**Methodologies for Scenario Generation and Dependency Structures in the Stress Testing of Credit Risk**

Michael Jacobs, Jr.[[1]](#footnote-1)

Accenture Consulting

Frank J. Sensenbrenner[[2]](#footnote-2)

Johns Hopkins University

Draft: September 22nd, 2017

**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. Furthermore, we find that the multiple equation VAR model outperforms the single equation *autoregressive* (“AR”) models according to various metrics across all modeling segments. .

Keywords: Stress Testing, CCAR, DFAST, Credit Risk, Financial Crisis, Model Risk, Vector Autoregression, Markov Switching Model, Scenario Generation

JEL Classification: C31, C53, E27, E47, E58, G01, G17, C54, G21, G28, G38.

**1 Introduction**

In the aftermath of the financial crisis (Acharya (2009), Demirguc-Kunt et al (2010)), regulators have utilized stress testing as a means to which to evaluate the soundness of financial institutions’ risk management procedures. The primary means of risk management, particularly in the field of credit risk (Merton, 1974), is through advanced mathematical, statistical and quantitative techniques and models, which leads to *model risk*. Model risk (Board of Governors of the Federal Reserve System, 2011) 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. *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 (Basel Committee for Banking Supervision 2006; 2009 a,b,c,d; 1010 a, b) 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 (Schuermann, 2014). 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. 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. There are two broad categories of model types in use. *Bottom-up models* are loan- or obligor-level models used by banks to forecast the expected losses of retail and wholesale loans for each loan. The expected loss is calculated for each loan, and then the sum of expected losses across all loans provides an estimate of portfolio losses, through conditioning on macroeconomic or financial / obligor specific variables. The primary advantages of bottom-up models are the ease of modeling heterogeneity of underlying loans and interaction of loan-level risk factors. The primary disadvantages of loan-level models are that while there are a variety of loan-level methodologies that can be used, these models are much more complex to specify and estimate. These models generally require more sophisticated econometric and simulation techniques, and model validation standards may more stringent. In contrast, *top-down models* are pool (or segment) level models used by banks to forecast charge-off rates by retail and wholesale loan types as a function of macroeconomic and financial variables. In most cases for these models, banks use only one to four macroeconomic and financial risk drivers as explanatory variables. These variables are usually determined by interaction between model development teams and line of business experts. The primary advantage of top-don models has been the ready availability of data and the simplicity of model estimation. The primary disadvantage of pool-level models is that borrower specific characteristics are generally not used as variables, except at the aggregate level using pool averages. Modeling challenges include determination of appropriate loss horizon (e.g., for CCAR it is a 9-quarter duration), determination of an appropriate averaging methodology, appropriate data segmentation and loss aggregation, as well as the annualization of loss rates. In this paper we consider top-down models.

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**

Since the dawn of modern risk management in the 1990s, ST has been a tool used to address the basic question of how exposures or positions behave under adverse conditions. Traditionally this form of ST has been in the domain of *sensitivity analysis* (e.g., shocks to spreads, prices, volatilities, etc.) or *historical scenario analysis* (e.g., historical episodes such as Black Monday 1987 or the post-Lehman bankruptcy period; or hypothetical situations such as modern version of the Great Depression or stagflation). These analyses are particularly suited to market risk, where data are plentiful, but for other risk types in data-scarce environments (e.g., operational, credit, reputational or business risk) there is a greater reliance on *hypothetical scenario analysis* (e.g., natural disasters, computer fraud, litigation events, etc.).

Regulators first introduced ST within the Basel I According, with the 1995 Market Risk Amendment (Basel Committee for Banking Supervision 1988, 1996). Around the same time, the publication of RiskMetricsTM 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 (CGFS, 2000). 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 (Mosser et al, 2001).

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 CreditMetricsTM (J.P. Morgan, 1997) and CreditRisk+TM (Wilde, 1997), ST was not a component of such models. The commonality of all such credit portfolio models was subsequently demonstrated (Koyluoglu and Hickman, 1998), 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 CreditMetricsTM (J.P. Morgan, 1997) framework through macroeconomic stress testing on credit portfolios using credit migration matrices (Bangia, et al, 2002).

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 (The Board of Governors of the Federal Reserve System, 1996), guidance on counterparty credit risk (The Board of Governors of the Federal Reserve System, 1999), as well as country risk management (The Board of Governors of the Federal Reserve System, 2002).

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 (Foglia, 2009). 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 (Inanoglu and Jacobs, Jr., 2009), 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 (Schuermann, 2014). 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 CreditMetricsTM (M Jacobs, Jr., 2013). 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. In another paper, the author proposes a Bayesian casual network model, for ST of a bank (Rebonato, 2010). 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 (Jacobs, Jr. et al, 2015). Jacobs (2015) presents an analysis of the impact of asset price bubbles on standard credit risk measures and provides evidence that asset price bubbles are a phenomenon that must be taken into consideration in the proper determination of economic capital for both credit risk management and measurement purposes. The author also calibrates the model to historical equity prices and in in ST exercise project credit losses on both baseline and stressed conditions for bubble and non-bubble parameter estimate settings. Jacobs (2017) extends Jacobs (2015) by performing a sensitivity analysis of the models with respect to key parameters, empirically calibrates the model to a long history of equity prices, and simulates the model under normal and stressed parameter settings. While the author find statistically significant evidence that the historical S&P index exhibits only mild bubble behavior, this translates in underestimation of potential extreme credit losses according to standard measures by an order of magnitude; however, the degree of relative underestimation of risk due to asset price bubbles is significantly attenuated under stressed parameter setting in the model.

The relative merits of various risk measures and the aggregation of varying risk types, classic examples being *Value-at-Risk* (“VaR”) and related quantities, have been discussed extensively by prior research (Jorion 1997, 2006). An important result in the domain of modeling dependency structures is a general result of mathematical statistics due to Sklar (1956), allowing the combination of arbitrary marginal risk distributions into a joint distribution while preserving a non-normal correlation structure, readily found an application in finance. Among the early academics to introduce this methodology is Embrechts et al. (1999, 2002, 2003). This was applied to credit risk management and credit derivatives by Li (2000). The notion of copulas as a generalization of dependence according to linear correlations is used as a motivation for applying the technique to understanding tail events in Frey and McNeil (2001). This treatment of tail dependence contrasts to Poon et al (2004), who instead use a data intensive multivariate extension of extreme value theory, which requires observations of joint tail events. Inanoglu and Jacobs (2010) developing a coherent approach to aggregating different risk types for a diversified financial institutions. The authors model the main risks faced - market, credit and operational – that have distinct distributional properties, and historically have been modeled in differing framework, contributing to the modeling effort by providing tools and insights to practitioners and regulators.

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.

Clark and Ryu (2015) note 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. 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 jurisdiction. 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 admitted 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 (Clark and Ryu, 2015).

The available public literature on scenario generation is rather limited to date. Bidder and McKenna (2015) of the San Francisco Federal Reserve Bank argue 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. 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. Frame et al (2015) of the Atlanta Federal Reserve Bank present 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. 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 Estimation and 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 (Stock and Watson, 2001). In the ST application, we are mainly concerned with the forecasting and policy advisory functions, as stressed loss projections help banking risk manager and banking 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 macro-econometric paradigm started to take hold, VAR, a simple yet flexible way to model and forecast macroeconomic relationships (Sims, 1980). In contrast to the univariate autoregressive model (Box and Jenkins (1970); Brockwell and Davis. (1991); Commandeur and Koopman (2007)), 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  be a -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, an-dimensional vector valued time series also referred as *exogenous variables*, and in our context representing a set of macroeconomic factors. This gives rise to the(“vector autoregressive-moving average with exogenous variables”) representation:

 (3.1)

Which is equivalent to:

 (3.2)

Where,  andare *autoregressive lag polynomials* of respective orders**, s** and****, respectively, and****is the *back-shift operator* that satisfies****for any process**.** It is common to assume that the input processis generated independently of the noise process [[3]](#footnote-3). The autoregressive parameter matrices represent sensitivities of output variables to their own lags and to lags of other output variables, while the corresponding matrices  are model sensitivities of output variables to contemporaneous and lagged values of input variables[[4]](#footnote-4). It follows that the dependency structure of the output variables , as given by the autocovariance function, is dependent upon the parameters , and hence the correlations amongst the  as well as the correlation amongst the that depend upon the parameters . In contrast, in a system of univariate (“ autoregressive-moving average with exogenous variables”) models, the correlations amongst the  is not taken into account, hence the parameter vectors  have a diagonal structure (Brockwell and Davis, 1991).

In this study we consider a *vector autoregressive model with exogenous variables* (“VARX”), denoted by, which restricts the *Moving Averag*e (“MA”) terms beyond lag zero to be zero, or:

 (3.3)

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 (R Development Core Team, 2017). 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  will be time-varying. However, the process might be time-invariant conditional on an unobservable regime variable, denoting the state at timeout offeasible states. In that case, then the conditional probability density of the observed time seriesis given by:

 (3.4)

Whereis the VAR parameter in regime and  are the observations . Therefore, given a regime, the conditional system in expectation form can be written as:

 (3.5)

We define the innovation term as:

 (3.6)

The innovation process is a Gaussian, zero-mean white noise process having variance-covariance matrix:

 (3.7)

If theprocess is defined conditionally upon an unobservable regimeas in equation (3.9), the description of the process generating mechanism should be made complete by specifying the stochastic assumption of the MS-VAR model. In this construct, thefollows a discrete state homogenous Markov chain:

 (3.8)

Wheredenotes the parameter vector of the regime generating process. We estimate the MS-VAR model using MSBVAR the package in R (R Development Core Team, 2017).

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 in the system, and the effects of regime change. The second foundation is the statistics of *probabilistic functions of Markov chai*ns introduced by Baum and Petrie (1966) and Baum et al (1970). Furthermore, the MS-VAR model also encompasses the even older traditions of *mixtures of normal distributions* (Pearson, 1984) and the hidden Markov-chain of Blackwell and Koopmans (1957) and Heller (1965). Finally, another root can be found in the construction of rather basic Markov-chain regression models in econometrics (Goldfeld and Quandt, 1973). The first holistic approach to the statistical analysis to the statistical analysis of the Markov-switching model can be found in Hamilton (1988, 1989). 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 (Tjostheim, 1986).

**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 project-ions 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”)

For CCAR 2013, the Federal Reserve System used the same set of variables to define the supervisory adverse scenario as in 2012. 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 the purposes of this research, let us consider the supervisory base and severely adverse scenario in 2015.

Our model selection process imposed the following criteria in selecting input and output variables across both multiple VARMAX and univariate ARMAX models[[5]](#footnote-5):

* Transformations of chosen variables should indicate stationarity
* Signs of coefficient estimates are economically intuitive
* Probability values of coefficient estimates indicate statistical significance at conventional confidence levels
* Residual diagnostics indicate white noise behavior
* Model performance metrics (goodness of fit, risk ranking and cumulative error measures) are within industry accepted thresholds of acceptability Scenarios rank order intuitively (i.e., severely adverse scenario stress losses exceeding scenario base expected losses)

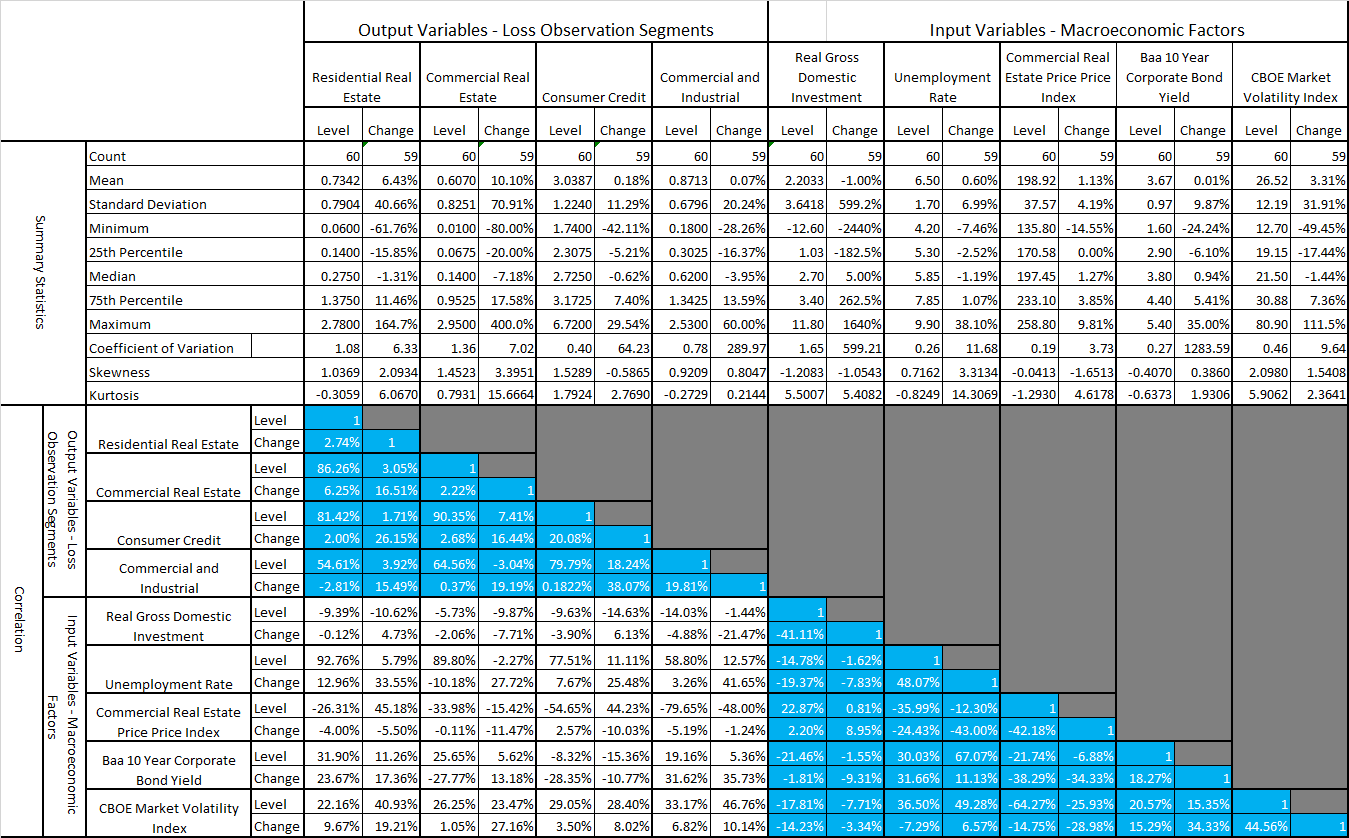
A diverse set of macroeconomic drivers representing varied dimensions of the economic environment, and a sufficient number of drivers) balancing the consideration of avoiding over-fitting) by industry standards (i.e., at least 2-3 and no more than 5-7 independent variables). According to these criteria, we identify the optimal set focusing of 5 of the 9 most commonly used national Fed CCAR MVs as input variables in the VARMAX model:

* Real Gross Domestic Investment (“RDIG”)
* Unemployment Rate (“UNEMP”)
* Commercial Real Estate Price Index (“CREPI”)
* BBB Corporate Credit Spread (“BBBCS”)
* CBOE’s Market Volatility Index (“VIX”)

Similarly, we identify the following loss segments (with loss measured by Gross Charge-offs – “GCOs”) according to the same criteria, in conjunction with the requirement that they cover the most prevalent portfolio types in typical traditional banking institutions:

* Residential Real Estate (“RESI”)
* Commercial Real Estate (“CRE”)
* Consumer Credit (“CONS”)
* Commercial and Industrial (“C&I”)

This historical data, 60 quarterly observations from 1Q01 to 4Q15, are summarized in Table 4.1 in terms of distributional statistics and correlations, as in Figures 4.1 through 4.9 of this section. A detailed description of these summary statistics are of Appendix 1, in this section we only provide a high level summary of the characteristics of the data. Across all series when looking at the time series dimension (in the left panels of the figures, in levels in the top and percent changes on the bottom), we observe that the credit cycle is clearly reflected, with indicators of economic or financial stress (health) and charge-off loss rates displaying peaks (troughs) in recession of 2001-2 and in the financial crisis of 2007-8, with latter episode dominating in terms of severity by **Table 4.1: Summary Statistics and Correlations of Historical Y9 Credit Loss Rates and Federal Reserve Macroeconomic Variables**

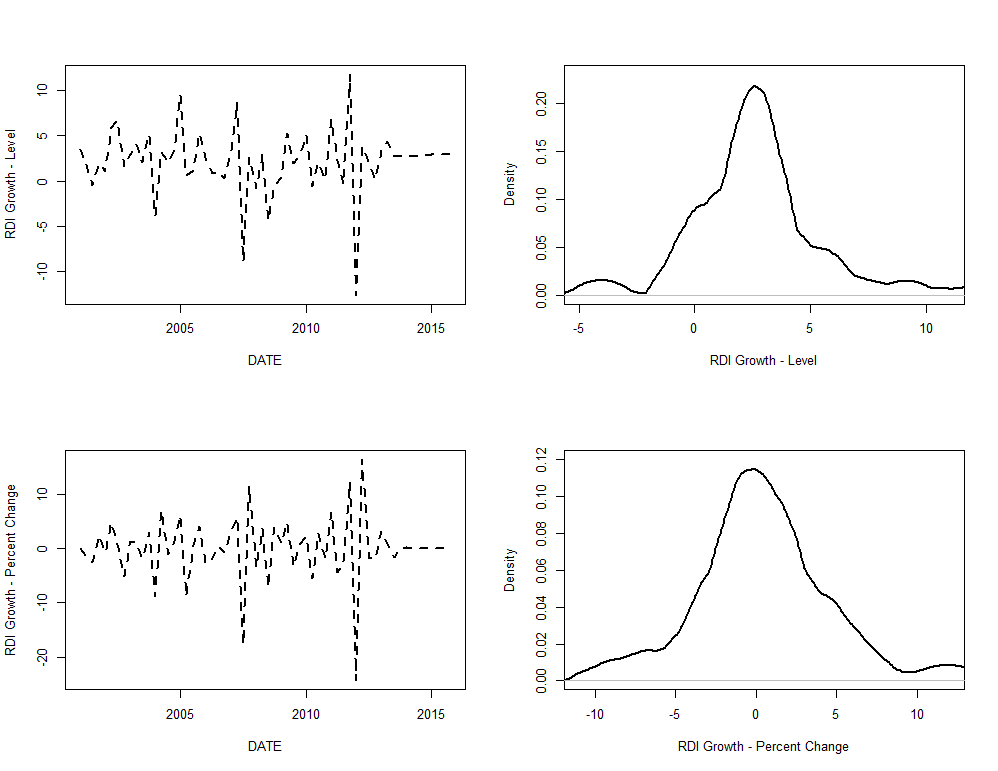


an order of magnitude. However, there are some differences in timing, extent and duration of these spikes across macroeconomic variables and loss rates. These patterns are reflected in the percent change transformations of the variables as well, with corresponding spikes in these series that correspond to the cyclical peaks and troughs, although there is also much more idiosyncratic variation observed when looking at the data in this form. Shifting focus to the smoothed histogram graphs (in the right panels of the figures, in levels in the top and percent changes on the bottom),we note that there are significant deviations from normality in terms of excess skewness and excess kurtosis relative to the Gaussian case, although the extent of these deviations exhibits significant variations across variables (e.g., in the case of the VIX, the non-normality is extreme, and obviously in the case of certain indices or loss rates the bounded domain are clear violations of normality). Furthermore, such deviations from normality are accentuated by an order of magnitude when examining these distributions of the variables in percent change form, which holds generally although with the extent of the deviations varying somewhat across variables. Finally, we not that in general the variation relative to the mean is an order of magnitude greater hen looking at percent changes relative to levels.

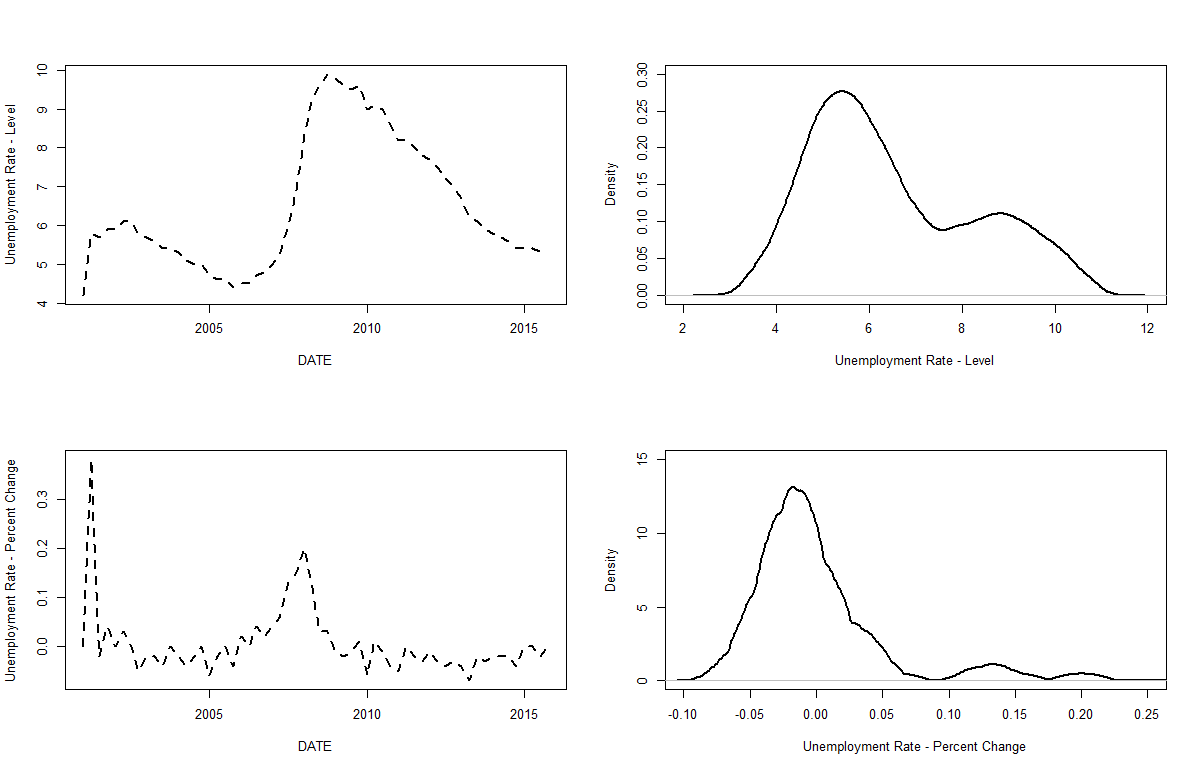
The correlations amongst all of the independent and dependent variables, in both their level and percentage change forms, are displayed I the bottom panel of Table 4.1. First, we will describe main features of the dependency structure within the group of input macroeconomic variables, then the same for the output loss rate variables, and finally the cross-correlations between these two groups. We observe that all correlations have intuitive signs and magnitudes that suggest significant relationships, although the latter are not large enough to suggest any issues with multicollinearity. While the correlations of the percent change transformations are generally lower, they are still intuitive and of reasonable magnitudes. We also note that percent changes of variables are negatively (positively) correlated with levels when indicators are those of economic strength (weakness). The correlation matrix amongst the macroeconomic variables appear in the lower right quadrant of the bottom panel of Table 4.1. For example, considering some of the stronger relationships amongst the levels, the correlations between UNEMP / VIX, CREP / UNEMP and BBBCY / RDIG are 36.5%, -36.0% and -21.5%, respectively. For example, considering some of the stronger relationships amongst the percent changes, the correlations between BBBCR / CREPI, UNEMP / RDIG, and VIX / CREP are 34.3%, -7.8% and 28.6%, respectively.

The correlation matrix amongst the credit loss rate variables appear in the upper left quadrant of the bottom panel of Table 4.1. For example, considering some of the stronger relationships of the levels, the correlations between CRE / RESI, CONS / CRE and CNI / CONS are 86.3%, 90.4% and 79.8%, respectively. For example, considering some of the stronger relationships amongst the percent changes, the correlations between CONS / CRE, CNI / CRE, and CNI / CONS are 26.2%, 15.5% and 38.1%, respectively.

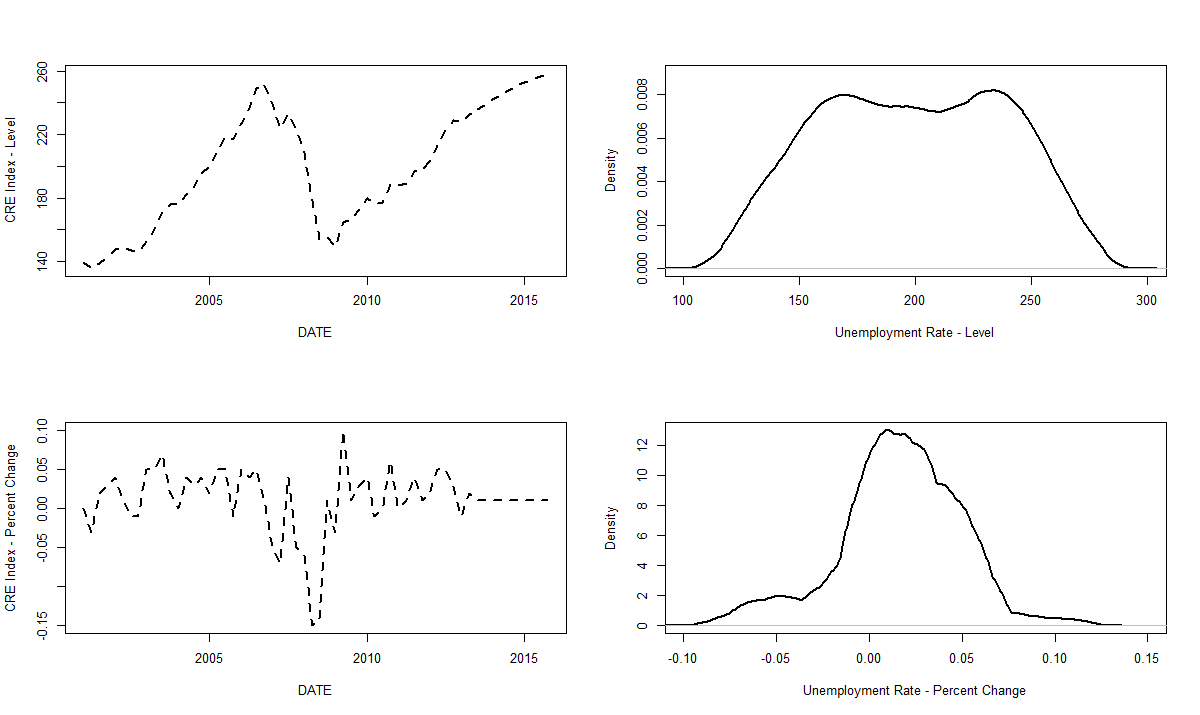
**Figure 4.1: Time Series and Kernel Density Plot – Real Domestic Investment Growth**



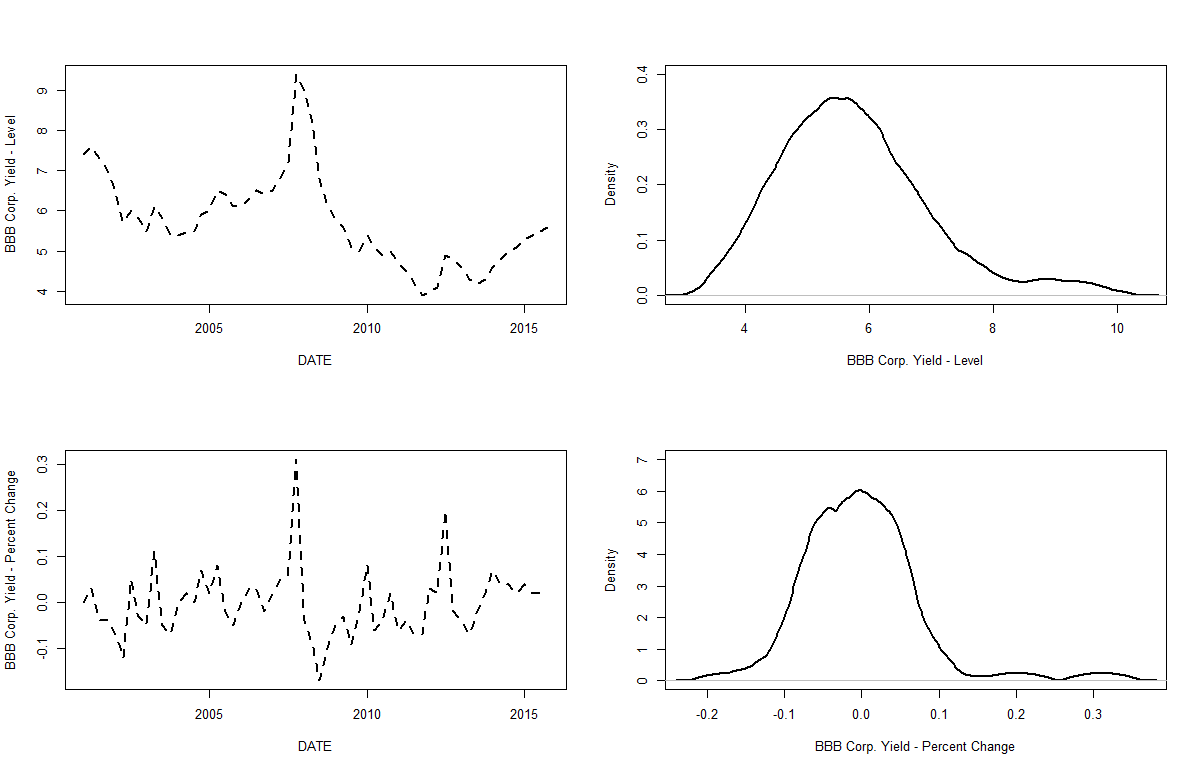
**Figure 4.2: Time Series and Kernel Density Plot – Unemployment Rate**



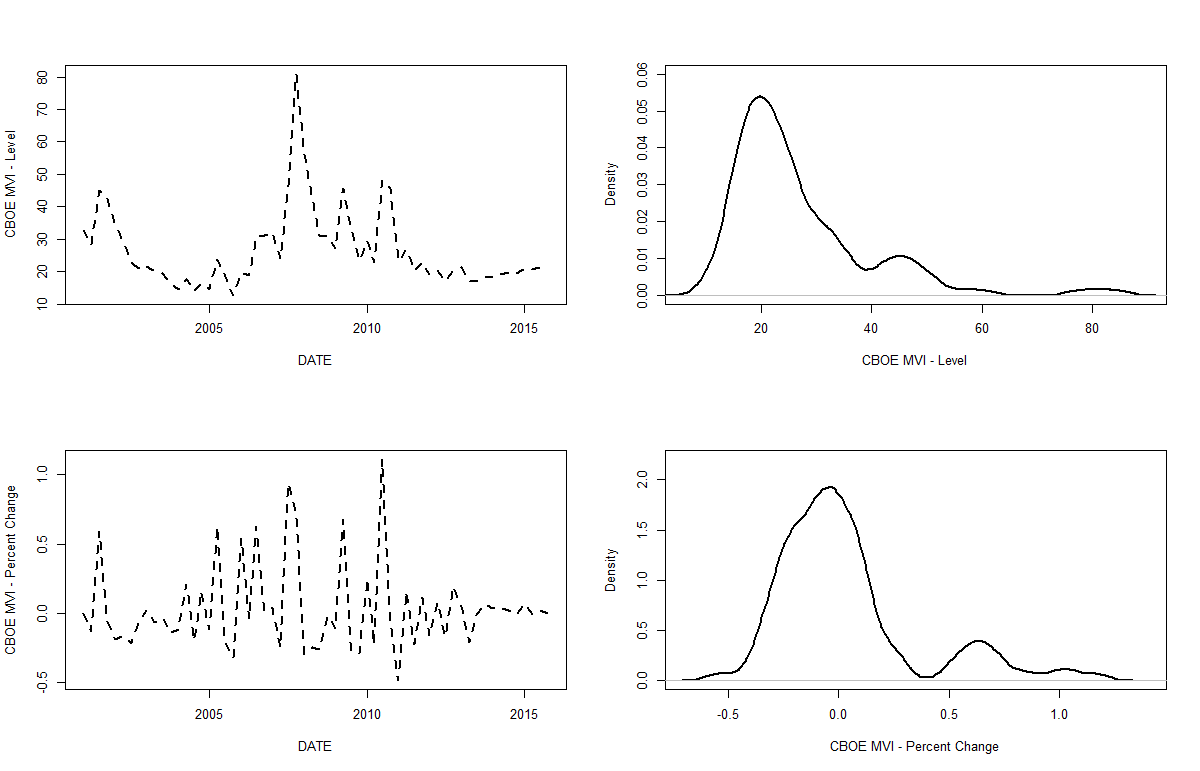
**Figure 4.3: Time Series and Kernel Density Plot – Commercial Real Estate Index**



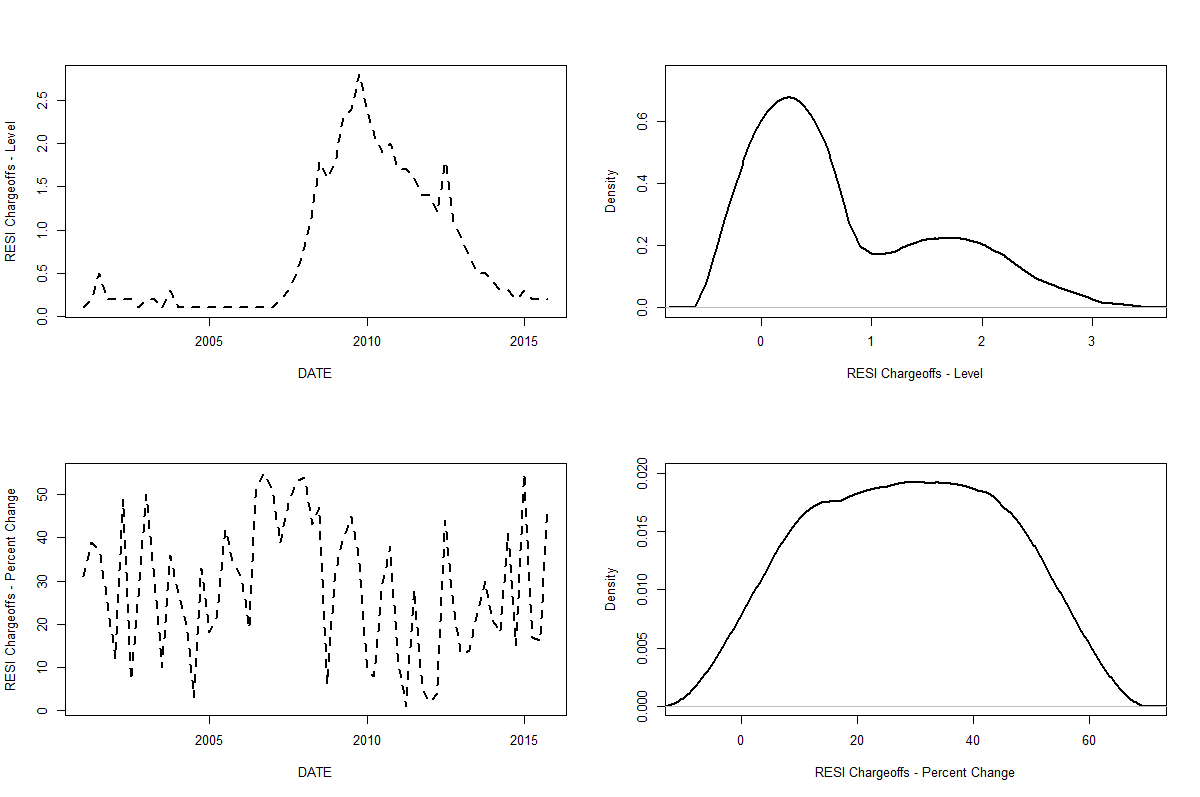
**Figure 4.4: Time Series and Kernel Density Plot – BBB Corporate Bond Yield**



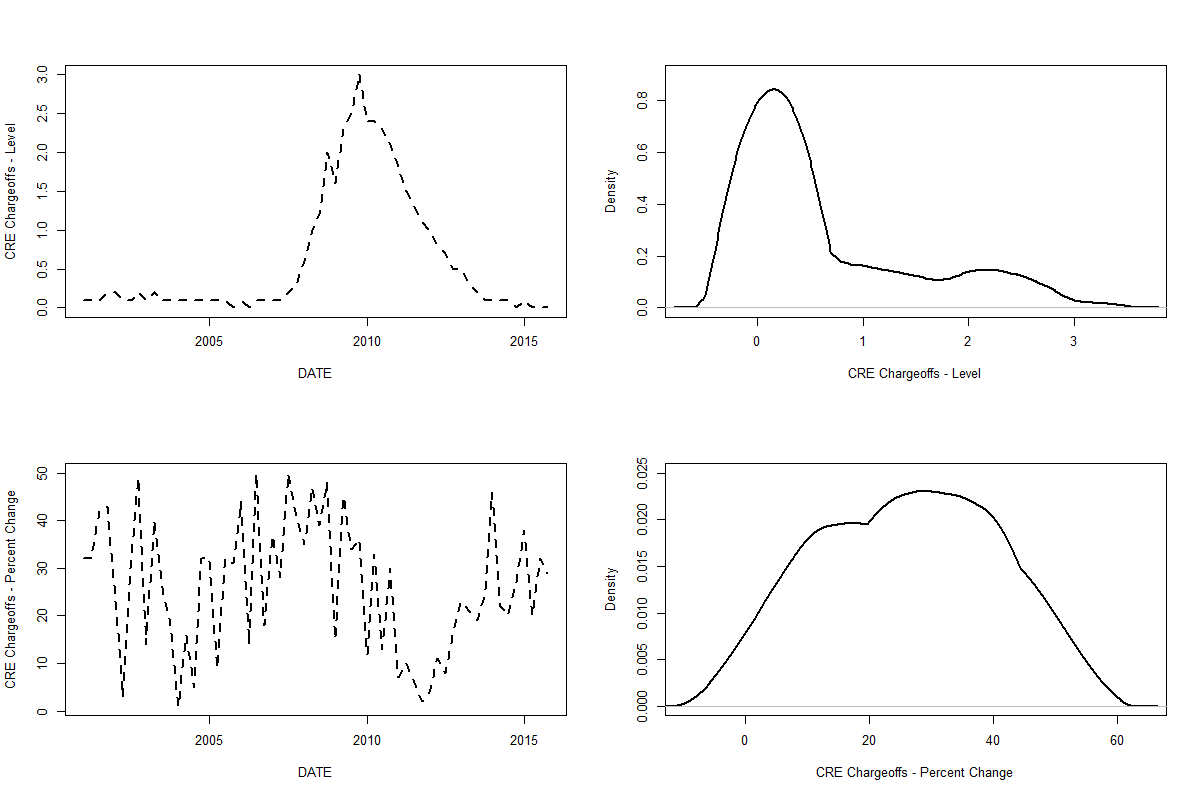
**Figure 4.5: Time Series and Kernel Density Plot – CBOE Market Volatility Index**



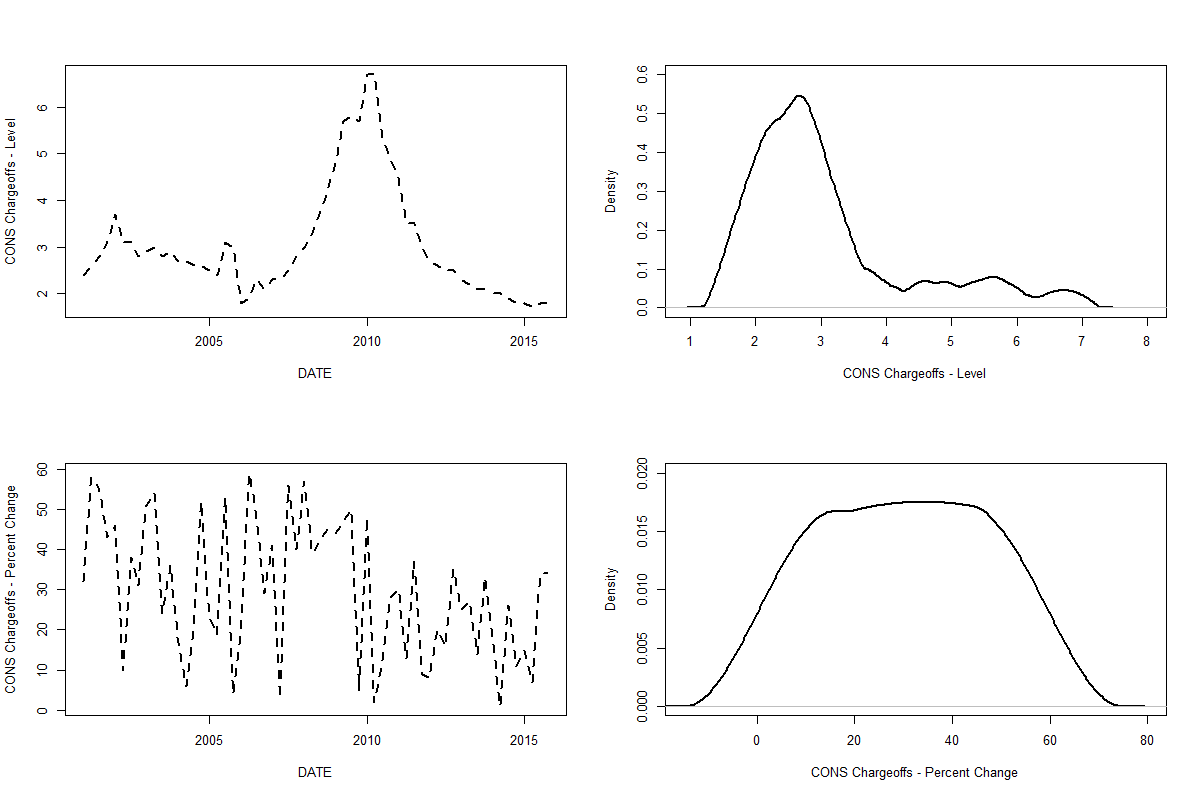
**Figure 4.6: Time Series and Kernel Density Plot – Residential Real Estate Loan Charge-off Rates**



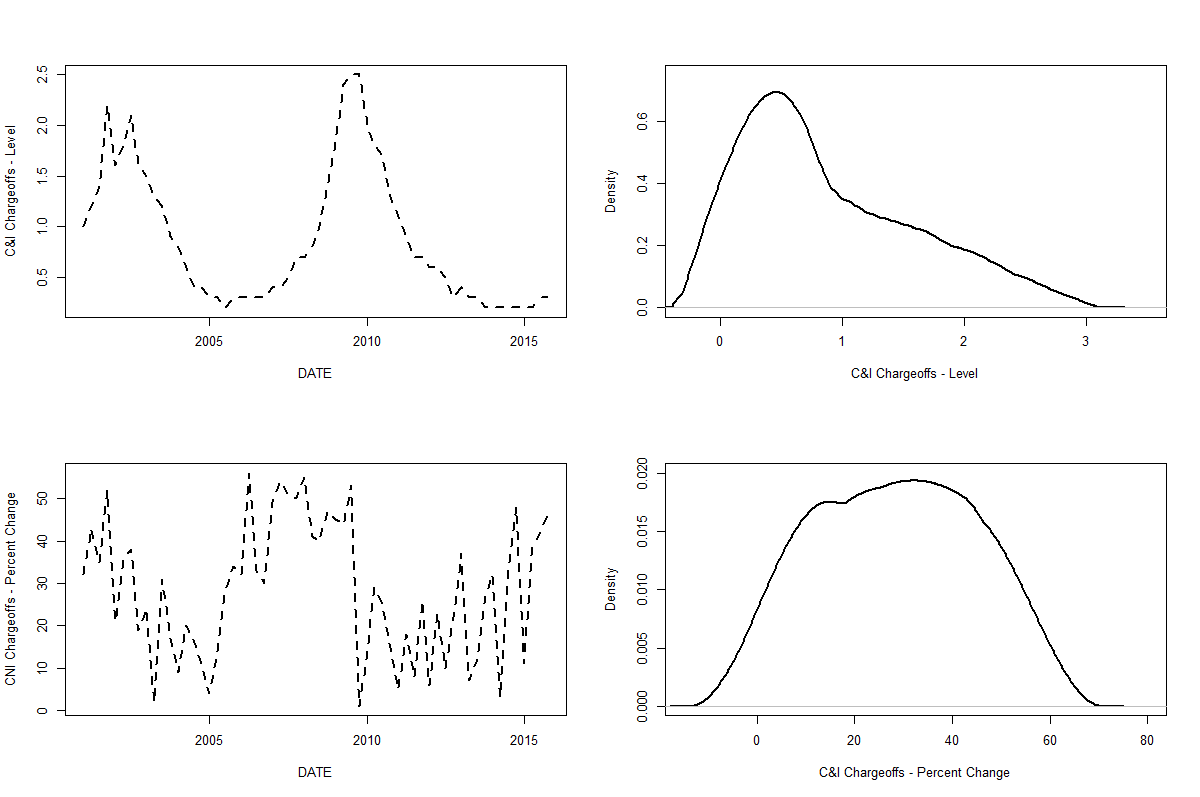
**Figure 4.7: Time Series and Kernel Density Plot – Commercial Real Estate Loan Charge-off Rates**



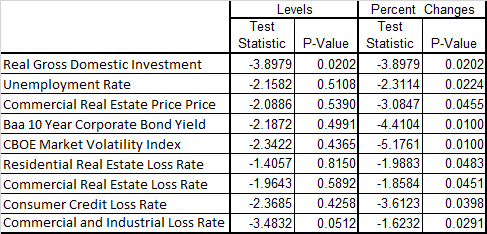
**Figure 4.8: Time Series and Kernel Density Plot – Consumer Loan Charge-off Rates**



**Figure 4.9: Time Series and Kernel Density Plot – Commercial and Industrial Loan Charge-off Rates**



**Table 4.2: Augmented Dickey-Fuller Stationarity Test Statistics of Credit Loss Rates and Macroeconomic Variables**



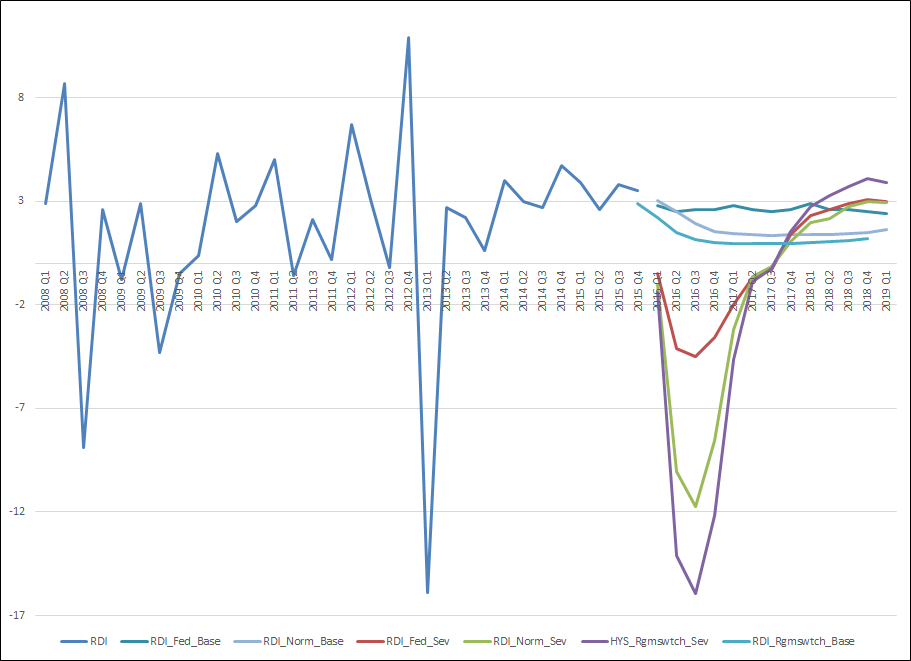
The correlation matrix amongst the credit loss rate and macroeconomic variables appear in the lower left quadrant of the bottom panel of Table 4.1. For example, considering some of the stronger relationships of the levels, the correlations between UNEMP / CRE, CREPI / CNI and UNEMP / RESI are 89.8%, 58.8% and 92.8%, respectively. For example, considering some of the stronger relationships amongst the percent changes, the correlations between UNEMP / CNI, UNEMP / CONS, and VIX / CRE are 41.7%, 25.5% and 27.2%, respectively.

In Table 4.2 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 4.2, 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.

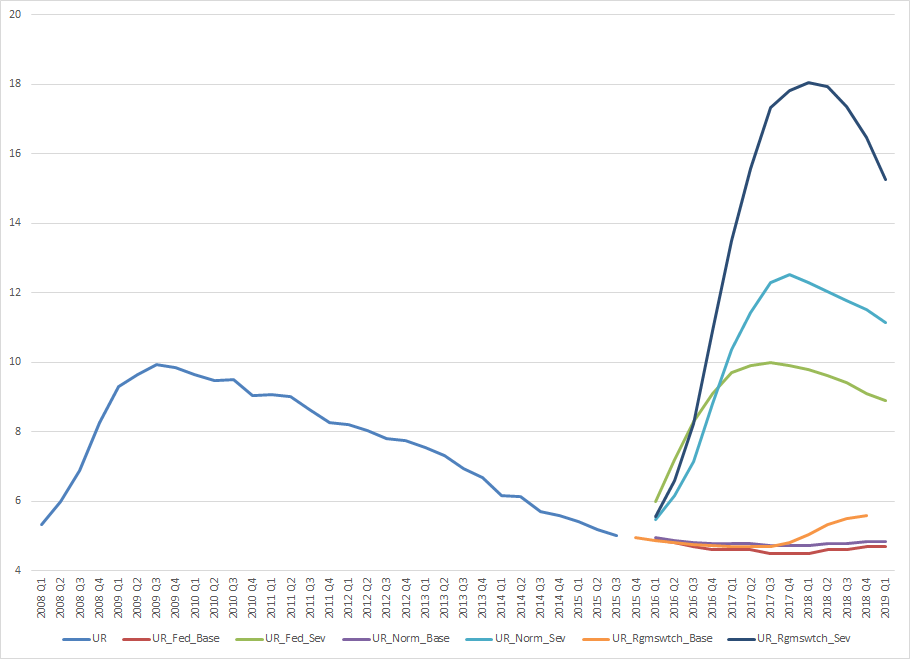
A critical modeling consideration for the MS-VAR estimation is the choice of process generation distributions for the normal and the stressed regimes. As described in the summary statistics of Table 4.1, we find that when analyzing the macroeconomic data in percent change form, there is considerable 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.[[6]](#footnote-6)

Another important modeling consideration with respect to scenario generation 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 99.0th 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.[[7]](#footnote-7) 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 4.10 through 4.14 where we show for each macroeconomic variable

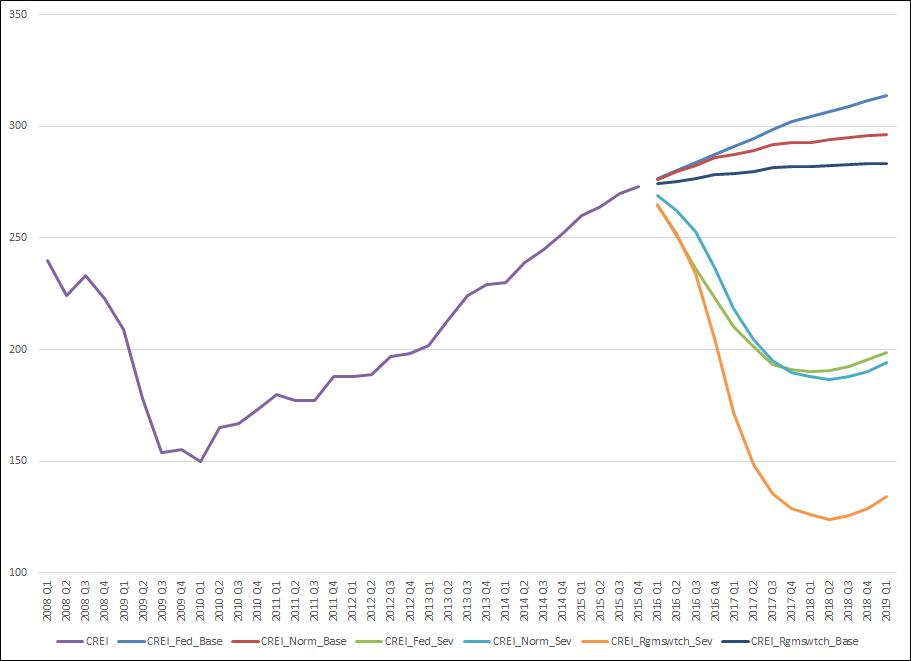
**Figure 4.10: Historical Time Series, Base and Severe Scenarios for the VAR, MS-VAR and Fed Models – Real Disposable Income Growth**



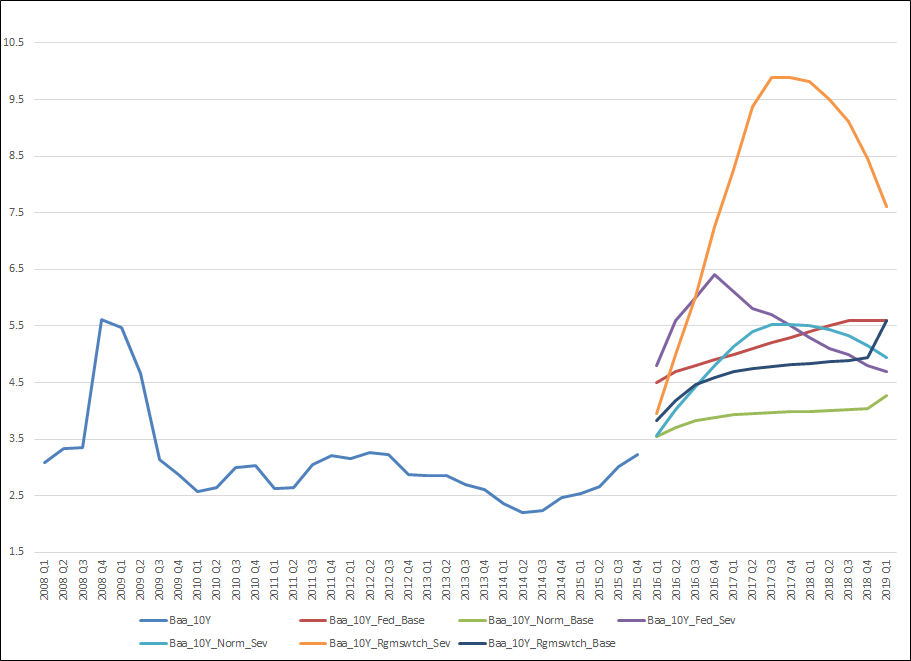
**Figure 4.11: Historical Time Series, Base and Severe Scenarios for the VAR, MS-VAR and Fed Models – Unemployment Rate**



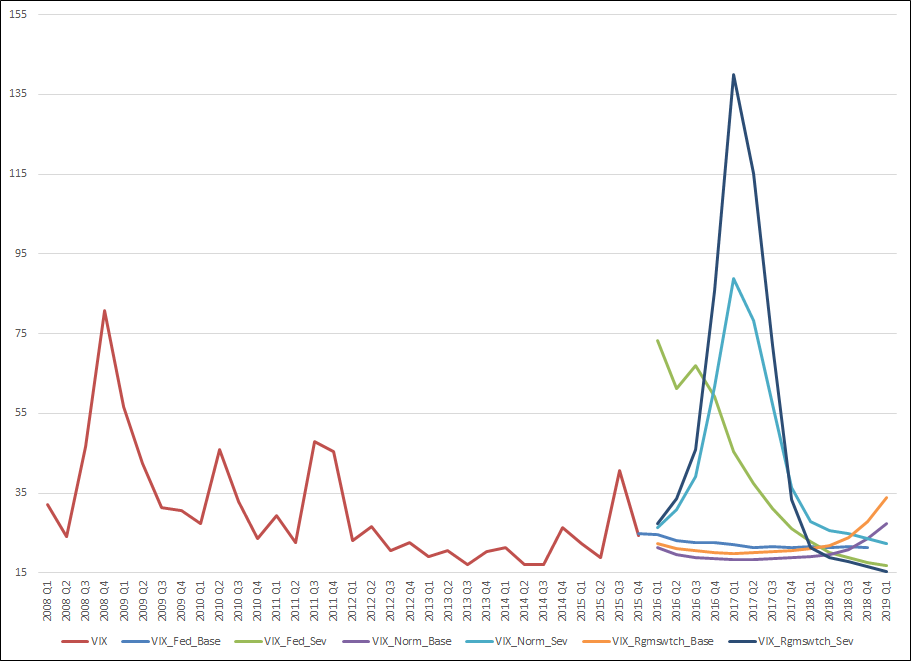
**Figure 4.12: Historical Time Series, Base and Severe Scenarios for the VAR, MS-VAR and Fed Models – Commercial Real Estate Price Index**



**Figure 4.13: Historical Time Series, Base and Severe Scenarios for the VAR, MS-VAR and Fed Models – Baa Corporate Credit Spread**



**Figure 4.14: Historical Time Series, Base and Severe Scenarios for the VAR, MS-VAR and Fed Models – VIX Equity Market Volatility Index**



the Base nd Severe scenarios for the VAR and MS-VAR models[[8]](#footnote-8), and also compare this to the corresponding Fed scenarios, along the historical time series. The alternative scenarios are summarized in Table 4.3 and figures 4.1 through 4.5. We make the following general conclusions regarding the different scenario generation methods (detailed description is given in Appendix 2):

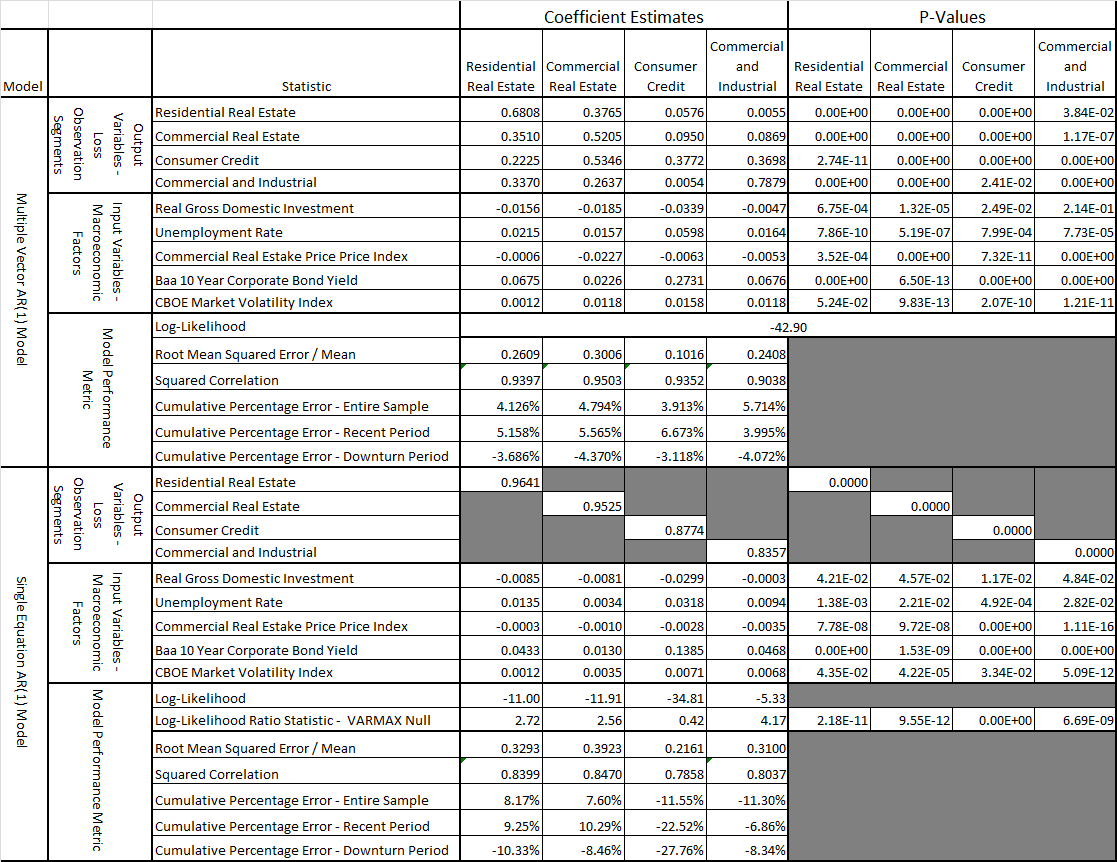
* In the Severe scenario, the MS-VAR model is far more conservative than the VAR model, and always at least match and in some cases can well exceed historical peaks or troughs in the adverse direction.
* In terms of magnitude, the VAR model is similar to the Fed scenarios, but the trajectories of either 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 estimation results are summarized in Table 4.3. The top panel tabulates the results of the VAR(1) estimation of a 4-equation system, while the bottom panel tabulates the results of the single equation AR(1) models for each portfolio segment separately. Before proceeding to discuss detailed results for each segment, we highlight the main conclusions of this analysis (a detailed discussion of the estimation results for each segment is given in Appendix 3):

* In both the VAR and AR models, all coefficient estimates are of intuitive sign, statistically significant at conventional confidence levels, although we note that the significance levels are generally at higher levels for the VAR as comparted to the AR models
* Residual diagnostics reveal lack of serial autocorrelation and a Gaussian distribution (refer to Figures 9.1 through 9.15 of Appendix 3) in both VAR and AR models, although we note that the quality of residuals if somewhat better for the VAR as comparted to the AR models
* Across all 4 segments, according to the likelihood ratio statistic, we reject the hypothesis that the restrictions of the single equation AR models are justified
* The results of the estimation are broadly consistent across the VAR and AR models, but with a few notable differences, such that the autocorrelation terms are larger in the AR models than in the VAR model
* The VAR models show greater sensitivity to macroeconomic factors than do the AR models
* The VAR models are generally more accurate according to standard measures of model fit with respect to each segment
* The VAR is more conservative than the AR as by measured by cumulative 9-quarter percentage error in the sense of under-predicting (over-predicting) to a lesser degree during the downturn (recent) period

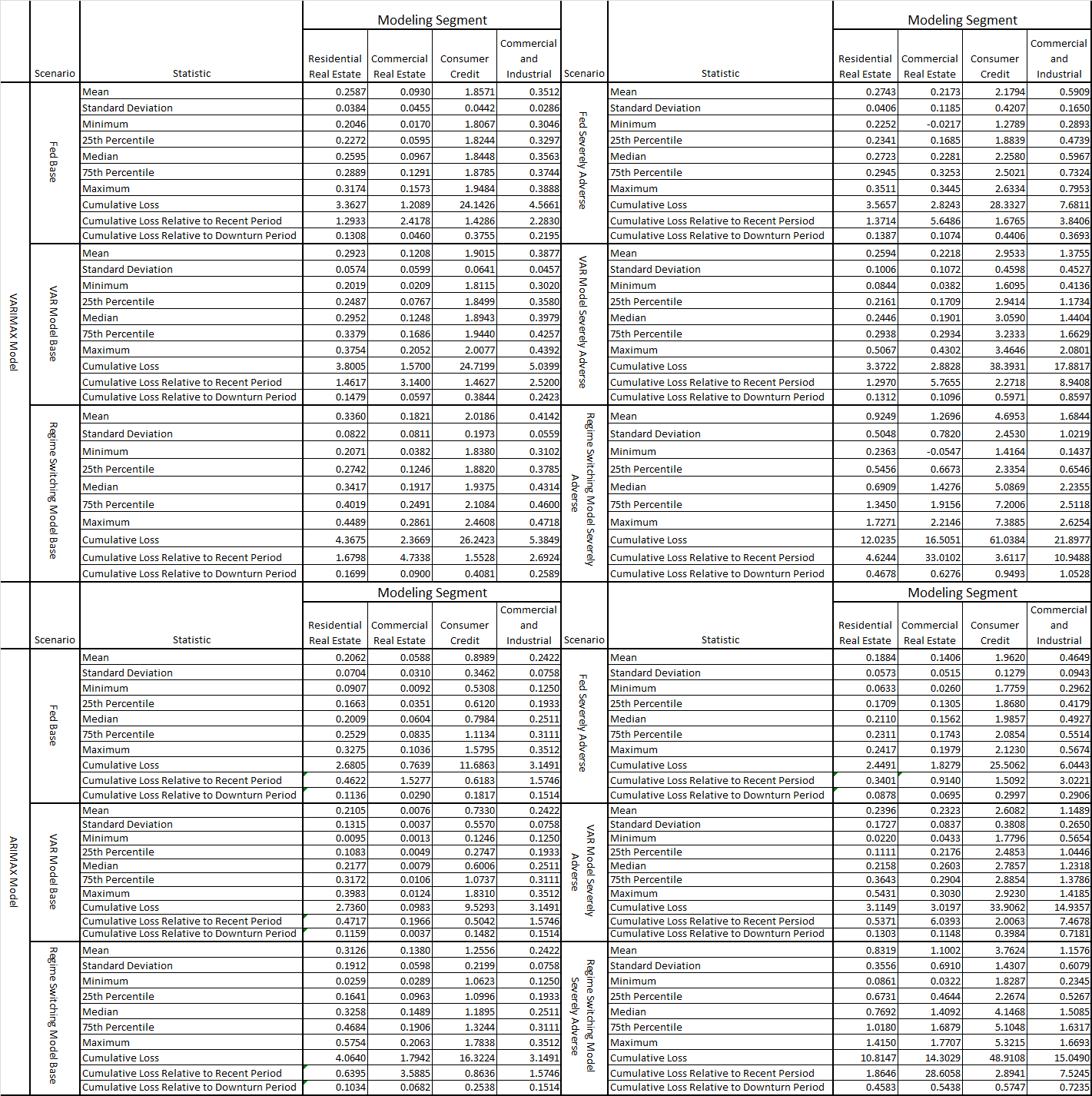
The results of the scenario analysis with respect to the credit loss segments, for both AR and VAR estimation, as well as for the three scenario generation methodologies (Fed, VAR and MS-VAR), are shown in Table 4.4 and ion Figures 4.15 through 4.18. The results across modeling segments are in line with the scenarios analysis as per macroeconomic variable as discussed in this section, but these results in terms of conser4vatism of the severe forecasts are accentuated in the VAR model and dampened in the AR models. In the Severe scenario, the MS-VAR model is far more conservative than the VAR model, reflecting the greater sensitivity to macroeconomic factors noted in the estimation results, and always at least match and in some cases can well exceed historical peaks or troughs in the adverse direction.

**Table 4.3: Vector Autoregressive vs. Single Equation Autoregressive Model Estimation Compared (Fed Macroeconomic Variables and Aggregate Y9 Bank Chargeoffs)**

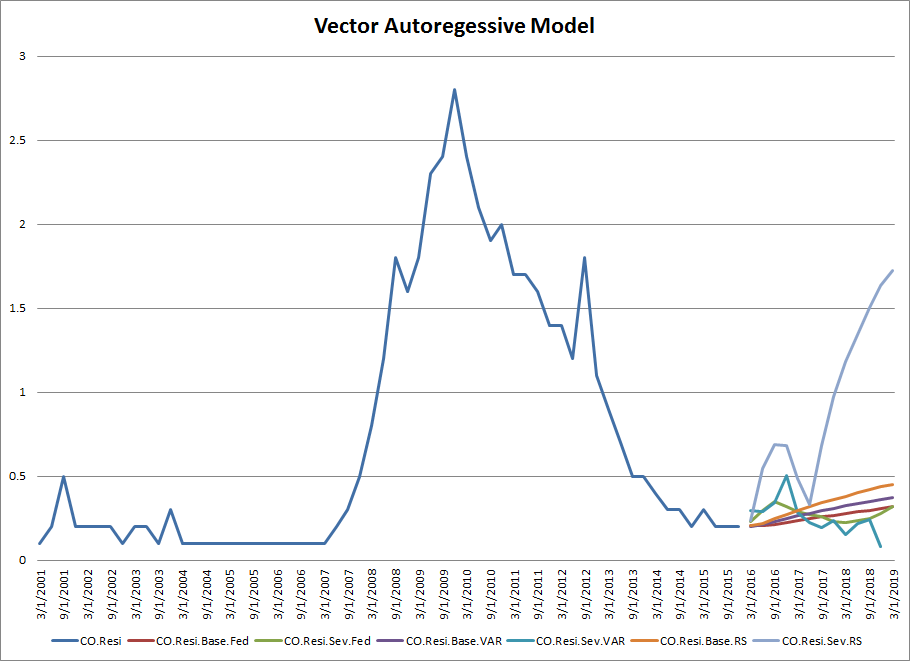
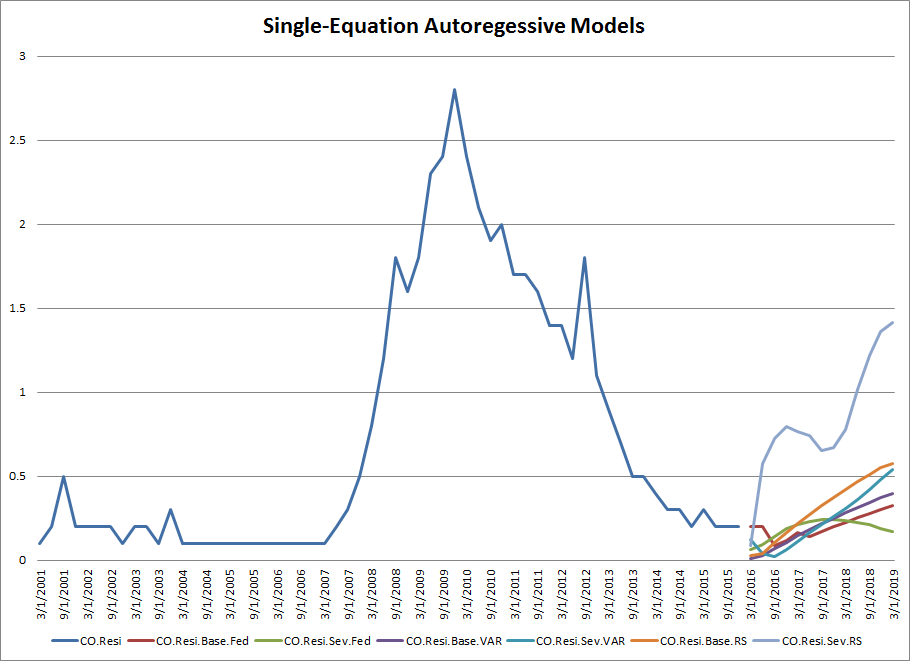


As an example, in the case of the C&I segment in the VAR estimation and as measured by the cumulative loss relative to that in the downturn period in VAR estimation, in the C&I segment this multiple is 1.05 in the MS-VAR model but only 0.85 (0.36) in the VAR (FED) scenario generation models. However, in the corresponding multiple is 0.75 in the MS-VAR model but only 0.71 (0.29) in the VAR (FED) scenario generation models.

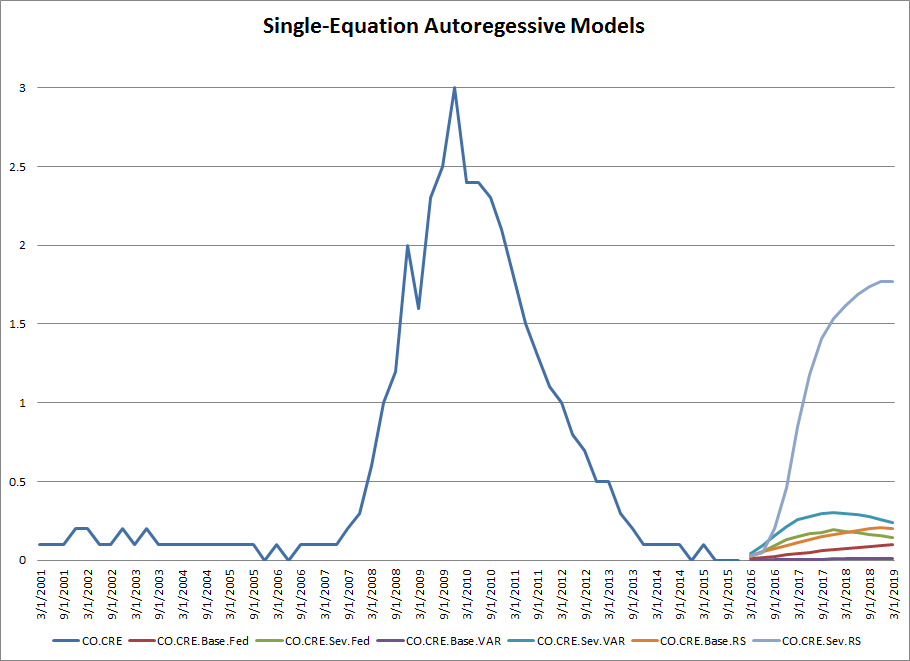
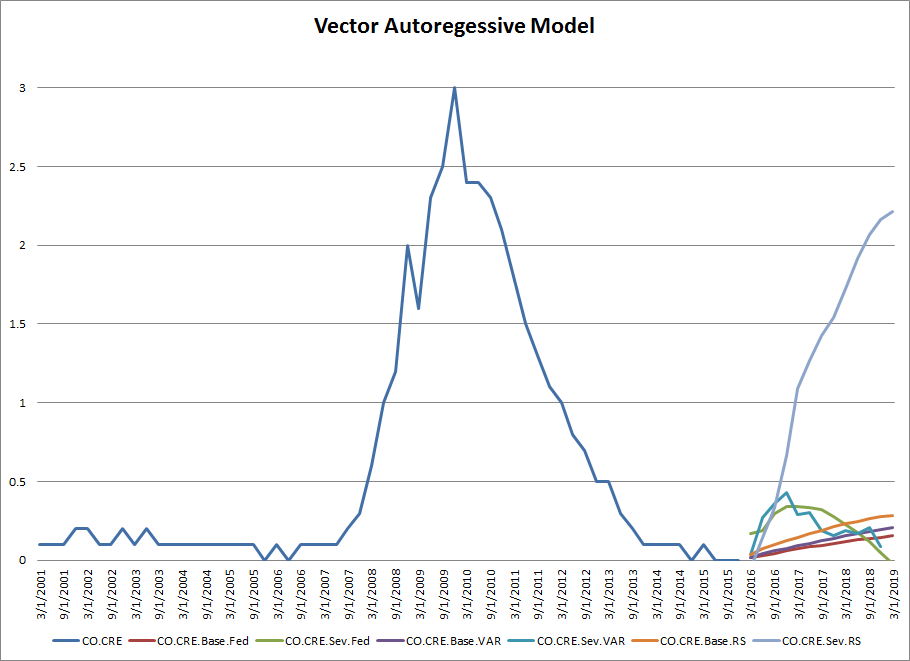
**Table 4.4: Vector Autoregressive vs. Single Equation Autoregressive Model Estimation and Gaussian Vector Autoregressive vs. Regime Switching Vector Autoregressive Scenario Generation Compared (Fed Macroeconomic Variables and Aggregate Y9 Bank Chargeoffs)**



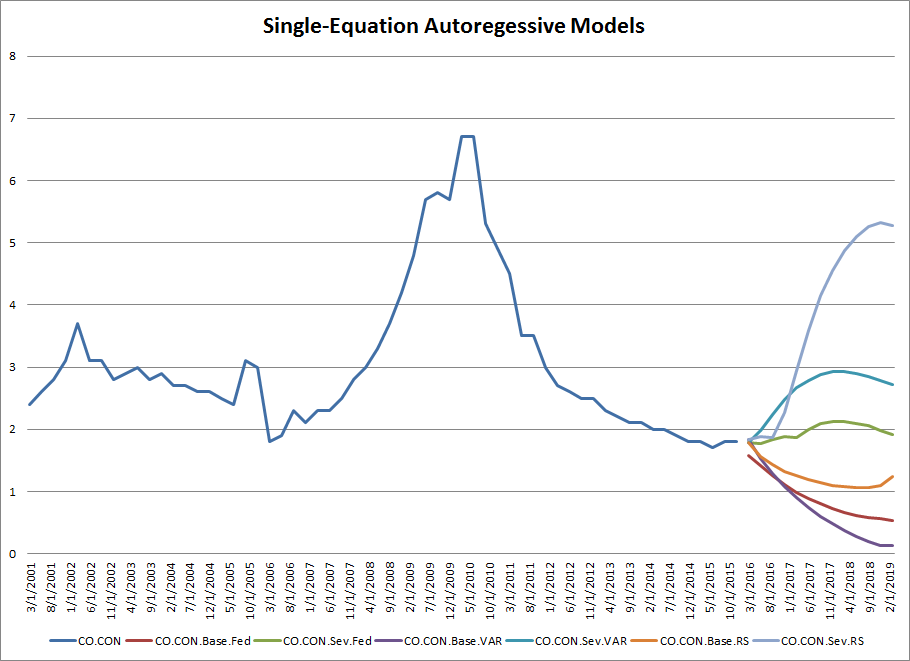
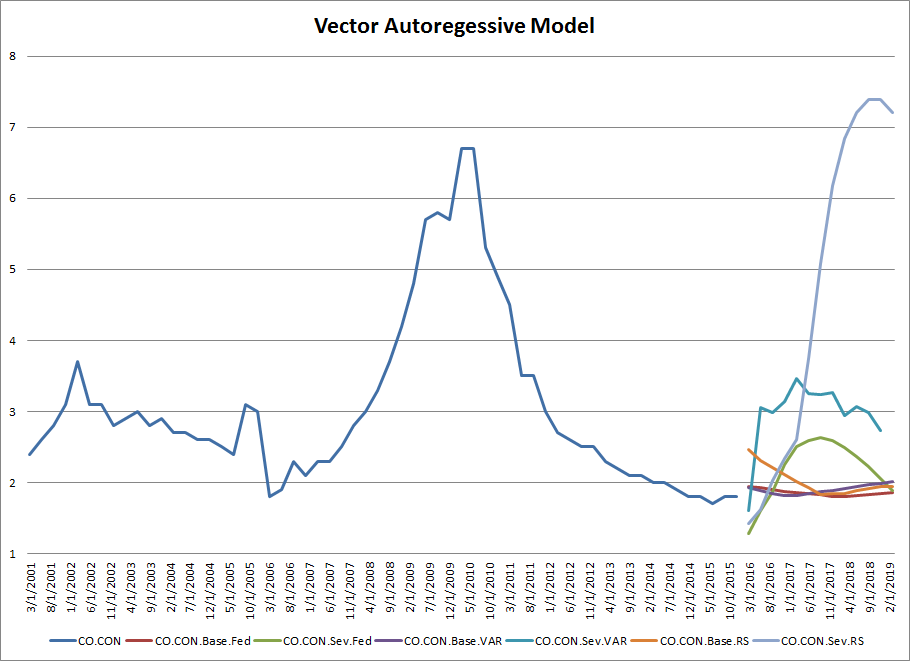
**Figure 4.15: Vector Autoregressive vs. Single Equation Autoregressive Model Estimation and Gaussian Vector Autoregressive vs. Regime Switching Vector Autoregressive Scenario Generation Compared - Residential Real Estate**

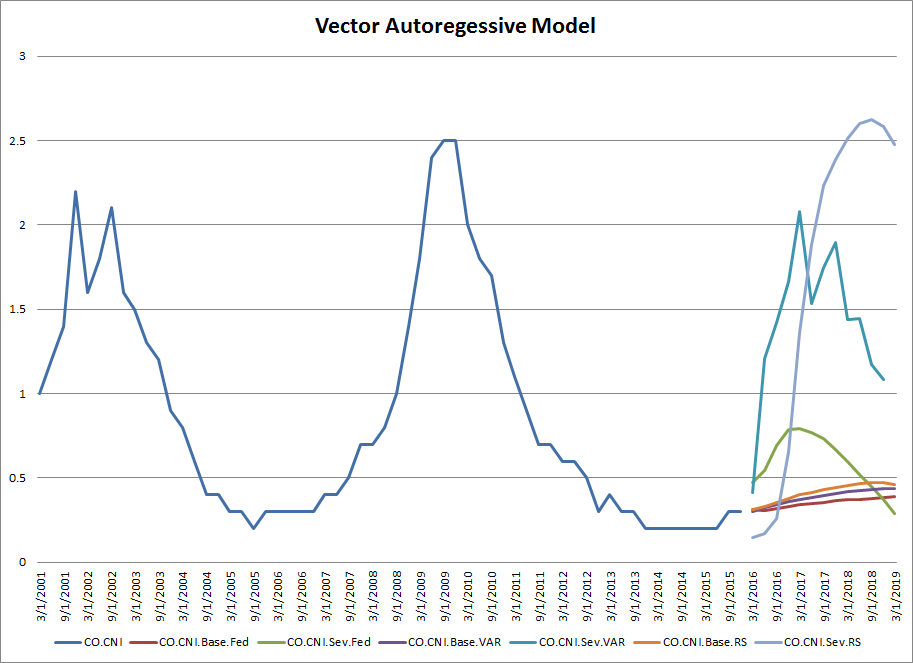
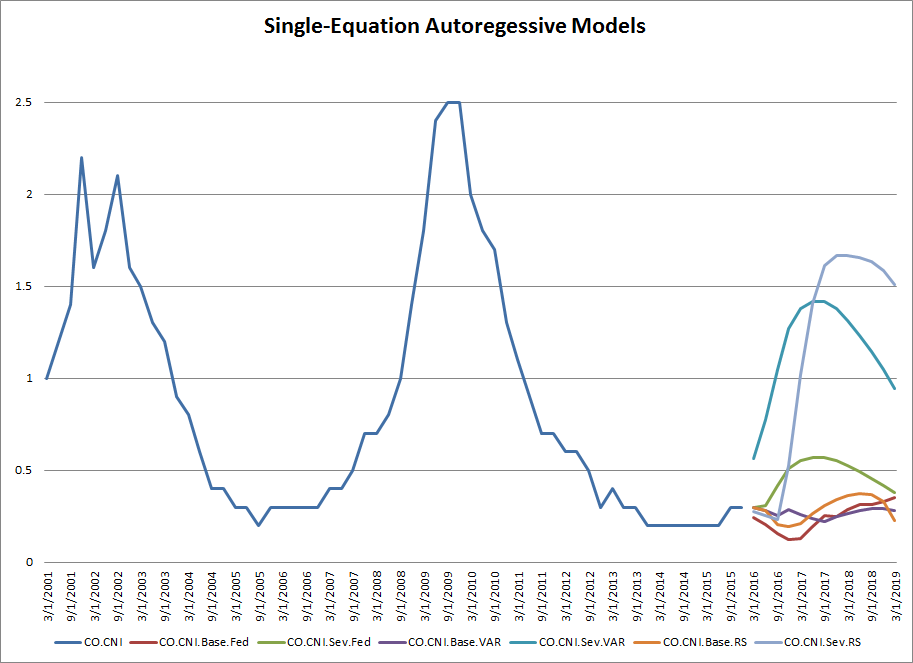
**Figure 4.16: Vector Autoregressive vs. Single Equation Autoregressive Model Estimation and Gaussian Vector Autoregressive vs. Regime Switching Vector Autoregressive Scenario Generation Compared - Commercial Real Estate**



**Figure 4.17: Vector Autoregressive vs. Single Equation Autoregressive Model Estimation and Gaussian Vector Autoregressive vs. Regime Switching Vector Autoregressive Scenario Generation Compared – Consumer Loans**



**Figure 4.18: Vector Autoregressive vs. Single Equation Autoregressive Model Estimation and Gaussian Vector Autoregressive vs. Regime Switching Vector Autoregressive Scenario Generation Compared – Commercial and Industrial Loans**

**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 mod-el, 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.

We have considered the case of banks that model the risk of their portfolios using top-of-the-house modeling techniques. We have addressed an issue of how to incorporate the correlation of risks amongst the different segments. An approach to incorporate this consideration of a dependency structure was proposed, and the bias that results from ignoring this aspect is quantified, through estimating a vector autoregressive (VAR) time series models for credit loss using Fed Y9 data. We found that the multiple equation VAR model outperforms the single equation autoregressive (AR) models according to various metrics across all modeling segments. The results of the estimation are broadly consistent across the VAR and AR models, but with a few notable differences (e.g., most segments exhibit significant but mild autocorrelation, and different subsets of the macro variables are significant across different segments). Across all 4 segments, according to the likelihood ratio statistic, we reject the hypothesis that the restrictions of the single equation AR models are justified. The VAR models are generally more accurate according to standard measures of model fit with respect to each segment. It is inconclusive whether the VAR or AR models are more or less conservative as measured by cumulative 9-quarter loss.

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

**6 References**

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.

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.

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

*Amendment to the capital accord to incorporate market risks* (1996).

. "The committee on global financial systems." *Stress testing by large financial institutions:: urrent practice and aggregation issues* (April, 2000).

*International convergence of capital measurement and capital standards:: a revised framework* (2006).

*Principles for sound stress testing practices and supervision - consultative paper (May – No. 155)* (2009a).

*Guidelines for computing capital for incremental risk in the trading book* (2009b).

*Revisions to the Basel II market risk framework* (2009c).

*Analysis of the trading book quantitative impact study (October)* (2009d).

*Strengthening the resilience of the banking sector - consultative document (December)* (2009e).

*Basel III:: a global regulatory framework for more resilient banks and banking systems (December)* (2010a).

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

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.

, 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.

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

R. Bidder and A. McKenna, Robust Stress Testing, Federal Reserve Bank of San Francisco *Working Paper* **September**, (2015).

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.

Board of Governors of the Federal Reserve System, 1996, “Joint Policy Statement on Interest Rate Risk”, Supervisory Letter 96-13, May 23rd.

, 1999, “Supervisory Guidance Regarding Counterparty Credit Risk Management”, Supervisory Letter 99-03, February 1st.

, 2002, “Interagency Guidance on Country Risk Management”, Supervisory Letter 02-05, March 8th.

, 2011, “Supervisory Guidance on Model Risk Management”, Supervisory Letter 11-7, April 4th.

G. Box and G. Jenkins, *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco, 1970.

P.J. Brockwell and R.A. Davis, *Time Series: Theory and Methods*, Springer-Verlag, New York, 1991.

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).

J. J. F. Commandeur and S.J. Koopman, *Introduction to State Space Time Series Analysis*, Oxford University Press, New York, 2007.

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).

Embrechts, P.,McNeil, A.J., Straumann, D., 1999, Correlation: pitfalls and alternatives, Risk 12,

69–71.

, McNeil, A.J., Straumann, D., 2002. Correlation and dependence in risk management: properties and pitfalls., in Dempster, M.A.H., ed.:, Risk Management: Value at Risk and Beyond, (Cambridge University Press, Cambridge, UK), 176–223.

, Lindskog, F., McNeil, A.J., 2003, Modeling dependence with copulas and applications to risk management, in: Rachev, S., ed.: Handbook of Heavy Tailed Distributions in Finance, (Elsevier, Rotterdam), 329–384.

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).

Foglia, A., Stress Testing Credit Risk: A Survey of Authorities’ Approaches, *International Journal of Central Banking*, **5**(3), (2009), 9-37.

Frey, R., McNeil, A.J., 2001, Modeling dependent defaults, Working paper, ETH Zurich.

Goldfeld, S.M., and R.E. Quandt. A Markov Model for Switching Regressions, *Journal of Econometrics*, **1**(1), (1973), 3-15.

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.

, A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle, *Econometrica*, **57** (2), (1989), 357-384.

Hanan, E.J., The Identification Problem for Multiple Equation Systems with Moving Average Errors, *Econometrica,* **39**, (1971), 751-766.

, *The Statistical Theory of Linear Systems*, John Wiley, New York, 1988.

Heller, A., On Stochastic Processes Derived from Markov Chains, *The Annals of Mathematical Statistics*, **36**(4), (1965), 1286-1291.

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.

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.

, Stress Testing Credit Risk Portfolios, *Journal of Financial Transformation*, **37**, (2013), 53-75.

, 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.

, 2015, The impact of asset price bubbles on credit risk measures, *The Journal of Financial Risk Management* **4**, 251–266.

, 2017, Asset price bubbles and the quantification of credit risk capital with sensitivity analysis, empirical implementation and an application to stress testing, *Journal of Risk Model Validation* **11**(4), 1–35.

J.P. Morgan, *RiskMetricsTM*, Second Edition, J.P. Morgan, New York, 1994.

, *CreditMetricsTM*, First Edition, J.P. Morgan, New York, 1997.

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

(1997).Value at Risk:The New Benchmark for Controlling Market Risk, volume 2.

McGraw-Hill, New York.

, 2006, Value at Risk: The Benchmark for Managing Financial Risk, Third Ed, (McGraw Hill, New York, NY).

Kohn, R., Asymptotic Results for ARMAX Structures, *Econometrica*, **47**, (1979), 1295–1304.

Koyluoglu, H, and A Hickman, Reconcilable Differences, *Risk*, **October**, (1998), 56-62.

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

Li, D.X., 2000, On default correlation: a copula function approach, Journal of Fixed Income 9, 43–54.

Loregian.A, and A., Meucci. Neither “Normal” nor “Lognormal”: Modeling

Interest Rates across All Regimes, *Financial Analysts Journal*, 72(3), (2016), 68-82.

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

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.

Pearson, K. Contribution to the Mathematical Theory of Evolution, Philosophical *Transactions of the Royal Society*, **185**, (1894). 71-110.

Poon, S.-H., Rockinger, M., Tawn, J., 2004, Extreme value dependence in financial markets: diagnostics, models and financial implications, Review of Financial Studies 17, 581–610.

R Development Core Team, 2017: “R: A Language and Environment for Statistical Computing.” R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.

Rebonato, R., *Coherent Stress Testing: A Bayesian Approach*, Wiley, New York. 2010/

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

Sims, C.A., Macroeconomics and Reality, *Econometrica*, **48**, (1980), 1-48.

Stock, J.H., and M.W. Watson, Vector Autoregressions, *Journal of Economic Perspectives*, **15**(4), (2001), 101-115.

Tjøstheim, D., Estimation in Nonlinear Time Series Models, Stochastic Processes and their Applications, **21**(2), (1986), 251-273.

Wilde, T., *CreditRisk+. A Credit Risk Management Framework*, Credit Suisse First Boston, 1997.

**7 Appendix 1: Detailed Description of Summary Statistics and Distributional Properties - Macroeconomic Variables and Gross Charge-off Loss Rates**

The summary statistics of RESI are shown in the 1st and 2nd columns of Table 4.1, respectively in levels and percent changes, with the corresponding time series plots and histograms shown in Figure 7.6. The level of the RESI over the historical period averages 0.73%, with a median of 0.28%, displaying significant departures from normality in terms of right skewness (1.04) and thin-tails (kurtosis of -0.31). The series is somewhat volatile, displaying a coefficient of variation of 1.08, and ranging from 0.06% to 2.78% over the sample period. On the other hand, the percent of the RESI over the historical period averages 6.43%, with a median of -1.31%, displaying significant departures from normality in terms of right skewness (2.09) and fat-tails (kurtosis of 6.07), well in excess of the raw variable. The percent change of the series is also rather volatile in also excess of the raw variable, displaying a coefficient of variation of 2.09 and ranging from -61.8% to 164.7% over the sample period.

The summary statistics of CRE are shown in the 3rd and 4th columns of Table 4.1, respectively in levels and percent changes, with the corresponding time series plots and histograms shown in Figure 7.7. The level of the CRE over the historical period averages 0.61%, with a median of 0.14%, displaying significant departures from normality in terms of right skewness (1.36) and thin-tails (kurtosis of 0.79). The series is somewhat volatile, displaying a coefficient of variation of 1.45, and ranging from 0.01% to 2.95% over the sample period. On the other hand, the percent of the CRE over the historical period averages 10.10%, with a median of -7.18%, displaying significant departures from normality in terms of right skewness (7.02) and fat-tails (kurtosis of 15.67), well in excess of the raw variable. The percent change of the series is also rather volatile in also excess of the raw variable, displaying a coefficient of variation of 7.02 and ranging from -80.0% to 400.0% over the sample period.

The summary statistics of CONS are shown in the 5th and 6th columns of Table 4.1, respectively in levels and percent changes, with the corresponding time series plots and histograms shown in Figure 7.8. The level of the CONS over the historical period averages 3.04%, with a median of 2.72%, displaying significant departures from normality in terms of right skewness (1.51) and thin-tails (kurtosis of 1.79). The series is somewhat volatile, displaying a coefficient of variation of 0.40, and ranging from 1.74% to 6.72% over the sample period. On the other hand, the percent of the CONS over the historical period averages 0.18%, with a median of -0.62%, displaying significant departures from normality in terms of right skewness (0.59) and fat-tails (kurtosis of 1.79), well in excess of the raw variable. The percent change of the series is also rather volatile in also excess of the raw variable, displaying a coefficient of variation of 64.23 and ranging from -42.11% to 29.54% over the sample period.

The summary statistics of CNI are shown in the 7th and 8th columns of Table 4.1, respectively in levels and percent changes, with the corresponding time series plots and histograms shown in Figure 7.9. The level of the CNI over the historical period averages 0.87%, with a median of 0.62%, displaying significant departures from normality in terms of right skewness (0.92) and thin-tails (kurtosis of -2.27). The series is somewhat volatile, displaying a coefficient of variation of 0.78, and ranging from 0.18% to 2.53% over the sample period. On the other hand, the percent of the CNI over the historical period averages 0.078%, with a median of -3.95%, displaying significant departures from normality in terms of right skewness (0.80) and fat-tails (kurtosis of 0.21), well in excess of the raw variable. The percent change of the series is also rather volatile in also excess of the raw variable, displaying a coefficient of variation of 0.80 and ranging from -28.26% to 60.00% over the sample period.

The.ummary statistics of RDIG are shown in the 9th and 10th columns of Table 4.1, respectively in levels and percent changes, with the corresponding time series plots and histograms shown in Figure 7.1. The level of the RDIG over the historical period averages 2.20%, with a median of 2.70, displaying significant departures from normality in terms of left skewness (-1.21) and fat-tails (kurtosis of 5.50). The series is somewhat volatile, displaying a coefficient of variation of 1.65, and ranging from -12.6% to 11.8% over the sample period. On the other hand, the percent of the RDIG over the historical period averages -1.00%, with a median of 5.00%, displaying significant departures from normality in terms of left skewness (-1.05) and fat-tails (kurtosis of 5.41), in excess of the raw variable. The percent change of the series is also extremely volatile, displaying a coefficient of variation of 599.2 and ranging from -2,440.0% to 1,640.0% over the sample period.

The summary statistics of UR are shown in the 10th and 11th columns of Table 4.1, in levels and percent changes, respectively, with the corresponding time series plots and histograms shown in Figure 7.2. The level of the UR over the historical period averages 6.50%, with a median of 5.85$, displaying some departures from normality in terms of right skewness (0.82) and thin-tails (kurtosis of -0.82). The series is only somewhat volatile, displaying a coefficient of variation of 0.26, and ranging from 4.20% to 9.90% over the sample period. On the other hand, the percent change of the UR over the historical period averages 0.60%, with a median of -1.19%, displaying significant departures from normality in terms of right skewness (3.31) and fat-tails (kurtosis of 14.31), in excess of the raw variable. The percent change of the series is also extremely volatile, displaying a coefficient of variation of 11.68, and ranging from -7.46% to 38.10% over the sample period.

The summary statistics of CREPI are shown in the 12th and 13th columns of Table 4.1, respectively levels and percent changes, with the corresponding time series plots and histograms shown in Figures 7.3. The level of the CREPI over the historical period averages 198.9, with a median of 197.5, displaying rather mild departures from normality in terms of left skewness (-0.01) and thin-tails (kurtosis of -1.29). The series is only somewhat volatile, displaying a coefficient of variation of 0.19, and ranging from 135.8 to 258.8 over the sample period. On the other hand, the percent of the CREPI over the historical period averages 1.13%, with a median of 1.27%, displaying some departures from normality in terms of left skewness (-1.65) and fat-tails (kurtosis of 4.62), in excess of the raw variable. The percent change of the series is also extremely volatile, displaying a coefficient of variation of 3.73, and ranging from -14.55% to 9.81% over the sample period.

The summary statistics of BBBCS are shown in the 14th and 15th columns of Table 4.1, respectively levels and percent changes, with the corresponding time series plots and histograms shown in Figure 7.4. The level of the BBBCS over the historical period averages 3.7%, with a median of 3.80%, displaying very mild departures from normality in terms of left skewness (-0.41), and only moderate deviations from a Gaussian distribution with respect to thin-tails (kurtosis of -0.63). The series is only somewhat volatile, displaying a coefficient of variation of 0.27, and ranging from 1.6% to 5.4% over the sample period. On the other hand, the percent of the BBBCS over the historical period averages 0.01%, with a median of 0.97%, displaying significant departures from normality in terms of right skewness (0.39) and fat-tails (kurtosis of 1.93), well in excess of the raw variable. The percent change of the series is also extremely volatile, displaying a coefficient of variation of -1283.6, and ranging from -24.24% to 35.00% over the sample period.

The summary statistics of VIX are shown in the 16th and 14th columns of Table 4.1, respectively levels and percent change, with the corresponding time series plots and histograms shown in Figure 7.5. The level of the VIX over the historical period averages 26.52, with a median of 21.80, displaying significant departures from normality in terms of right skewness (2.10), as well as material deviations from a Gaussian distribution with respect to fat-tails (kurtosis of 5.91). The series is only somewhat volatile, displaying a coefficient of variation of 0.46, 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 3.31%, with a median of -1.44%, displaying significant departures from normality in terms of right skewness (1.54) and fat-tails (kurtosis of 2.36), far in excess of the raw variable. The percent change of the series is also extremely volatile, displaying a coefficient of variation of 9.64, and ranging from -49.45% to 111.5% over the sample period.

**8 Appendix 2: Detailed Description of Summary Statistics and Distributional Properties - Macroeconomic Variables and Gross Charge-off Loss Rates**

The scenario projections for the HYS are shown in Figure 4.1. 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 4.2. 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 4.3. 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 4.4. 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 4.5. 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 4.6. 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 4th 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.

**9 Appendix 3: Detailed Description of Vector Autoregressive vs. Single Equation Autoregressive Model Estimation Compared (Fed Macroeconomic Variables and Aggregate Y9 Bank Chargeoffs) - Model Estimation MLE Results and In-Sample Performance Metrics**

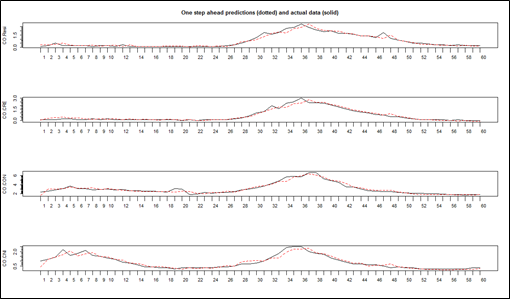
First, we discuss the results for the RRE segment. In the VAR model the autoregressive term is significantly positive and of substantial magnitude, a parameter estimate 0.68, which is also significant but rather somewhat in the AR model having an estimate 0.96. The cross autoregressive terms on the CRE, CONS and CNI are significantly positive in the VAR model and of lower of magnitude than the autocorrelation term, parameter estimates ranging in 0.22 to 0.35. The coefficient estimates of the macroeconomic sensitivities are of acceptable magnitude with respect to model sensitivity and range in absolute vales of 0.0012 to 0.068; this compares to a range of 0.0003 to 0.0433 in the AR model, thereby showing that the VAR model exhibits greater macroeconomic sensitivity that the AR model. According to the likelihood ratio test for the RRE segment, a value of -11.00 rejects the 35 parameter restrictions of the AR with respect to the VAR model at a very high level of confidence. The VAR model outperforms the AR model for the RRE segment according to the RMSE, SC and CPE measures with values of 0.2609, 0.9397 and -3.69% in the former as compared to 0.33, 0.84 and -10.33% in the latter. The VAR model is also more conservative than the AR model, having a CPE over the downturn (recent) period of -3.69% (5.16%) in the former versus -10.33% (9.26%) in the latter.

Second, we discuss the results for the CRE segment. In the VAR model the autoregressive term is significantly positive and of substantial magnitude, a parameter estimate 0.52, which is also significant but rather somewhat in the AR model having an estimate 0.95. The cross autoregressive terms on the CRE, CONS and CNI are significantly positive in the VAR model and of lower of magnitude than the autocorrelation term, parameter estimates ranging in 0.26 to 0.53. The coefficient estimates of the macroeconomic sensitivities are of acceptable magnitude with respect to model sensitivity and range in absolute vales of 0.01 to 0.02; this compares to a range of 0.004 to 0.01 in the AR model, thereby showing that the VAR model exhibits greater macroeconomic sensitivity that the AR model. According to the likelihood ratio test for the RRE segment, a value of -11.91 rejects the 35 parameter restrictions of the AR with respect to the VAR model at a very high level of confidence. The VAR model outperforms the AR model for the RRE segment according to the RMSE, SC and CPE measures with values of 0.30, 0.95 and 4.79% in the former as compared to 0.39, 0.85 and 7.6% in the latter. The VAR model is also more conservative than the AR model, having a CPE over the downturn (recent) period of -4.37% (5.57%) in the former versus -8.46% (10.29%) in the latter.

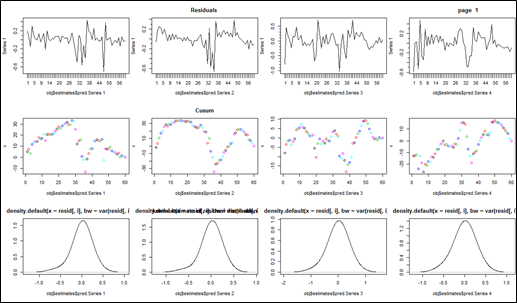
Next, we discuss the results for the CC segment. In the VAR model the autoregressive term is significantly positive and of substantial magnitude, a parameter estimate 0.37, which is also significant but rather somewhat in the AR model having an estimate 0.87. The cross autoregressive terms on the CRE, CONS and CNI are significantly positive in the VAR model and of lower of magnitude than the autocorrelation term, parameter estimates ranging in 0.26 to 0.53. The coefficient estimates of the macroeconomic sensitivities are of acceptable magnitude with respect to model sensitivity and range in absolute vales of 0.006 to 0.27; this compares to a range of 0.003 to 0.14 in the AR model, thereby showing that the VAR model exhibits greater macroeconomic sensitivity that the AR model. According to the likelihood ratio test for the RRE segment, a value of -34.81 rejects the 35 parameter restrictions of the AR with respect to the VAR model at a very high level of confidence. The VAR model outperforms the AR model for the RRE segment according to the RMSE, SC and CPE measures with values of 0.10, 0.94 and 3.91% in the former as compared to 0.21, 0.79 and -11.55% in the latter. The VAR model is also more conservative than the AR model, having a CPE over the downturn (recent) period of -3.12% (27.76%) in the former versus 6.67% (-22.5%) in the latter.

Finally, we discuss the results for the C&I segment. In the VAR model the autoregressive term is significantly positive and of substantial magnitude, a parameter estimate 0.79, which is also significant but rather somewhat in the AR model having an estimate 0.84. The cross autoregressive terms on the CRE, CONS and CNI are significantly positive in the VAR model and of lower of magnitude than the autocorrelation term, parameter estimates ranging in 0.01 to 0.40. The coefficient estimates of the macroeconomic sensitivities are of acceptable magnitude with respect to model sensitivity and range in absolute vales of 0.005 to 0.07; this compares to a range of 0.003 to 0.05 in the AR model, thereby showing that the VAR model exhibits greater macroeconomic sensitivity that the AR model. According to the likelihood ratio test for the RRE segment, a value of -5.33 rejects the 35 parameter restrictions of the AR with respect to the VAR model at a very high level of confidence. The VAR model outperforms the AR model for the RRE segment according to the RMSE, SC and CPE measures with values of 0.24, 0.90 and -11.30% in the former as compared to 0.31, 0.80 and 11.55% in the latter. The VAR model is also more conservative than the AR model, having a CPE over the downturn (recent) period of -4.07% (4.00%) in the former versus -8.34% (-6.86%) in the latter.

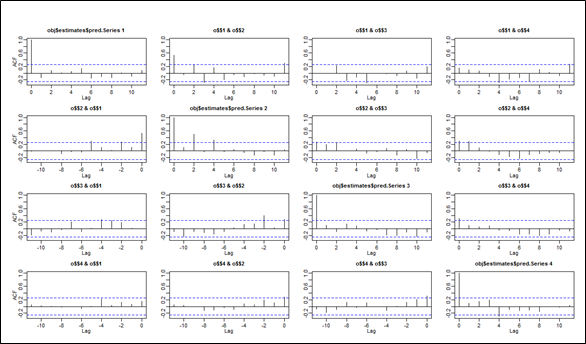
**Figure 9.1: Vector Autoregressive Maximum Likelihood Estimation (Fed Macroeconomic Variables and Aggregate Y9 Bank Chargeoffs) – 1-Step Ahead Model Predictions vs. Historical Data**



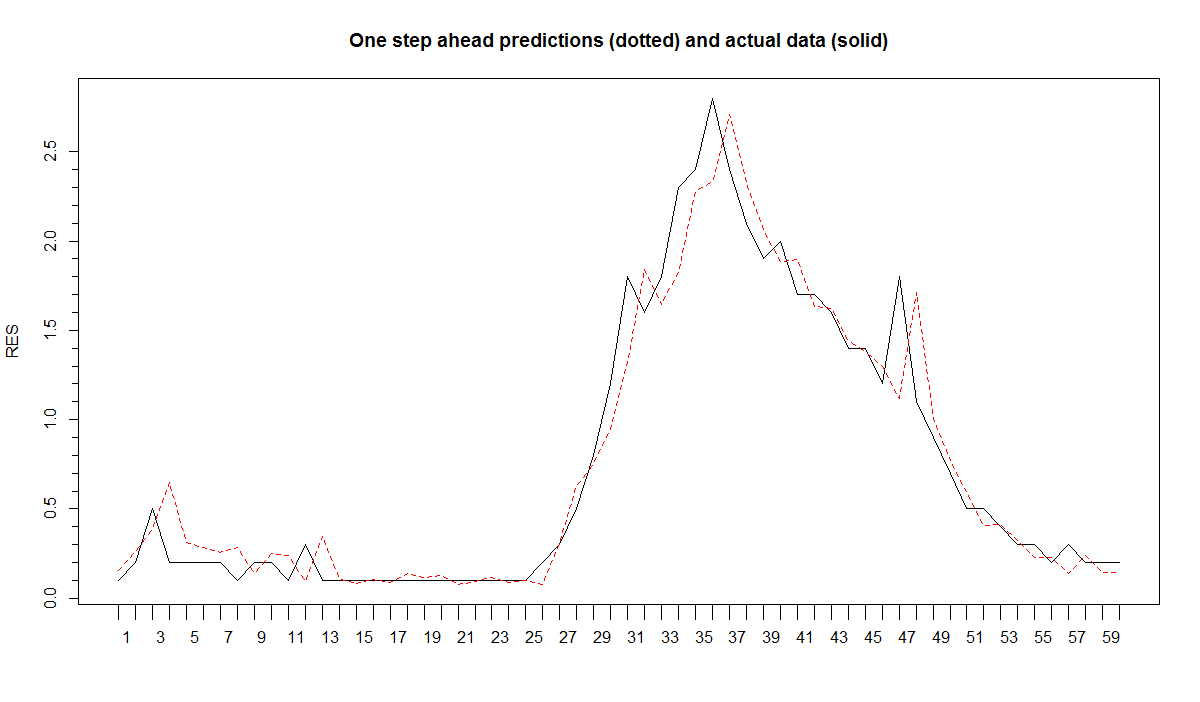
**Figure 9.2: Vector Autoregressive Maximum Likelihood Estimation (Fed Macroeconomic Variables and Aggregate Y9 Bank Chargeoffs) – Residual Time Series, Cumulative Sum and Histogram Plots**



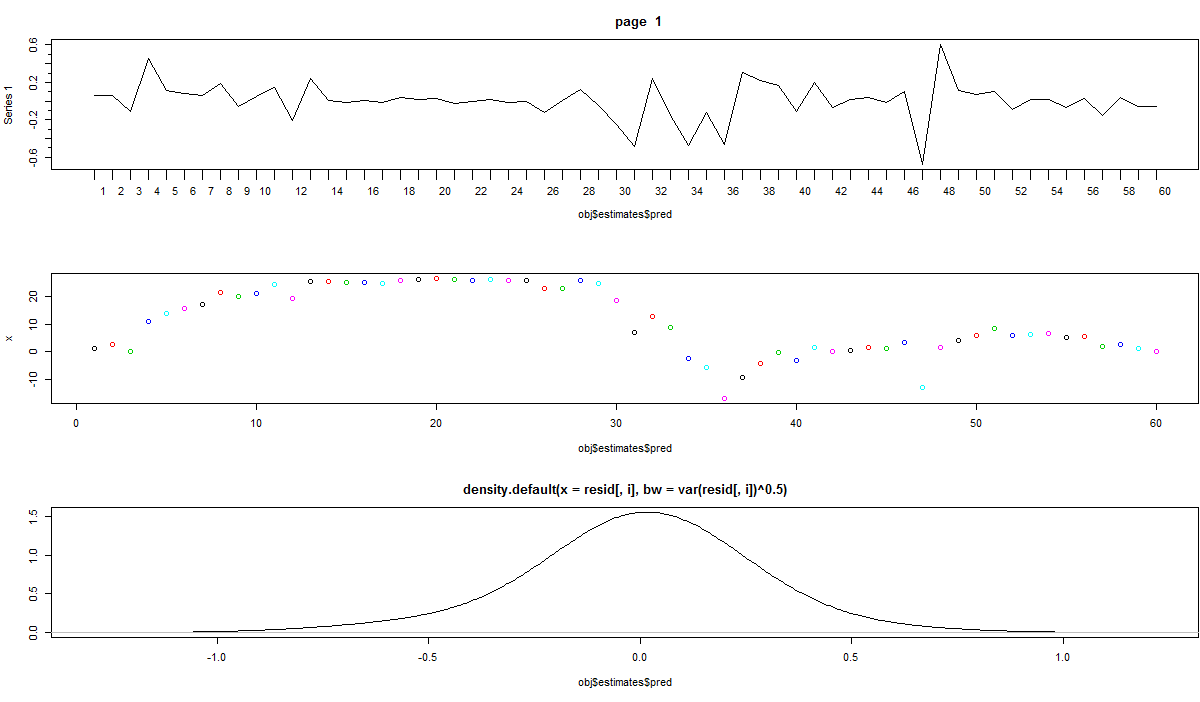
**Figure 9.3: Vector Autoregressive Maximum Likelihood Estimation (Fed Macroeconomic Variables and Aggregate Y9 Bank Chargeoffs) – Residual Autocorrelation Plots**



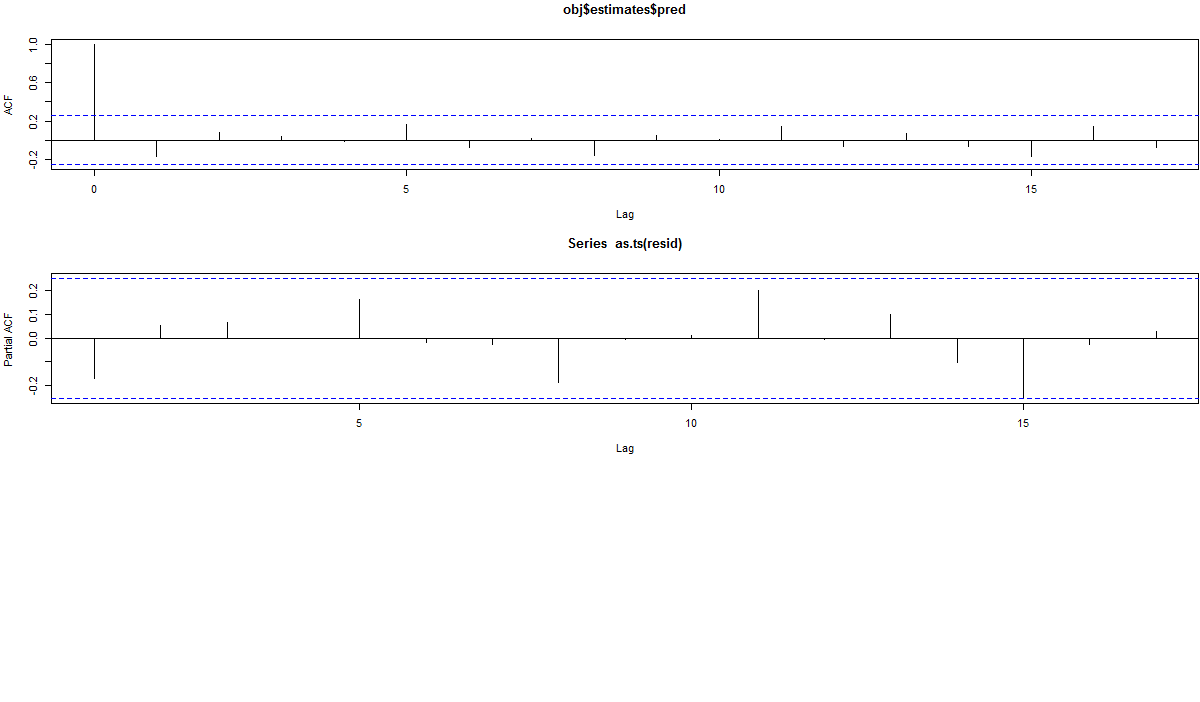
**Figure 9.4: Univariate Autoregressive Maximum Likelihood Estimation (Fed Macroeconomic Variables and Aggregate Y9 Bank Chargeoffs) – 1-Step Ahead Model Predictions vs. Historical Data (Residential Real Estate)**



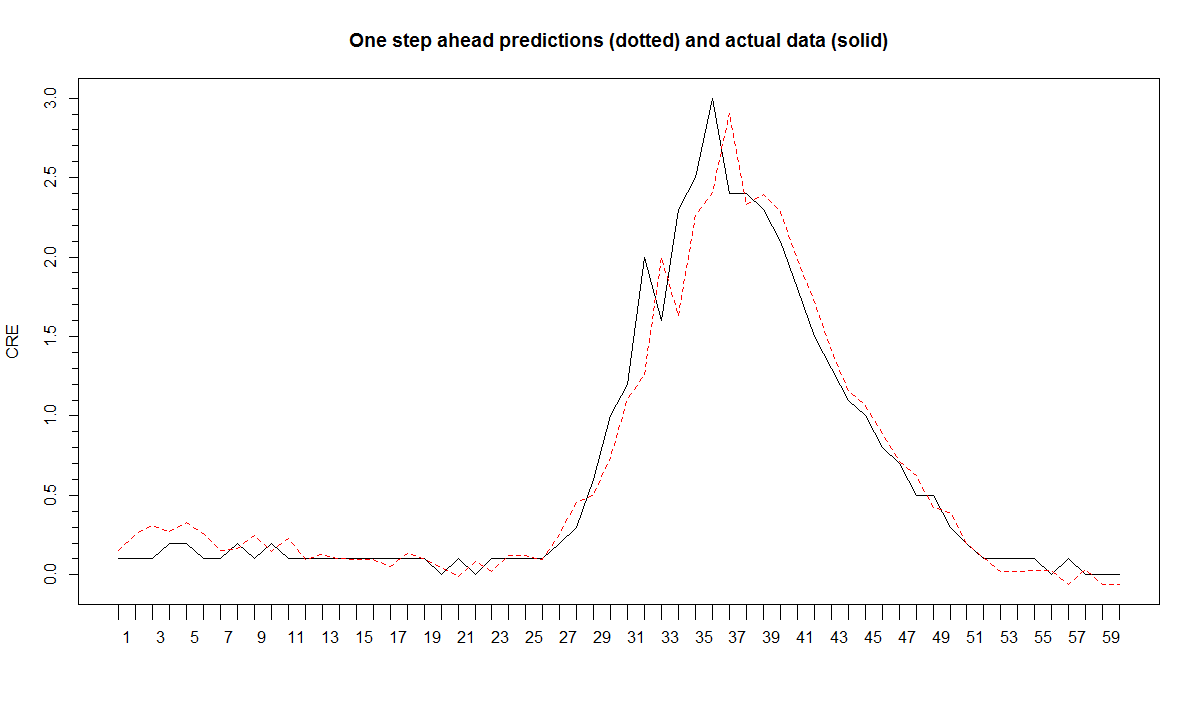
**Figure 9.5: Univariate Autoregressive Maximum Likelihood Estimation (Fed Macroeconomic Variables and Aggregate Y9 Bank Chargeoffs) – Residual Time Series, Cumulative Sum and Histogram Plots (Residential Real Estate)**



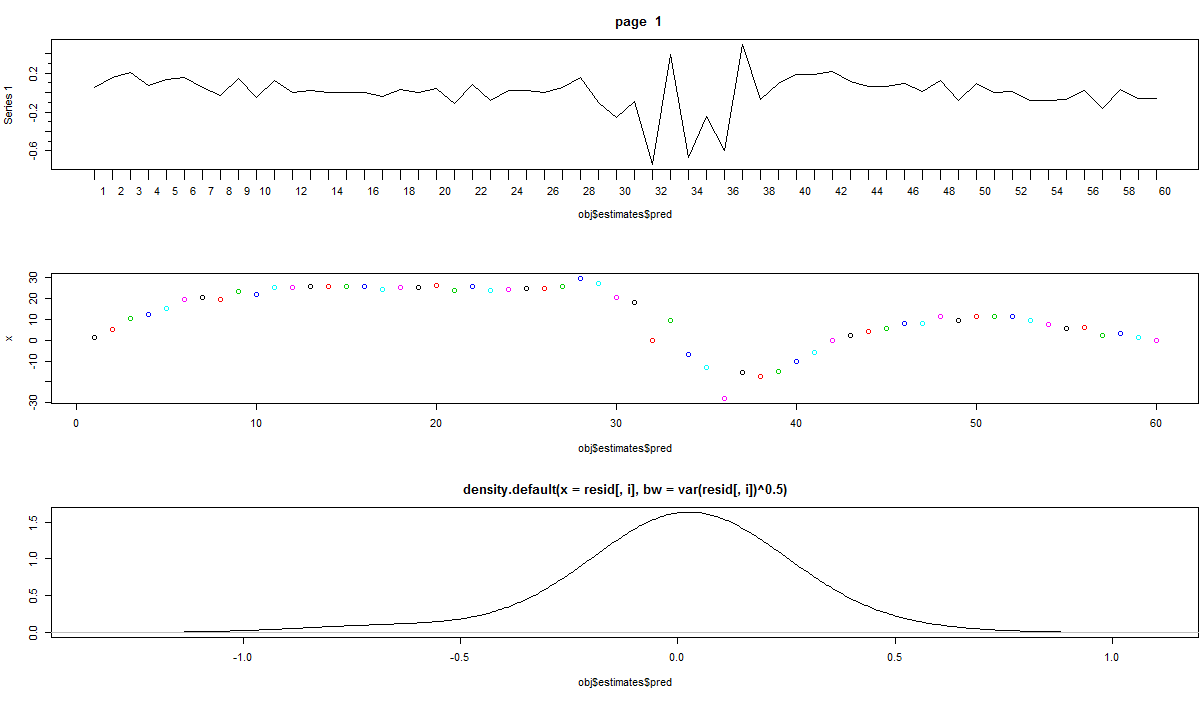
**Figure 9.6: Univariate Vector Autoregressive Maximum Likelihood Estimation (Fed Macroeconomic Variables and Aggregate Y9 Bank Chargeoffs) – Residual Autocorrelation Plots (Residential Real Estate)**



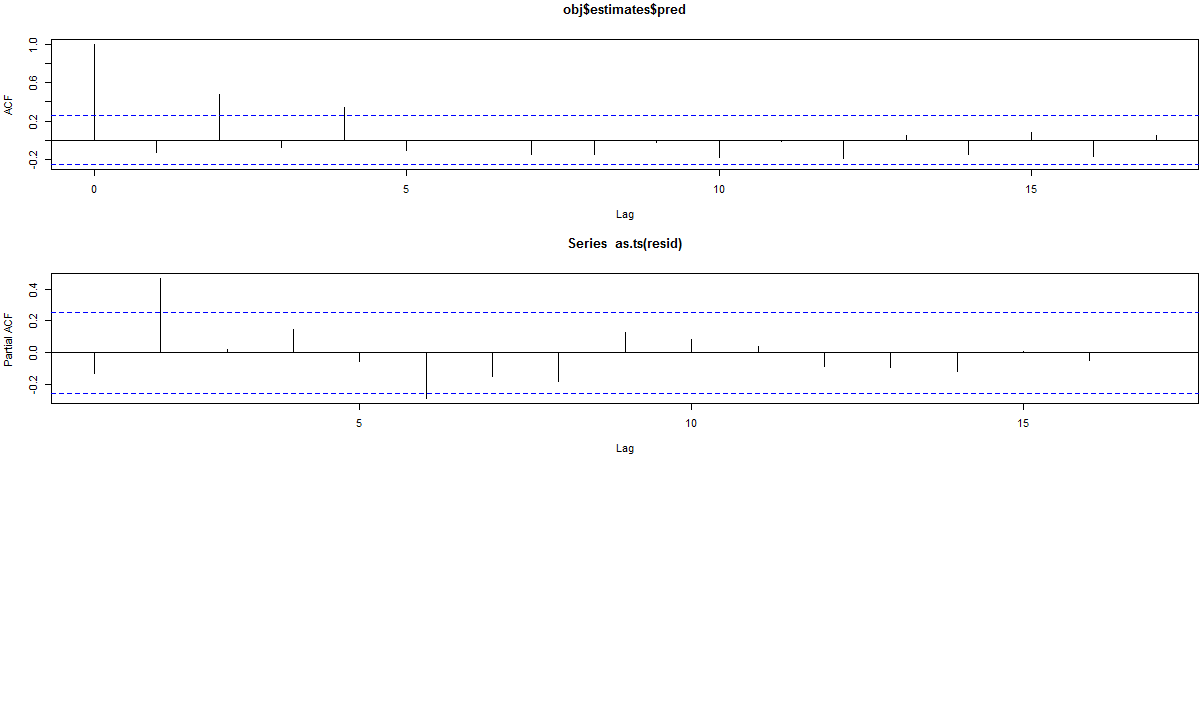
**Figure 9.7: Univariate Autoregressive Maximum Likelihood Estimation (Fed Macroeconomic Variables and Aggregate Y9 Bank Chargeoffs) – 1-Step Ahead Model Predictions vs. Historical Data (Commercial Real Estate)**



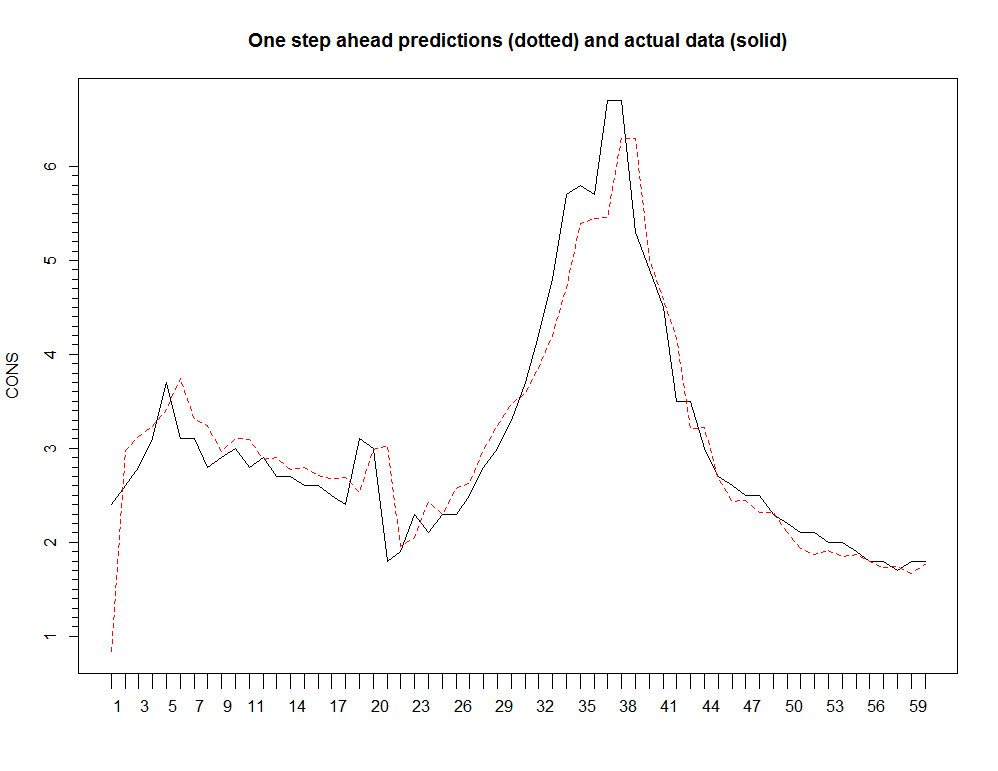
**Figure 9.8: Univariate Autoregressive Maximum Likelihood Estimation (Fed Macroeconomic Variables and Aggregate Y9 Bank Chargeoffs) – Residual Time Series, Cumulative Sum and Histogram Plots (Commercial Real Estate)**



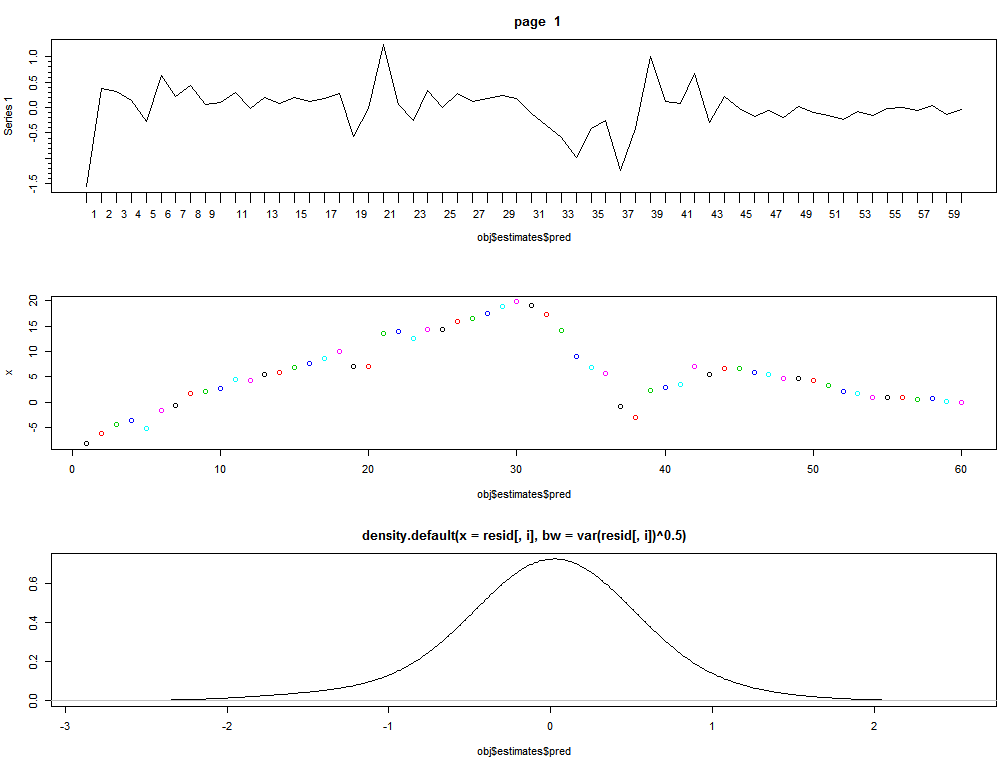
**Figure 9.9: Univariate Vector Autoregressive Maximum Likelihood Estimation (Fed Macroeconomic Variables and Aggregate Y9 Bank Chargeoffs) – Residual Autocorrelation Plots (Commercial Real Estate)**



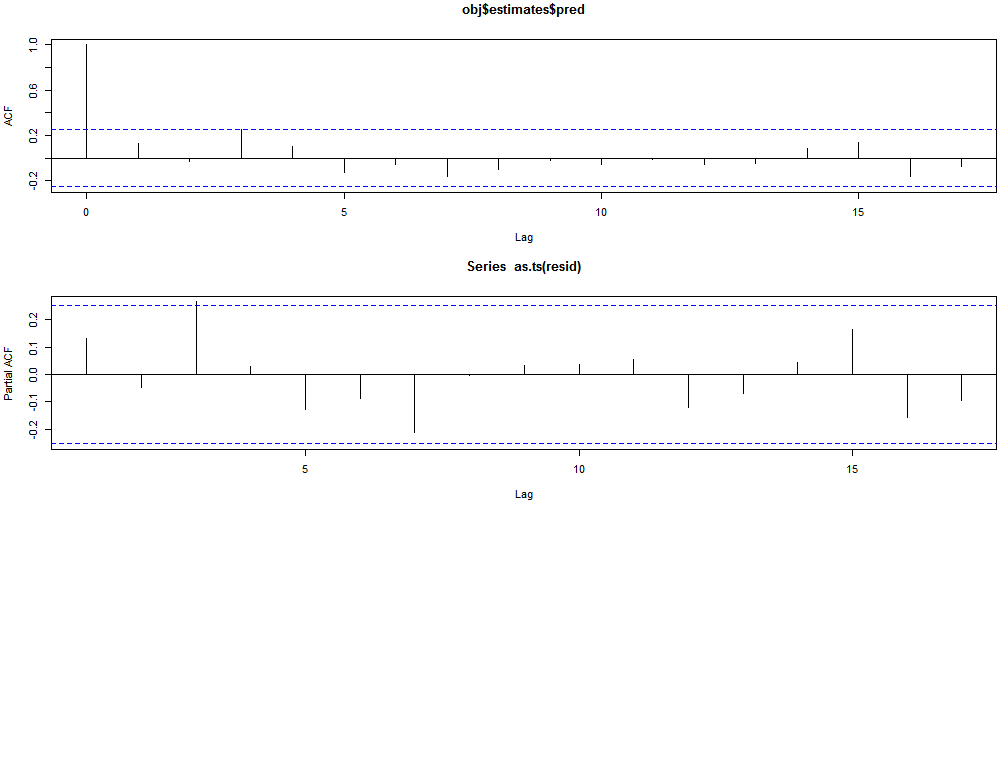
**Figure 9.10: Univariate Autoregressive Maximum Likelihood Estimation (Fed Macroeconomic Variables and Aggregate Y9 Bank Chargeoffs) – 1-Step Ahead Model Predictions vs. Historical Data (Consumer Credit)**



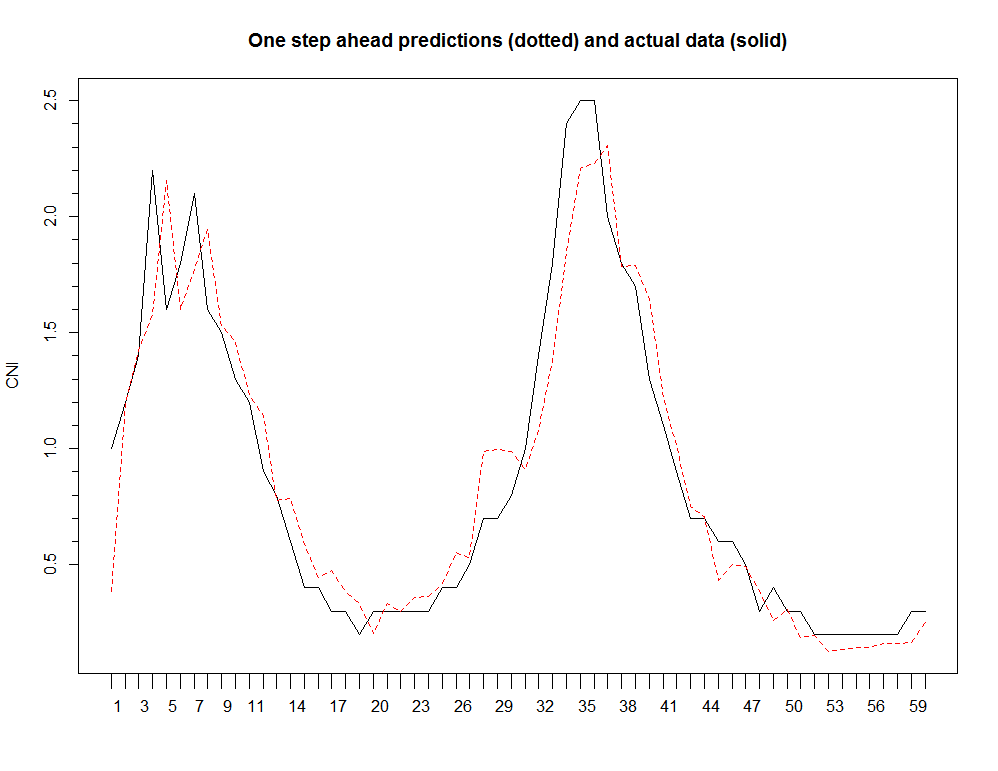
**Figure 9.11: Univariate Autoregressive Maximum Likelihood Estimation (Fed Macroeconomic Variables and Aggregate Y9 Bank Chargeoffs) – Residual Time Series, Cumulative Sum and Histogram Plots (Consumer Credit)**



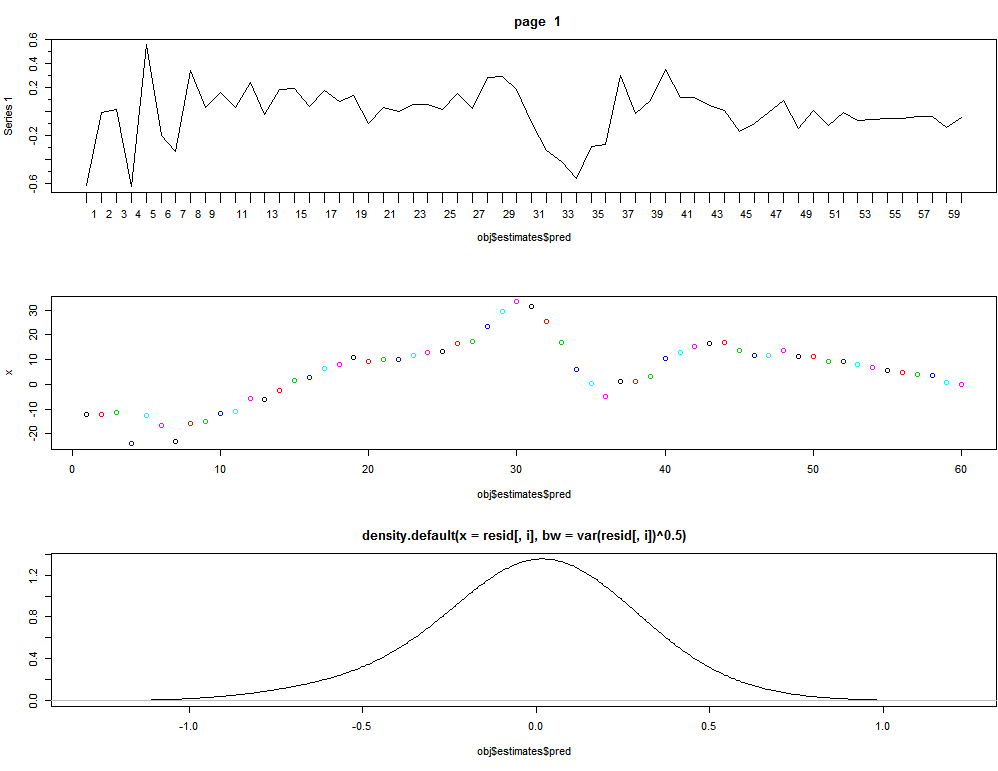
**Figure 9.12: Univariate Vector Autoregressive Maximum Likelihood Estimation (Fed Macroeconomic Variables and Aggregate Y9 Bank Chargeoffs) – Residual Autocorrelation Plots (Consumer Credit)**



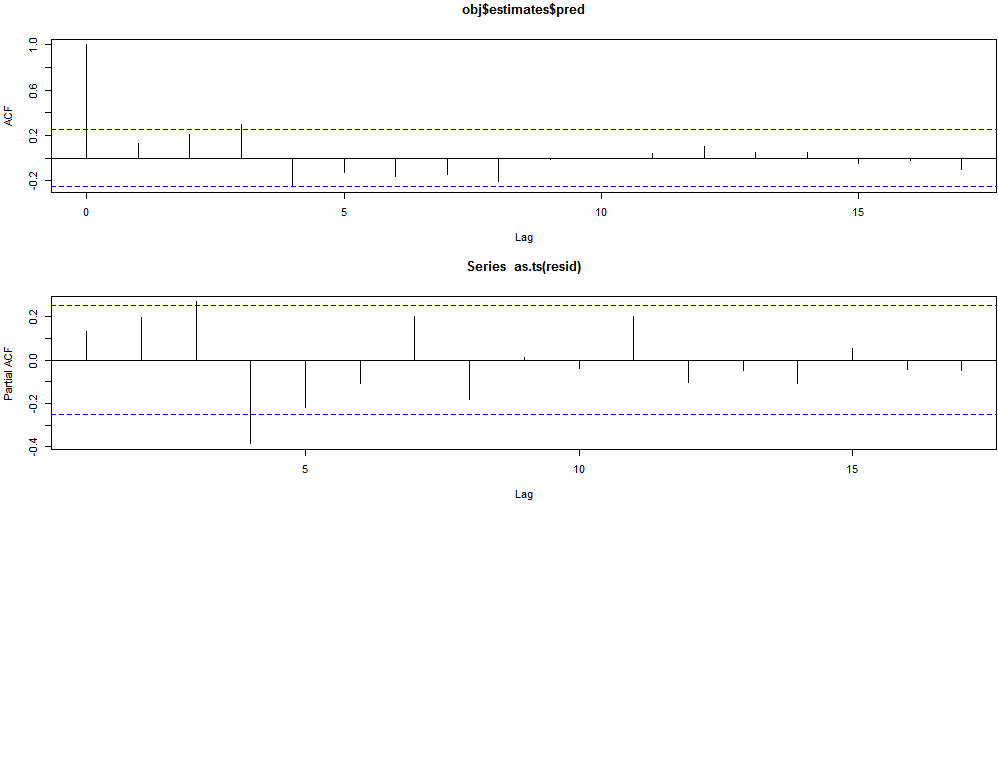
**Figure 9.13: Univariate Autoregressive Maximum Likelihood Estimation (Fed Macroeconomic Variables and Aggregate Y9 Bank Chargeoffs) – 1-Step Ahead Model Predictions vs. Historical Data (Commercial and Industrial)**



**Figure 9.14: Univariate Autoregressive Maximum Likelihood Estimation (Fed Macroeconomic Variables and Aggregate Y9 Bank Chargeoffs) – Residual Time Series, Cumulative Sum and Histogram Plots (Commercial and Industrial)**



**Figure 9.15: Univariate Vector Autoregressive Maximum Likelihood Estimation (Fed Macroeconomic Variables and Aggregate Y9 Bank Chargeoffs) – Residual Autocorrelation Plots (Commercial and Industrial)**



1. Corresponding author: Michael Jacobs, Jr., Ph.D., CFA, 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, [michael.a.jacobs@accenture.com](mailto:michael.a.jacobs@accenture.com). The views expressed herein are those of the authors and do not necessarily represent a position taken either by Accenture nor any affiliated firms. [↑](#footnote-ref-1)
2. Frank J Sensenbrenner, Ph.D., Non-Visiting Fellow, Johns Hopkins School of Advanced International Studies, SAIS Center for Transatlantic Relations, 1717 Massachusetts Avenue, Washington, DC 20008, 202-714-2659, frank.sensenbrenner@gmail.com. The views expressed herein are those of the authors and do not necessarily represent a position taken by Johns Hopkins SAIS Center for Transatlantic Relations. [↑](#footnote-ref-2)
3. In fact, the exogenous variables 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, 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. [↑](#footnote-ref-3)
4. Note that the VARMAX model (3.1)-(3.2) could be written in various equivalent forms, involving a lower triangular coefficient matrix forat lag zero, or a leading coefficient matrix forat lag zero, or even a more general form that contains a leading (non-singular) coefficient matrix forat lag zero that reflects instantaneous links amongst the output variables that are motivated by theoretical considerations (provided that the proper identifiability conditions are satisfied (Hanan (1971), Kohn (1979)). 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 (3.4), 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, (3.3) has a state space representation of the form (Hanan, 1988). [↑](#footnote-ref-4)
5. We perform this model selection in an R script designed for this purpose, using the libraries “dse” and “tse” to estimate and evaluate VARMAX and ARMAX models (R Core Development Team, 2016). [↑](#footnote-ref-5)
6. This is similar to the finding of Loregian and Meucci (2016) in the context of modeling the U.S. Treasury yields. We observe that this mixture well characterizes the empirical distributions of the data. [↑](#footnote-ref-6)
7. We have performed a sensitivity analysis, available upon eques, using the 95th and 99.9th percentiles, and the results are not greatly changed. [↑](#footnote-ref-7)
8. 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 (R Development Core Team, 2016) [↑](#footnote-ref-8)