with the variables in level form – in 6 out of 15 (e.g., negative and positive for CREPI vs.

RDIG and UNEMP, respectively) pairs in the latter as opposed to 2 pairs in the former (HYS/CREP and RDIG/BBBCS are positive). These considerations of magnitude and direction of relationships will have a bearing on the cross-equation restrictions that we will impose in the VAR estimation.

In Table 3 we display the *Augmented Dickey-Fuller* ("ADF") statistics of the macroeconomic variables under consideration. We observe that we only reject the null hypothesis of a unit root process (or of non-stationarity) in one case for the variables in level for, whereas in percent change for we are able to reject this in all cases at the 5% confidence level or better. Taken in combination with the observations regarding the correlation analysis of Table 2, this leads to the choice of modeling the percent changes in the macroeconomic variables in order to generate base and stress scenarios. As a practice, when modeling in a time series framework, it is preferable to work with data that are jointly stationary.

A critical modeling consideration for the MS-VAR estimation is the choice of process generation distributions for the normal and the stressed regimes. s described in the summary statistics of Table 1, we find that when analyzing the macroeconomic data in percent change form, there is considerable skewness in the direction of adverse changes (i.e., right skewness for variables where increases denote deteriorating economic conditions such as UNEMP, and left skewness in variables where declines are a sign of weakening conditions such as RDIG). Furthermore, in normal regimes where percent changes are small we find a normal distribution to adequately describe the error distribution, whereas when such changes are at extreme levels in the adverse direction we find that a log-normal distribution does a good job of characterizing the data generating process. This is similar to the findings of in the context of modeling the U.S. Treasury yields [56]. We observe that this mixture well characterizes the empirical distributions of the data.

The final modeling consideration is the methodology for partitioning the space of scenario paths across our 6 macroeconomic variables for a Base and for a Severe Scenario. In the case of the Severe scenario, we choose to identify such a path in which *all six* macroeconomic variables exceed their historical 99.0^{th} percentile in at least a single quarter, and then in that set for each variable we take an average across such paths in each quarter. It is our view that this is a reasonable definition of a Severe scenario, and in our risk advisory practice we have observed similar definitions in the industry.[2] In the case of the Base scenario, we take an average across all paths in a given quarter for a given variable.

The main findings of our study are shown in Figures 1 through 6 where we show for each macroeconomic variable the Base and Severe scenarios for the VAR and MS-VAR models[3], and also compare this to the corresponding Fed scenarios, along the historical time series. We make the following general conclusions:

---

[2] We have performed a sensitivity analysis, available upon eques, using the 95^{th} and 99.9^{th} percentiles, and the results are not greatly changed.

[3] Estimation results for the VAR and MS-VAR model are available upon request. The models are all convergent and goodness of fit metrics in with industry standards. Signs of coefficient estimates are in line with economic intuition and estimates are all significant at conventional levels. We use the dse, tseries and MSBVAR libraries in R in order to perform the estimations [46].

- In the Severe scenario, the MS-VAR model is far more conservative than the VAR model, and always at least match and in some cases can well exceed historical peaks or troughs in the adverse direction.
- In terms of magnitude, the VAR model is similar to the Fed scenarios, but the trajectories of either the VAR or MS-VAR model tend to be more regular, rising at a more gradual pace into the forecast period.
- In the Base scenarios, the Fed model is rather similar to the VAR model, but in all cases the MS-VAR model produces a higher base, which is driven by the skewness of the mixture error distribution.

The scenario projections for the HYS are shown in Figure 1. The HYS has ranged in around 3% to 8% in benign periods, and peaked during the financial crisis at around 16%. The Fed severe scenario for the HYS peaks just below the latter historical spike at around 14% at a year into the forecast period, declining linearly to around the level of the jump-off period in the 4th quarter of 2015 of about 5%. The VAR model does not quite achieve this level of stress but is close, peaking at just below 14% 2 years into the forecast period and then declining in a steeper and similarly linear fashion to around the pre-stress levels. In contrast, the MS-VAR model in the severe scenario exhibits far more extreme levels, peaking at just over 23% (about 50% greater than either the VAR or the Fed model) a year into the forecast period, and remaining elevated for another year before reverting to the pre-forecast period levels. Finally, with regard to the Base scenario, the VAR and Fed models are similar in trajectory, rising somewhat in the first year and then reverting to the levels of the period just prior to the forecast, although the former is somewhat higher peaking at around 10% vs. around 9% for the latter.

The scenario projections for the RDIG are shown in Figure 2. The RDIG has ranged in around -5% to 5% in benign periods, and unlike the other macroeconomic variables, had its most adverse levels twice, during the financial crisis (in early 2013) at around -9% (-16%). The Fed severe scenario for the RDIG reaches a trough quite below the first and milder historical trough at around -4% at around a year into the forecast period, then increasing in an s-shape to around the level of the jump-off period in the 4th quarter of 2015 of about 4%. The VAR model does not quite achieve the most level of stress for RDIG but exceeds the financial crisis value in absolute terms, a trough at around -12% 1 year into the forecast period and then increasing smoothly to around the pre-stress levels. In contrast, the MS-VAR model in the severe scenario exhibits far more extreme levels, declining to just about the historical trough of -15% (about 20% and 5 times greater in magnitude than the VAR and Fed models, respectively) a year into the forecast period. Finally, with regard to the Base scenario, the VAR and MS-VAR models are similar in trajectory, dipping somewhat in the first year and then stabilizing to levels somewhat below that of the period just prior to the forecast, although the latter is somewhat lower. On the other hand, the Fed model appears to be rather different, dipping slightly and remaining flat at levels somewhat below the levels of 2015.

The scenario projections for the UNEMP are shown in Figure 3. The UNEMP has ranged in around 4% to 8% in benign periods, and peaked during the financial crisis at around 10%, exactly where the Fed severe scenario peaks around six weeks into the forecast period and then declining linearly to around the level of 2011 of about 9% and well short of recovery by the end of the forecast period. The VAR model exceeds this level of stress, peaking at just above 12% 2 years into the forecast period, and then declining in a similarly linear fashion to a still distressed level of around 11% by the end of the forecast period. In contrast, the MS-VAR model in the severe scenario exhibits far more extreme levels, peaking at about 18% (about 50% and about 80% greater than the VAR or the Fed models, respectively) a year into the forecast period, and remaining elevated at no less than about 15% by the end of the forecast period. Finally, with regard to the Base scenario, the VAR and Fed models are rather similar in trajectory, dipping somewhat and then flattening out throughout the forecast period to a level about 1% lower than in the late 2015, with the VAR model being slight below the Fed model. In contrast, the MS-VAR model starts off with a similar trajectory to the VAR and Fed models, but diverges toward the end of the forecast period to end slightly higher.

The scenario projections for the CREPI are shown in Figure 4. The CREPI has ranged in around 200 to 250 in benign periods and had its most adverse levels at 150 during the financial crisis. The Fed severe scenario for the RDIG fails to reaches a trough at this level and only goes as low as around 180 at around a year and a half into the forecast period, then increasing gradually to around 200, far from the level of the jump-off period in the 4th quarter of 2015 of about 270. The VAR model has a rather similar trajectory, albeit both declining and recovering at a slower pace. In contrast, the MS-VAR model in the severe scenario exhibits far more extreme levels, declining to somewhat below the historical trough at around 125, or about 15% lower than the VAR and Fed models, two years into the forecast period. Finally, with regard to the Base scenario, the VAR and Fed models are similar in trajectory, rising linearly throughout the forecast period from about 275 to around 300 and 320, with the VAR model lower. On the other hand, the MS-VAR model appears to be rather different, increasing only slightly by the end of the forecast period to about 280.

The scenario projections for the BAACR are shown in Figure 5. The BAACR has ranged in around 2.5% to 3.5% in benign periods, and peaked during the financial crisis at around 5.5%. The Fed severe scenario for the HYS peaks above the latter historical spike at around 6.5% at a year into the forecast period, declining linearly to 4.6%, quite above the 3.3% level of the jump-off period in the 4th quarter of 2015. The VAR model does not quite achieve this level of stress, rising more gradually peaking at 5.5% 2 years into the forecast period, and then declining to about the level where the Fed model ends up quite above the pre-stress levels. In contrast, the MS-VAR model in the severe scenario exhibits far more extreme levels, peaking nearly 10% (about 35% and 45% greater than the Fed and VAR models, respectively) a little over a year into the forecast period, and remaining elevated for the remainder of the forecast period with a decline to a little over 7.5%. Finally, with regard to the Base scenario, the MS-VAR and Fed models are similar in trajectory, rising throughout the forecast period to about 5.5%, although the Fed model remains above the MS-VAR model forecast until the end. In contrast, the VAR model rises far more gradually than either the MS-VAR or Fed models, reaching around only 4.4% by the end of the forecast period.

The scenario projections for the VIX are shown in Figure 6. The VIX has ranged in around 15 to 30 in benign periods, and peaked during the financial crisis at just above 75. The Fed severe scenario for the VIX peaks just below 75 at the start of the forecast period, declining linearly thereafter to the level of the jump-off period in the 4^{th} quarter of 2015. The VAR model slightly this level of stress, rising more gradually peaking at around 80 at 6 quarters into the forecast period, and then declining to about the level where the Fed model ends up near the pre-stress levels. In contrast, the MS-VAR model in the severe scenario exhibits far more extreme levels, peaking nearly just over 135 (about 80-90% greater than either the Fed and VAR models) in about the same timeframe of the VAR model and then declining to slightly lower levels as during the jump-off period. Finally, with regard to the Base scenario, the MS-VAR and VAR models are similar in trajectory, rising gradually throughout the forecast period to about 30-35, with the MS-VAR model remaining above the VAR model forecast until the end. In contrast, the Fed model gradually declines throughout the projection, ending up at somewhat lower by the end of the forecast period.

5 Conclusion and Future Directions

In this study we have examined a critical input into the process stress testing process, the macroeconomic scenarios provided by the prudential supervisors to institutions for exercises such as the Federal Reserve's CCAR program. In particular, we have addressed the supervisory requirements that banks develop their own macroeconomic scenarios. We have analyzed a common approach of a VAR statistical model that exploit the dependency structure between both macroeconomic drivers, as well between modeling segments, and addressed the well-known phenomenon that linear models such as VAR are unable to explain the phenomenon of fat-tailed distributions that deviate from normality, an empirical fact that has been well documented in the empirical finance literature. We have proposed a challenger approach, widely used in the academic literature, but not commonly employed in practice, the MS-VAR model. We empirically tested these models using Federal Reserve macroeconomic data, gathered and released by the regulators for CCAR purposes, respectively.

We find the MS-VAR model to be more conservative than the VAR model, and also to exhibit greater accuracy in model testing, as the latter model can better capture extreme events observed in history. In the Severe scenario, the MS-VAR model is far more conservative than the VAR model, and always at least match and in some cases can well exceed historical peaks or troughs in the adverse direction. In terms of magnitude, the VAR model is similar to the Fed scenarios, but the trajectories of either the VAR or MS-VAR model tend to be more regular, rising at a more gradual pace into the forecast period. In the Base scenarios, the Fed model is rather similar to the VAR model, but in all cases the MS-VAR model produces a higher base, which is driven by the skewness of the mixture error distribution.

There are several directions in which this line of research could be extended, including but not limited to the following:

- More granular classes of credit risk models, such as ratings migration or PD / LGD scorecard / regression
- Alternative data-sets, for example bank or loan level data
- More general classes of regression model, such as logistic or semi-parametric
- Applications related to ST, such as RC or EC

References

[1] Acharya, V. V. and P. Schnabl, P., How Banks Played the Leverage Game, 44-78 in: Acharya, V. V. and M. Richardson (eds), Restoring Financial Stability, Wiley Finance, New York, 2009.

[2] Demirguc-Kunt, A., Detragiache, E., and O. Merrouche, Bank Capital: Lessons from the Financial Crisis, World Bank Policy Research Working Paper No. 5473, (November, 2010)

[3] Merton, R., On the Pricing of Corporate Debt: The Risk Structure of Interest Rates, Journal of Finance, 29(2), (1974), 449–470.

[4] Board of Governors of the Federal Reserve System, 2011, "Supervisory Guidance on Model Risk Management", Supervisory Letter 11-7, April 4th.

[5] BANK FOR INTERNATIONAL SETTLEMENTS. "Basel committee on banking supervision (BCBS)." Internal convergence of capital measurement and capital standards: a revised framework (1988).

[6] ___________International convergence of capital measurement and capital standards: a revised framework (2006).

[7] ___________Principles for sound stress testing practices and supervision - consultative paper (May – No. 155) (2009a).

[8] ___________Guidelines for computing capital for incremental risk in the trading book (2009b).

[9] ___________Revisions to the Basel II market risk framework (2009c).

[10] ___________Analysis of the trading book quantitative impact study (October) (2009d).

[11] ___________Basel III:: a global regulatory framework for more resilient banks and banking systems (December) (2010a).

[12] ___________ An Assessment of the Long-Term Economic Impact of Stronger Capital and Liquidity Requirements (August) (2010b).

[13] Schuermann, T., Stress Testing Banks, International Journal of Forecasting, 30(3), (2014), 717-728.

[14] BANK FOR INTERNATIONAL SETTLEMENTS. "Basel committee on banking supervision (BCBS)." Amendment to the capital accord to incorporate market risks (1996).

[15] ________"The committee on global financial systems."Stress testing by large financial institutions: current practice and aggregation issues (April, 2000).

[16] J.P. Morgan, RiskMetricsTM, Second Edition, J.P. Morgan, New York, 1994.

[17] Jorion, P., Risk2: Measuring the Value at Risk, Financial Analyst Journal, 52(6), (1996), 47-56.

[18] Kupiec, P, Risk Capital and VaR, The Journal of Derivatives, 7(2), (1999),41-52.

[19] J. Berkowitz and A. Jeremy, A Coherent Framework for Stress-Testing, FEDS Working Paper No. 99-29, (1999).

[20] BANK FOR INTERNATIONAL SETTLEMENTS. "The committee on global financial systems." Stress testing by large financial institutions: current practice and aggregation issues (April, 2000).
[21] Mosser, P.C., I., Fender, and M.S. Gibson, An International Survey of Stress Tests, Current Issues in Economics and Finance, 7(10). (2001), 1-6.
[22] J.P. Morgan, Credit Metrics TM, First Edition, J.P. Morgan, New York, 1997.
[23] Wilde, T., Credit Risk+. A Credit Risk Management Framework, Credit Suisse First Boston, 1997.
[24] Koyluoglu, H, and A Hickman, Reconcilable Differences, Risk, October, (1998), 56-62.
[25] Bangia, A. Diebold, F.X., Kronimus, A., Schagen, C., and T. Schuermann, Ratings Migration and the Business Cycle, with Application to Credit Portfolio Stress Testing, Journal of Banking and Finance, 26(2), (2002), 445-474.
[26] Board of Governors of the Federal Reserve System, 1996, “Joint Policy Statement on Interest Rate Risk”, Supervisory Letter 96-13, May 23rd.
[27] ________1999, “Supervisory Guidance Regarding Counterparty Credit Risk Management”, Supervisory Letter 99-03, February 1st.
[28] _______ 2002, “Interagency Guidance on Country Risk Management”, Supervisory Letter 02-05, March 8th.
[29] Foglia, A., Stress Testing Credit Risk: A Survey of Authorities’ Approaches, International Journal of Central Banking, 5(3), (2009), 9-37.
[30] Inanoglu, H., and M. Jacobs, Jr., Models for Risk Aggregation and Sensitivity Analysis: An Application to Bank Economic Capital, The Journal of Risk and Financial Management, 2, 2009, 118-189.
[31] Jacobs, Jr., M., Validation of Economic Capital Models: State of the Practice, Supervisory Expectations and Results from a Bank Study, Journal of Risk Management in Financial Institutions, 3(4), (2010), 334-365.
[32] ________Stress Testing Credit Risk Portfolios, Journal of Financial Transformation, 37, (2013), 53-75.
[33] Rebonato, R., Coherent Stress Testing: A Bayesian Approach, Wiley, New York. 2010.
[34] Jacobs, Jr., M., Karagozoglu, A.K., and F.J. Sensenbrenner, Stress Testing and Model Validation: Application of the Bayesian Approach to a Credit Risk Portfolio, The Journal of Risk Model Validation, 9(3), (2015), 1-70.
[35] T. Clark and L. R. Ryu, 2015, CCAR and Stress Testing as Complementary Supervisory Tools, Board of Governors of the Federal Reserve Supervisory Staff Reports June 24, (2015).
[36] R. Bidder and A. McKenna, Robust Stress Testing, Federal Reserve Bank of San Francisco Working Paper September, (2015).
[37] R.S Frame, A. Fuster, J. Tracy and J. Vickery, The Rescue of Fannie Mae and Freddie Mac, Federal Reserve Bank of New York Staff Reports No. 79 (March 2015).

[38] Stock, J.H., and M.W. Watson, Vector Autoregressions, Journal of Economic Perspectives, 15(4), (2001), 101-115.
[39] Sims, C.A., Macroeconomics and Reality, Econometrica, 48, (1980), 1-48.
[40] G. Box and G. Jenkins, Time Series Analysis: Forecasting and Control, Holden-Day, San Francisco, 1970.
[41] P.J. Brockwell and R.A. Davis, Time Series: Theory and Methods, Springer-Verlag, New York, 1991.
[42] J. J. F. Commandeur and S.J. Koopman, Introduction to State Space Time Series Analysis, Oxford University Press, New York, 2007.
[43] Hanan, E.J., The Identification Problem for Multiple Equation Systems with Moving Average Errors, Econometrica, 39, (1971), 751-766.
[44] Kohn, R., Asymptotic Results for ARMAX Structures, Econometrica, 47, (1979), 1295–1304.
[45] Hanan, E.J., The Statistical Theory of Linear Systems, John Wiley, New York, 1988.
[46] R Development Core Team, 2016: “R: A Language and Environment for Statistical Computing.” R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.
[47] Baum, L.E., and T. Petrie, Statistical Inference for Probabilistic Functions of Finite State Markov Chains, The Annals of Mathematical Statistics, 37(6), (1966), 1554-1563.
[48] ___________, Soules, G. and N, Weiss, A Maximization Technique Occurring in the Statistical Analysis of Probabilistic Functions of .Markov Chains, Annals of Mathematical Statistics, 41(1), (1970), 164-171.
[49] Pearson, K. Contribution to the Mathematical Theory of Evolution, Philosophical Transactions of the Royal Society, 185, (1894). 71-110.
[50] Blackwell, D., and L. Koopmans, On the Identifiable Problem for Functions of Finite Markov Chains, The Annals of Mathematical Statistics, 28(4), (1957), 1011-1015.
[51] Heller, A., On Stochastic Processes Derived from Markov Chains, The Annals of Mathematical Statistics, 36(4), (1965), 1286-1291.
[52] Goldfeld, S.M., and R.E. Quandt. A Markov Model for Switching Regressions, Journal of Econometrics, 1(1), (1973), 3-15.
[53] Hamilton. J.D., Rational-Expectations Econometric Analysis of Changes in Regime: An Investigation of the Term Structure of Interest Rates, Journal of Economic Dynamics and Control, 12, (1988), 385-423.
[54] ________, A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle, Econometrica, 57 (2), (1989), 357-384.
[55] Tjøstheim, D., Estimation in Nonlinear Time Series Models, Stochastic Processes and their Applications, 21(2), (1986), 251-273.
[56] Loregian.A, and A., Meucci. Neither “Normal” nor “Lognormal”: Modeling Interest Rates across All Regimes, Financial Analysts Journal, 72(3), (2016), 68-82.

7 Appendix 1: Time Series Plots and Smoothed Histograms of Macroeconomic Variables

Table 1: Summary Statistics of Historical Macroeconomic Variables

		High Yield Spread	Real Disposable Income Growth	Unemployment Rate	Commercial Real Estate Index	BBB Corporate Credit Spread	Market Volatility Index
Levels	Count	87	87	87	87	87	87
	Mean	5.20	2.80	5.97	166.66	6.35	27.26
	Median	4.69	2.90	5.50	160.00	6.50	23.90
	Std Dev	2.48	3.74	1.65	52.36	1.35	11.31
	Coefficient of Variation	47.62%	133.57%	27.66%	31.42%	21.31%	41.49%
	Skewness	2.09	(1.61)	1.04	0.23	-0.04	1.73
	Kurtosis	6.48	7.45	0.02	-0.99	-0.81	4.92
	75th Percentile	6.33	4.25	6.45	209.00	7.35	31.95
	25th Percentile	3.38	1.30	4.75	124.50	5.40	19.80
	Interquatile Range	2.96	2.95	1.70	84.50	1.95	12.15
	Interquatile Range/Median	63.07%	101.72%	30.91%	52.81%	30.00%	50.84%
	Minimum	2.39	(15.90)	3.90	88.00	3.90	12.70
	Maximum	16.27	10.90	9.90	273.40	9.40	80.90
	Range	13.88	26.80	6.00	185.40	5.50	68.20
Percent Changes	Count	87	87	87	87	87	87
	Mean	2.17%	-0.92%	-0.13%	1.40%	-0.45%	4.50%
	Median	-0.81%	0.00%	-1.75%	1.54%	-1.37%	-1.42%
	Std Dev	19.12%	594.18%	4.91%	4.09%	6.63%	33.39%
	Coefficient of Variation	881.11%	-64617.54%	-3660.30%	292.86%	-1460.32%	741.29%
	Skewness	2.23	(0.87)	1.92	-1.00	1.32	1.37
	Kurtosis	8.11	5.93	4.49	4.13	5.21	1.82
	75th Percentile	7.32%	250.00%	0.88%	3.75%	2.99%	16.14%
	25th Percentile	-8.80%	-190.00%	-2.69%	0.00%	-4.73%	-18.88%
	Interquatile Range	16.12%	440.00%	3.56%	3.75%	7.72%	35.03%
	Interquatile Range/Median	-1984.66%	N/A	-203.19%	243.94%	-563.53%	-2469.29%
	Minimum	-26.37%	-2680.00%	-7.46%	-14.83%	-17.07%	-49.45%
	Maximum	101.92%	1860.00%	20.29%	12.09%	30.56%	115.34%
	Range	1.28	45.40	0.28	0.27	0.48	1.65

Table 2: Historical Correlations of Macroeconomic Variables

		Levels						Percent Changes					
		High Yield Spread	Real Disposable Income Growth	Unemployment Rate	Commercial Real Estate Index	BBB Corporate Credit Spread	Market Volatility Index	High Yield Spread	Real Disposable Income Growth	Unemployment Rate	Commercial Real Estate Index	BBB Corporate Credit Spread	Market Volatility Index
Levels	High Yield Spread	1	-	-	-	-	-	-	-	-	-	-	-
	Real Disposable Income Growth	-18.56%	1	-	-	-	-	-	-	-	-	-	-
	Unemployment Rate	37.41%	-23.63%	1	-	-	-	-	-	-	-	-	-
	Commercial Real Estate Index	16.76%	-16.03%	16.07%	1	-	-	-	-	-	-	-	-
	BBB Corporate Credit Spread	26.80%	9.55%	-36.74%	-62.19%	1	-	-	-	-	-	-	-
	Market Volatility Index	78.83%	-9.50%	19.06%	18.42%	30.22%	1	-	-	-	-	-	-
Percent Changes	High Yield Spread	30.20%	-4.63%	-21.40%	18.42%	17.78%	55.41%	1	-	-	-	-	-
	Real Disposable Income Growth	-1.27%	-5.75%	9.66%	-4.22%	-2.48%	-11.39%	-26.28%	1	-	-	-	-
	Unemployment Rate	71.92%	-25.29%	10.66%	4.91%	44.54%	56.82%	21.50%	-8.909%	1	-	-	-
	Commercial Real Estate Index	-45.08%	10.45%	-22.35%	5.59%	-27.63%	-19.65%	7.63%	5.848%	-42.94%	1	-	-
	BBB Corporate Credit Spread	5.69%	5.12%	-25.52%	16.36%	17.97%	18.72%	39.10%	1.19%	2.20%	12.60%	1	-
	Market Volatility Index	6.34%	-6.13%	-7.64%	12.44%	2.84%	45.60%	54.69%	-0.348%	5.23%	21.28%	27.37%	1

Table 3: Augmented Dickey-Fuller Stationarity Test Statistics of Macroeconomic Variables

	Levels		Percent Changes	
	Test Statistic	P-Value	Test Statistic	P-Value
High Yield Spread	-2.914	0.2004	-3.3448	0.0348
Real Disposable Income Growth	-4.2137	< 0.001	-7.1875	< 0.001
Unemployment Rate	-2.4774	0.3798	-3.2771	0.0276
Commercial Real Estate Index	-2.7076	0.2852	-3.4552	0.0257
BBB Corporate Credit Spread	-2.5600	0.3459	-4.8234	< 0.001
Market Volatility Index	-2.3952	0.4136	-4.4797	< 0.001

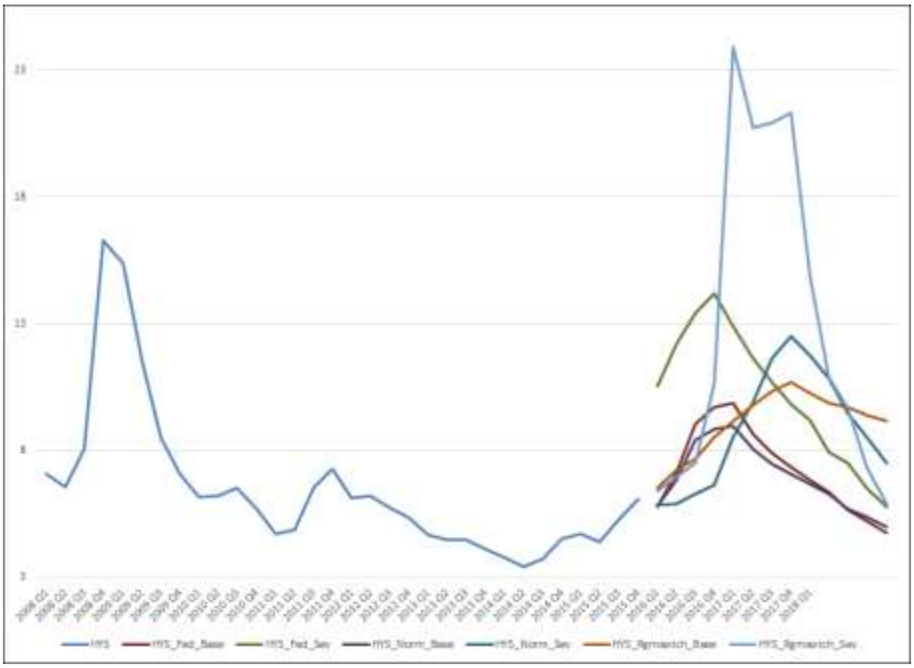

Figure 1: Historical Time Series, Base and Severe Scenarios for the VAR, MS-VAR and Fed Models – High Yield Spread

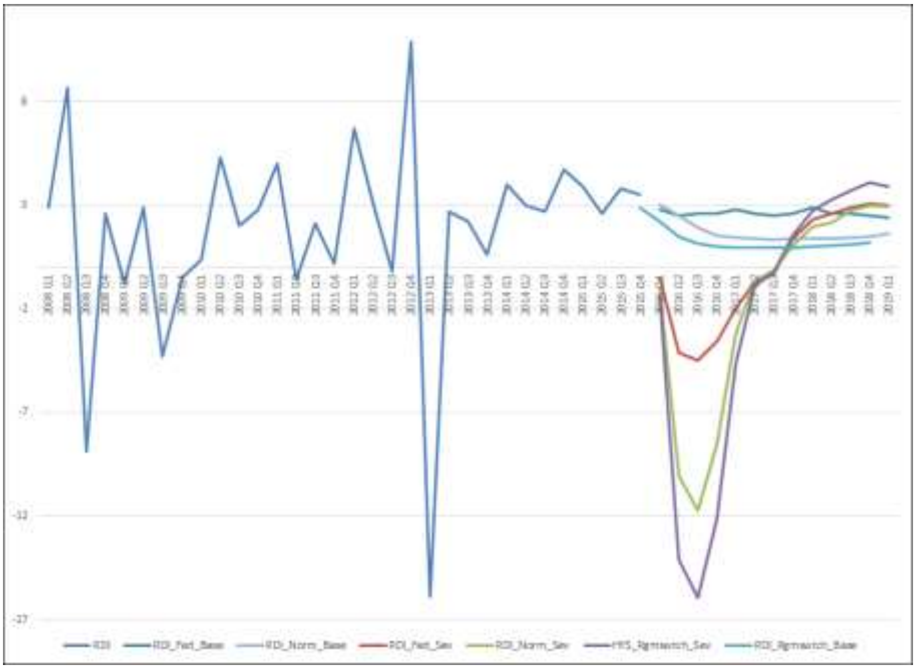

Figure 2: Historical Time Series, Base and Severe Scenarios for the VAR, MS-VAR and Fed Models – Real Disposable Income Growth

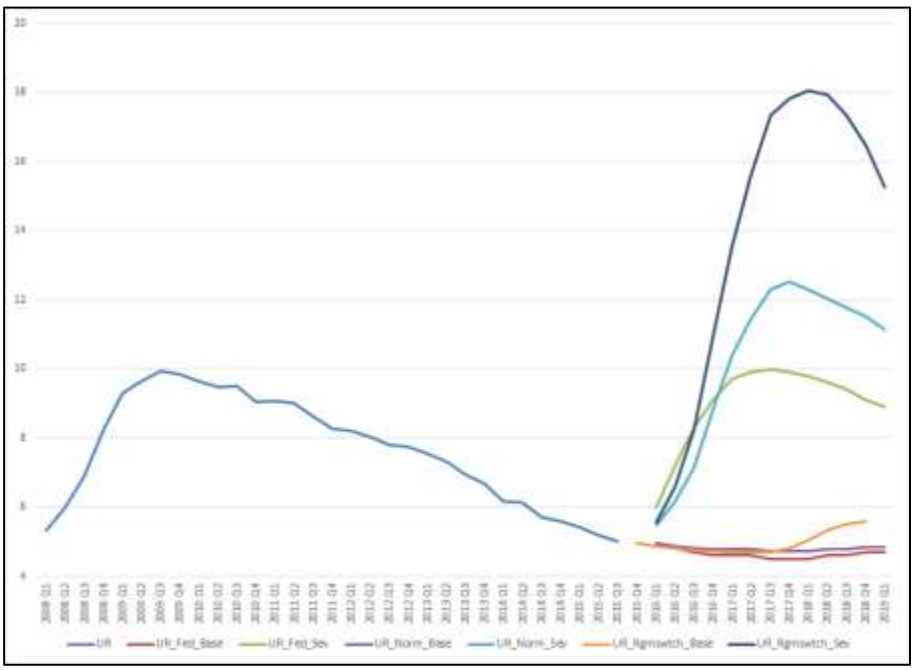

Figure 3: Historical Time Series, Base and Severe Scenarios for the VAR, MS-VAR and Fed Models – Unemployment Rate

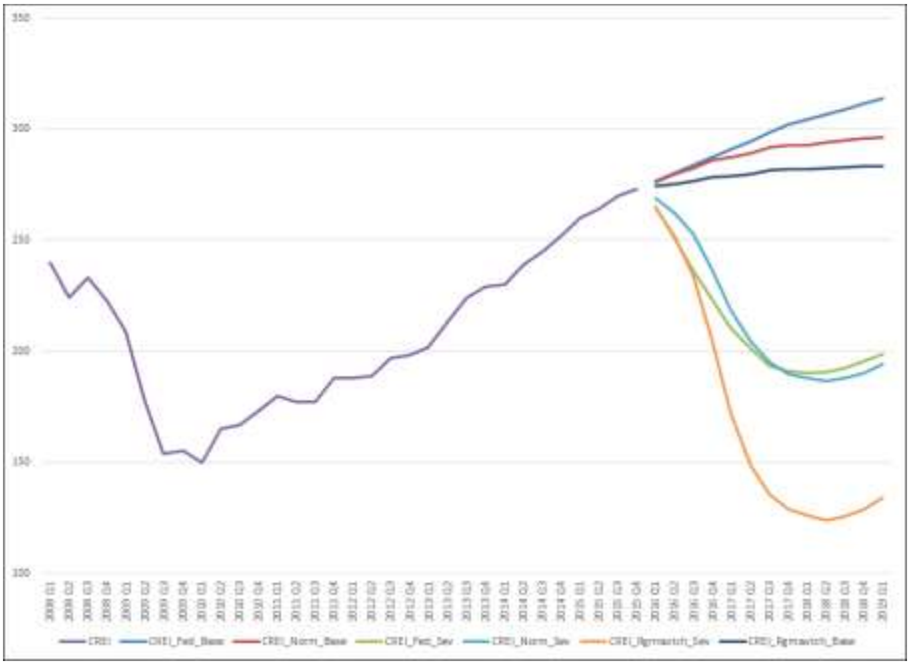

Figure 4: Historical Time Series, Base and Severe Scenarios for the VAR, MS-VAR and Fed Models – Commercial Real Estate Price Index

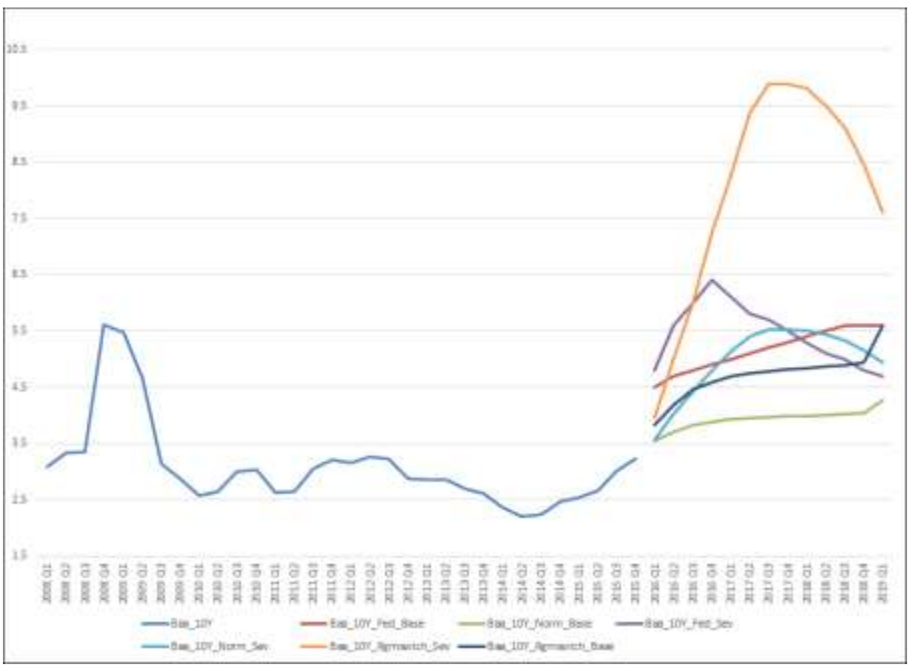

Figure 5: Historical Time Series, Base and Severe Scenarios for the VAR, MS-VAR and

Fed Models – Baa Corporate Credit Spread

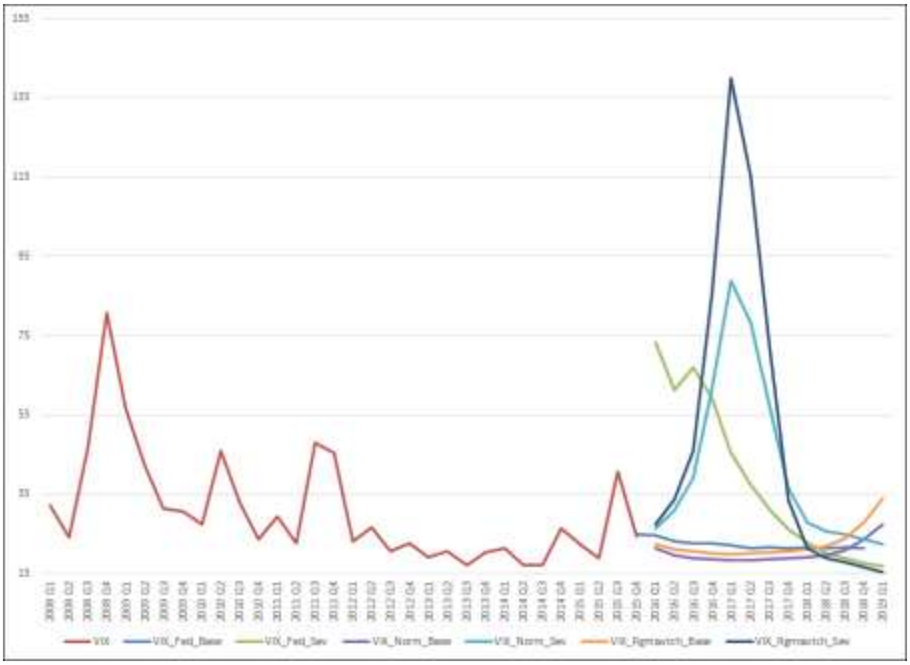

Figure 6: Historical Time Series, Base and Severe Scenarios for the VAR, MS-VAR and Fed Models – VIX Equity Market Volatility Index

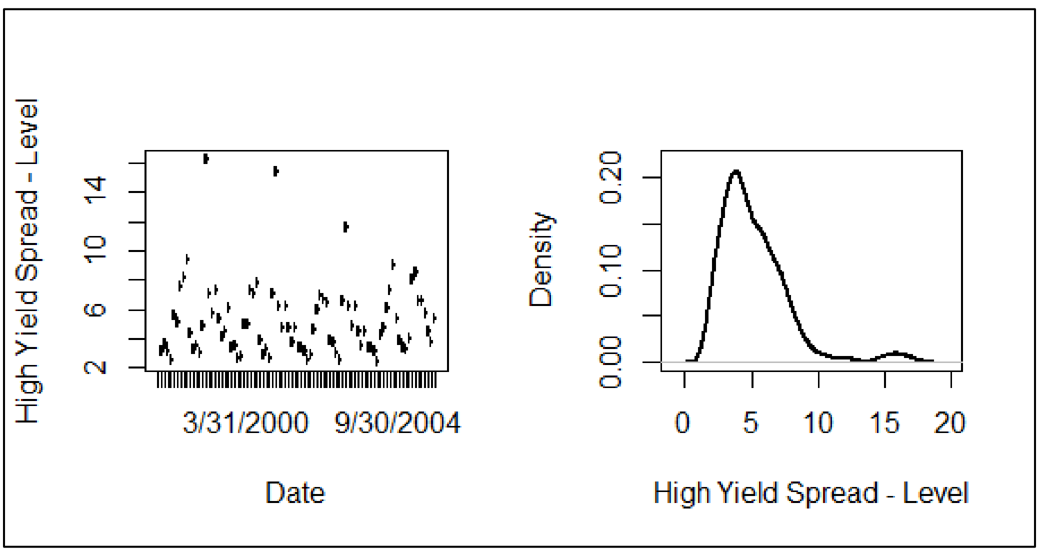

Figure 7: Time Series and Kernel Density Plot – High Yield Spread (Levels)

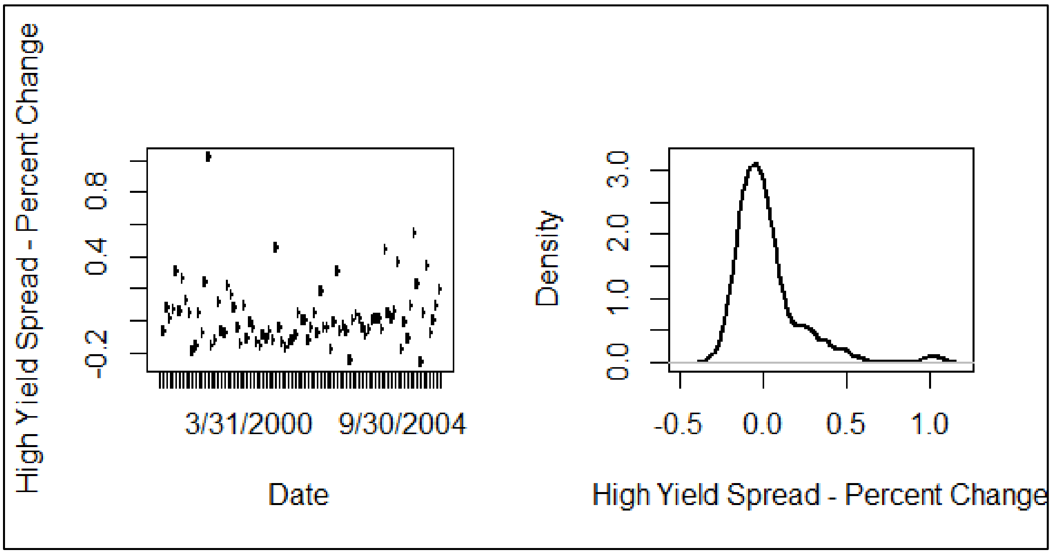

Figure 8: Time Series and Kernel Density Plot – High Yield Spread (Percent Changes)

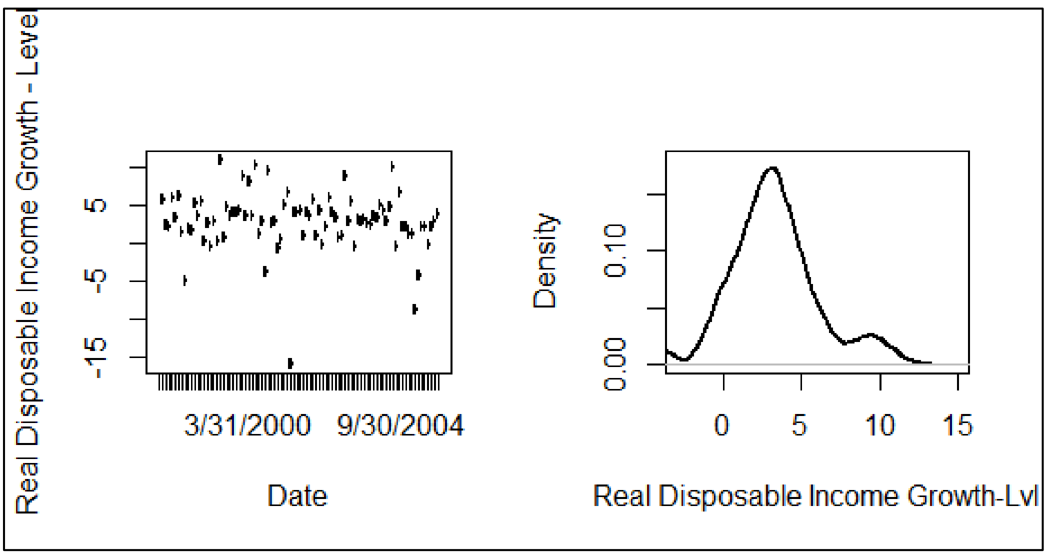

Figure 9: Time Series and Kernel Density Plot – Real Disposable Income Growth (Levels)

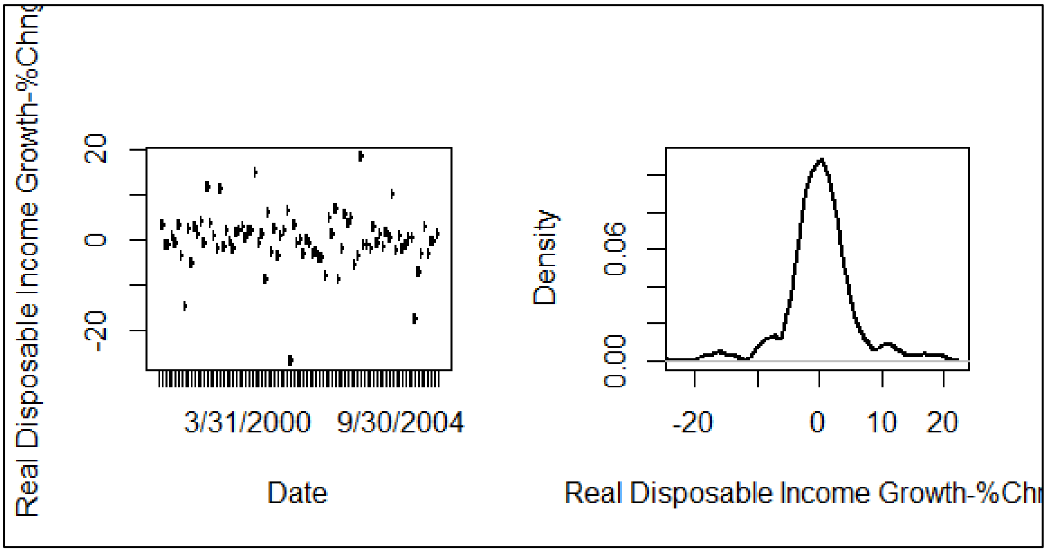

Figure 10: Time Series and Kernel Density Plot – Real Disposable Income Growth (Percent Changes)

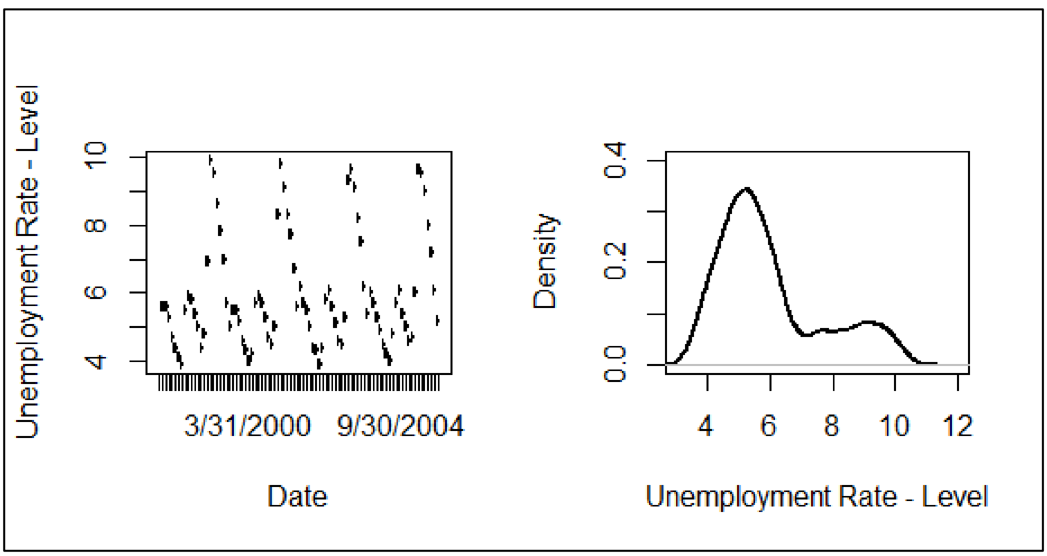

Figure 11: Time Series and Kernel Density Plot – Unemployment Rate (Levels)

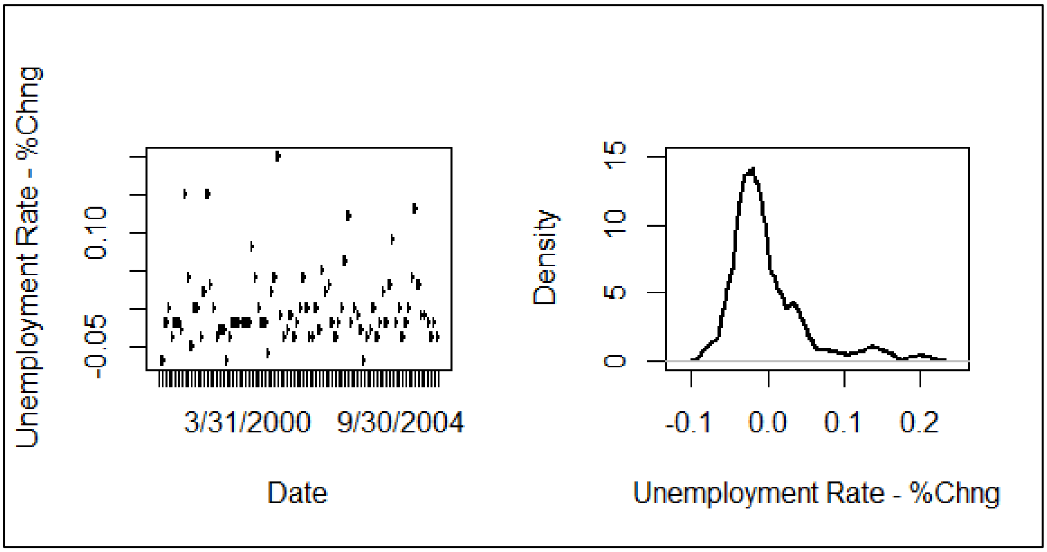

Figure 12: Time Series and Kernel Density Plot – Unemployment Rate (Percent Changes)

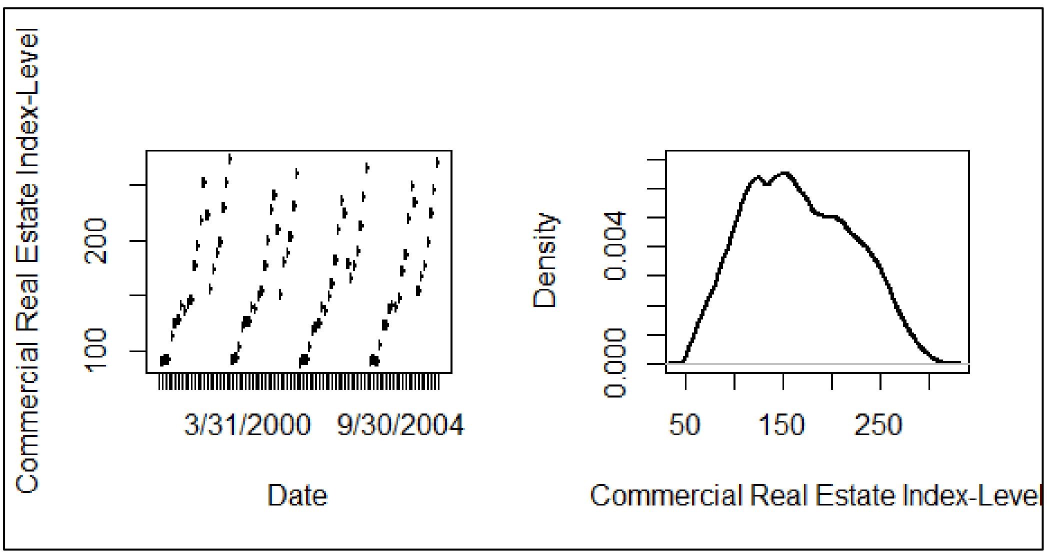

Figure 13: Time Series and Kernel Density Plot – Commercial Real Estate Index (Levels)

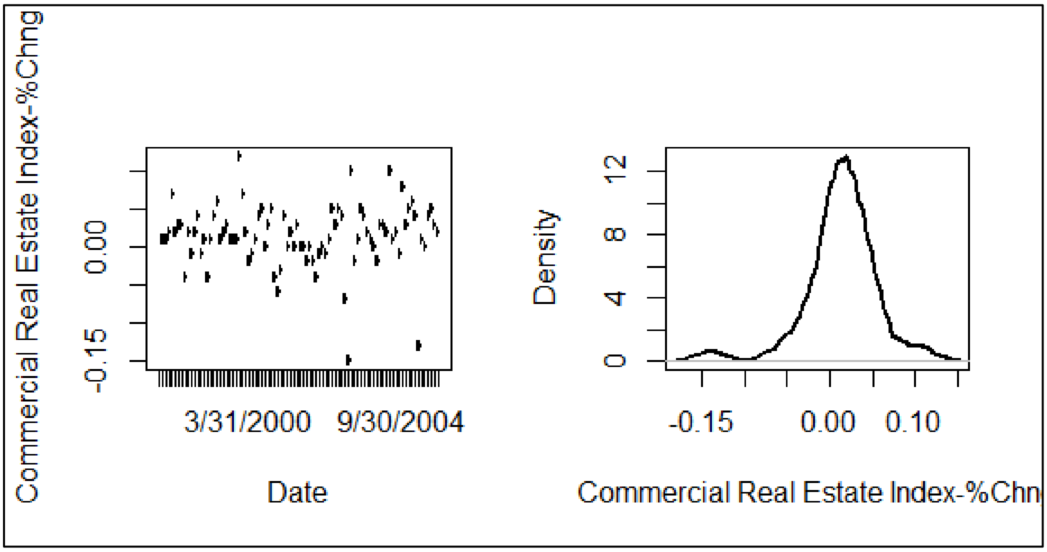

Figure 14: Time Series and Kernel Density Plot – Commercial Real Estate Index (Percent Changes)

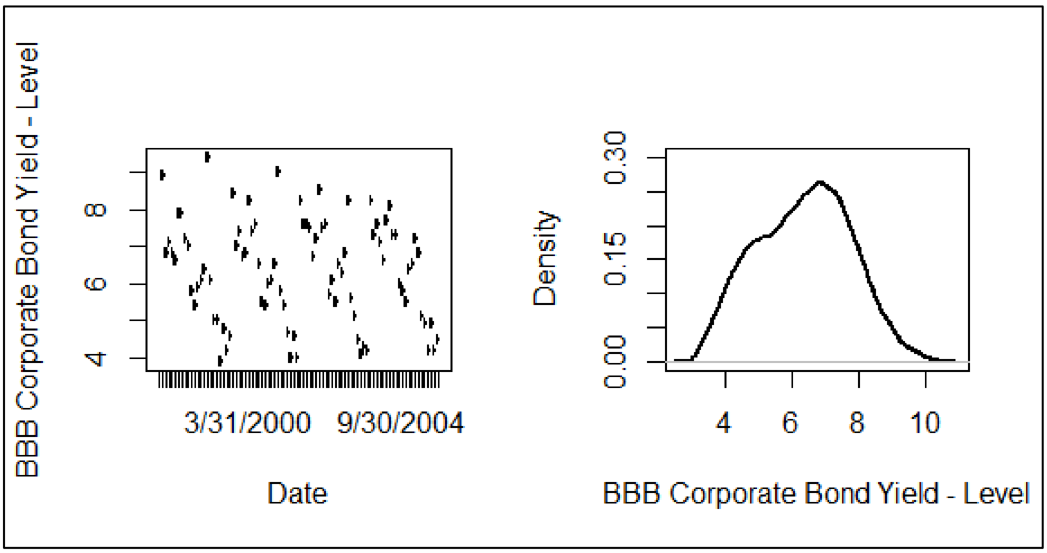

Figure 15: Time Series and Kernel Density Plot – BBB Corporate Bond Yield (Levels)

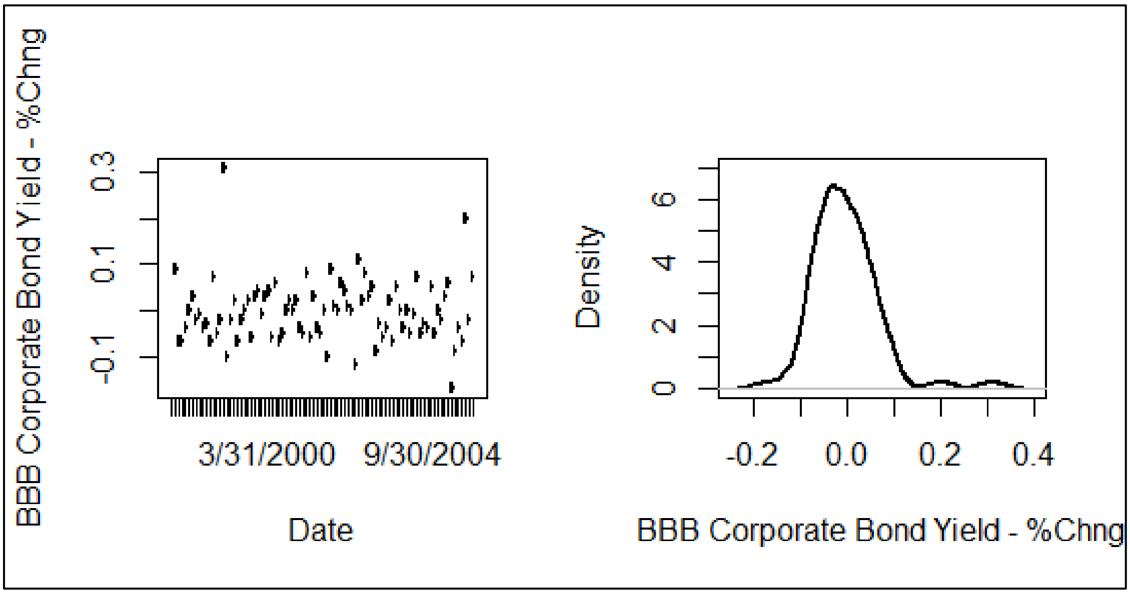

Figure 16: Time Series and Kernel Density Plot – BBB Corporate Bond Yield (Percent Changes)

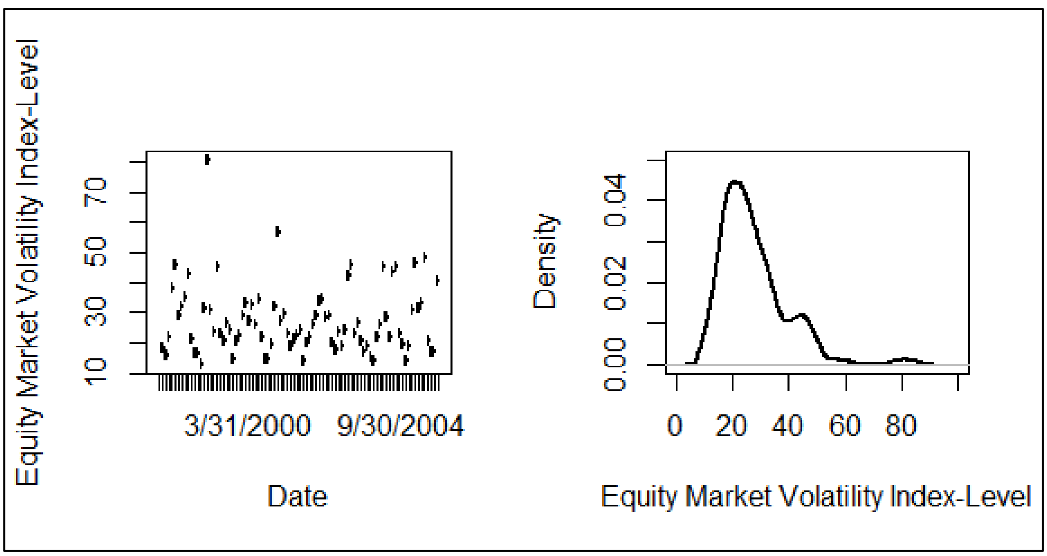

Figure 17: Time Series and Kernel Density Plot – Equity Market Volatility Index (Levels)

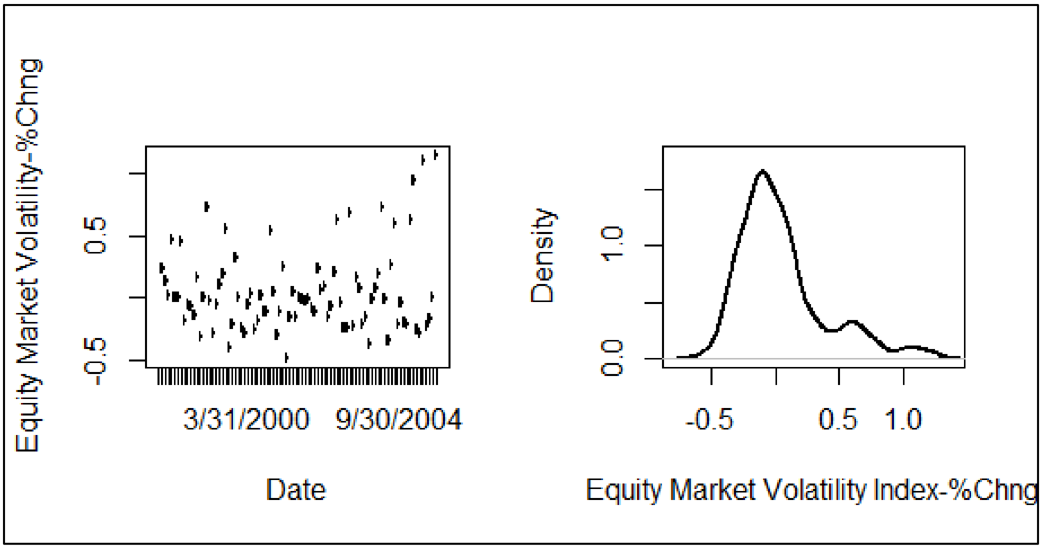

Figure 18: Time Series and Kernel Density Plot – Equity Market Volatility Index (Percent Changes)