

# The Validation of Machine Learning Models for the Stress Testing of Credit Risk

Michael Jacobs, Jr.<sup>1</sup>  
Accenture Consulting

Draft: March 18<sup>th</sup>, 2018

## Abstract

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 common approach to modeling is for stress testing statistical regression model, such as a *Vector Autoregression* (“VAR”). 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 in the machine learning class of models, widely used in the academic literature, but not commonly employed in practice, the *Multivariate Adaptive Regression Splines* (“MARS”) 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 validate our champion MARS model through a rigorous horse race against the VAR model, and find it to exhibit greater accuracy in model testing, as well as superior out-of-sample performance, according to various metrics across all modeling segments. Furthermore, we find that the MARS model produces more reasonable forecasts, from the perspective of quality and conservatism in severe scenarios.

Keywords: Stress Testing, CCAR, DFAST, Credit Risk, Financial Crisis, Model Risk, Vector Autoregression, Multivariate Adaptive Regression Splines, Model Validation

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

---

<sup>1</sup> 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. The views expressed herein are those of the author and do not necessarily represent a position taken either by Accenture or any affiliated firms.

# 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.

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. 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-down

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 alternative estimation techniques. Section 3 presents the competing econometric methodologies for generating scenarios, a time series *Vector Autoregressive* (“VAR”) and *Multivariate Adaptive Regression Splines* (“MARS”) 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 Accord, with the 1995 Market Risk Amendment (Basel Committee for Banking Supervision 1988, 1996). Around the same time, the publication of RiskMetrics™ in 1994 (J.P. Morgan, 1994) marked risk management as a separate technical discipline, and therein all of the above mentioned types of ST are referenced. The seminal handbook on *Value-at-Risk* (“VaR”), also had a part devoted to the topic of ST (Jorion, 1996), while other authors (Kupiec (1999), Berkowitz and Jeremy (1999)) provided detailed discussions of VaR-based stress tests as found largely in the trading and treasury functions. The *Committee on Global Financial Systems* (“CGFS”) conducted a survey on stress testing in 2000 that had similar findings (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 CreditMetrics™ (J.P. Morgan, 1997) and CreditRisk+™ (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 CreditMetrics<sup>TM</sup> (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 CreditMetrics<sup>TM</sup> (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 (2016) 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 (2016) through a model validation of the latter model using sensitivity analysis, and an empirical implementation of the model with an application to stress testing, wherein it is found that standard credit risk models understate capital measures by an order or magnitude, but that under stressed parameter settings this understatement is greatly attenuated. Jacobs (2018) propose a challenger approach for ST and macroeconomic scenario generation, widely used in the academic literature but not commonly employed in practice, the Markov Switching VAR (“MS-VAR”) model. He empirically test the model using Federal Reserve Y-9 filing and macroeconomic data, finding the

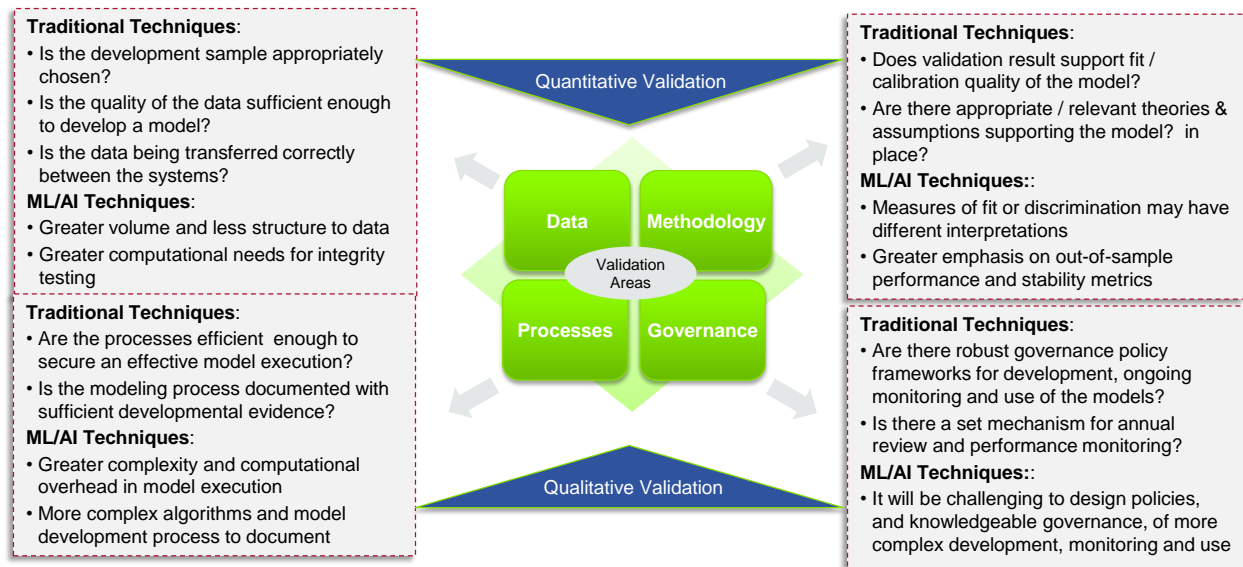
MS-VAR model to be more conservative than the VAR model, and also to exhibit greater accuracy in model testing.

There is extensive academic literature in the field of *machine learning* and *artificial intelligence* (“ML / AI”). In a classic work by Bellman (1957), the father of dynamic programming theory and optimal control, it is asserted that high *dimensionality* of data (i.e., many features associated with what we are trying to model) is a fundamental hurdle in many scientific applications, particularly in the context of pattern classification applications where learning complexity grows at a much faster space than the degree of this dimensionality of the data (so so-called “the curse of dimensionality”). Wallis et al (1997, 1999) discuss the motivation behind the emergence of the subfield of *deep machine learning*, which focuses on computational models for information representation that exhibit similar characteristics to that of the neocortex (a part of the human brain governing for complex thought), finding that in addition to the spatial aspect of real-life data a temporal component also plays a key role (i.e., an observed sequence of patterns often conveys a meaning to the observer, whereby independent fragments of this sequence would be hard to decipher in isolation, so that meaning is often inferred from events or observations that are received close in time). Lee and Mumford (2003) and Lee et al (1998) examine neuroscience findings for insight into the principles governing information representation in the brain, leading to ideas for designing systems that represent information, and find that the neocortex associated with many cognitive abilities does not explicitly pre-process sensory signals, but rather allows them to propagate through a complex hierarchy that learns to represent observations based on the regularities they exhibit. Duda et al (2000) point out that the standard approach of dealing with this phenomenon has been data pre-processing or feature extraction algorithms that shift to a complex and rather application-dependent human-engineered process. Itamar et al (2010) provides an overview of the mainstream deep learning approaches and research directions proposed, emphasizing the strengths and weaknesses of each approach, and present a summary on the current state of the deep ML field and some perspective into how it may evolve.

Several vendors, consultancies and other financial practitioners have published white papers in this area of AI/AL. Bucheli and Thompson (2014) of the SAS Institute introduce key ML concepts and describes new SAS solutions that allow data scientists to perform machine learning at scale, and shares its experiences using machine learning to differentiate a new customer loyalty program. Jones (2017) of IBM presents an overview of the history of AI as well as the latest in neural network and deep learning approaches, and demonstrates how such new approaches like deep learning and cognitive computing have significantly raised the bar in these disciplines. Jacobs et al (2018) of Accenture highlights some of the challenges the financial services industry will face on the road to ML/AI adoption, in terms of gaining comfort with the robustness of modeling techniques through meaningful models that can withstand internal and external scrutiny.

ML / AI has in fact many had several applications in finance well before the advent of this modern era. The high volume and accurate nature of the historical data, coupled with the quantitative nature of the finance fields, has made this industry a prime candidate for the application of these techniques. The proliferation of such applications has been driven by more powerful capabilities in computing power and more accessible ML / AI methodologies. The fields of financial (i.e., credit, market, business and model) as well as that of non-financial (i.e., operational, compliance, fraud and cyber) risk modeling is and has been a natural domain of application for ML / AI techniques. Indeed, many work-horse modeling techniques in risk modeling (e.g., logistic regression,

**Figure 2.1 – The Model Validation Function and Challenges in ML/AI Modeling Methodologies**



Copyright © 2016 Accenture All rights reserved.

1

discriminant analysis, classification trees, etc. can be viewed in fact as much more basic versions of the merging ML / AI modeling techniques of the recent period. That said, there are risk types for which ML / AI has a greater degree of applicability than others – for example, one would more likely find this application in data-rich environments such a retail credit risk scoring (e.g., credit card, mortgages), as compared to relatively data poor domains such as low default credit portfolios for highly rated counterparties (e.g., sovereigns, financials, investment grade corporates). In the non-financial realm, we are seeing fruitful application in areas such as fraud analytics, where there is ample data to support ML / AL estimations. Later in this paper, we will present an example for Anti-Money Laundering (AML) in terms of both model development and model validation.

Finally, we discuss elements of the model validation process and focus on elements that may or may differ in the context of ML/AI modeling methodologies and techniques. In Figure 4.1 (Accenture Consulting, 2017) we depict the model validation function as the nexus of four core components (data, methodology, processes and governance), and two dimensions in the spectrum from quantitative to qualitative validation methodology. In the graphic highlight examples of some differences in the context of ML/AI modeling methodology. It is clear that many of the validation elements that have been the practice for traditional models will carry over to the ML / AI context, and the differences will be in emphasis or extentions of existing techniques.

### 3 Time Series VAR and MARS Methodologies for Stressed Loss Estimation

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 macroeconometric 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  $\mathbf{Y}_t = (Y_{1t}, \dots, Y_{kt})^T$  be a  $k$ -dimensional vector valued time series, the output variables of interest, in our application with the entries representing some loss measure in a particular segment, that may be influenced by a set of observable *input variables* denoted by  $\mathbf{X}_t = (X_{1t}, \dots, X_{rt})^T$ , an  $r$ -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  $VARMAX(p, q, s)$  ("vector autoregressive-moving average with exogenous variables") representation:

$$\mathbf{Y}_t \Phi(B) = \mathbf{X}_t \Theta(B) + \mathbf{E}_t \Theta^*(B) \quad (3.1)$$

Which is equivalent to:

$$\mathbf{Y}_t - \sum_{j=1}^p \Phi_j \mathbf{Y}_{t-j} = \sum_{j=0}^s \Theta_j \mathbf{X}_{t-j} + \mathbf{E}_t - \sum_{j=1}^q \Theta_j^* \mathbf{E}_{t-j} \quad (3.2)$$

Where  $\Phi(B) = \mathbf{I}_k - \sum_{j=1}^p \Phi_j B^j$ ,  $\Theta(B) = \sum_{j=0}^s \Theta_j B^j$  and  $\Theta^*(B) = \mathbf{I}_r - \sum_{j=1}^q \Theta_j^* B^j$  are *autoregressive lag polynomials* of respective orders  $p$ ,  $s$  and  $q$ , respectively, and  $B$  is the *back-shift operator* that satisfies  $B^i X_t = X_{t-i}$  for any process  $\{X_t\}$ . It is common to assume that the input process  $\mathbf{X}_t$  is

generated independently of the noise process  $\mathbf{E}_t = (E_{1t}, \dots, E_{kt})^T$ <sup>2</sup>. The autoregressive parameter matrices  $\Phi_j$  represent sensitivities of output variables to their own lags and to lags of other output variables, while the corresponding matrices  $\Theta_j$  are model sensitivities of output variables to contemporaneous and lagged values of input variables<sup>3</sup>. It follows that the dependency structure of the output variables  $\mathbf{Y}_t$ , as given by the autocovariance function, is dependent upon the parameters  $\mathbf{X}_t$ , and hence the correlations amongst the  $\mathbf{Y}_t$  as well as the correlation amongst the  $\mathbf{X}_t$  that depend upon the parameters  $\Theta_j$ . In contrast, in a system of univariate *ARMAX* ( $p, q, s$ ) (“autoregressive-moving average with exogenous variables”) models, the correlations amongst the  $\mathbf{Y}_t$  is not taken into account, hence the parameter vectors  $\Theta_j$  have a diagonal structure (Brockwell and Davis, 1991).

In this study we consider a *vector autoregressive model with exogenous variables* (“VARX”), denoted by *VARX* ( $p, s$ ), which restricts the *Moving Average* (“MA”) terms beyond lag zero to be zero, or  $\Theta_j^* = \mathbf{0}_{k \times k} \quad j > 0$ :

$$\mathbf{Y}_t - \sum_{j=1}^p \Phi_j \mathbf{Y}_{t-j} = \sum_{j=1}^s \Theta_j \mathbf{X}_{t-j} + \mathbf{E}_t \quad (3.3)$$

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.

---

<sup>2</sup> In fact, the exogenous variables  $\{\mathbf{X}_t\}$  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  $\{\mathbf{Y}_t\}$ , 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.

<sup>3</sup> Note that the VARMAX model (3.1)-(3.2) could be written in various equivalent forms, involving a lower triangular coefficient matrix for  $\mathbf{Y}_t$  at lag zero, or a leading coefficient matrix for  $\mathbf{E}_t$  at lag zero, or even a more general form that contains a leading (non-singular) coefficient matrix for  $\mathbf{Y}_t$  at lag zero that reflects instantaneous links amongst the output variables that are motivated by theoretical considerations (provided that the proper identifiability conditions are satisfied (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).



We now consider a machine learning methodology, MARS, an adaptive procedure for regression that is well suited for high-dimensional problems (Friedman, 1991). MARS can be viewed as a generalization of stepwise linear regression or a modification of the *classification and regression tree* (“CART”) method to improve the latter’s performance in the regression setting. MARS uses expansions in piecewise linear basis functions of the form:

$$(x-t)^+ = \begin{cases} x-t & \text{if } x > t \\ 0 & \text{otherwise} \end{cases} , \quad (3.4)$$

and:

$$(x-t)^- = \begin{cases} x-t & \text{if } x < t \\ 0 & \text{otherwise} \end{cases} . \quad (3.5)$$

Therefore, each function is piecewise linear, with a knot at the value which is known as a *linear spline*. We call the two functions (3.4) and (3.5) a *reflected pair* and the idea to form such pairs for input  $X_j$  with knots at each observed value  $X_{ij}$  of that input. The collection of basis functions is given by:

$$\mathbb{C} = \left\{ (X_j - t)^+, (t - X_j)^+ \right\}_{\substack{t \in \{x_{1j}, \dots, x_{Nj}\} \\ j=1, \dots, p}} . \quad (3.6)$$

If all of the input values are distinct, then in total there are  $2Np$  basis functions altogether. Note that although each basis function depends only on a single  $X_j$ , we consider it as a function over the entire input space  $\mathbb{R}^p$ . Therefore, while the modeling strategy is similar a forward stepwise linear regression, in lieu of using the original inputs we have the flexibility use functions from the set  $\mathbb{C}$  and their products and the model is of the form:

$$f(X) = \beta_0 + \sum_{m=1}^M \beta_m h_m(X) , \quad (3.7)$$

where each  $h_m(X)$  is a function in  $\mathbb{R}^{p \times N}$ , or a product of two or more such functions, and given a choice for the such functions the coefficients  $\beta_m$  are estimated by standard linear regression. However, the power of this method lies in the design of the basis functions. The algorithm is initiated with a constant function  $h_0(X) = 1$  and the set functions in the set  $\mathbb{C}$  are considered as to be candidate functions. At each stage we consider as a new basis function pair all products of a function  $h_m(X)$  in the model set  $M$  with one of the reflected pairs in  $\mathbb{C}$ . We add to the model set  $M$  the term of the following form that produces the largest decrease in training error:

$$\hat{\beta}_{M+1}h_l(X)(X_j - t)^+ + \hat{\beta}_{M+2}h_l(X)(t - X_j)^+, h_l \in M \quad (3.8)$$

In equation (3.8)  $M + 1$  and  $M + 2$  are coefficients estimated by least squares, along with all the other  $M + 1$  coefficients in the model. Then the winning products are added to the model and the process is continued until the model set  $M$  contains some preset maximum number of terms. Since at the end of this process we have a large model of this form that typically over-fits the data, a backward deletion procedure is applied. The term whose removal causes the smallest increase in residual squared error is deleted from the model at each stage, producing an estimated best model of each size (i.e., number of terms)  $A$ , which could be optimally estimated through a cross-validation procedure. However, for the sake of computational savings, the MARS procedure instead uses a *generalized cross-validation* (“GCV”) criterion, that is defined as:

$$\text{GCV}(\lambda) = \frac{\sum_{i=1}^N (y_i - \hat{f}_\lambda(x_i))^2}{\left(1 - \frac{M(\lambda)}{N}\right)^2} \quad (3.9)$$

The value  $M(\lambda)$  is the *effective number of parameters* in the model, accounting both for the number of terms in the models, as well as the number of parameters used in selecting the optimal positions of the knots<sup>4</sup>. Since some mathematical and simulation results suggest that one should pay a price of three parameters for selecting a knot in a piecewise linear regression (Hastie et al, 2009), it follows that if there are  $r$  linearly independent basis functions in the model, and  $K$  knots were selected in the forward process, the formula is  $M(\lambda) = r + CK$ , where  $C = 3$ . Using this, we choose the model along the backward sequence that minimizes  $\text{GCV}(\lambda)$ .

The modeling strategy rationale for these piecewise linear basis functions is linked to the key property of this class of functions, namely their ability to operate locally. That is, they are zero over part of their range, but when they are multiplied together the result is nonzero only over the small part of the feature space where both component functions are nonzero, which is to say that the regression surface is built up parsimoniously using nonzero components locally only where they are needed. The importance of this lies in the principle that we should economize on parameters in high dimensions, as we can quickly face the curse of dimensionality in an exploding parameter space, as in the use of other basis functions such as polynomials that produce nonzero product everywhere become quickly intractable. Another key advantage of the piecewise linear basis function concerns a significantly reduced computational overhead. Consider the product of a function in  $M$  with each of the  $N$  reflected pairs for an input  $X_j$ , which appears to require the

---

<sup>4</sup> Some mathematical and simulation results suggest that one should pay a price of three parameters for selecting a knot in a piecewise linear regression (Hastie et al, 2009).

fitting of  $N$  single-input linear regression models, each of which uses  $O(N)$  operations, making a total of  $O(N^2)$  operations. However, we can exploit the simple form of the piecewise linear function by first fitting the reflected pair with rightmost knot, and as the knot is moved successively one position at a time to the left, the basis functions differ by zero over the left part of the domain and by a constant over the right part. Therefore, after each such move we can update the fit in  $O(1)$  operations, which allows us to try every knot in only  $O(N)$  operations. The forward modeling strategy in MARS is hierarchical, in the sense that multiway products are built up from products involving terms already in the model. While the theory that a high-order interaction will likely only exist if some of its lower-order versions exist as well need not be true, nevertheless this is a reasonable working assumption and avoids the search over an exponentially growing space of alternatives. Finally, note that here one restriction put on the formation of model terms, namely each input can appear at most once in a product, which prevents the formation of higher-order powers of an input that may increase or decrease too sharply near the boundaries of the feature space. Such powers can be approximated in a more stable way with piecewise linear functions. In order to implement this, we may exploit a useful option in the MARS procedure that sets an upper limit on the order of interaction, which can aid in the interpretation of the final model (e.g., an upper limit of one results in an additive model).

## 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 be-

ginning with data as of September 30, spanning the nine-quarter planning horizon. The projections begin in the fourth quarter of the current year and conclude at the end of the fourth quarter two years forward. Hence, for defining BHC stress scenarios, firms are asked to project the movements of key MVs over the planning horizon of nine quarters. As for determining the severity of the global market shock components for trading and counterparty credit losses, it will not be discussed in this paper, because it is a one-time shock and the evaluation will be on the movements of the market risk factors rather the MVs. In the 2011 CCAR, the Federal Reserve defined the stress supervisory scenario using nine MVs:

- Real GDP (“RGDG”)
- Consumer Price Index (“CPI”)
- Real Disposable Personal Income (“RDPI”)
- Unemployment Rate (“UNEMP”)
- Three-month Treasury Bill Rate (“3MTBR”)
- Five-year Treasury Bond Rate (“5YTBR”)
- 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 15. In addition to the original nine variables, the added variables were:

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

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 scenarios in 2016, as well as the following calculated interest rate spreads:

- 10 Year Treasury minus 3 Month Treasury Spread or Term Spread (“TS”)
- BBB Corporate Rate minus 5 Year Treasury Spread or Corporate Spread (“CS”)

Therefore, we consider 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) are considered. Our model selection process imposed the following criteria in select-

ing input and output variables across both multiple VARMAX / univariate ARMAX and MARS models<sup>5</sup>:

- 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)

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:

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

This historical data, 74 quarterly observations from 2Q98 to 3Q16, are summarized in Table 4.1 in terms of distributional statistics and correlations. 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.

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 / CS, RGDPG / VIX and DOW / CREPI are 65.3%, -45.1% and 77.5%, respectively. The correlation matrix amongst the credit loss rate variables appear in the upper left quadrant of the bottom panel of Table 4.1. The correlations between CRE / CNI, CONS / CNI and CRE / CONS are 51.0%, 78.9% and 54.1%, respectively.

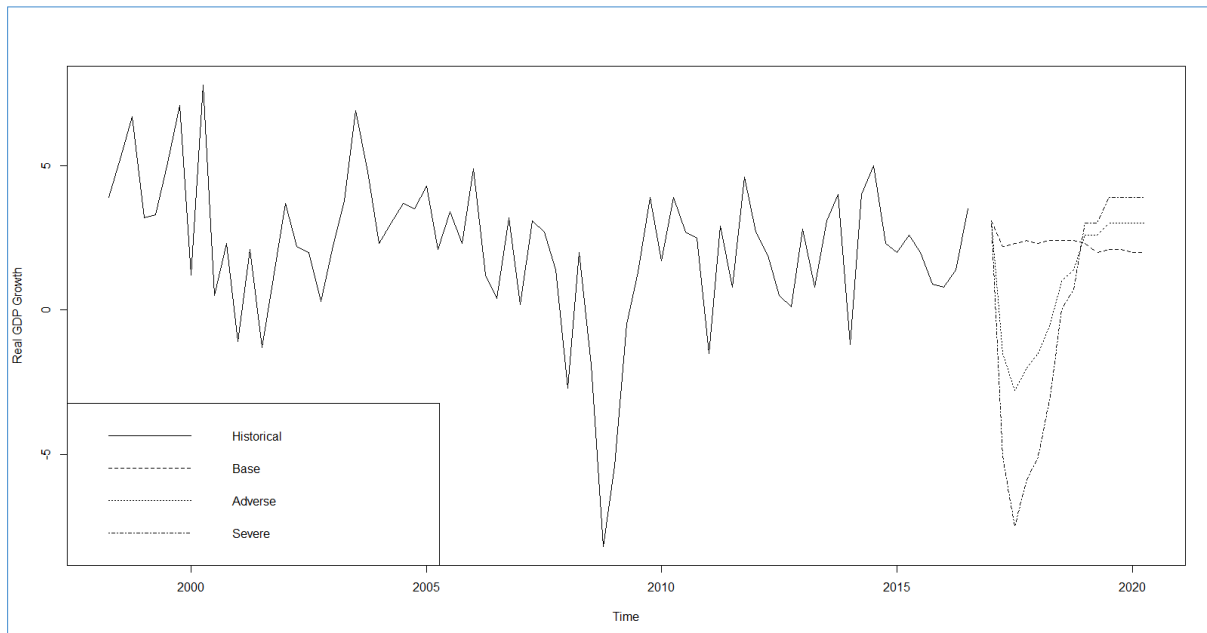
---

<sup>5</sup> 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, and the “Earth” package FOR THE mars MODELS (R Core Development Team, 2016).

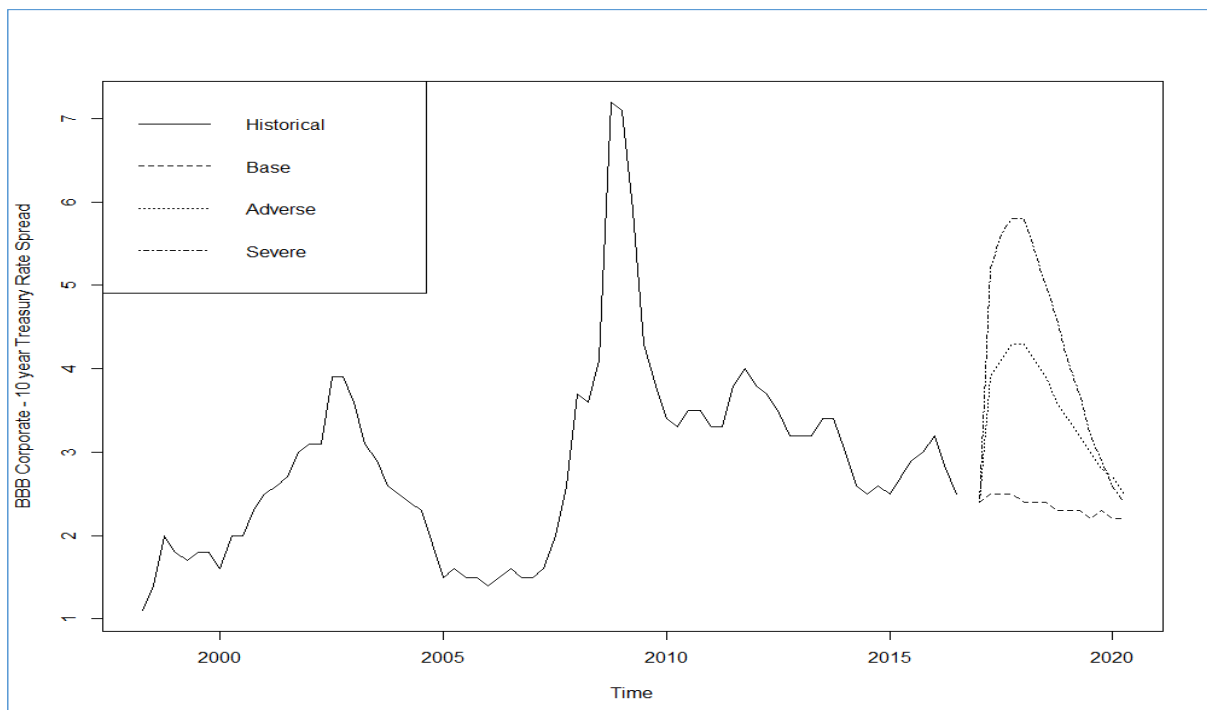
**Table 4.1: Summary Statistics and Correlations of Historical Y9 Credit Loss Rates and Federal Reserve Macroeconomic Variables (Q298—Q316)**

		Output Variables - Loss		Input Variables - Macroeconomic Factors																
		Observation Segments																		
		Commercial and Industrial	Commercial Real Estate	Consumer Credit	Real GDP Growth	Real Disposable Income Growth	Unemployment Rate	CPI Inflation Rate	3 Month Treasury Yield	5 Year Treasury Yield	10 Year Treasury Yield	BBB Corporate Bond Yield	Mortgage Rate	Prime Rate	Dow Jones Equity Index	Residential Housing Price Index	Commercial Real Estate Price Index	CSO7 Equity Market Volatility Index	10 Year - 3 Month Treasury Term Spread	BBB Corporate - 5 Year Treasury Spread
Summary Statistics	Count	74	74	74	74	74	74	74	74	74	74	74	74	74	74	74	74	74	74	74
	Mean	1.19%	1.72%	4.66%	2.1622	2.5743	6.0527	2.1797	1.9294	3.1568	3.9473	6.0068	5.5068	5.2230	13,502.19	144.66	186.82	28.457	2.0189	2.6500
	Standard Deviation	0.57%	1.57%	1.34%	2.5614	3.3902	1.7705	2.2043	2.0369	1.6592	1.3779	1.3657	1.3527	2.1445	3,948.07	28.67	47.55	11.557	1.1359	1.1524
	Minimum	0.38%	0.44%	2.74%	-8.2000	-15.7000	3.9000	-8.9000	0.0000	0.7000	1.6000	3.7000	3.4000	3.3000	7,773.60	89.20	118.20	12.7000	-0.2000	1.1000
	25th Percentile	0.69%	0.73%	3.47%	0.5500	1.1000	4.7000	1.4250	0.1000	1.6000	2.7250	4.9000	4.3250	3.3000	10,990.20	129.83	143.30	20.3500	1.2250	2.0000
	Median	1.19%	0.92%	4.58%	2.3000	2.7500	5.5500	2.4000	1.1500	3.1000	4.1500	6.0000	5.8000	4.2500	12,695.25	140.70	179.05	26.2500	2.1000	2.7500
	75th Percentile	1.59%	2.29%	5.56%	3.6500	4.2750	7.2000	3.3750	4.1750	4.6000	4.9000	6.5000	6.5000	7.4750	14,814.58	167.20	223.90	32.8750	2.9000	3.4000
	Maximum	2.79%	5.84%	7.89%	7.8000	10.9000	9.9000	6.3000	6.0000	6.6000	6.7000	9.4000	8.3000	9.5000	22,468.60	193.90	290.30	80.9000	3.8000	7.2000
	Coefficient of Variation	48.11%	91.65%	28.81%	118.46%	152.67%	29.25%	101.13%	105.63%	52.56%	34.91%	22.74%	24.33%	41.06%	28.33%	19.82%	25.67%	40.58%	56.26%	40.43%
	Skewness	0.5857	1.4376	0.4414	-1.0693	-1.4760	0.8672	-1.9415	0.6308	0.2590	-0.0652	0.2801	0.0512	0.6534	0.8934	-0.1123	0.3988	1.6816	-0.2659	1.4692
Kurtosis	-0.4208	0.7000	-0.4762	3.5572	6.4203	-0.5036	8.0348	-1.1973	-1.1015	-0.9640	-0.6506	-1.0064	-1.2063	-0.0519	-0.8184	-0.9274	4.7096	-0.9789	3.9862	
Output Variables	Commercial and Industrial																			
	Commercial Real Estate	50.99%																		
	Consumer Credit	76.65%	54.12%																	
	Real GDP Growth	7.10%	-27.97%	-12.53%																
Correlations	Real Disposable Income Growth	4.23%	-18.83%	-8.40%	13.28%															
	Unemployment Rate	37.37%	89.62%	35.26%	-19.31%	-20.33%														
	CPI Inflation Rate	1.35%	-11.46%	-2.42%	27.98%	-3.74%	-16.59%													
	3 Month Treasury Yield	3.00%	-48.57%	0.34%	21.05%	15.34%	-74.17%	30.82%												
	5 Year Treasury Yield	18.77%	-44.76%	22.83%	24.94%	17.16%	-70.12%	29.25%	93.31%											
	10 Year Treasury Yield	33.20%	-31.27%	40.99%	19.92%	14.64%	-56.73%	29.25%	84.75%	97.11%										
	BBB Corporate Bond Yield	38.23%	0.10%	61.08%	-20.88%	3.42%	-30.07%	0.42%	58.46%	72.60%	80.65%									
	Mortgage Rate	28.60%	-34.44%	37.75%	15.31%	14.53%	-60.99%	27.64%	83.81%	96.62%	98.81%	83.56%								
	Prime Rate	1.84%	-46.71%	-0.05%	17.54%	14.38%	-74.13%	29.33%	99.44%	91.72%	83.24%	59.12%	85.00%							
	Dow Jones Equity Index	-59.36%	-23.57%	-23.57%	8.24%	3.54%	-9.39%	-8.53%	-28.95%	-43.63%	-58.14%	-68.22%	-57.47%	-27.05%						
Residential Housing Price Index	-70.80%	-17.89%	-47.08%	-13.81%	-6.93%	-3.90%	2.67%	-21.69%	-33.70%	-43.19%	-49.58%	-44.06%	-22.60%	47.55%						
Commercial Real Estate Price Index	-75.08%	-8.56%	-62.92%	-28.43%	-9.99%	1.38%	-15.42%	-39.98%	-56.25%	-67.47%	-56.74%	-64.49%	-36.71%	77.45%	79.52%					
CSO7 Equity Market Volatility Index	33.99%	38.05%	49.81%	-45.08%	-11.20%	18.78%	-40.55%	-7.93%	-0.68%	8.59%	51.95%	13.74%	-3.44%	-43.76%	-44.43%	-19.83%				
10 Year - 3 Month Treasury Term Spread	34.90%	49.16%	49.11%	-13.59%	-9.75%	64.17%	-19.79%	-76.51%	-49.53%	-30.66%	-7.02%	-34.02%	-77.35%	-18.60%	-10.14%	24.63%				
BBB Corporate - 5 Year Treasury Spread	18.28%	64.57%	39.52%	-59.79%	-20.66%	65.32%	-42.29%	-65.07%	-57.95%	-44.26%	13.98%	-40.08%	-62.00%	-18.03%	-10.23%	13.74%	38.52%	67.34%		

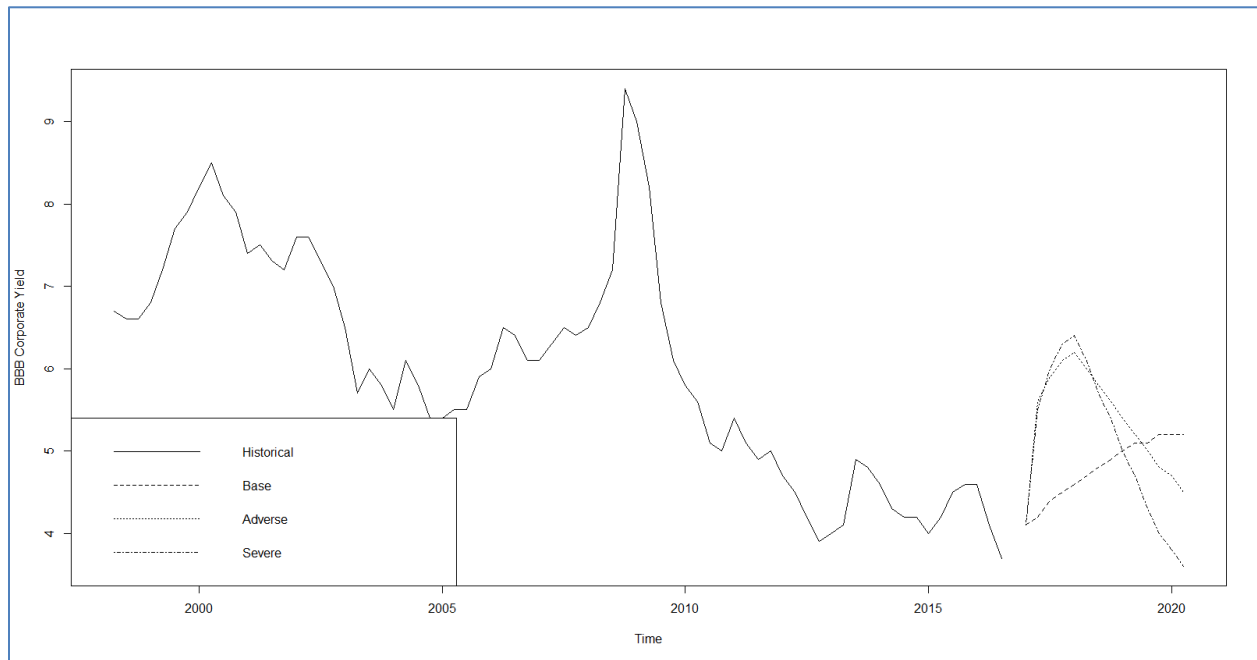
**Figure 4.1: Real GDP Growth – Time Series and Fed Scenarios (Historical Y9 Credit Loss Rates and Federal Reserve Macroeconomic Variables 2Q98—3Q14)**



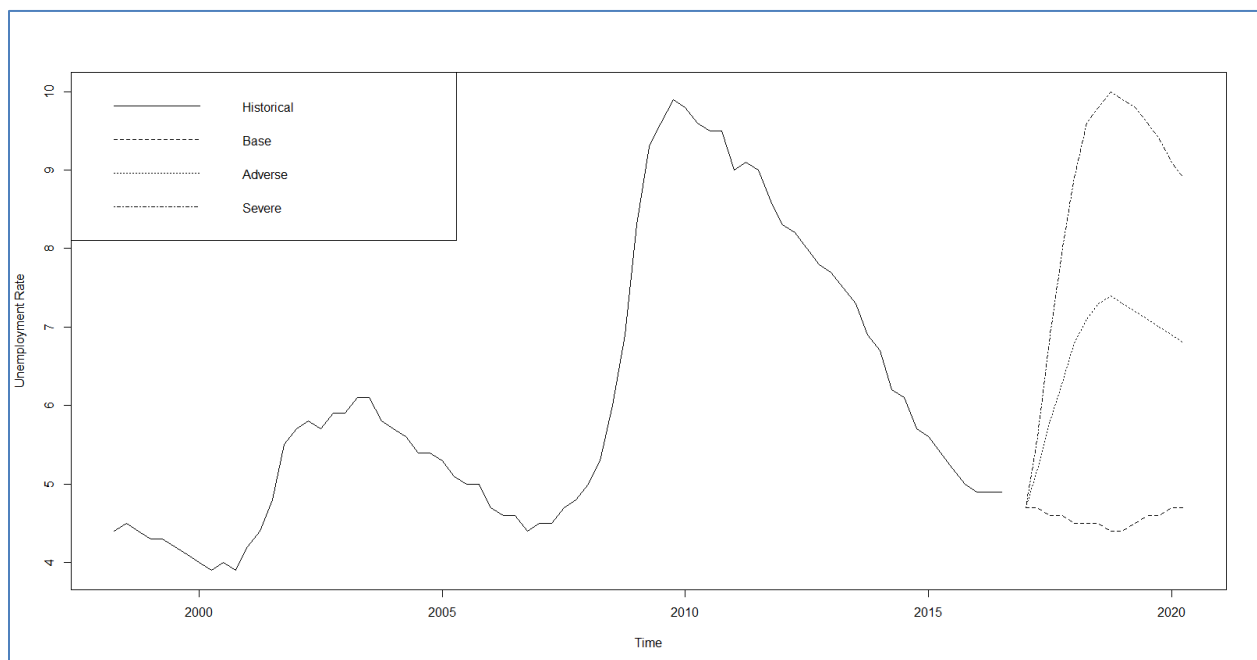
**Figure 4.2: BBB minus 5 Year Treasury Rate Spread – Time Series and Fed Scenarios (Historical Y9 Credit Loss Rates and Federal Reserve Macroeconomic Variables 2Q98—3Q16)**



**Figure 4.3: BBB Corporate Yield – Time Series and Fed Scenarios (Historical Y9 Credit Loss Rates and Federal Reserve Macroeconomic Variables 2Q98—3Q16)**



**Figure 4.4: Unemployment Rate – Time Series and Fed Scenarios (Historical Y9 Credit Loss Rates and Federal Reserve Macroeconomic Variables 2Q98—3Q16)**





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 BBBCY / CONS are 89.6%, -75.1% and 61.1%, respectively.

In the case of C&I, the best model for the quarterly change in charge-off rates was found, according to our model selection process, to contain the transformations of the following macroeconomic variables:

- RGDP: lagged 4 quarters
- CS: 2 quarter change lagged 4 quarters

In the case of CRE, the best corresponding model is found to be:

- BBB Corporate Yield: 4 quarter change lagged 2 quarters
- Unemployment Rate: lagged 1 quarters

In the case of CONS, the best corresponding model is found to be:

- BBB Corporate Yield: 3 quarter change lagged 4 quarters
- Unemployment Rate: lagged 1 quarters

The time series and Fed scenario forecasts of RDGP, CS, BBBCR and UNEMP are shown in Figures 4.1 through 4.4, and the estimation results (parameter estimates and p-values) for the VAR and MARS estimation are shown in Table 4.2 (the model estimation diagnostic plots are shown in the Appendix Section 7, Figures 7.1 through 7.6). We note that all residuals exhibit Gaussian behavior, and in the case of the VAR models zero autocorrelation, which indicates that the models are well-specified.

In the case of the C&I model, we note in Table 4.2 that the optimal VAR model contained only a single autoregressive terms and that the intercept was statistically significant and therefor dropped from the model, having a point estimate of 0.15 and therefore showing while there is positive autoregression there is yet a rather low level of persistence in the dependent variable. The signs on the coefficient estimates are intuitive and of reasonable magnitudes (-0.0002 and 0.0016 for RGDP and CS, respectively) and they are all statistically significant. On the other hand, the optimal MARS model has 3 terms including an intercept (the so-called effective number of parameters is 3, and as in the VAR model the signs on the coefficient estimates are intuitive (note that that for MARS they are inverted), of reasonable magnitudes (0.0007 and -0.0015 for RGDP and CS, respectively) and they are all statistically significant. In the case of the CRE model, we note in Table 4.2 that the optimal VAR model contained only a single autoregressive terms and that the intercept was statistically significant and therefor dropped from the model, having a point estimate of 0.11 and therefore showing while there is positive autoregression there is yet a rather low level of persistence in the dependent variable.

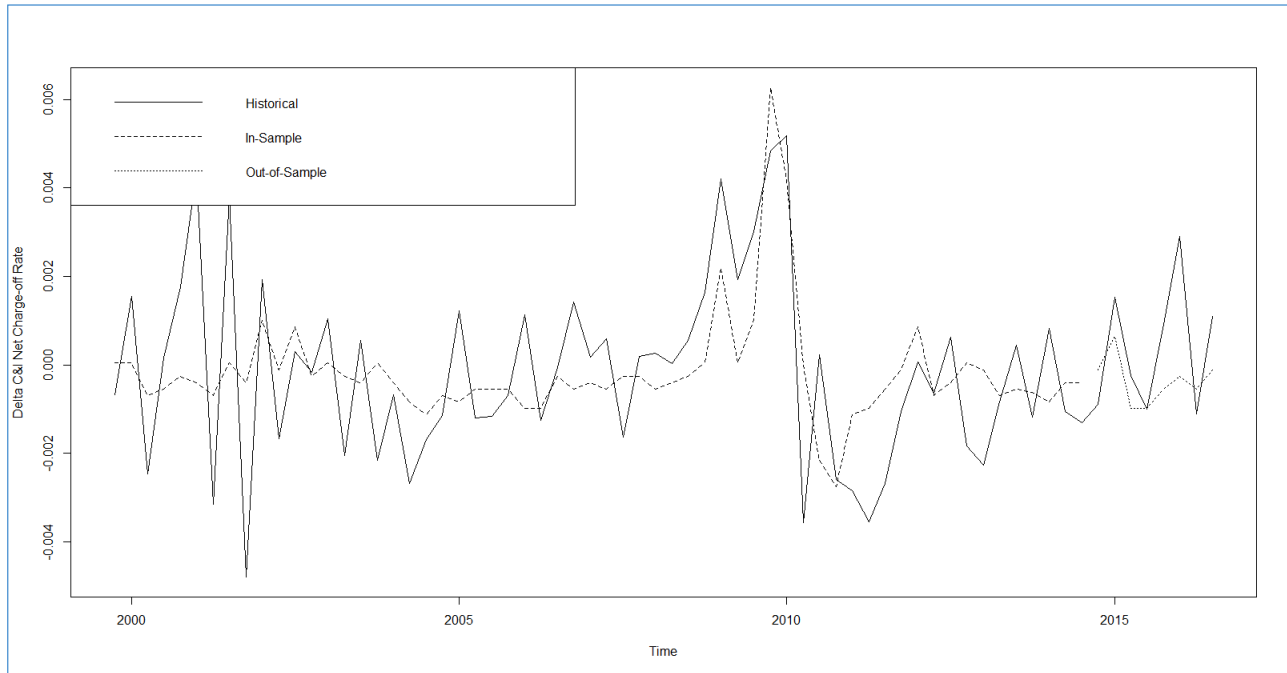
**Table 4.2: Credit Loss Estimation Results – VAR and MARS Models (Historical Y9 Credit Loss Rates and Federal Reserve Macroeconomic Variables 4Q99—3Q14)**

Segment	Parameters	VAR		MARS	
		Estimate	P-Value	Estimate	P-value
Commercial and Industrial	Intercept	N/A	N/A	1.61E-04	3.62E-05
	Autoregressive Term	1.5E-01	7.3E-02	N/A	N/A
	Real GDP Growth	-2.0E-04	8.4E-02	7.4E-04	3.6E-05
	BBB Corporate-5 Year Treasury Spread	1.6E-03	8.8E-05	-1.5E-03	2.4E-02
Commercial Real Estate	Intercept	N/A	N/A	-5.94E-04	4.03E-02
	Autoregressive Term	1.1E-01	2.9E-02	N/A	N/A
	BBB Corporate Yield	5.1E-03	1.3E-03	1.2E-03	-2.4E+00
	Unemployment Rate - 1	1.3E-03	1.1E-04	3.1E-03	2.5E-04
	Unemployment Rate - 2	N/A	N/A	-3.8E-03	1.0E-03
Consumer Credit	Intercept	N/A	N/A	3.29E-04	5.22E-02
	Autoregressive Term -1	2.6E-01	1.1E-03	N/A	N/A
	Autoregressive Term -2	-1.3E-01	7.2E-02	N/A	N/A
	BBB Corporate Yield -1	4.4E-03	4.4E-06	-2.0E-03	8.2E-06
	BBB Corporate Yield -2	1.3E-03	9.6E-02	N/A	N/A
	Unemployment Rate -1	1.3E-03	7.4E-05	5.2E-03	4.3E-05
	Unemployment Rate - 2	2.8E-04	5.0E-02	N/A	N/A

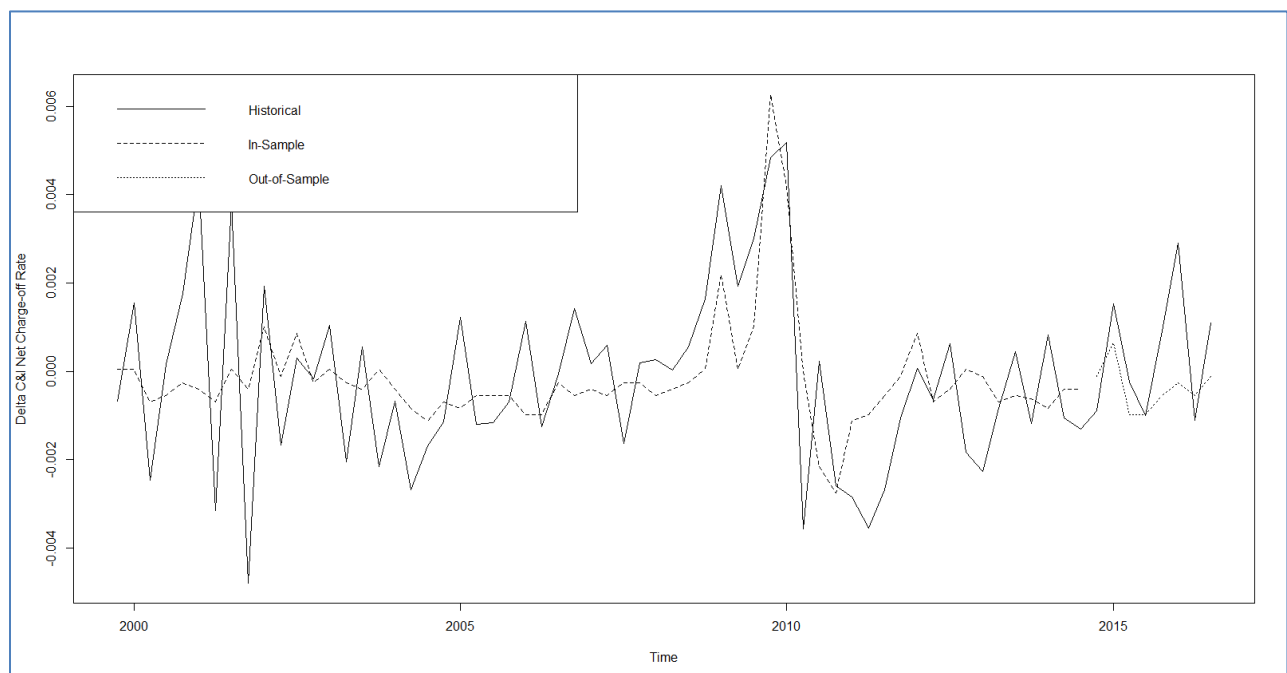
The signs on the coefficient estimates are intuitive and of reasonable magnitudes (0.0051 and 0.0013 for BBBCY and UNEMP, respectively) and they are all statistically significant. On the other hand, the optimal MARS model has 4 terms including an intercept, and as in the VAR model the signs on the coefficient estimates are intuitive (note that for MARS they are inverted), of reasonable magnitudes (0.0012 for BBBCY; and 0.031 and -0.0038 for UNEMP; note that the flip in signs for the UNEPM is still admissible due to the relative magnitudes, as the leading negative term dominates) and they are all statistically significant. In the case of the CONS model, we note in Table 4.2 that the optimal VAR model contained two autoregressive terms and that the intercept was statistically significant and therefor dropped from the model, having a point estimates of 0.26 and -0.13 on the 1<sup>st</sup> and 2<sup>nd</sup> orders, respectively, therefore showing while there is positive autoregression (as the 1<sup>st</sup> autoregressive term dominates the second in magnitude) there is yet a rather low level of persistence in the dependent variable. The signs on the coefficient estimates are intuitive and of reasonable magnitudes (0.0044 and 0.0013).

In Figures 4.5 through 4.10 we show the accuracy plot for the models of fit vs. history, for in-sample (4Q99-3Q14) vs. out-of-sample (4Q14-3Q16) periods, for the MARS and VAR estimations of the C&I, CRE and CONS models.

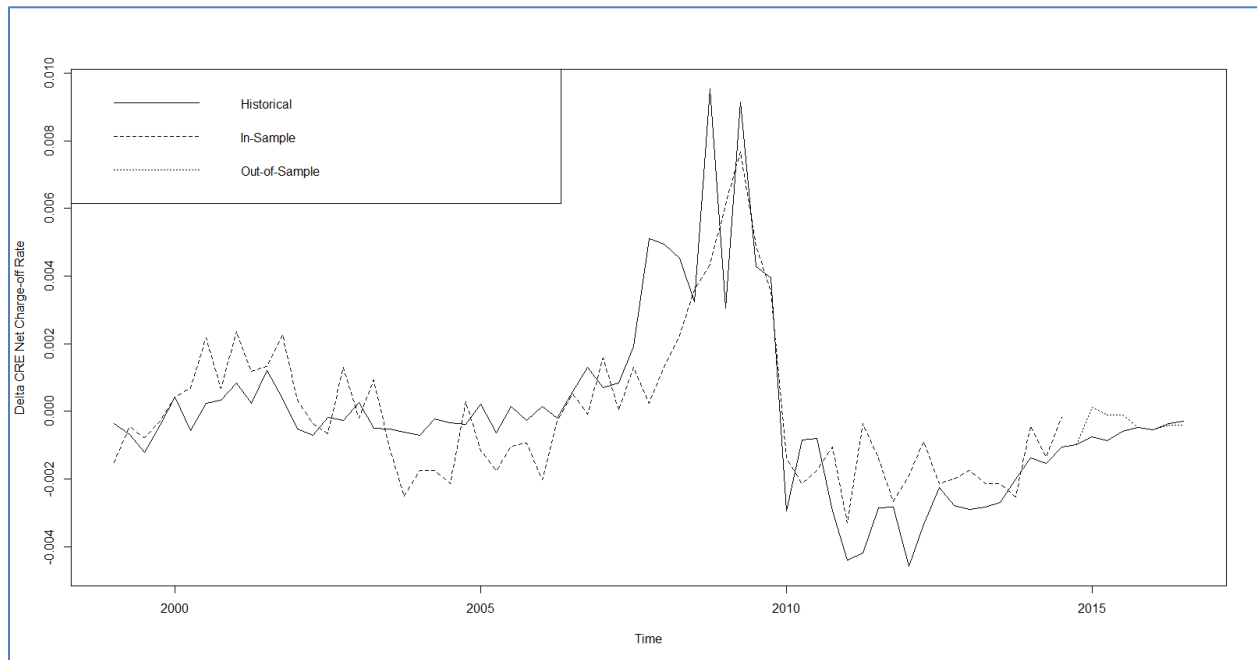
**Figure 4.5: C&I MARS Model Accuracy Plot – Fit vs. History for In-Sample (4Q99-3Q14) vs. Out-of-Sample (4Q14-3Q16) Periods**



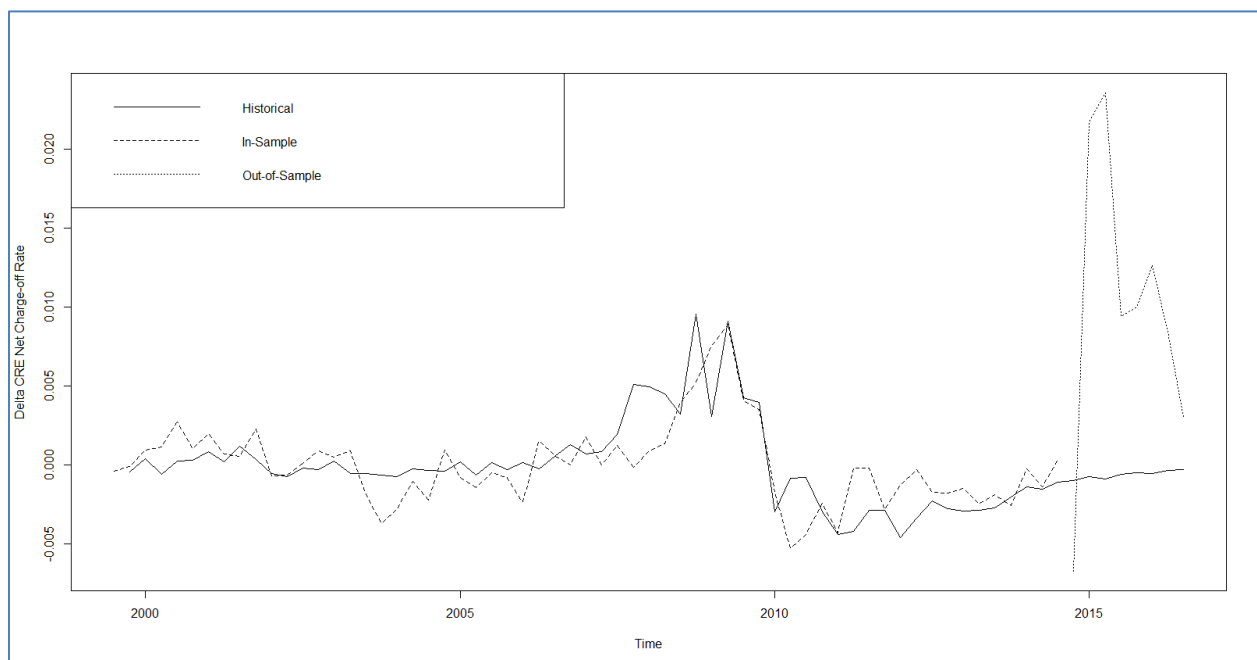
**Figure 4.6: C&I VAR Model Accuracy Plot – Fit vs. History for In-Sample (4Q99-3Q14) vs. Out-of-Sample (4Q14-3Q16) Periods**



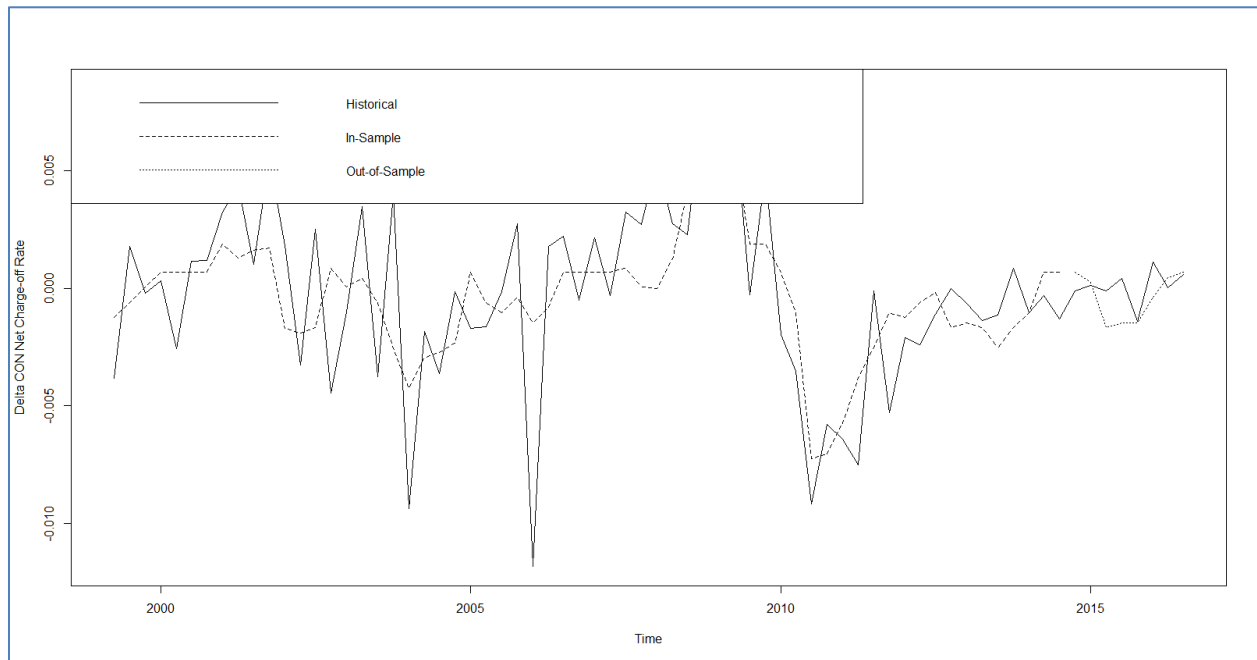
**Figure 4.7: CRE MARS Model Accuracy Plot – Fit vs. History for In-Sample (4Q99-3Q14) vs. Out-of-Sample (4Q14-3Q16) Periods**



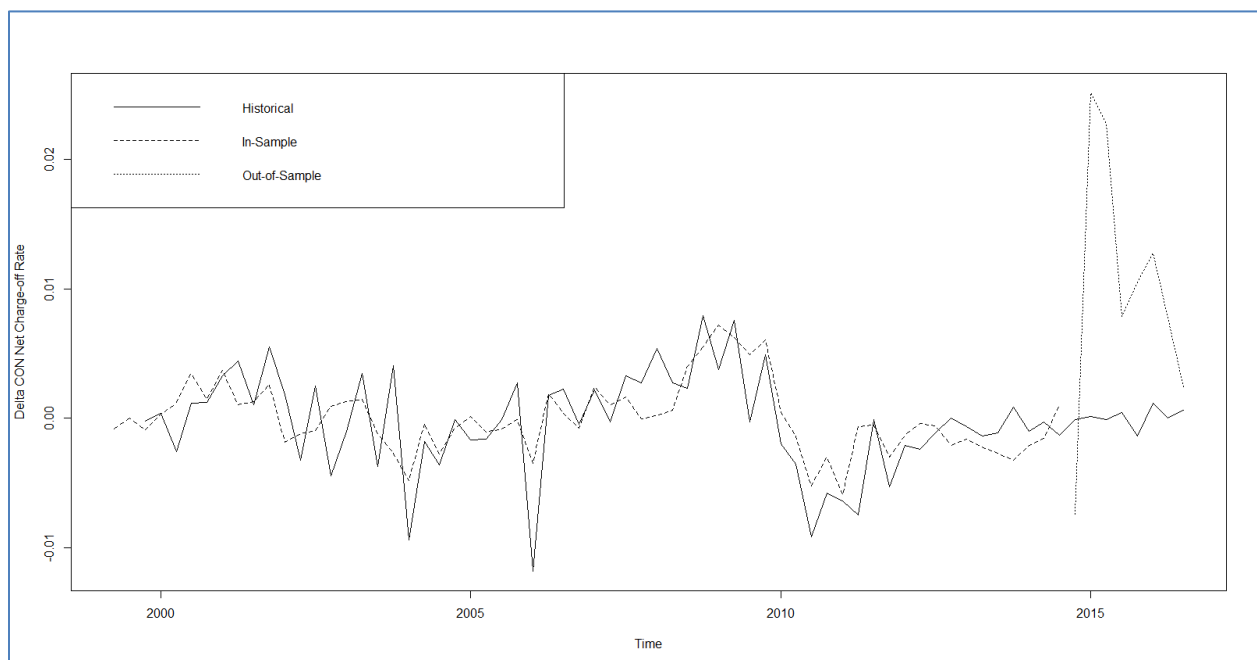
**Figure 4.8: CRE VAR Model Accuracy Plot – Fit vs. History for In-Sample (4Q99-3Q14) vs. Out-of-Sample (4Q14-3Q16) Periods**



**Figure 4.9: CONS MARS Model Accuracy Plot – Fit vs. History for In-Sample (4Q99-3Q14) vs. Out-of-Sample (4Q14-3Q16) Periods**



**Figure 4.10: CONS VAR Model Accuracy Plot – Fit vs. History for In-Sample (4Q99-3Q14) vs. Out-of-Sample (4Q14-3Q16) Periods**



**Table 4.3: Loss Estimation Results for C&I, CRE and CONS Segments– VAR and MARS Model Performance Metrics Comparison (Historical Y9 Credit Loss Rates and Federal Reserve Macroeconomic Variables 4Q99—3Q14)**

		Model Performance Metrics	Development Sample	Full Sample	Downturn Period	Out-of-Time Sample
Multivariate Adaptive Regression	Commercial and Industrial	Generalized Cross Validation	3.47E-06	2.93E-06	1.87E-05	3.18E-05
		Squared Correlation	34.16%	39.28%	79.68%	27.37%
		Root Mean Squared Error	1.80E-03	1.58E-03	2.49E-03	1.41E-02
		Cumulative Percentage Error	-1.55E-07	-1.34E-08	-4.20E-02	-9.53E-02
		Aikaike Information Criterion	-5.07E+02	-6.29E+02	-8.60E+01	-9.41E+01
Vector Autoregression		Generalized Cross Validation	5.11E-06	3.12E-06	2.02E-05	4.07E-05
		Squared Correlation	24.05%	33.26%	12.68%	24.88%
		Root Mean Squared Error	2.20E-03	1.73E-03	3.67E-03	1.78E-02
		Cumulative Percentage Error	-2.12E+00	2.33E+00	6.58E-01	-1.84E+00
		Aikaike Information Criterion	-4.86E+02	-5.75E+02	-7.88E+01	-8.24E+01
Multivariate Adaptive Regression	Commercial Real Estate	Generalized Cross Validation	3.41E-06	3.14E-06	9.61E-04	7.23E-05
		Squared Correlation	63.43%	62.57%	4.19%	0.61%
		Root Mean Squared Error	1.77E-03	2.08E-03	1.20E-02	2.16E-03
		Cumulative Percentage Error	-1.07E-15	1.58E-15	-2.93E-01	-4.83E-01
		Aikaike Information Criterion	-5.30E+02	-5.52E+02	-6.74E+01	-8.31E+01
Vector Autoregression		Generalized Cross Validation	4.82E-06	3.76E-06	1.50E-03	5.14E-04
		Squared Correlation	42.73%	52.43%	3.68%	0.52%
		Root Mean Squared Error	3.03E-03	2.84E-03	1.62E-02	1.79E-02
		Cumulative Percentage Error	-1.03E+00	4.00E-01	-3.77E-01	-1.80E+01
		Aikaike Information Criterion	-4.54E+02	-4.75E+02	-6.37E+01	-4.55E+01
Multivariate Adaptive Regression	Comsumer Credit	Generalized Cross Validation	9.83E-06	8.35E-06	5.07E-05	2.92E-05
		Squared Correlation	44.74%	46.23%	24.95%	2.66%
		Root Mean Squared Error	3.03E-03	2.46E-03	4.50E-03	2.70E-03
		Cumulative Percentage Error	1.16E-15	-9.71E-17	-2.59E-01	8.58E-01
		Aikaike Information Criterion	-4.57E+02	-5.38E+02	-8.22E+01	-7.71E+01
Vector Autoregression		Generalized Cross Validation	1.78E-05	1.68E-05	5.44E-05	5.11E-04
		Squared Correlation	39.23%	39.83%	20.00%	1.16%
		Root Mean Squared Error	3.79E-03	2.77E-03	5.55E-03	1.79E-02
		Cumulative Percentage Error	-9.89E-01	-5.26E-01	-4.73E-01	1.22E+02
		Aikaike Information Criterion	-2.33E+02	-2.61E+02	-6.94E+01	-4.55E+01

in the 1<sup>st</sup> order, and 0.0013 and 0.0003 in the 2<sup>nd</sup> order, for BBBCY and UNEMP, respectively) and they are all statistically significant. On the other hand, the optimal MARS model has 4 terms including an intercept, and as in the VAR model the signs on the coefficient estimates are intuitive (note that for MARS they are inverted), of reasonable magnitudes (0.0012 for BBBCY; and 0.031 and -0.0038 for UNEMP; note that the flip in signs for the UNEPM is still admissible due to the

relative magnitudes, as the leading negative term dominates) and they are all statistically significant.

In Table 4.3 we present a comparison of model performance metrics (Generalized Cross Validation – GCV, Squared Correlation - SC, Root Mean Squared Error – RMSE, Cumulative Percentage Error – CPE and Aikake Information Criterion - AIC) across key time periods (development sample- 3Q99-4Q14, full sample: 3Q99-3Q16, downturn period: 1Q08-1Q10, out-of-sample: 4Q14-3Q16) for the MARS and VAR estimations. We make the following conclusions in comparing the model estimation results:

- We observe that, generally, across metrics and time periods, the MARS model outperforms the VAR model.
- There are some notable differences across segments – in particular, for the CRE and CONS portfolios, the out-of-sample performance of the VAR model is much worse than the MARS model.
- Furthermore, the MARS model is generally more accurate over the downturn period than the VAR model, showing more accuracy and less underprediction.
- Finally, according to the CPE measure of model accuracy – one preferred by regulators – the MARS model performs better by several orders of magnitude.

In the case of the C&I model, we observe that in the development sample, all model performance metrics are superior in the MARS as compared to the VAR model: GCV is about 25% lower, SC about 40% higher, RMSE about 5% lower and AIC about 20% lower in the former as compared to the latter; and strikingly, the CPE is several order of magnitudes lower, nearly nil in the MARS model, as compared to underprediction of about -10% in the VAR model. However, in the full sample recalibration of the model, the superiority of the MARS model narrows somewhat, as compared to VAR in MARS we have that: GCV is about 15% lower, SC about 20% higher, RMSE about still 5% lower and AIC about 10% lower in the former as compared to the latter; however, , the CPE is still several order of magnitudes lower, nearly nil in the MARS model, as compared to underprediction of about -23% in the VAR model. Turning to out-of-sample performance, superiority of the MARS model narrows but holds across metrics, as compared to VAR in MARS we have that: GCV is about 20% lower, SC about 10% higher, RMSE about still 20% lower and AIC about 15% lower in the former as compared to the latter; however, the CPE is an order of magnitudes lower, about -10% in the MARS model, as compared to underprediction of about -184% in the VAR model. Finally considering downturn sample performance, we see once again that the superiority of the MARS model across metrics and in some cases the margin is greater, as compared to VAR in MARS we have that: GCV is about 8% lower, SC about 6 times higher, RMSE about still 30% lower and AIC about 10% lower in the former as compared to the latter; and the

CPE is an order of magnitudes lower, about -4% in the MARS model, as compared to over-prediction of about 66% in the VAR model, the latter being a troubling results from a supervisory perspective regarding conservatism of the VAR model.

In the case of the CRE model, we observe that in the development sample, all model performance metrics are superior in the MARS as compared to the VAR model: GCV is about 30% lower, SC about 50% higher, RMSE about 40% lower and AIC about 17% lower in the former as compared to the latter; and strikingly, the CPE is several order of magnitudes lower, nearly nil in the MARS model, as compared to underprediction of about -100% in the VAR model. However, in the full sample recalibration of the model, the superiority of the MARS model narrows somewhat, as compared to VAR in MARS we have that: GCV is about 17% lower, SC about 20% higher, RMSE about still 30% lower and AIC about 16% lower in the former as compared to the latter; however, , the CPE is still several order of magnitudes lower, nearly nil in the MARS model, as compared to underprediction of about -23% in the VAR model. Turning to out-of-sample performance, superiority of the MARS model is in fact much amplified across metrics, as compared to VAR in MARS we have that: GCV is about 86% lower, SC about 18% higher, RMSE about still 20% lower and AIC about 98% lower in the former as compared to the latter; however, the CPE is an order of magnitudes lower, about -5% in the MARS model, as compared to overprediction of about 180% in the VAR model. Finally considering downturn sample performance, we see once again that the superiority of the MARS model across metrics and in some cases the margin is greater, as compared to VAR in MARS we have that: GCV is about 36% lower, SC about 13% higher, RMSE about still 26% lower and AIC about 10% lower in the former as compared to the latter; and the CPE is lower, about -29% in the MARS model, as compared to underprediction of about -38% in the VAR model.

In the case of the CONS model, we observe that in the development sample, all model performance metrics are superior in the MARS as compared to the VAR model: GCV is about 45% lower, SC about 14% higher, RMSE about 20% lower and AIC about twice as low in the former as compared to the latter; and strikingly, the CPE is several order of magnitudes lower, nearly nil in the MARS model, as compared to underprediction of about -99% in the VAR model. However, in the full sample recalibration of the model, the superiority of the MARS model narrows somewhat, as compared to VAR in MARS we have that: GCV is about 50% lower, SC about 16% higher, RMSE about 10% lower and AIC about half as large in the former as compared to the latter; however, the CPE is still several order of magnitudes lower, nearly nil in the MARS model, as compared to underprediction of about -53% in the VAR model. Turning to out-of-sample performance, superiority of the MARS model is in fact much amplified across metrics, as compared to VAR in MARS we have that: GCV is about 94% lower, SC over twice as high, RMSE about 85% lower and AIC about 170% low in the former as compared to the latter; however, the CPE is an order of magnitudes lower, about -9% in the MARS model, as compared to overprediction of about 122% in the VAR model. Finally considering downturn sample performance, we see once again that the superiority of the MARS model across metrics and in some cases the



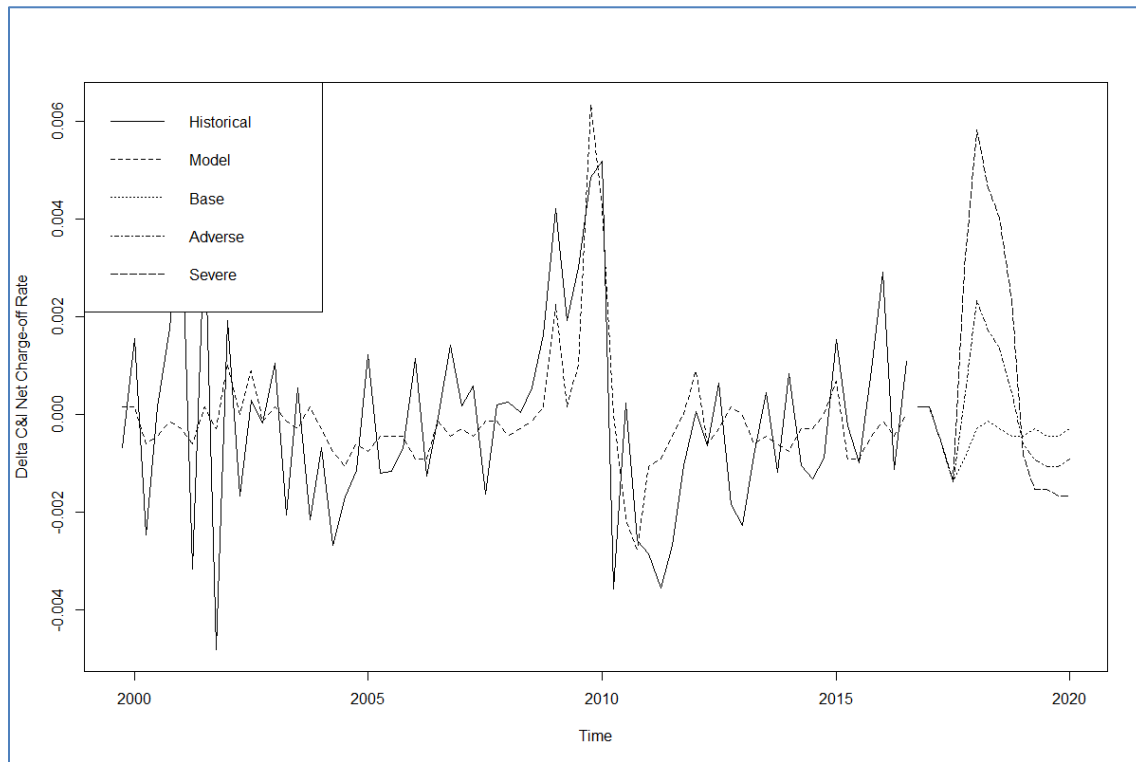
**Table 4.4: Loss Estimation Results for C&I, CRE and CONS Segments– VAR and MARS Model Scenario Comparison (Historical Y9 Credit Loss Rates and Federal Reserve Macroeconomic Variables Estimation History 4Q99—3Q16 and Forecast Period 4Q16-1Q20)**

		Multivariate Adaptive Regression Splines	Vector Autoregression
Commercial and Industrial	Cumulative Loss Base Scenerio	0.2848%	0.5440%
	Cumulative Loss Adverse Scenerio	0.6599%	0.6370%
	Cumulative Loss Severe Scenerio	1.0871%	0.7744%
Commercial Real Estate	Cumulative Loss Base Scenerio	0.4085%	0.6112%
	Cumulative Loss Adverse Scenerio	0.9807%	0.7901%
	Cumulative Loss Severe Scenerio	1.4999%	0.9716%
Consumer Credit	Cumulative Loss Base Scenerio	0.5886%	0.7966%
	Cumulative Loss Adverse Scenerio	1.1452%	0.9772%
	Cumulative Loss Severe Scenerio	1.8108%	1.3004%

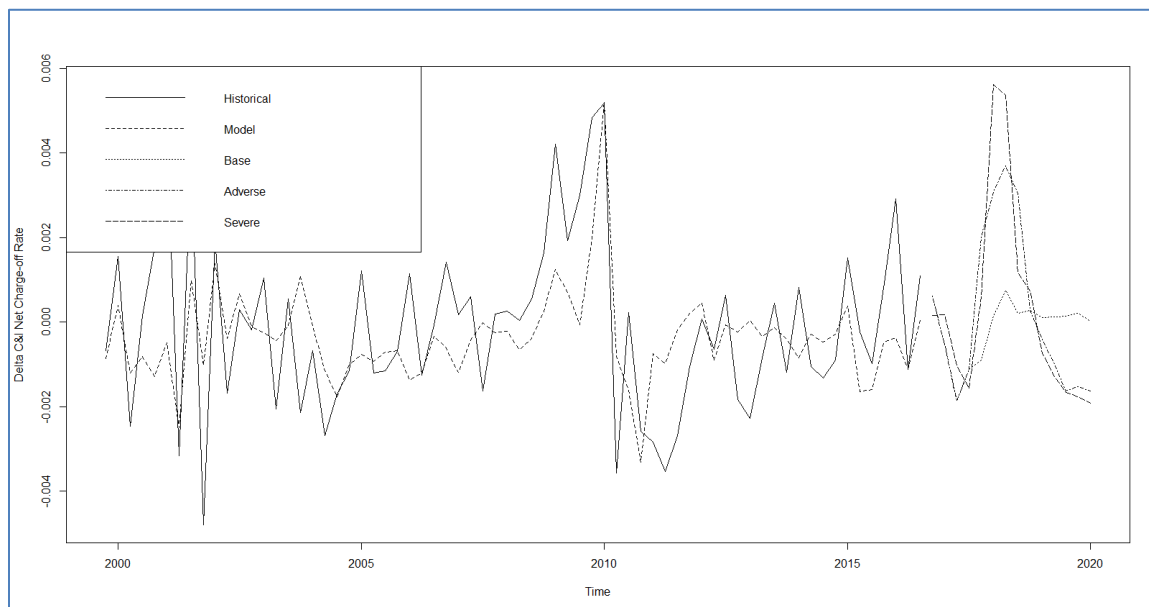
margin is greater, as compared to VAR in MARS we have that: GCV is about 7% lower, SC about 25% higher, RMSE about still 20% lower and AIC about 20% lower in the former as compared to the latter; and the CPE is lower, about -0.03% in the MARS model, as compared to underprediction of about -0.05% in the VAR model.

In Table 4.4 we show the cumulative losses for the VAR and MARS models in each of the Fed scenarios. The scenario forecasts are also shown graphically in Figures 4.11 through 4.16. Across modeling segments, we observe in Table 4.4 that the MARS model exhibits greater separation in cumulative loss between Base and either Adverse or Severe scenarios. Furthermore, while the Base scenarios are lower in the MARS than in the VAR model, in the former model the Adverse and Severe scenarios have much higher cumulative losses than in the latter model. In the case of the C&I segment, the cumulative loss in the Severe (Adverse) scenario is 1.09% (0.66%) in the MARS model, as compared to 0.77% (0.64%) in the VAR model, respectively. Furthermore, the spread in cumulative losses between the Severe (Adverse) and Base scenarios is 0.80% (0.38%) in the MARS model, versus 0.23% (0.09%) in the VAR model, respectively. In the case of the CRE segment, the cumulative loss in the Severe (Adverse) scenario is 1.50% (0.98%) in the MARS model, as compared to 0.97% (0.79%) in the VAR model, respectively.

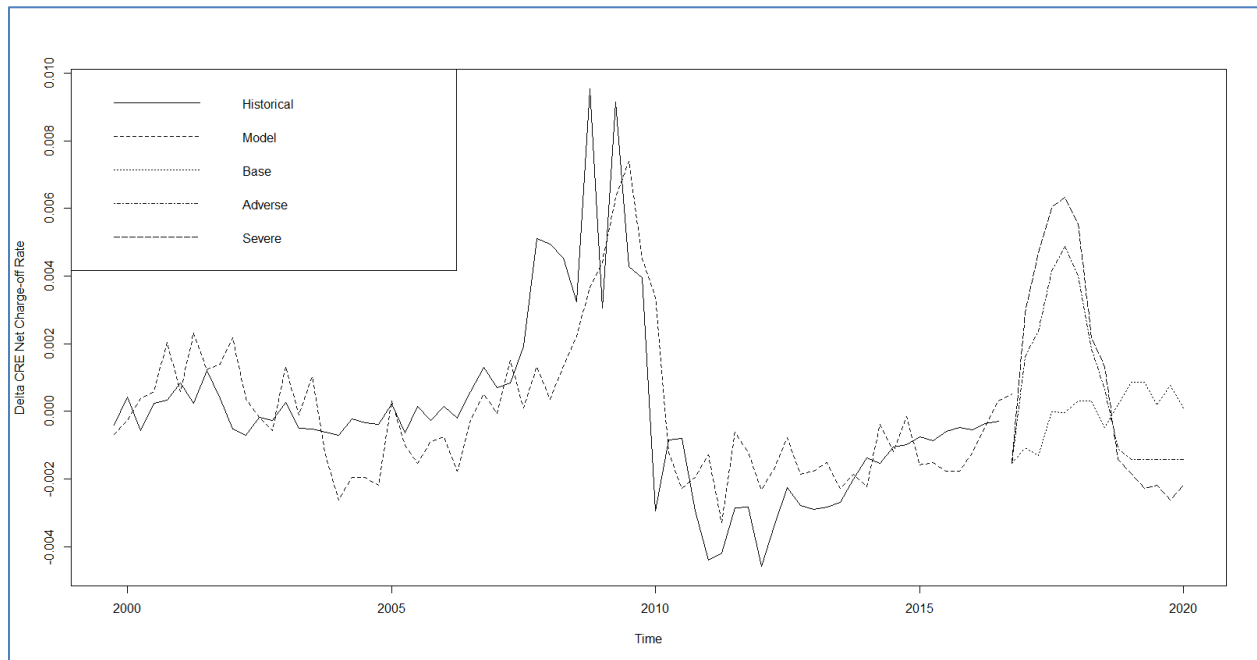
**Figure 4.11: C&I MARS Model Scenario Plot – Fit vs. History for In-Sample (4Q99-3Q16) vs. Forecast Period (4Q16-1Q20)**



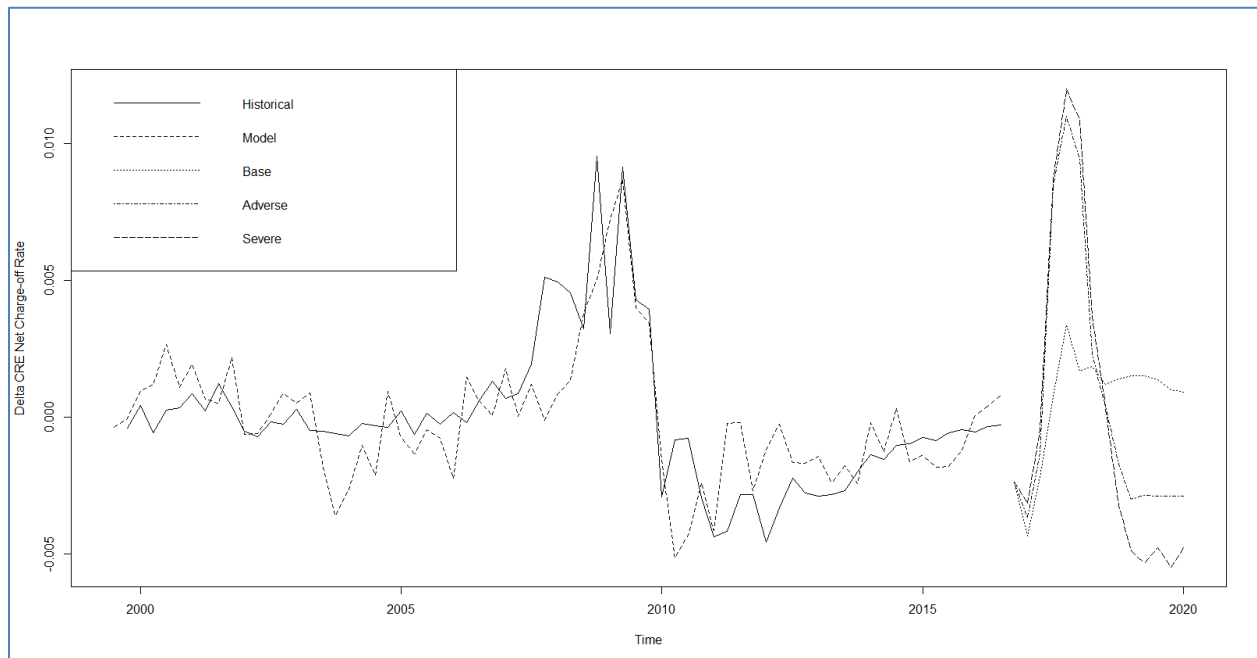
**Figure 4.12: C&I VAR Model Scenario Plot – Fit vs. History for In-Sample (4Q99-3Q16) vs. Forecast Period (4Q16-1Q20)**



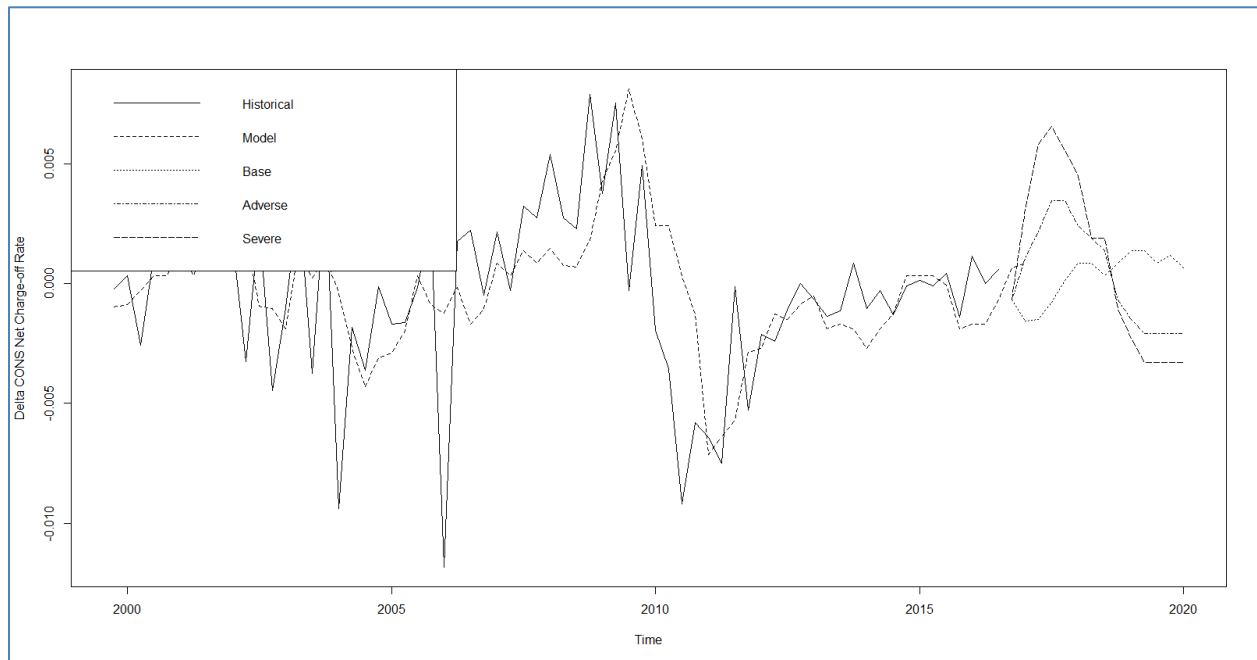
**Figure 4.13: CRE MARS Model Scenario Plot – Fit vs. History for In-Sample (4Q99-3Q16) vs. Forecast Period (4Q16-1Q20)**



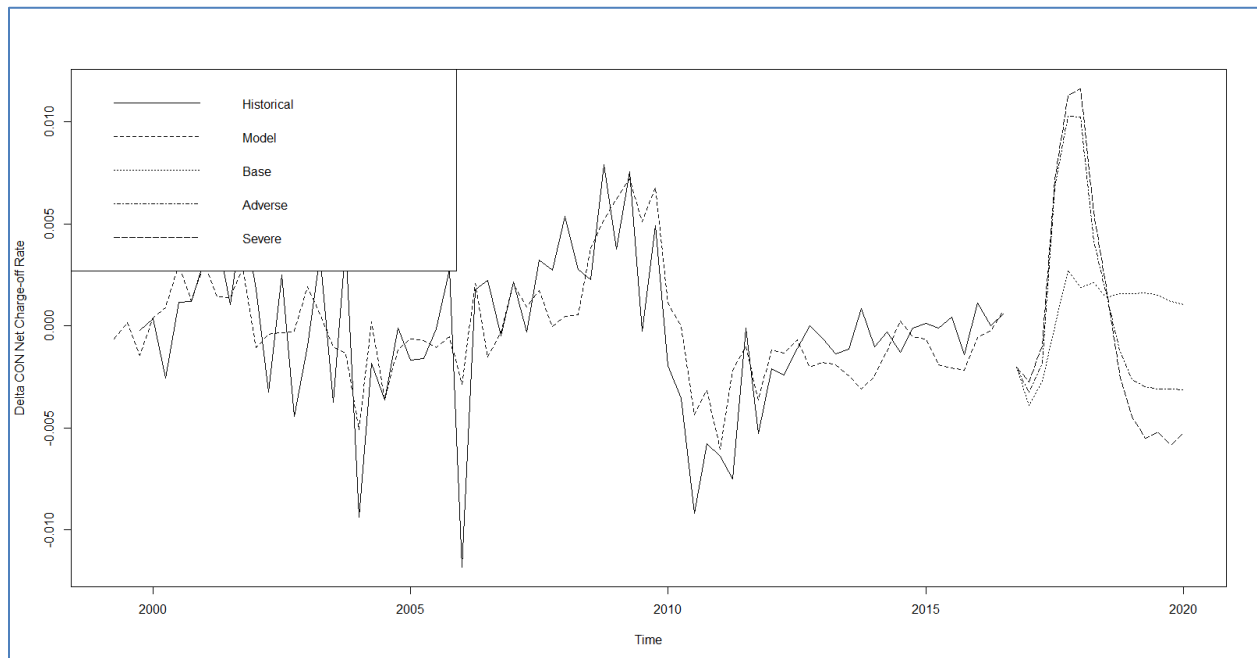
**Figure 4.14: CRE VAR Model Scenario Plot – Fit vs. History for In-Sample (4Q99-3Q16) vs. Forecast Period (4Q16-1Q20)**



**Figure 4.15: CONS MARS Model Scenario Plot – Fit vs. History for In-Sample (4Q99-3Q16) vs. Forecast Period (4Q16-1Q20)**



**Figure 4.14: CONS VAR Model Scenario Plot – Fit vs. History for In-Sample (4Q99-3Q16) vs. Forecast Period (4Q16-1Q20)**



Furthermore, the spread in cumulative losses between the Severe (Adverse) and Base scenarios is 1.09% (0.57%) in the MARS model, versus 0.36% (0.18%) in the VAR model, respectively. In the case of the CONS segment, the cumulative loss in the Severe (Adverse) scenario is 1.81% (1.15%) in the MARS model, as compared to 1.30% (0.98%) in the VAR model, respectively. Furthermore, the spread in cumulative losses between the Severe (Adverse) and Base scenarios is 1.22% (0.56%) in the MARS model, versus 0.50% (0.18%) in the VAR model, respectively.

## 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. We have considered the case of banks that model the risk of their portfolios using top-of-the-house modeling techniques. 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 machine learning approach, widely used in the academic literature, but not commonly employed in practice, the MARS model. We empirically tested these models using Federal Reserve macroeconomic data, gathered and released by the regulators for CCAR purposes, respectively. While the results of the estimation are broadly consistent across the VAR and MARS models, we find that the MARS model generally outperforms, across model performance metrics and time periods, the MARS model outperforms the VAR model. There are some notable differences across segments – in particular, for the CRE and CONS portfolios, the out-of-sample performance of the VAR model is much worse than the MARS model. Furthermore, the MARS model is generally more accurate over the downturn period than the VAR model, showing more accuracy and less underprediction. Finally, according to the CPE measure of model accuracy – one preferred by regulators – the MARS model performs better by several orders of magnitude. Furthermore, we find that the MARS model produces more reasonable forecasts, from the perspective of quality and conservatism in severe scenarios. Across modeling segments, we observe that the MARS model exhibits greater separation in cumulative loss between Base and either Adverse or Severe scenarios. Furthermore, while the Base scenarios are lower in the MARS than in the VAR model, in the former model the Adverse and Severe scenarios have much higher cumulative losses than in the latter model. .

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
- 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.
- Arel, Itamar, Derek C. Rose, and Thomas P. Karnowski. "Deep machine learning-a new frontier in artificial intelligence research [research frontier]." *IEEE computational intelligence magazine* 5.4 (2010): 13-18.
- 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.

R. Bellman, *Dynamic Programming*. Princeton, NJ: Princeton Univ. Press, 1957.

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 23<sup>rd</sup>.

\_\_\_\_\_, 1999, “Supervisory Guidance Regarding Counterparty Credit Risk Management”, Supervisory Letter 99-03, February 1<sup>st</sup>.

\_\_\_\_\_, 2002, “Interagency Guidance on Country Risk Management”, Supervisory Letter 02-05, March 8<sup>th</sup>.

\_\_\_\_\_, 2011, “Supervisory Guidance on Model Risk Management”, Supervisory Letter 11-7, April 4<sup>th</sup>.

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.

Bucheli and Thompson (2014), *Statistics and Machine Learning at Scale : New Technologies Apply Machine Learning to Big Data*, The SAS Institute, Analytics 2014 Conference.

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).

R. Duda, P. Hart, and D. Stork, *Pattern Recognition*, 2nd ed. New York: Wiley-Interscience, 2000.

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.

Friedman Jerome H. , Multivariate Adaptive Regression Splines. Author(s):. The Annals of Statistics, Vol. 19, No. 1 (Mar., 1991), pp. 1-67

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.



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

\_\_\_\_\_, S., Kazmi, M., Klein, L. and S. Regan, 2018 (January), "Emerging Trends in the Validation of Machine Learning and Artificial Intelligence Models", Accenture Consulting / Finance & Risk Services / Risk Model, Methodologies & Analytics.

\_\_\_\_\_, and F.J. Sensenbrenner, 2018 (February), Methodologies for scenario generation and dependency structures in the stress testing of credit risk, Working paper, *Forthcoming Quantitative Finance & Economics*.

J.P. Morgan, *RiskMetrics<sup>TM</sup>*, Second Edition, J.P. Morgan, New York, 1994.

\_\_\_\_\_, *CreditMetrics<sup>TM</sup>*, 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).

Tim Jones , “A beginner's guide to artificial intelligence, machine learning, and cognitive computing”, IBM.

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.

T. Lee and D. Mumford, “Hierarchical Bayesian inference in the visual cortex,” *J. Opt. Soc. Amer.*, vol. 20, pt. 7, pp. 1434–1448, 2003.

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

F. Rosenblatt, “Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms”, Spartan Books, 1962

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

Sergios Theodoridis , “Machine Learning: A Bayesian and Optimization Perspective”, First Edition, Elsevier, New York 2017.

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.

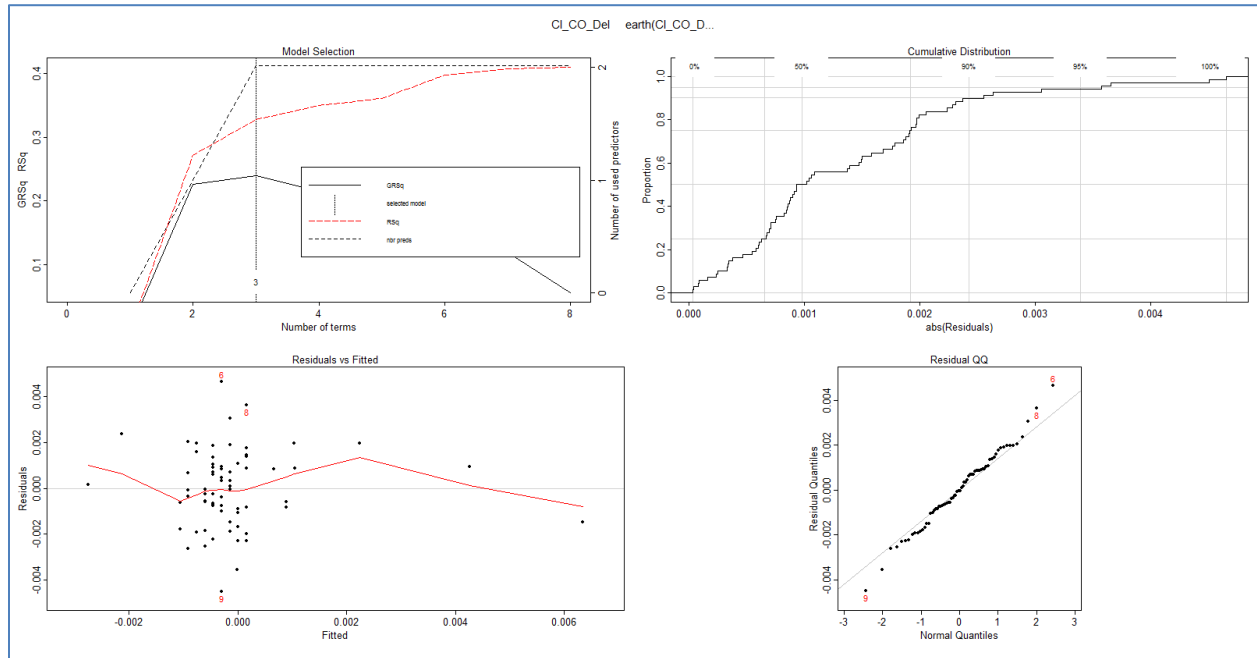
Tjøstheim, D., Estimation in Nonlinear Time Series Models, *Stochastic Processes and their Applications*, **21**(2), (1986), 251-273.

G. Wallis and E. Rolls, “Invariant face and object recognition in the visual system,” *Prog. Neurobiol.*, vol. 51, pp. 167–194, 1997

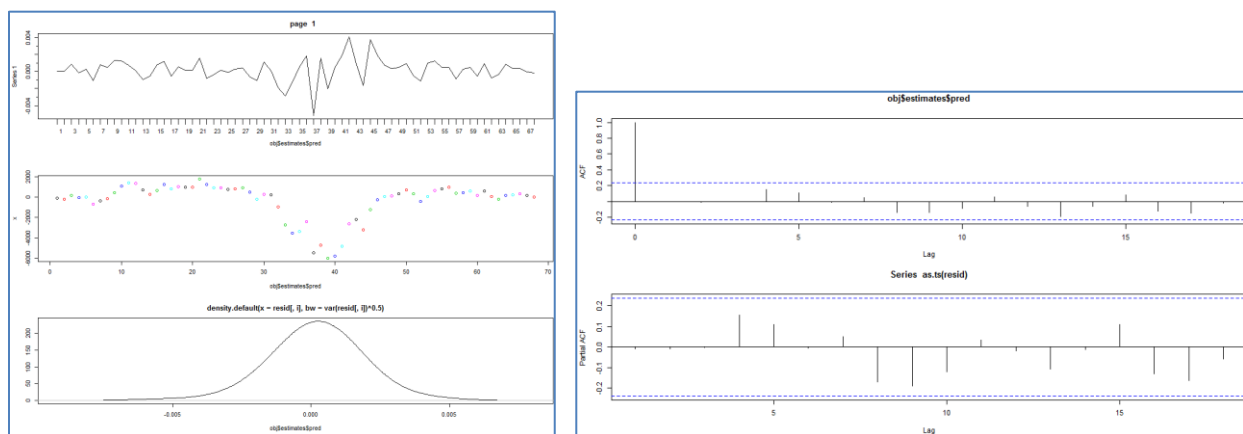
Wilde, T., *CreditRisk+. A Credit Risk Management Framework*, Credit Suisse First Boston, 1997.

## 7 Appendix

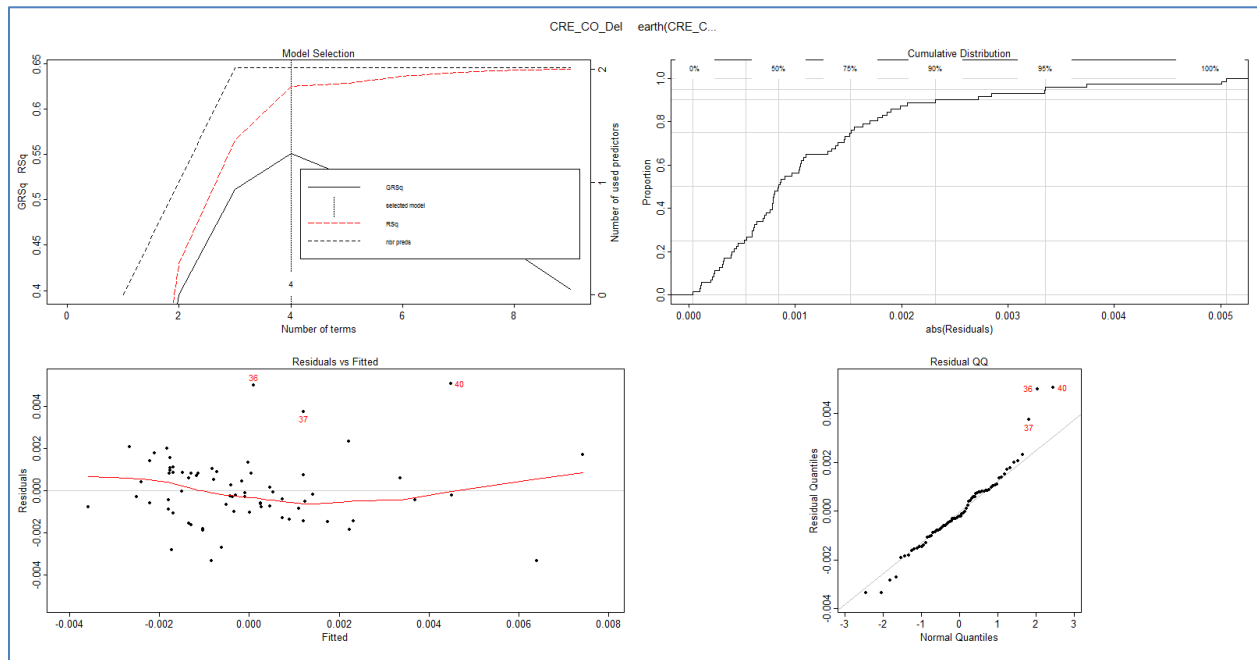
**Figure 7.1: C&I MARS Model Estimation Diagnostic Plots Scenarios (Historical Y9 Credit Loss Rates and Federal Reserve Macroeconomic Variables 4Q99—3Q14)**



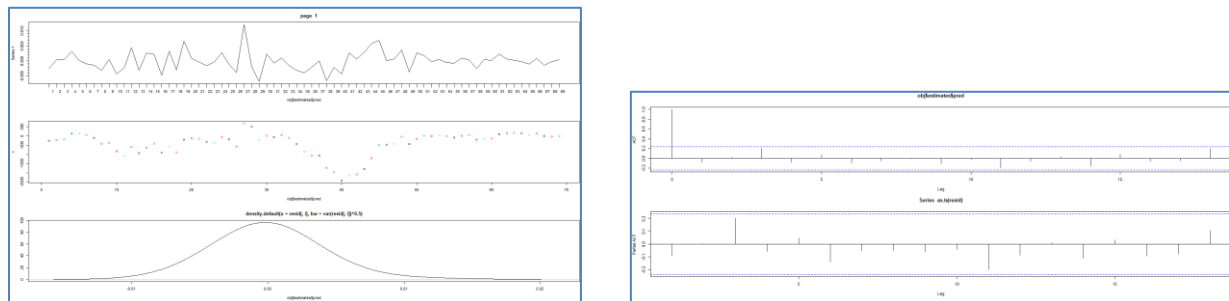
**Figure 7.2: C&I Model Estimation Diagnostic Plots Scenarios (Historical Y9 Credit Loss Rates and Federal Reserve Macroeconomic Variables 4Q99—3Q14)**



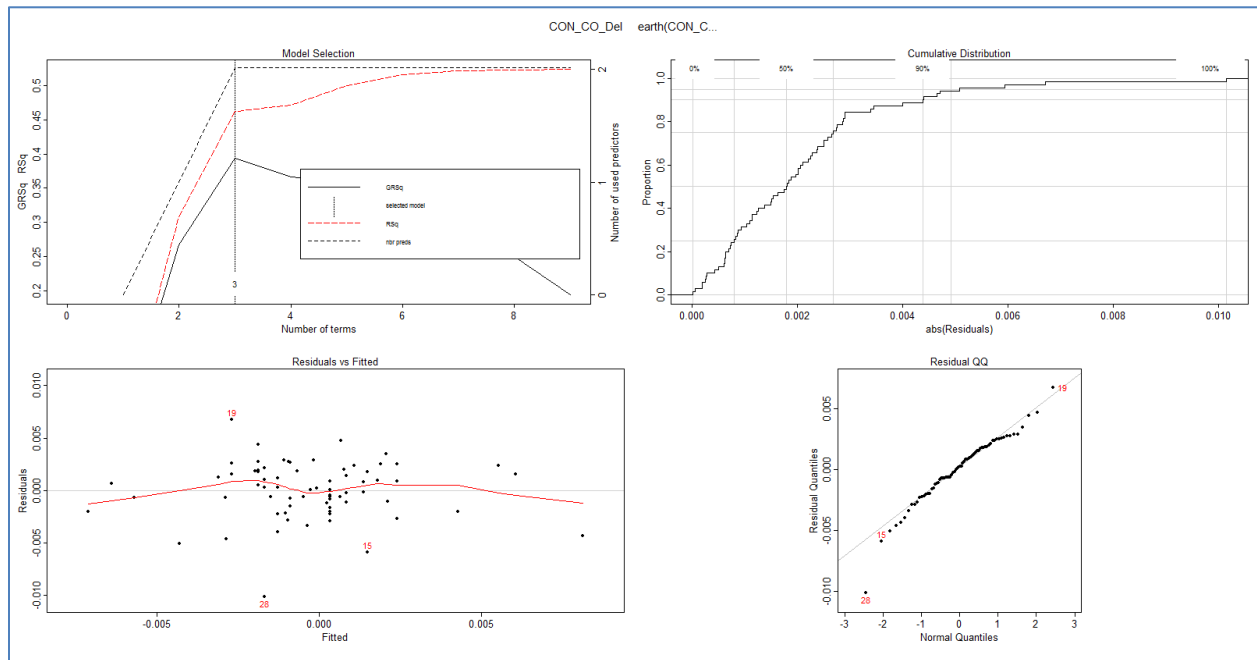
**Figure 7.3: CRE MARS Model Estimation Diagnostic Plots Scenarios (Historical Y9 Credit Loss Rates and Federal Reserve Macroeconomic Variables 4Q99—3Q14)**



**Figure 7.4: CRE Model Estimation Diagnostic Plots Scenarios (Historical Y9 Credit Loss Rates and Federal Reserve Macroeconomic Variables 4Q99—3Q14)**



**Figure 7.5: CRE MARS Model Estimation Diagnostic Plots Scenarios (Historical Y9 Credit Loss Rates and Federal Reserve Macroeconomic Variables 4Q99—3Q14)**



**Figure 7.6: CRE Model Estimation Diagnostic Plots Scenarios (Historical Y9 Credit Loss Rates and Federal Reserve Macroeconomic Variables 4Q99—3Q14)**

