Essays in Forecasting Macroeconomic Indicators

Pawin Siriprapanukul

A thesis submitted for the degree of Doctor of Philosophy
of The Australian National University
August 2009

School of Economics, College of Business and Economics
The Australian National University

The research and material presented in this thesis are original and solely conducted by the author, except where acknowledged.

Pawin Siriprapanukul

Pawin Siriprapanukul

ACKNOWLEDGEMENT

I am indebted to my supervisor, Professor Farshid Vahid. Without him, I could not finish this thesis. I also would like to thank Dr. Andrew Weiss and Professor Heather Anderson for their valuable comments and suggestions. Additionally, I am grateful to Dr. Ronald Stauber, who had patiently involved in the early development of my research writing.

On the other part of my Ph.D. life, I am indebted to my mom, Miss Yupared Rattanamateedon, and Dr. Nada Wasi, for their support and encouragement. Without them, I would not finish this work either. I am also grateful to my friends and colleagues at the ANU, who help provide enjoyable moments in the previous four and a half years of my life. These include Lili, Jing, Satoshi, Kim, Emilio, Aum, Jo, Tune, Paul, Jon, Pim, Pong, Rune, Tang, Aim, Poom, Wow, Yot, Shuhei, Ying, and many others.

Many other people provide other supports to me during this time. These include, among others, all the members of my family, many other staffs at the College of Business and Economics, ANU, many staffs at the University House, all staffs at the Office of Educational Affairs, Royal Thai Embassy, many workers at McDonald's in Braddon and Oporto in Civic, and many colleagues at Thammasat University. I want to thank all of them.

The financial support from the Royal Thai Government is gratefully acknowledged.

CONTENTS

ABSTRACT

This study comprises three papers each considering macroeconomic forecasting with non-structural models. Chapter 2 assesses forecasting performances of financial variables in comparison with non-financial variables in predicting Australian recessions. With rare recession sessions, a re-sampling assessment scheme is designed to help provide robust predictors. This involves repeatedly re-shuffling the data set and re-applying the out-of-sample assessment on this re-shuffled set. It is shown that the results from re-sampling assessment may be quite different from that of the usual out-of-sample results. We believe that the results from our re-sampling assessment are more robust, which make the best predictors from this assessment perform better in predicting Australian recessions in the future.

Chapter 3 re-examines the notion that larger Bayesian VARs usually perform better than smaller ones. Since the performances of these Bayesian VARs can be affected by a hyperparameter governing the overall dispersion of the prior distribution, we assess the performances with careful consideration on this hyperparameter value. Our results still support the idea that larger Bayesian VARs perform better than smaller ones. However, when the hyperparameter of a smaller model is carefully chosen, the improvement in performances of larger models is not as impressive as reported in the literature.

Chapter 4 compares forecasting performances between a Bayesian VAR and some other shrinkage regressions applied on the VAR. Since there is a close relationship between the Bayesian VAR and the ridge regression, some other shrinkage regressions, which have records to outperform the ridge regression in the literature, may outperform the Bayesian VAR as well. We choose the LASSO, the elastic net, and a procedure that uses the LASSO as the variable selector before applying the Bayesian VAR as our selected alternatives. We also modify these alternative shrinkage regressions to admit the decaying rate of effects of lags as in the case of the Bayesian VAR. Our empirical study shows that the LASSO outperforms the Bayesian VAR, which signals the redundancy problem in the VAR with 3 endogenous variables and 13 lags.

Chapter 1

Introduction

1.1 Overview

Economists have long been forecasting macroeconomic indicators. The demand for these forecasts comes partly from the forward-looking household sector that wants to plan its economic activities, and partly from the academic arena, where forecasting is used as a way to observe patterns in actual data, as well as to link theoretical models back to the real world. In addition, policy makers consider forecasts in order to implement policies that require a period of time to take effect. With these broad interests, a lot of effort is devoted to proposing and developing forecasting methods.

Methods of macroeconomic forecasting can be divided into two broad categories: structural and nonstructural. Structural forecasting methods try to establish relationships between macroeconomic variables according to a particular economic theory, in order to provide some conditional forecasts of the target variables. Conditional forecasts are ones that depend on the values of some exogenous variables. For example; what will be the level of Australian GDP next year, if the interest rate is slashed by one percentage point from the current level? Nonstructural methods, in contrast, employ reduced-form relationships to investigate movements of the data, with little reliance on economic theories. Their forecasts are unconditional, meaning the forecasts are under the assumption of no change in the economic environment.

In this way, structural forecasting has risen and fallen with macroeconomic theory. During the golden age of Keynesian theory, structural Keynesian econometric models were widely celebrated. However, these forecasting models started to decline after the rise of real business cycle theory in the 1980s. During the time, most economists became dissatisfied with the postulation of many ad hoc decision rules in Keynesian theory. For example, consumption functions in Keynesian models did not follow what predicted by consumer theory. Moreover, the theory cannot explain the simultaneous appearance of high inflation and unemployment in the 1970s. In contrast, the development of nonstructural forecasting methods is less disturbed by any change in direction of macroeconomic theory.

In this study, we focus on nonstructural forecasting methods. These methods can be traced back to 1920s. However, arguably, they become an important part of economics and econometrics after the book of Box and Jenkins (1970). The centerpiece of Box and Jenkins' work is the autoregressive moving average (ARMA) model. An ARMA model is a combination of autoregressive model, which is a simple linear difference equation, and moving average model, which is a weighted average of current and lagged random shocks. There is a series of studies showing that these simple dynamic models, without any involvement of economic structure, often forecast macroeconomic indicators just as well as, or even better than, large-scale Keynesian macro-econometric models. See, for example, Nelson (1972), Naylor et al. (1972), and Cooper (1972). This seminal work of Box and Jenkins started a strand of literature that has grown explosively up to the present time.

An ongoing line of research involves multivariate extension of the Box-Jenkins framework. The approach of Box and Jenkins uses only the past values of an economic variable to forecast itself in the future. A straightforward extension on this framework is the vector autoregression (VAR), which is advocated by Sims (1980) as an alternative to traditional econometric system-of-equations models. The framework employs past values of vector of endogenous variables to forecast itself in the future. This framework is relatively simple in contrast to multivariate ARMA models. Moreover, it can accommodate many other extensions of the Box-Jenkins framework, e.g. Granger causality [Granger (1969)], co-integration and error-correction [Engle and Granger (1987)], and dynamic factor model [Sargent and Sims (1977)]. All of these extensions have become standard frameworks in econometrics in the present time. Note that all of these extensions are based on linear models.

Nonlinear extensions of the Box-Jenkins framework have also attached increasing attention in recent years. However, in most part, they are applied to areas outside macroeconomic forecasting. Only one strand of nonlinear models is mentioned here. It has a close relationship to macroeconomic forecasting in general, and to this thesis in particular. It is the line of works following the idea of business cycle leading indicators, pioneered by the seminal work of Burns and Mitchell (1946). Early papers in this line employed a non-parametric graphical method as in the work of Burns and Mitchell. This method was frequently criticized as an exercise in measurement without theory. Later, Harding and Pagan (2002) showed that there is a close relationship between this line of literature and the Box-Jenkins framework. Moreover, model-based methods were developed to address several criticisms on the graphical methodology.

One of the model-based methods employed by Stock and Watson (1989) applies the dynamic factor model of Sargent and Sims (1977) to construct composite coincident and composite leading indices. The rationale of this approach is that a set of variables is driven by a limited number of common forces, and by idiosyncratic components that are uncorrelated across the variables. We can extract important information of the common forces out of the variables under analysis. Another model-based method is the Markov-switching model of Hamilton (1989). This model allows the growth rate of the variables to depend on the status of the business cycle, which is modeled as a Markov chain. Other non-linear models applied to macroeconomic forecasting include, for example, smooth-transition model [Teräsvirta (1994)], threshold autoregressive model [Potter (1995)], and binary probit/logit models for the case of predicting the binary expansion/recession indicator [Estrella and Mishkin (1998)].

With the advance of computer technology, there is an increasing interest in accommodating more data into a forecasting model. This interest follows the fact that decision-makers in reality consider a large number of economic variables in their decision frameworks. A forecasting model that can incorporate a large number of variables is the approximate factor model, popularized by Stock and Watson (2002a). The model employs principal components analysis to construct a small number of factors to put into the forecasting model. One can also see each factor as a composite leading indicator as it is hard to provide economic rationales behind the factor. Forecast combination is another popular method to combine a lot of information into a forecasting framework. This is just a simple combination of two or more individual forecasts to produce a single, pooled forecast. The theory of forecast combination was originally developed by Bates and Granger (1969). Another method of incorporating a large number of variables in a forecasting model relevant to this thesis is applying Bayesian estimation, demonstrated in De Mol et al. (2008). This method use real observations to updates a prior distribution to form a posterior distribution, of which mean and dispersion can be used to form forecasts. The evidence in the literature so far supports that these techniques improve forecasting. See, for example, Stock and Watson (2002b), De Mol et al. (2008), and Groen and Kapetanios (2008).

In evaluating a forecasting model, we need a way to measure its performance. There are two methods to evaluate the performance of forecasting models: in-sample and out-of-sample methods. In-sample method involves fitting the model with the full sample. The performance of the model can then be indicated by, for example, the regression R^2 or the in-sample mean squared forecast error. In-sample analysis can be easily biased by over-fitting due to the use of the same data for estimation and evaluation. Stock and Watson (2003) further argue that if the coefficients of the model change over time, the in-sample method can be misleading for out-of-sample forecasting performance.

In contrast, the (pseudo) out-of-sample method mimics real time forecasting practice, in the sense that each forecast will be constructed with information available up to that point in time. For example, constructing a one-month ahead forecast for *January 1981* will use information up to *December 1980* only. The performance can be measured, for example, by the mean squared forecast errors or mean absolute forecast error from a sequence of recursive or rolling regressions. West (2006) provides a comprehensive survey on tests for equal forecast accuracy. When the target variable is a binary expansion/recession indicator while the forecast is a probability of recession, similar techniques to mean squared forecast errors and mean absolute forecast error can be used to measure out-of-sample forecast accuracy. According to Diebold and Rudebusch (1989), forecast accuracy in this latter case can be measured by quadratic probability score (QPS), which is the counterpart of the mean squared forecast error. A similar loss function that assigns more weight to larger forecast errors is the log probability score (LPS). The loss functions that weight errors asymmetrically as in Elliott and Timmermann (2004) or contingency tables can be applied as well if one consider difference in importance between false alarms (prediction of recession when it does not take place) and missed signals (no prediction of recession when it takes place).

A problem with out-of-sample method is that we cannot use the full sample as our "evaluation period". In assessing a model, we need a "pre-evaluation period" in selecting parameters of the model, and fitting the first forecasting equation in order to construct the first forecast. This can become a serious problem, if there are a limited amount of sample observations. We follow the common practice in the recent forecasting literature, focusing more on the out-of-sample evaluation method.

With the nature of nonstructural models that are not based on a particular macroeconomic theory, we are not concerned about the consistency of the model. More weight is put on finite-sample forecasting performances. Moreover, under the situation where there is a suspected tradeoff between the consistency and the forecasting performances, we choose a model in favour of the latter.

1.2 Contributions

One interesting question in the forecasting literature is whether asset prices can be used to predict the real economy. The movement in asset prices is forward-looking. Hence, they have a strong potential to predict economic activities in the future. Chapter 2 of this research considers this question for the case of Australia. We compare forecasting performances between financial and non-financial variables in predicting Australian recessions, and figure out the best predictors. We follow Estrella and Mishkin (1998) in using probit models.

In determining the best predictors, it is more common in the recent forecasting literature to base our judgement on the out-of-sample performances. This is challenging for the case of predicting Australian recessions. The fact that there have been no recessions in the recent fifteen years in Australia leaves only a few recessionary sessions in the out-of-sample evaluation period. If these recessions are the results of some unusual events, the variables that perform well in this evaluation period may not perform well for the general case or in the future. This problem has not been considered before in the previous studies.

In this chapter, we design a re-sampling scheme that helps ameliorate this problem. We repeatedly reshuffle the data and apply the out-of-sample performance assessment, which should make the results more robust to an unusual event. To preserve dependency of consecutive data, we tie dependent and independent variables together and relocate a block of consecutive dates in forming each re-sampled data set, instead of randomly reshuffling the data. We believe that the results of our re-sampling scheme are more robust than ones of the usual out-of-sample assessment that consider only a few recessions in the evaluation period.

Chapter 3 and 4 consider the Bayesian VAR with Litterman prior. There is a lot of evidence in the literature on the impressive forecasting performance of the method. See, for example, Litterman (1986) and Robertson and Tallman (1999). We investigate the practice of adding large numbers of endogenous variables into the model, recently demonstrated to be possible and satisfactory by Bańbura et al. (2008). This practice can be linked with the literature on the approximated factor model, popularized by Stock and Watson (2002a) and Bai and Ng (2002). For the case of the approximate factor model, Stock and Watson (2002b) and Bernanke and Boivin (2003), for example, show that employing a large number of predictors, usually more than 100 predictors, helps improve the forecasting performances.

The way to set appropriate values for the parameters of the model is an open question in applying the Bayesian VAR. Chapter 3 considers the parameter that governs the overall dispersion of the Litterman prior. We first show that this parameter affects the forecasting performance of even the smallest model that employs only 3 endogenous variables to forecast themselves. If we set the parameter in a different way, the finding that employing large numbers of endogenous variables in the model helps improve the forecasting performance, as in Bańbura et al. (2008), may not be correct.

We propose a method that can be used to figure out the appropriate value for this parameter. This method is close to the cross validation method in statistics[1]. However, with the nature of time-series data and the amount of observations available at the present time, we just use a part of the pre-evaluation period to be our test period. This test period is used to figure out the appropriate value of the parameter to be applied in the evaluation period. We also show that we can figure out the optimum value of the parameter in the test period.

Chapter 4 looks at the Bayesian VAR with Litterman prior from a different perspective. Our interest originated from the results of Bai and Ng (2008), who show that applying a variable selection scheme before applying the procedure of the approximate factor model helps improve its forecasting performance. This means the smaller approximate factor model with only selected variables can outperform the larger model that employs every predictor in the data set. We interpret their finding as a result of the bias-variance tradeoff, well-known in the statistical literature[2]. Next, we show the close relationship between the Bayesian VAR with Litterman prior and ridge regression.

[1] See, for example, Frank and Friedman (1993) or Chapter 7 of Hastie et al. (2001) for the details of the cross validation method.

[2] See Chapter 2 of Hastie et al. (2001) for more details of this tradeoff.

With this information, alternative shrinkage regressions, which are developed on the basis of bias-variance tradeoff, that can outperform ridge regression under some circumstances may outperform Bayesian VAR as well. An obvious alternative is the LASSO proposed by Tibshirani (1996). In this chapter, we go further to the elastic net proposed by Zou and Hastie (2005), and the procedure that use the LASSO as the variable selector before applying the Bayesian VAR. We apply these alternative regressions on the VAR framework, and modify them to allow declining effects from the lags of endogenous variables. Behind this big picture, we also design a way to choose various parameters of these regressions on the VAR.

An interesting feature of the LASSO is that it shrinks some estimated coefficients to zeros. This is different from the ridge regression or the Bayesian VAR with Litterman prior, which keeps all the estimated coefficients nonzero. Putting the estimated coefficients at zeros is similar to throwing some independent variables out of the model. In this way, this study also has the implication that incorporating larger numbers of endogenous variables or longer lags in the Bayesian VAR may not always be satisfactory. This point should be investigated more thoroughly in the future.

1.3 Thesis Organization

There are five chapters in this study. Chapter 2, 3, and 4 are written in an article style; each independent and self-contained. Following each of these chapters is its appendix. Chapter 5 concludes this study. All the references appear at the end of the thesis.

Chapter 2

Predicting Australian Recessions Using Financial Variables

This chapter assesses the forecasting performances of financial and non-financial variables in predicting Australian recessions, using probit models. With rare recession periods, we design a re-sampling exercise to provide robust predictors. This exercise captures the idea that actual recessions in our evaluation period may be well predicted by a specific factor, which may not perform well in general. We find that our re-sampling results may be different from the usual out-of-sample results. According to our re-sampling assessment, the inflation rate and short-term interest rate perform best as the individual predictors for short forecast horizons, while the U.S. short-term interest rate and U.S. interest rate spread are best for longer horizons. The combinations of two predictors turn out to perform better than each individual predictor.

2.1 Introduction

The task of forecasting recessions is widely agreed as one that receives strong attention from the policy makers and the business sector. Recessions are the periods of sharp declines in real economic activity, which usually cover shorter time spans than the expansionary sessions. Policy related authorities are usually required to take some actions in these periods in order to improve the well-being of the people. Players in the market also want to plan some actions in order to avoid the strong negative effects during recessionary periods. In this study, we are interested in predicting Australian recessions using financial variables.

Financial variables have received strong attention as leading indicators for real economic activity in recent years. They are usually associated with expectations over future economic events and available promptly without any major revisions. There is also a huge body of empirical research in the last two decades that demonstrates strong predictive performances of financial variables. Stock and Watson (2003) comprehensively review this literature, covering researches that use interest rates, interest rate spreads, stock indices, dividend yields, and exchange rates to predict GDP growth and recessions in the U.S. and some European countries. One of the main conclusions from this review is that some of these financial variables have substantial marginal predictive content for real economic activity at some times in some countries. Estrella and Mishkin (1998) show that financial variables obviously perform better than popular non-financial leading indicators in predicting U.S. recessions during the period 1971 – 1995. Dotsey (1998) provides a detailed review on researches that focuses on the predictive performances of interest rate spread. Combined with his own analysis on the U.S. economy during the period 1970 – 1997 in the same piece of work, the author concludes that the spread contains useful information about future real economic activity not contained in past economic activity or past monetary policy, even though over more recent periods the spread has not been as informative as it had been in the past. For additional evidences, see, for example, Ang et al. (2006), Estrella and Hardouvelis (1991), Zellner and Hong (1989), Zellner and Min (1999), Estrella and Mishkin (1997), and Bernard and Gerlach (1998).

There are some previous studies that investigate the performances of financial variables in predicting Australian real economic activities. Lowe (1992), Alles (1995), and Karunaratne (2002) conclude that the Australian interest rate spreads have significant power in forecasting Australian real GDP growth during the period 1972 – 1997. The in-sample evaluation method is employed in these studies. Lowe (1992) and Karunaratne (2002) do not compare these performances to ones of the other variables, but Alles (1995) show that performances of the spreads are better than that of a popular composite leading index. Poke and Wells (2007) also investigate the in-sample forecasting power of Australian interest rate spread in predicting Australian real GDP growth. However, the authors argue that this power has been obviously reduced since the adoption of inflation targeting policy in 1993. Smith (2005) compares forecasting performances of various economic indicators, including financial and U.S. indicators, in predicting the growth rate of Australian GDP. This is the only study in this group that employ the out-of-sample assessment method. The author concludes that indicators reflecting status of external demand (e.g. U.S. financial indicators) perform best in predicting the Australian growth, during the period 1995 – 2004. At last, Karunaratne (2002) is the only one in this group that assesses the performance of interest rate spread in predicting Australian recessions. The author employs probit models to forecast the binary recession indicator, and concludes that the spread has significant power in predicting the recessions during 1972 – 1997. However, the assessment method is in-sample, and there is no comparison of this performance to ones of the other indicators.

We see that the out-of-sample evaluation method is more appropriate in assessing forecasting performances of various indicators from reduced-form models. Suppose we want to predict a macroeconomic indicator next quarter, the best we can do is to use the data available up to today to construct the forecast. Similarly, in producing a 1-quarter ahead forecast for 1st Quarter of 1980, for example, we should use the data available up to 4th Quarter of 1979 only. The out-of-sample assessment measures the performance of each predictor from forecasts constructed in this way. Therefore, in this study we concern more on the out-of-sample predictive performances of financial variables in predicting Australian recessions. These performances will be compared with ones of non-financial leading indicators that have been proposed in the literature, e.g.

the value of retail trade, the dwelling approvals, the non-residential approvals, and the terms of trade. Since Australia is a small open economy with the U.S. as its major trading partner, the economic situation in the U.S. may affect Australian economic activity. We also assess the out-of-sample performances of the U.S. financial variables in predicting Australian recessions.

The task of determining the best predictors for Australian recessions is challenging, especially when one wants to base his judgement on the out-of-sample forecasting performances. In an out-of-sample assessment, we need a part of the data set to estimate a forecasting model in order to produce the first forecast for a future date. Hence, the "out-of-sample evaluation period" must start at the date that leaves enough data for estimating the model the first time. In our case, the out-of-sample evaluation period starts at 1st Quarter of 1981. However, there have been no recessions in the recent 16 years in Australia. This fact leaves only 2 or 3 recessionary sessions in our out-of-sample evaluation period. Thus if we try to find good predictors on the basis of their out-of-sample performances, we will be basing our decision on their performances in forecasting these 2 or 3 recessions only. This problem is more pronounced, since the last recession in Australia was in 1990 – 1991. This was a world-wide recession triggered by the incident of the Gulf War. It is documented by some studies including Stock and Watson (1992) and Estrella and Mishkin (1998) to be the recession that many predictors, which perform well in the past, fail to predict. Unfortunately, we have to include this recession as one of a few in our out-of-sample evaluation period.

In response to these problems, we design an additional re-sampling evaluation scheme. Our assumption is that although each recession may be caused by a different factor, there is a predictor that behaves in a particular fashion before every recession. We intend to detect this predictor, which we expect to be more robust than the predictor that performs well in the period that contains only 2 or 3 recession periods. Under this scheme, we repeatedly reshuffle the data set, before applying the procedure of out-of-sample assessment. This makes it possible to find the predictor that forecasts well recessions in the whole data set. To preserve dependency of consecutive data, we tie independent and dependent variables together and relocate a block of 50 consecutive dates each time, instead of randomly rearranging the data.

We employ static probit model to forecast the binary recession/expansion variable. Our decision and the details of this model will be explained in section 2.3. Apart from this, section 2.2 provides the information about the data employed and defines our recession periods. Section 2.4 reports the in-sample and out-of-sample assessments. Section 2.5 provides the results of our re-sampling assessment. The final section of the chapter contains the conclusion.

2.2 Data Used and Recession Periods

2.2.1 Data

We compare the forecasting performances of financial variables with ones of non-financial variables that are usually used as leading indicators. Financial variables assessed in this paper are the interest rate spread, short-term and long-term interest rates, growth rate of the stock index, growth rate of nominal and real M3 money aggregate, and the inflation rate.

According to Moore (1990), non-financial variables assessed here are those related to market expectations, economic policies, or the initiation of economic activities that impacts the economy with a delay. The representatives of these data, which also have long enough series, are growth rate of the real value of retail trade, growth rate of housing approvals, growth rate of service exports, growth rate of terms of trade, and growth rate of the gold price.

Since Australia is a small open economy and the U.S. is its major trading partner, the economic situation in the U.S. may affect Australian economic activity. This is evidenced in Smith (2005). Estrella and Mishkin (1998) show that the interest rate spread and growth rate of the stock index perform best in predicting U.S. recessions. Ang et al. (2006) report that U.S. interest rates perform best in forecasting U.S. business cycles during the 1990s. Hence, we also assess the forecasting performances of the U.S. interest rate spread, U.S. interest rates, and growth rate of the U.S. stock index in predicting Australian recessions.

All data used in this study are quarterly data. Any data formally published as monthly data are quarterly averaged. An interest rate spread is the long-term interest rate minus the short-term interest rate. All other variables except interest rates, interest rate spreads, and the inflation rate are in quarterly growth rates. Australian data come from the Reserve Bank of Australia. U.S. interest rates are from FRED database of the Federal Reserve Bank of St. Louis. The series of stock indices are from the website Yahoo! Finance. Gold prices are from the Reserve Bank of Australia and the website Kitco.com. The descriptions of the data are in Table 2.1.

Our data set covers the period from 1st Quarter of 1970 to 2nd Quarter of 2008 (154 observations). Most of the non-financial variables are published with one-quarter lag by the related authorities. Because of the information lag in the consumer price index, the real monetary aggregates are with one-quarter lag as well. We move these data forward by one period. This means the variable

Table 2.1: **Description of data and their information lags**

Variables	Description	Information Lag (Quarter)
	Financial Variables	
SPREAD	10-year Australian Gov Bond rate minus 90-day bank accepted bill rate	0
SH_R	90-day bank accepted bill rate	0
LG_R	10-year Australian Gov Bond rate	0
D_SH_R	90-day bank accepted bill rate minus 3-month U.S. Treasury Bill rate	0
D_LG_R	10-year Australian Gov Bond rate minus 10-year U.S. Treasury Bond rate	0
M3_GR	Growth rate of M3	0
CUR_GR	Growth rate of currency	0
R_M3_GR	Growth rate of M3 deflated by CPI	1
R_CUR_GR	Growth rate of currency deflated by CPI	1
STOCK_GR	Growth rate of Australian All Ord Index	0
EXCH_GR	Growth rate of exchange rate (U.S. Dollar/Aus Dollar)	0
INF	Inflation (growth rate of CPI)	1
	Non-Financial Variables	
RTT_GR	Growth rate of value of retail trade deflated by CPI	1
DWELL_GR	Growth rate of dwelling approvals	1
NONRES_GR	Growth rate of non-residential approvals	1
SV_EX_GR	Growth rate of volume of service export	1
TOT_GR	Growth rate of terms of trade	1
GOLD_GR	Growth rate of gold price	0
	U.S. Financial Variables	
US_SPR	10-year U.S. Treasury Bond rate minus 3-month U.S. Treasury Bill rate	0
US_SH_R	3-month U.S. Treasury Bill rate	0
US_LG_R	10-year U.S. Treasury Bond rate	0
US_STK_GR	Growth rate of "Dow Jones Industrial Average index"	0

x_t below is the latest available data of variable x at time t. This is to mimic the real forecasting practice.

2.2.2 Defining Recession Periods

Since there is no universally accepted definition of recessions in Australia, we begin by defining what is meant by "recession" in this paper. It is widely agreed that a recession may not be just a period of two consecutive quarters of decline in real GDP. Rather it must be a period of real contraction. This makes the definition of recession depend on subjective judgment as well.

In this study, we define a recession as a period from a peak to the subsequent trough of a classical cycle in the Australian real GDP. The turning points of classical cycles are used in accordance with Harding and Pagan (1999) that it is normally the interest of policymakers to focus on classical cycles, rather than growth cycles. We use two sets of turning points, the first of which is

Figure 2.1: **Australian Real GDP and R1**

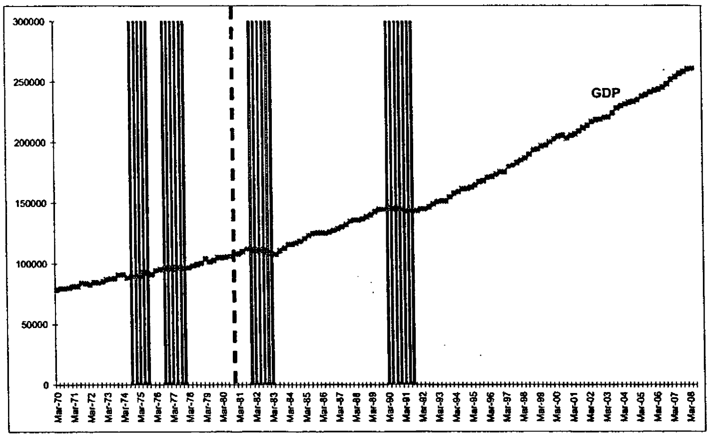
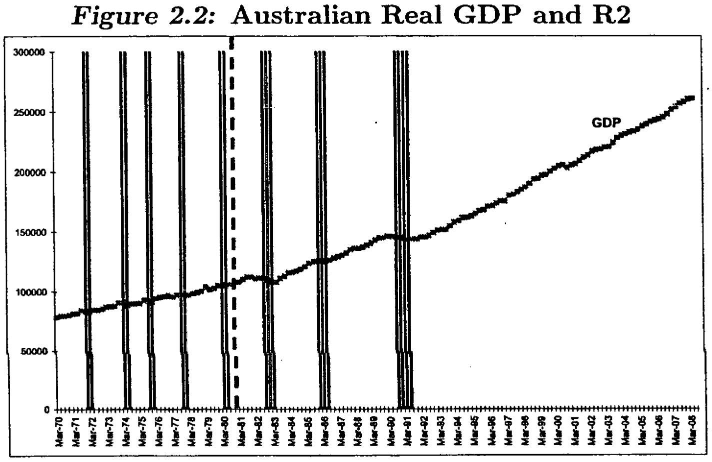


Figure 2.2: **Australian Real GDP and R2**



from the Melbourne Institute, and the second from applying GAUSS programs of Harding and Pagan (2002) to Australian real GDP.

We denote these two sets of recessions as R1 and R2, respectively. Figure 2.1 and Figure 2.2 show the data of Australian Real GDP and our recession periods R1 and R2. Bars in both figures represent these recession periods. We can see that there are more recession periods with the R2 definition. R2 is also closer to the definition that uses two consecutive quarters of decline in real GDP.

2.3 Model and Assessment Criteria

2.3.1 The Model

Following Estrella and Mishkin (1998), we use static probit model to predict the probability of recessions. The dependent variable is a dummy variable indicating quarterly recession period R_t, where:

$$R_t = \begin{cases} 1, & \text{if the economy is in recession in quarter } t, \\ 0, & \text{otherwise.} \end{cases}$$

The estimated equation is:

$$P(R_{t+h} = 1) = F(\beta' x_t), \tag{2.1}$$

where $P(R_{t+h} = 1)$ is the probability of recession in period $t + h$ with h representing the forecast horizon, ranging from 1-quarter ahead to 8-quarters ahead, $F(.)$ is the cumulative normal distribution function, β is the vector of parameters to be estimated, and x_t is the vector of independent variables. The equation is estimated using the method of maximum likelihood.

There are other alternatives to forecast recessions. Hamilton and Perez-Quiros (1996) apply Markov-switching model of Hamilton (1989) to investigate the practical usefulness of the composite leading index in forecasting. A structure in their model can be used to forecast probability of recession in the future period. However, the authors find that their model provides only a weak signal of recession in 1960, 1970 and 1990. Moreover, better forecasts can be constructed from a linear error-correction model. Smooth-transition model of

Teräsvirta (1994) can be used to forecast recessions as well. The transition function of the model is related to the probability of recession next period (See Teräsvirta, 2005, and section 8.1 of Marcellino, 2006, for more details). However, forecasting for longer forecast horizon requires the use of simulations.

We can also use econometric models that forecast turning points or growth rate of GDP to construct forecasts for recessions. This can be done through simulation techniques. See section 2.1 of Anderson and Vahid (2001) for an example of these simulations. This way, the linear AR and VAR, or the non-linear Markov-switching model [Hamilton (1989)], smooth-transition model [Teräsvirta (1994)], and threshold autoregressive model [Potter (1995)] can be used to forecast recessions in the future. Anderson and Vahid (2001) apply the aforementioned non-linear models to forecast recessions in the U.S. during 1960 – 1996 and find that these models provide just weak signals for the recessions. Marcellino (2006) also find unsatisfactory performances of linear models in forecasting recessions of the U.S. during 1990 – 2002.

With problems of their own for each alternative, we choose to apply the more direct and simpler static probit model (2.1) as in Estrella and Mishkin (1998) for this study. The model is widely adopted worldwide. See, for example, Bernard and Gerlach (1998), Birchenhall et al. (2001), Estrella et al. (2003), for European countries, and Karunaratne (2002) for Australia. Moreover, we see that section 9.2 of Marcellino (2006) shows that this probit model with appropriate regressor does not perform worse than any other linear or non-linear alternatives.

There are some later developments upon the application of probit model in predicting recessions as well. Dueker (1997) and Karunaratne (2002), for example, put lagged values of the recession indicator into the probit function.

Kauppi and Saikkonen (2008) further extend probit model to have conditional probability of the binary response depend on its lagged values, not only on lagged values of the binary response. The results of this study show that the dynanic probit model that includes the first lagged value of the binary response R_{t-1} yields the most accurate out-of-sample forecasts of U.S. recessions during 1978 – 2004, in comparison to its static counterpart and the other extensions. We see that this finding may be caused by the censoring rules applied upon constructing the binary response R_t, as discussed by Harding and Pagan (2009). Normally, when one constructs the binary indicator, he sets some censoring rules, such as a recessionary period must possess at least two consecutive quarters of negative growth in GDP. With these rules, the value of R_{t-1} becomes important in predicting the value of R_t as there is a lot higher probability that next quarter will be in recession again if this quarter falls into recession. The same effect also happens in expansionary periods. We have not considered dynamic probit models in this study yet. However, as there are substantial delays in the formal announcements of GDP data, it is not known in period t whether the economy is in recession or not. We are still not certain whether including R_{t-1} in probit models is useful in practical forecasting. Chauvet and Potter (2005) expands probit model to cover structural instabilities that the authors found in their previous study, Chauvet and Potter (2002). This extension requires the authors to use Bayesian estimation techniques, which consume much higher resources especially in an out-of-sample assessment. We have not considered this extension either. However, with limited recessionary dates in Australia, we do not expect a lot of gain in forecasting accuracy from structural instability, especially when compared with the increase in costs.

We start from using only one explanatory variable and a constant in the vector x_t in (2.1). The forecasting performances of each variable are assessed and compared. In the remainder of the paper, the "in-sample regressions" refers to when we employ the whole data set to estimate each model. Since the total number of observations is 154, the numbers of observations included in the estimations of 1-quarter ahead to 8-quarters ahead range from 153 to 146.

In "out-of-sample assessments", we fix our evaluation period to be from 1st Quarter of 1981 to 2nd Quarter of 2008 (110 quarters). This is the period after the dashed line in each of Figure 2.1 and Figure 2.2 in section 2.2. We call this period "out-of-sample evaluation period". Let t_0 and t_1 represent the positions of 1st Quarter of 1981 and 2nd Quarter of 2008 in the data set. We first use the observations from 1st Quarter of 1970 (The first observation) to period $t_0 - h$ to estimate a model and produce an h-step ahead forecast. This is the forecast for 1st Quarter of 1981. Next, we add one more observation and make a forecast for 2nd Quarter of 1981. We repeat this practice until we reach the period $t_1 - h$, when the h-step ahead forecast produces the forecast for 2nd Quarter of 2008.

The best performer in individual assessments are paired up with each of the other variables. We allow up to 5 lags of the additional variable to be combined with the best individual predictor. For example, suppose the short-term interest rate turns out to be the best 2-quarter ahead predictor of a recession. We combine the short rate at time t with each of the remaining predictors at time $t, t-1, \ldots, t-5$ to again predict a 2-quarters ahead recession. We do this because if, say, the growth rate of M3 at time t is very informative for predicting recessions 4-quarters ahead then the growth rate of M3 at time $t-2$ is likely to be useful for predicting recessions 2-quarters ahead. This is something that Estrella and Mishkin (1998) and the papers that repeat Estrella and Mishkin (1998) for other countries do not consider. Let $l_2 = 0, 1, \ldots, 5$ represent the lag of this additional variable. We are left with $154 - h - l_2$ for the in-sample estimations. We still fix the same evaluation period in the out-of-sample assessment and the forecast is still based on the data from the first observation up to the time we make the forecast (Expanding window).

2.3.2 Test Criteria

The principle measure for forecasting performances implemented in this paper is the *pseudo* R^2 used in Estrella and Mishkin (1998). The formula for this measure is:

$$pseudo\ R^2 = 1 - \left(\frac{L_m}{L_c}\right)^{-(2/n)L_c},$$

where n is the number of observations, L_m is the maximum value of log likelihood function of the model, for in-sample prediction, or the value of log likelihood function of the forecasts, for out-of-sample prediction, and L_c is the associated maximum value of log likelihood of the model with only a constant term.

Let $\hat{P}^m_{t+h}$ be the forecasted probability of recession in period $t+h$ from a specific model m. For the case of out-of-sample prediction, the log likelihood function of the forecasts L_m can be expressed as:

$$L_m = \sum_{t=t_0-h}^{t_1-h} [(1 - R_{t+h}) \ln(1 - \hat{P}^m_{t+h}) + R_{t+h} \ln \hat{P}^m_{t+h}].$$

For the case of in-sample assessment, the sum in the above formula is over all the dates in the data set. Note that while L_m varies along the model used in making the forecasts, the value L_c does not.

There is a direct relationship between the value L_m and the log probability score (LPS), which is commonly used to measure the accuracy of probability forecasts. The LPS measures the closeness, on average, between the predicted probabilities and the actual realizations represented by the binary 0/1 variable R_{t+h}. See Diebold and Rudebusch (1989) for more details. We have $LPS = -(1/n)L_m$, where $n = t_1 - t_0$ is the number of observations in the out-of-sample evaluation period.

Since the values of L_m and L_c are always non-positive, the lower the value L_m/L_c the higher the value of LPS, which suggests the more accurate are the forecasts from the model. For an in-sample assessment, the *pseudo* R^2 ranges from 0, for the case that the predictor is uninformative (L_m is equal to L_c), to 1, for the case that the predictor is very informative (L_m is zero, which is the highest possible value). For an out-of-sample assessment, the *pseudo* R^2 may be below 0. A value below zero means that the model with predictor(s) produces worse forecasts than the model with only a constant term.

In an in-sample regression, t-statistics provide additional information. The t-statistics are used to perform the statistical hypothesis testing, which is not possible when considering only *pseudo* R^2. Since our forecasting models are dynamically incomplete, we follow Estrella and Rodrigues (1998) in calculating Newey-West HAC covariance matrix (See Estrella and Mishkin, 1998, or Estrella and Rodrigues, 1998, for more details). This covariance matrix is used in constructing our t-statistics.

2.3.3 The Re-Sampling Exercise

Since recession periods are rare in Australia, the out-of-sample assessment from actual data may not be robust. Consider our out-of-sample evaluation period (After the dashed line) in Figure 2.1, for example, there are only two recessionary sessions with our R1 definition. Moreover, the session in 1990 – 1991 is considered to be triggered by an unusual circumstance, which is the invasion of Kuwait by Iraq. The practice of the usual out-of-sample assessment will base on just a few, and maybe unusual, recession dates.

With rare recessions, we do not see the problem of structural instability as important as the problem of small sample size. Our underlying assumption here is that although each recession may be caused by a different specific factor, there is a predictor that behaves in a particular fashion before every recession, regardless of the cause or the trigger of the recession. That is, we do not deny that recessions in the seventies and in the nineties had completely different flavours, but we are assuming that there are variables that behave similarly in both instances and these variables provide robust indication of the imminence of a recession. Our re-sampling exercise is designed to find such predictors.

With this thought in mind, we consider applying a re-sampling exercise to provide robust out-of-sample assessments. The idea of the exercise is to choose random blocks of the observed sample to evaluate a predictor or a combination of predictors, instead of depending on just one specific evaluation period. The exact procedure is as follows. Firstly, we tie the data of dependent variable R_{t+h} to the independent variable(s) x_t in each date t. Note that x_t is a vector. For the case that we combine two predictors together, an element of x_t may be a lag of one variable. This is to make sure that we use independent variables at the appropriate date to forecast the dependent variable, even after we move them to the other position in the data set. Secondly, we move a block of 50 consecutive dates to the end of the data set. At last, we perform out-of-sample assessment, using the relocated block as our evaluation period. We remove a block of 50 dates, instead of randomly rearranging the new data set, to preserve the dependence structure of consecutive data in our re-sampled sets. The number 50 is chosen, because it is large enough to be an evaluation period (About one-third of the total sample observations), while it leaves every pre-evaluation period some recession dates.

For example, in an exercise we may move (R_{t+h}, x_t) from 1st quarter of 1972 to 2nd quarter of 1984 (50 quarters) to the end of the data set. Next, we start estimating a forecasting model from the unmoved data and make a forecast for 1st quarter of 1972. After this, we estimate another model from the unmoved data and the data of 1st quarter of 1972 to make a forecast for 2nd quarter of 1972. We repeat this process until we make a forecast for 2nd quarter of 1984, which is the last data moved to the end of the data set. We use these forecasts to compute the *pseudo* R^2 as in the case of out-of-sample assessment.

Given a forecast horizon h, we are left with $154-h$ observations to do individual re-sampling assessment. We run our re-sampling exercise $104 - h$ times to deplete all possibilities of moving a block of 50 consecutive dates from the data set. The reported *pseudo* R^2 is the simple average of the statistics from these repetitions.

For the case of the combinations of two variables, we allow up to $l_2 = 5$ lags of the additional variable. We are left with $154 - h - l_2$ observations in performing this exercise. Hence, we repeat it $104 - h - l_2$ times to deplete all the possibilities. The reported *pseudo* R^2 is also the simple average of the statistics from these repetitions.

Table 2.2: **In-Sample Statistics: LHS is R1, RHS are Constant and the Variable as Stated**

Dependent Variable is R1		h = Quarters Ahead							
		1	2	3	4	5	6	7	8
		Financial Variables							
SPREAD	pseudo R^2	0.079	0.128	0.146	0.127	0.101	0.057	0.035	0.020
	t-stat	-3.99*	-3.61*	-2.97*	-2.48*	-2.16**	-1.56	-1.28	-1.31
SH_R	pseudo R^2	**0.183**	0.210	**0.210**	0.177	0.133	0.082	0.050	0.029
	t-stat	**6.14***	4.92*	**4.19***	3.45*	2.80*	1.98**	1.52	1.20
M3_GR	pseudo R^2	0.011	0.031	0.101	**0.183**	**0.160**	**0.153**	0.127	0.089
	t-stat	1.14	1.68	3.23*	**4.21***	**2.19****	**2.35***	2.42*	2.38*
CUR_GR	pseudo R^2	0.098	0.091	0.080	0.072	0.113	0.135	**0.168**	**0.158**
	t-stat	3.61*	2.85*	2.34*	2.00**	2.55*	3.07*	**4.37***	**5.43***
INF	pseudo R^2	0.181	**0.263**	0.197	0.166	0.142	0.135	0.143	0.099
	t-stat	4.24*	**5.71***	4.61*	4.49*	4.54*	4.77*	4.31*	2.98*
		Non-Financial Variables							
RTT_GR	pseudo R^2	0.049	0.054	0.000	0.006	0.016	0.031	0.004	0.004
	t-stat	-1.85	-2.17**	-0.21	1.08	1.70	2.47*	1.12	1.04
DWELL_GR	pseudo R^2	0.111	0.103	0.041	0.008	0.001	0.009	0.008	0.015
	t-stat	-3.55*	-2.95*	-2.04**	-0.91	0.36	1.00	0.84	0.99
		U.S. Financial Variables							
US_SPR	pseudo R^2	0.000	0.010	0.055	0.099	0.094	0.097	0.094	0.081
	t-stat	0.28	-1.00	-1.95	-2.19**	-2.09**	-2.13**	-2.67*	-3.28*
US_SH_R	pseudo R^2	0.068	0.106	0.150	0.179	0.157	0.134	0.116	0.095
	t-stat	3.55*	3.32*	3.50*	3.51*	3.43*	3.38*	3.24*	2.71*
US_LG_R	pseudo R^2	0.093	0.102	0.103	0.102	0.087	0.069	0.057	0.047
	t-stat	4.71*	3.73*	3.29*	2.95*	2.55*	2.18**	1.87	1.62

Note: * significant at 1% ** significant at 5%

2.4 In-Sample and Out-of-Sample Assessments

2.4.1 Individual Assessment

Table 2.2 and Table 2.3 show the values of *pseudo* R^2 and t-statistics from the model with a constant and an individual variable as stated in the tables. The left-hand side (LHS) variable is R1 in Table 2.2, and R2 in Table 2.3. The tables show only variables that perform best in each group, leaving the full results shown in Table A.1 and Table A.2 of the Appendix A. Each bold-faced number indicates the maximum *pseudo* R^2 in each corresponding column.

Looking at *pseudo* R^2 and t-statistics from Table 2.2 and Table 2.3, the inflation rate (INF) and short-term interest rate (SH_R) are the best performers for

Table 2.3: **In-Sample Statistics: LHS is R2, RHS are Constant and the Variable as Stated**

Dependent Variable is R2		h = Quarters Ahead 1	2	3	4	5	6	7	8
		Financial Variables							
SPREAD	pseudo R^2	0.029	0.044	0.050	0.051	0.089	0.076	0.024	0.003
	t-stat	-2.21**	-2.25*	-2.21**	-2.52*	-2.54*	-2.35*	-1.41	-0.5
SH_R	pseudo R^2	0.087	0.099	0.090	0.077	0.083	0.072	0.037	0.021
	t-stat	4.08*	3.50*	3.08*	2.75*	2.50*	2.36*	1.8	1.42
M3_GR	pseudo R^2	0.023	0.023	0.034	0.049	0.022	0.044	0.052	0.055
	t-stat	1.92	1.94	1.51	2.40*	1.47	1.76	2.12**	2.31*
CUR_GR	pseudo R^2	0.035	0.076	0.069	0.055	0.050	0.034	0.069	**0.084**
	t-stat	2.52*	2.85*	2.40*	2.20**	2.20**	2.46*	3.58*	**3.21***
INF	pseudo R^2	**0.105**	**0.112**	**0.109**	0.057	0.017	0.020	0.010	0.013
	t-stat	**3.92***	**3.97***	**4.47***	3.19*	1.91	1.55	1.21	1.46
		Non-Financial Variables							
RTT_GR	pseudo R^2	0.037	0.048	0.006	0.004	0.006	0.041	0.022	0.013
	t-stat	-2.77*	-2.93*	-1.22	-1.09	1.41	3.11*	2.54*	1.89
DWELL_GR	pseudo R^2	0.011	0.014	0.030	0.018	0.009	0.006	0.014	0.047
	t-stat	-1.43	-1.79	-1.87	-1.55	-1.11	-0.94	1.6	2.64*
		U.S. Financial Variables							
US_SPR	pseudo R^2	0.001	0.001	0.003	0.012	0.030	0.041	0.037	0.014
	t-stat	-0.44	-0.31	-0.54	-1.05	-1.67	-1.71	-1.63	-1.20
US_SH_R	pseudo R^2	0.050	0.057	0.061	**0.081**	**0.109**	**0.116**	**0.100**	0.071
	t-stat	3.17*	2.59*	2.47*	**3.06***	**3.67***	**3.61***	**3.32***	2.87*
US_LG_R	pseudo R^2	0.054	0.064	0.062	0.071	0.082	0.082	0.071	0.062
	t-stat	3.36*	2.90*	2.73*	3.1*	3.54*	3.35*	2.87*	2.60*

Note: * significant at 1%   ** significant at 5%

1-quarter to 3-quarters ahead in-sample forecasting. The growth rate of M3 (M3_GR) and the growth rate of currency (CUR_GR) are best for 4-quarters ahead to 8-quarters ahead with R1, while the U.S. short-term interest rate (US_SH_R) is outstanding for these longer forecast horizons with R2.

Table 2.4 and Table 2.5 show the values of *pseudo* R^2 resulted from the out-of-sample assessment. The LHS variable is R1 and R2 for Table 2.4 and Table 2.5, respectively, while the RHS variables are a constant and an individual variable as stated in the tables. The missing numbers are the statistics with negative values. Each bold-faced number is the maximum value of *pseudo* R^2 in the corresponding forecast horizon (Column). Full results are in Table A.3 and Table A.4 of the Appendix A.

From both tables, the best performers for short-horizon (1-quarter to 3-quarters

Table 2.4: **Out-of-Sample *pseudo* R^2: LHS is R1, RHS are Constant and the Variable as Stated**

Dep Var = R1	h = Quarters Ahead							
	1	2	3	4	5	6	7	8
	Financial Variables							
SPREAD	0.079	0.150	0.180	0.159	0.131	0.012	...	...
SH_R	**0.148**	**0.198**	**0.190**	0.164	0.147	0.085	0.025	0.008
M3_GR	0.005	0.033	0.092	0.177	0.164	0.165	0.144	0.108
CUR_GR	0.028	0.026	0.002	0.015	0.030	0.071	0.102	**0.177**
INF	0.147	0.183	0.151	0.169	0.148	0.145	**0.156**	0.121
	Non-Financial Variables							
RTT_GR	0.058	0.062	...	...	...	0.001	...	0.000
DWELL_GR	0.047	0.034	0.029	0.003	...	...	...	0.013
	U.S. Financial Variables							
US_SPR	...	...	0.067	0.130	0.120	0.123	0.115	0.090
US_SH_R	0.049	0.113	0.190	**0.234**	**0.198**	**0.167**	0.131	0.090
US_LG_R	0.097	0.109	0.115	0.113	0.087	0.064	0.031	0.015

Table 2.5: **Out-of-Sample *pseudo* R^2: LHS is R2, RHS are Constant and the Variable as Stated**

Dep Var = R2	h = Quarters Ahead							
	1	2	3	4	5	6	7	8
	Financial Variables							
SPREAD	0.026	0.049	0.057	0.062	0.091	0.081	0.020	...
SH_R	0.104	0.119	0.106	0.088	0.092	0.079	0.025	0.000
M3_GR	0.018	0.026	0.039	0.062	0.019	0.052	0.067	0.070
CUR_GR	0.040	0.050	0.063	0.056	0.057	0.029	0.084	**0.108**
INF	**0.130**	**0.123**	**0.126**	0.075	0.009	0.021	...	0.005
	Non-Financial Variables							
RTT_GR	0.036	0.049	...	...	0.000	0.000	0.006	0.013
DWELL_GR	0.007	0.013	0.026	0.018	0.005	0.000	0.006	0.027
	U.S. Financial Variables							
US_SPR	...	...	...	0.002	0.020	0.041	0.036	0.004
US_SH_R	0.055	0.071	0.075	**0.095**	**0.118**	**0.133**	**0.110**	0.073
US_LG_R	0.059	0.067	0.047	0.059	0.072	0.074	0.058	0.049

ahead) out-of-sample forecasting are inflation rate (INF) and short-term interest rate (SH_R). This is similar to the in-sample results. However, the U.S. short-term interest rate (US_SH_R) obviously outperforms other variables in 4-quarters ahead to 6-quarters ahead forecasting and the growth rate of currency (CUR_GR) seems to be the best performer for 8-quarters ahead forecasting, regardless of the definition of recession.

Figure 2.3 plots the predicted probability of recessions R1 from the out-of-sample exercise with some selected variables. The forecast horizon is $h = 1$. The shaded areas represent recession periods. It is obvious that the short-term interest rate (SH_R) provides a strong signal to the recession during $1981 - 1983$, but fails to predict one in $1990 - 1991$. Only the growth rate of dwelling approvals (DWELL_GR) predicts recession in $1990 - 1991$. However, it also provides false signals in many periods. The U.S. financial indicators provide only weak signals on both recessions.

Figure 2.4 provides the same information as in Figure 2.3, but with forecast horizon $h = 4$. Only the short-term interest rate (SH_R) and the U.S. short-term interest rate (US_SH_R) correctly predict the $1981-1983$ recession. Similar to other variables, they provide only weak signals to the $1990 - 1991$ recession. Overall, the U.S. short-term interest rate may be better than the domestic short-term interest rate as it does not give a false signal during 1986.

Even the best performers for some forecast horizons are different in our in-sample and out-of-sample assessments, we can say that both results are close. The U.S. short-term interest rate (US_SH_R) is among the best performers in our in-sample assessment for both recession definitions, using forecast horizons $h = 4, 5, 6$. Up to this point, we see that it is the best leading indicator for these forecast horizons in predicting Australian recessions. For shorter horizons, the

Figure 2.3: **Probability of Recession: R1 and $h = 1$**

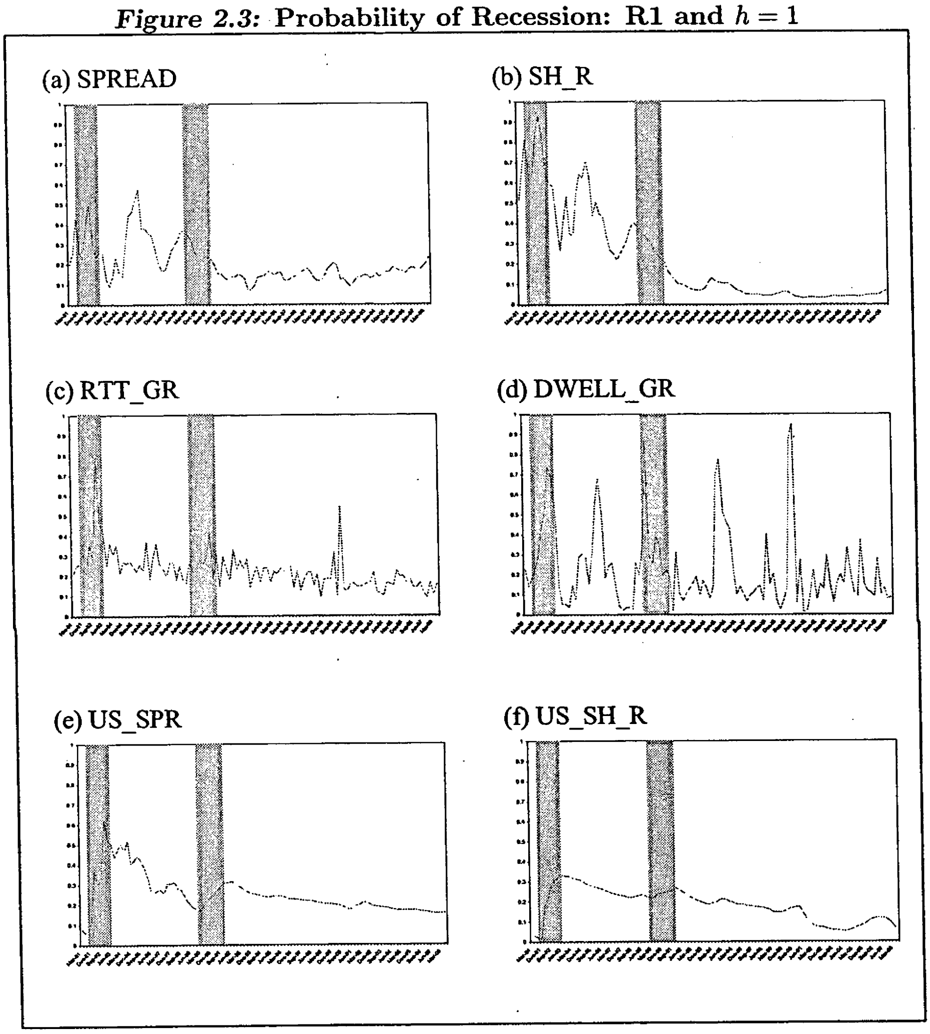
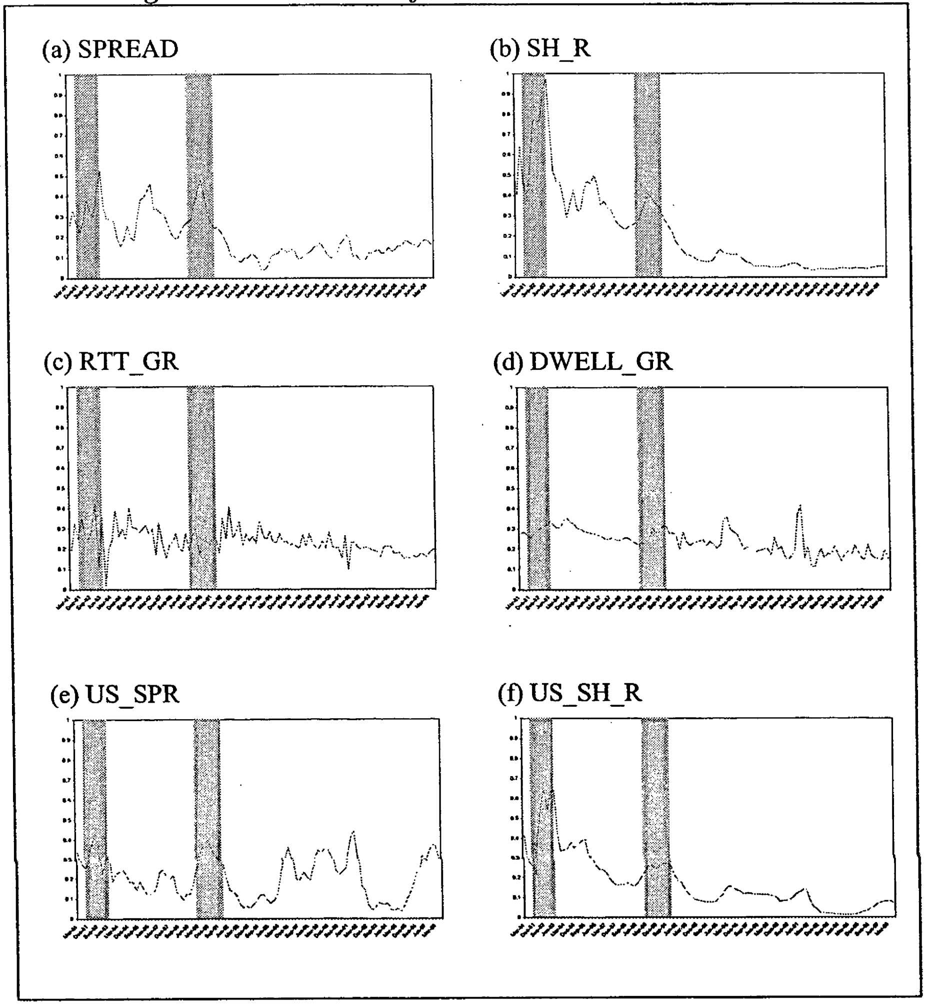


Figure 2.4: Probability of Recession: R1 and $h = 4$



short-term interest rate (SH_R) and the inflation rate (INF) perform best for both R1 and R2 recessions.

Both in-sample and out-of-sample assessments also show that financial variables outperform non-financial variables in predicting Australian recessions. The best performer among non-financial variables seems to be the growth rate of dwelling approvals (DWELL_GR). However, its forecasting performances cannot match those of the top performers in the group of financial variables.

2.4.2 Combination of Two Predictors

In this section, we assess combinations of two predictors. For short forecast horizons (1-quarter to 3-quarters ahead) we choose inflation rate (INF) to be combined with each of the other variables, and for the other horizons we choose U.S. short-term interest rate (US_SH_R). Up to 5 lags of the additional variable are allowed. As before, the bold-faced numbers indicate the maximum values of *pseudo* R^2 in each forecast horizon.

Table 2.6 presents the values of in-sample t-statistics and *pseudo* R^2 for the models using a constant, inflation rate (INF) and another variable stated in the table, as the independent variables. The LHS variable is R1 and R2 for the left and right panels of the table, respectively. Full results are in table A.5 of the Appendix A.

When R1 is used, it is obvious that the combination of inflation rate (INF) and short-term interest rate (SH_R) is the best performer for short forecast horizons (1-quarter to 3-quarters ahead). However, this is not the same when R2 is used.

Table 2.7 presents the values of in-sample t-statistics and *pseudo* R^2 for the models using a constant, U.S. short-term interest rate (US_SH_R), and another variable as independent variables. They are used with 4-quarters to 8-quarters ahead forecast horizons. The LHS variable is R1 and R2 for the left and right panels of the table, respectively. Full results are in table A.6 of the Appendix A.

When R1 is used, the combination of U.S. short-term interest rate and the growth rate of currency (CUR_GR), and the combination of U.S. short-term interest rate and the growth rate of M3 (M3_GR) perform best in long forecast horizons. However, when we switch to R2, it is quite obvious that the

Table 2.6: **In-Sample Statistics: RHS are Constant, INF, and the Variable as Stated**

Dependent Variable is R1	lag		h = Quarters Ahead 1	2	3	Dependent Variable is R2	lag		h = Quarters Ahead 1	2	3
SPREAD		pseudo R^2	0.216	0.330	0.281	SPREAD		pseudo R^2	0.114	0.130	0.132
	0	t-stat	-2.66*	-2.82*	-2.42*		0	t-stat	-1.20	-1.45	-1.45
		t-stat INF	3.86*	5.12*	4.15*			t-stat INF	3.58*	3.49*	3.90*
		pseudo R^2	0.245	0.336	0.257			pseudo R^2	0.117	0.127	0.124
	1	t-stat	-2.98*	-2.56*	-1.92		1	t-stat	-1.17	-1.13	-1.11
		t-stat INF	3.60*	5.04*	3.85*			t-stat INF	3.30*	3.12*	3.52*
		pseudo R^2	0.260	0.322	0.239			pseudo R^2	0.120	0.125	0.149
	2	t-stat	-3.11*	-2.12**	-1.56		2	t-stat	-1.19	-1.08	-1.67
		t-stat INF	3.60*	5.06*	3.71*			t-stat INF	3.22*	3.05*	3.39*
		pseudo R^2	0.241	0.295	0.208			pseudo R^2	0.122	0.154	0.141
	3	t-stat	-2.57*	-1.40	-0.73		3	t-stat	-1.33	-1.83	-1.52
		t-stat INF	3.59*	4.66*	3.58*			t-stat INF	3.23*	2.93*	3.27*
		pseudo R^2	0.246	0.282	0.207			pseudo R^2	0.164	0.161	0.116
	4	t-stat	-2.60*	-0.96	-0.75		4	t-stat	-2.70*	-2.04**	-0.75
		t-stat INF	3.97*	5.02*	3.99*			t-stat INF	3.76*	3.46*	3.87*
		pseudo R^2	0.217	0.277	0.204			pseudo R^2	0.161	0.123	0.110
	5	t-stat	-1.57	-0.87	-0.80		5	t-stat	-2.62*	-0.99	0.02
		t-stat INF	4.03*	5.23*	4.21*			t-stat INF	3.90*	3.61*	4.29*
SH_R		pseudo R^2	0.277	0.374	0.314	SH_R		pseudo R^2	0.144	0.158	0.149
	0	t-stat	4.29*	3.45*	3.05*		0	t-stat	2.36*	2.23**	1.91
		t-stat INF	3.12*	4.16*	3.32*			t-stat INF	2.94*	2.75*	3.12*
		pseudo R^2	0.295	0.365	0.281			pseudo R^2	0.150	0.150	0.136
	1	t-stat	4.37*	3.12*	2.46*		1	t-stat	2.50*	1.90	1.49
		t-stat INF	3.08*	4.43*	3.32*			t-stat INF	2.74*	2.59*	3.06*
		pseudo R^2	0.300	0.344	0.252			pseudo R^2	0.145	0.141	0.141
	2	t-stat	4.31*	2.55*	1.92		2	t-stat	2.34*	1.63	1.52
		t-stat INF	3.19*	4.44*	3.33*			t-stat INF	2.79*	2.71*	3.15*
		pseudo R^2	0.283	0.319	0.225			pseudo R^2	0.141	0.151	0.138
	3	t-stat	3.77*	2.12**	1.32		3	t-stat	2.14**	1.78	1.51
		t-stat INF	3.53*	4.68*	3.52*			t-stat INF	3.08*	2.94*	3.33*
		pseudo R^2	0.267	0.300	0.215			pseudo R^2	0.155	0.153	0.123
	4	t-stat	3.52*	1.68	1.14		4	t-stat	2.55*	1.96**	1.12
		t-stat INF	3.90*	4.94*	3.84*			t-stat INF	3.48*	3.35*	3.78*
		pseudo R^2	0.238	0.288	0.209			pseudo R^2	0.154	0.132	0.118
	5	t-stat	2.78*	1.46	1.01		5	t-stat	2.68*	1.43	0.91
		t-stat INF	4.04*	5.32*	4.15*			t-stat INF	3.69*	3.57*	4.04*
GOLD_GR		pseudo R^2	0.192	0.271	0.239	GOLD_GR		pseudo R^2	0.114	0.112	0.120
	0	t-stat	-1.41	-1.07	-1.91		0	t-stat	1.13	0.07	1.12
		t-stat INF	4.16*	5.52*	4.47*			t-stat INF	3.81*	3.91*	4.35*
		pseudo R^2	0.188	0.306	0.265			pseudo R^2	0.105	0.119	0.109
	1	t-stat	-0.89	-1.65	-1.79		1	t-stat	-0.05	1.00	-0.28
		t-stat INF	4.13*	5.51*	4.56*			t-stat INF	3.87*	3.97*	4.29*
		pseudo R^2	0.225	0.356	0.240			pseudo R^2	0.112	0.112	0.140
	2	t-stat	-2.12**	-3.03*	-2.09**		2	t-stat	0.93	-0.34	-2.36*
		t-stat INF	3.85*	5.10*	4.50*			t-stat INF	3.86*	4.07*	3.94*
		pseudo R^2	0.236	0.299	0.216			pseudo R^2	0.105	0.138	0.116
	3	t-stat	-2.34*	-1.99**	-1.39		3	t-stat	-0.18	-2.14**	-0.84
		t-stat INF	3.70*	4.97*	4.39*			t-stat INF	3.86*	3.77*	4.35*
		pseudo R^2	0.202	0.274	0.199			pseudo R^2	0.125	0.115	0.122
	4	t-stat	-1.34	-1.06	-0.52		4	t-stat	-1.78	-0.64	-1.08
		t-stat INF	4.26*	5.44*	4.47*			t-stat INF	3.60*	3.88*	4.45*
		pseudo R^2	0.214	0.293	0.201			pseudo R^2	0.120	0.144	0.171
	5	t-stat	-1.86	-1.68	-0.62		5	t-stat	-1.20	-1.68	-2.48*
		t-stat INF	4.35*	5.73*	5.01*			t-stat INF	3.79*	4.68*	5.14*
US_SH_R		pseudo R^2	0.197	0.293	0.267	US_SH_R		pseudo R^2	0.119	0.129	0.130
	0	t-stat	1.59	1.80	2.47*		0	t-stat	1.54	1.49	1.49
		t-stat INF	3.62*	4.76*	3.60*			t-stat INF	3.27*	3.17*	3.52*
		pseudo R^2	0.218	0.320	0.287			pseudo R^2	0.122	0.130	0.142
	1	t-stat	2.42*	2.48*	2.78*		1	t-stat	1.69	1.46	1.96**
		t-stat INF	3.44*	4.55*	3.41*			t-stat INF	3.16*	3.05*	3.32*
		pseudo R^2	0.240	0.325	0.256			pseudo R^2	0.120	0.138	0.157
	2	t-stat	3.00*	2.46*	2.24**		2	t-stat	1.48	1.78	2.35*
		t-stat INF	3.14*	4.16*	3.17*			t-stat INF	2.99*	2.70*	2.76*
		pseudo R^2	0.269	0.319	0.247			pseudo R^2	0.136	0.163	0.168
	3	t-stat	3.78*	2.39*	2.18**		3	t-stat	2.12**	2.46*	2.45*
		t-stat INF	3.25*	4.34*	3.39*			t-stat INF	2.99*	2.64*	2.88*
		pseudo R^2	0.255	0.306	0.239			pseudo R^2	0.158	0.172	0.157
	4	t-stat	3.57*	2.25**	2.06**		4	t-stat	2.83*	2.63*	2.14**
		t-stat INF	3.34*	4.44*	3.50*			t-stat INF	2.86*	2.59*	3.02*
		pseudo R^2	0.236	0.293	0.224			pseudo R^2	0.164	0.159	0.137
	5	t-stat	3.10*	1.77	1.50		5	t-stat	2.93*	2.19**	1.62
		t-stat INF	3.29*	4.45*	3.48*			t-stat INF	2.69*	2.60*	3.22*

Note: * significant at 1% ** significant at 5%

Table 2.7: **In-Sample Statistics: RHS are Constant, US_SH_R, and the Variable as Stated**

Dependent Variable is R1	lag		h = Quarters Ahead 4	5	6	7	8
SH_R		pseudo R^2	0.220	0.179	0.139	0.116	0.097
	0	t-stat	1.71	1.11	0.44	0.03	-0.28
		t-stat US_SH_R	1.79	1.86	2.08**	2.43*	2.42*
		pseudo R^2	0.203	0.163	0.135	0.120	0.097
	1	t-stat	1.18	0.46	0.05	-0.31	-0.19
		t-stat US_SH_R	2.25**	2.38*	2.63*	3.01*	2.62*
		pseudo R^2	0.188	0.160	0.137	0.119	0.098
	2	t-stat	0.58	0.11	-0.22	-0.14	-0.26
		t-stat US_SH_R	2.67*	2.79*	2.94*	2.91*	2.73*
		pseudo R^2	0.183	0.161	0.137	0.119	0.099
	3	t-stat	0.21	-0.14	-0.03	-0.18	-0.24
		t-stat US_SH_R	2.97*	2.98*	2.82*	2.99*	2.74*
		pseudo R^2	0.183	0.160	0.137	0.120	0.098
	4	t-stat	-0.08	0.02	-0.12	-0.22	-0.19
		t-stat US_SH_R	3.09*	2.81*	2.91*	3.04*	2.74*
		pseudo R^2	0.182	0.160	0.137	0.119	0.097
	5	t-stat	0.07	-0.09	-0.19	-0.2	-0.06
		t-stat US_SH_R	3.02*	2.93*	2.98*	3.09*	2.69*
M3_GR		pseudo R^2	0.331	0.293	0.259	0.218	0.163
	0	t-stat	4.46*	2.24**	2.49*	2.54*	2.50*
		t-stat US_SH_R	3.16*	3.30*	3.18*	3.11*	2.50*
		pseudo R^2	0.311	0.277	0.230	0.179	0.125
	1	t-stat	2.20**	2.39*	2.31**	2.36*	2.01**
		t-stat US_SH_R	3.24*	3.12*	3.00*	2.93*	2.46*
		pseudo R^2	0.294	0.249	0.193	0.144	0.100
	2	t-stat	2.31**	2.17**	2.17**	1.94	0.97
		t-stat US_SH_R	3.18*	3.11*	3.08*	3.00*	2.56*
		pseudo R^2	0.269	0.214	0.160	0.121	0.099
	3	t-stat	2.02**	1.98**	1.73	0.88	-0.50
		t-stat US_SH_R	3.20*	3.15*	3.16*	3.12*	2.64*
		pseudo R^2	0.235	0.182	0.139	0.121	0.098
	4	t-stat	1.79	1.55	0.77	-0.50	-0.07
		t-stat US_SH_R	3.19*	3.21*	3.26*	3.16*	2.59*
		pseudo R^2	0.205	0.162	0.138	0.118	0.098
	5	t-stat	1.48	0.80	-0.40	-0.03	0.34
		t-stat US_SH_R	3.29*	3.33*	3.26*	3.02*	2.51*
CUR_GR		pseudo R^2	0.203	0.216	0.214	0.229	0.200
	0	t-stat	1.06	1.69	2.19**	3.26*	4.26*
		t-stat US_SH_R	2.92*	2.81*	2.65*	2.40*	1.80
		pseudo R^2	0.237	0.235	0.244	0.217	0.195
	1	t-stat	1.67	2.13**	3.17*	4.01*	3.68*
		t-stat US_SH_R	2.95*	2.82*	2.67*	2.40*	1.88
		pseudo R^2	0.251	0.261	0.230	0.210	0.203
	2	t-stat	2.05**	3.01*	3.56*	3.25*	2.97*
		t-stat US_SH_R	2.87*	2.78*	2.59*	2.38*	1.82
		pseudo R^2	0.295	0.263	0.237	0.232	0.185
	3	t-stat	2.97*	3.55*	3.25*	2.82*	2.76*
		t-stat US_SH_R	3.22*	3.11*	2.93*	2.84*	2.08**
		pseudo R^2	0.295	0.269	0.259	0.212	0.163
	4	t-stat	3.45*	3.31*	2.85*	2.67*	2.49*
		t-stat US_SH_R	3.39*	3.33*	3.21*	2.85*	2.29**
		pseudo R^2	0.291	0.283	0.230	0.183	0.166
	5	t-stat	3.22*	2.83*	2.54*	2.24**	2.80*
		t-stat US_SH_R	3.45*	3.44*	3.17*	2.86*	2.21**
DWELL_GR		pseudo R^2	0.182	0.163	0.152	0.132	0.118
	0	t-stat	-0.53	0.65	1.24	1.04	1.13
		t-stat US_SH_R	3.50*	3.62*	3.60*	3.41*	2.87*
		pseudo R^2	0.183	0.170	0.143	0.132	0.110
	1	t-stat	0.37	0.93	0.71	0.87	0.70
		t-stat US_SH_R	3.54*	3.50*	3.40*	3.30*	2.76*
		pseudo R^2	0.189	0.164	0.146	0.127	0.098
	2	t-stat	0.71	0.50	0.72	0.58	0.19
		t-stat US_SH_R	3.51*	3.44*	3.44*	3.30*	2.69*
		pseudo R^2	0.187	0.169	0.144	0.119	0.098
	3	t-stat	0.45	0.67	0.53	0.14	0.08
		t-stat US_SH_R	3.52*	3.51*	3.53*	3.36*	2.69*
		pseudo R^2	0.188	0.165	0.137	0.119	0.098
	4	t-stat	0.51	0.41	0.02	-0.02	-0.10
		t-stat US_SH_R	3.57*	3.63*	3.73*	3.39*	2.69*
		pseudo R^2	0.187	0.160	0.136	0.118	0.102
	5	t-stat	0.40	0.00	-0.04	-0.13	0.77
		t-stat US_SH_R	3.67*	3.76*	3.67*	3.39*	2.61*
US_SPR		pseudo R^2	0.195	0.174	0.155	0.139	0.115
	0	t-stat	-0.95	-0.92	-1.03	-1.28	-1.38
		t-stat US_SH_R	3.24*	3.09*	2.58*	2.00**	1.56
		pseudo R^2	0.196	0.177	0.157	0.135	0.103
	1	t-stat	-0.94	-1.02	-1.37	-1.45	-0.73
		t-stat US_SH_R	3.44*	2.97*	2.50*	2.16**	1.90
		pseudo R^2	0.201	0.182	0.155	0.125	0.099
	2	t-stat	-1.11	-1.43	-1.59	-0.75	-0.32
		t-stat US_SH_R	3.27*	2.83*	2.51*	2.40*	2.12**
		pseudo R^2	0.201	0.176	0.142	0.120	0.106
	3	t-stat	-1.36	-1.47	-0.68	-0.24	-0.56
		t-stat US_SH_R	3.06*	2.71*	2.51*	2.41*	1.87
		pseudo R^2	0.197	0.165	0.138	0.127	0.121
	4	t-stat	-1.38	-0.64	-0.25	-0.54	-0.85
		t-stat US_SH_R	2.96*	2.70*	2.45*	2.21**	1.70
		pseudo R^2	0.187	0.161	0.145	0.142	0.114
	5	t-stat	-0.68	-0.26	-0.57	-0.84	-0.76
		t-stat US_SH_R	2.92*	2.66*	2.35*	2.08**	1.84

Dependent Variable is R2	lag		h = Quarters Ahead 4	5	6	7	8
SH_R		pseudo R^2	0.096	0.120	0.120	0.100	0.073
	0	t-stat	1.17	0.82	0.52	-0.17	-0.36
		t-stat US_SH_R	1.50	2.02**	2.37*	2.77*	2.57*
		pseudo R^2	0.104	0.118	0.117	0.103	0.073
	1	t-stat	1.25	0.77	-0.15	-0.54	-0.57
		t-stat US_SH_R	1.54	2.44*	3.00*	3.23*	2.87*
		pseudo R^2	0.101	0.111	0.121	0.105	0.075
	2	t-stat	1.22	0.14	-0.47	-0.74	-0.82
		t-stat US_SH_R	1.91	3.01*	3.38*	3.56*	3.10*
		pseudo R^2	0.086	0.114	0.126	0.107	0.073
	3	t-stat	0.57	-0.21	-0.65	-0.98	-0.57
		t-stat US_SH_R	2.44*	3.32*	3.56*	3.69*	3.00*
		pseudo R^2	0.089	0.124	0.128	0.104	0.071
	4	t-stat	0.13	-0.42	-0.90	-0.75	-0.28
		t-stat US_SH_R	2.79*	3.46*	3.64*	3.52*	2.92*
		pseudo R^2	0.103	0.124	0.124	0.101	0.073
	5	t-stat	-0.13	-0.66	-0.72	-0.38	0.54
		t-stat US_SH_R	3.06*	3.58*	3.60*	3.29*	2.56*
M3_GR		pseudo R^2	0.116	0.118	0.142	0.134	0.109
	0	t-stat	1.94	1.08	1.60	2.06**	2.29**
		t-stat US_SH_R	2.79*	3.63*	3.57*	3.22*	2.67*
		pseudo R^2	0.091	0.134	0.148	0.134	0.107
	1	t-stat	1.07	1.54	1.97**	2.22**	1.99**
		t-stat US_SH_R	2.95*	3.57*	3.48*	3.08*	2.59*
		pseudo R^2	0.105	0.139	0.150	0.134	0.073
	2	t-stat	1.54	1.83	2.03**	1.82	0.65
		t-stat US_SH_R	2.73*	3.41*	3.35*	2.95*	2.74*
		pseudo R^2	0.113	0.145	0.155	0.102	0.078
	3	t-stat	1.76	1.90	1.66	0.47	-0.76
		t-stat US_SH_R	2.66*	3.42*	3.29*	3.18*	2.82*
		pseudo R^2	0.123	0.156	0.123	0.111	0.072
	4	t-stat	1.89	1.68	0.39	-0.86	-0.34
		t-stat US_SH_R	2.89*	3.55*	3.50*	3.42*	2.76*
		pseudo R^2	0.140	0.123	0.133	0.102	0.079
	5	t-stat	1.81	0.49	-0.90	-0.41	0.71
		t-stat US_SH_R	3.23*	3.64*	4.02*	3.24*	2.62*
CUR_GR		pseudo R^2	0.106	0.126	0.122	0.130	0.118
	0	t-stat	1.42	1.29	1.00	2.77*	2.78*
		t-stat US_SH_R	2.37*	3.09*	3.22*	2.91*	2.23**
		pseudo R^2	0.103	0.116	0.145	0.141	0.121
	1	t-stat	1.40	1.01	2.37*	2.69*	3.59*
		t-stat US_SH_R	2.43*	3.25*	3.20*	2.74*	2.25**
		pseudo R^2	0.091	0.139	0.155	0.144	0.151
	2	t-stat	1.09	2.20**	2.56*	3.68*	3.76*
		t-stat US_SH_R	2.50*	3.09*	3.08*	2.70*	2.11**
		pseudo R^2	0.121	0.157	0.170	0.185	0.120
	3	t-stat	2.40*	2.72*	3.49*	3.62*	2.85*
		t-stat US_SH_R	2.48*	3.27*	3.17*	2.72*	2.43*
		pseudo R^2	0.140	0.175	0.212	0.152	0.082
	4	t-stat	2.91*	3.65*	3.51*	2.97*	1.26
		t-stat US_SH_R	2.72*	3.39*	3.15*	2.96*	2.62*
		pseudo R^2	0.158	0.212	0.172	0.110	0.079
	5	t-stat	3.80*	3.56*	2.86*	1.15	1.15
		t-stat US_SH_R	2.99*	3.37*	3.39*	3.00*	2.55*
DWELL_GR		pseudo R^2	0.095	0.115	0.119	0.122	0.133
	0	t-stat	-1.32	-0.88	-0.66	1.86	2.89*
		t-stat US_SH_R	3.06*	3.64*	3.56*	3.57*	3.12*
		pseudo R^2	0.093	0.117	0.133	0.153	0.174
	1	t-stat	-1.22	-0.95	1.64	2.88*	3.17*
		t-stat US_SH_R	3.17*	3.60*	3.73*	3.56*	3.00*
		pseudo R^2	0.093	0.122	0.165	0.197	0.128
	2	t-stat	-1.15	1.41	2.84*	3.28*	1.94
		t-stat US_SH_R	3.13*	3.63*	3.75*	3.44*	2.85*
		pseudo R^2	0.094	0.160	0.220	0.156	0.071
	3	t-stat	1.40	2.93*	3.37*	1.87	-0.08
		t-stat US_SH_R	3.00*	3.80*	3.70*	3.29*	2.89*
		pseudo R^2	0.134	0.217	0.173	0.102	0.073
	4	t-stat	2.93*	3.53*	1.80	-0.22	-0.34
		t-stat US_SH_R	3.13*	3.63*	3.45*	3.42*	2.81*
		pseudo R^2	0.208	0.179	0.123	0.102	0.076
	5	t-stat	3.73*	1.86	-0.23	-0.41	-0.82
		t-stat US_SH_R	3.37*	3.47*	3.84*	3.30*	2.82*
US_SPR		pseudo R^2	0.082	0.109	0.118	0.101	0.072
	0	t-stat	0.25	-0.21	-0.41	-0.37	0.21
		t-stat US_SH_R	2.95*	3.24*	2.93*	2.50*	2.39*
		pseudo R^2	0.084	0.113	0.118	0.100	0.080
	1	t-stat	-0.48	-0.51	-0.32	0.35	1.16
		t-stat US_SH_R	2.50*	3.01*	3.09*	3.00*	2.87*
		pseudo R^2	0.090	0.113	0.120	0.112	0.085
	2	t-stat	-0.81	-0.49	0.36	1.22	1.50
		t-stat US_SH_R	2.44*	3.08*	3.67*	3.42*	2.96*
		pseudo R^2	0.088	0.114	0.138	0.123	0.081
	3	t-stat	-0.68	0.29	1.37	1.76	0.99
		t-stat US_SH_R	2.44*	3.73*	4.07*	3.43*	2.78*
		pseudo R^2	0.089	0.137	0.148	0.115	0.081
	4	t-stat	0.02	1.32	1.90	1.31	0.80
		t-stat US_SH_R	3.09*	4.41*	3.76*	3.23*	2.63*
		pseudo R^2	0.113	0.147	0.138	0.114	0.086
	5	t-stat	1.03	1.88	1.38	0.95	0.91
		t-stat US_SH_R	3.80*	3.99*	3.34*	2.96*	2.61*

Note: * significant at 1% ** significant at 5%

Table 2.8: **Out-of-Sample *pseudo* R^2: RHS are Constant, INF, and the Variable as Stated**

Dep Var = R1	lag	h = Quarters Ahead 1	2	3	Dep Var = R2	lag	h = Quarters Ahead 1	2	3
SH_R		0.148	0.198	0.190	INF		0.130	0.123	0.126
		Financial Variables					*Financial Variables*		
SPREAD	0	0.156	0.232	0.246	SPREAD	0	0.128	0.137	0.146
	1	0.216	0.270	0.227		1	0.132	0.129	0.137
	2	0.244	0.252	0.201		2	0.137	0.134	0.159
	3	0.230	0.220	0.042		3	0.145	0.162	0.156
	4	0.241	0.138	0.088		4	0.187	0.176	0.128
	5	0.144	0.057	0.160		5	**0.189**	0.134	0.104
SH_R	0	0.200	**0.271**	**0.259**	SH_R	0	0.157	0.160	0.149
	1	0.258	0.269	0.231		1	0.165	0.138	0.129
	2	0.250	0.220	0.205		2	0.161	0.140	0.152
	3	0.246	0.244	0.134		3	0.160	0.163	0.153
	4	0.251	0.207	0.139		4	0.183	0.170	0.122
	5	0.217	0.189	0.162		5	0.185	0.133	0.108
		Non-Financial Variables					*Non-Financial Variables*		
GOLD_GR	0	0.141	0.189	0.193	GOLD_GR	0	0.135	0.119	0.107
	1	0.155	0.231	0.217		1	0.126	0.099	0.115
	2	0.197	0.247	0.192		2	0.110	0.115	0.151
	3	0.193	0.209	0.178		3	0.122	0.143	0.127
	4	0.181	0.206	0.156		4	0.141	0.119	0.145
	5	0.191	0.127	0.145		5	0.144	0.170	0.171
		U.S. Financial Variables					*U.S. Financial Variables*		
US_SH_R	0	0.084	0.115	0.205	US_SH_R	0	0.137	0.141	0.145
	1	0.112	0.175	0.254		1	0.147	0.139	0.156
	2	0.171	0.216	0.212		2	0.140	0.150	0.168
	3	**0.259**	0.237	0.206		3	0.157	0.174	**0.187**
	4	0.247	0.224	0.187		4	0.177	**0.192**	0.171
	5	0.221	0.176	0.165		5	0.188	0.168	0.141

combination of U.S. short-term interest rate and the growth rate of dwelling approval (DWELL_GR) is the best.

Table 2.8 reports the values of out-of-sample *pseudo* R^2 for specifications using a constant, inflation rate (INF), and another variable as explanatory variables. The forecast horizons are 1-quarter ahead to 3-quarters ahead. The LHS variables are R1 and R2, respectively, for the left and right panels of the table. Full results are in Table A.7 of the Appendix A. The first row of the table shows the best individual out-of-sample performances. Our best combination for each recession definition and each forecast horizon varies here. It is indicated by bold type in the table.

Table 2.9 shows the values of out-of-sample *pseudo* R^2 for specifications using a constant, U.S. short-term interest rate (US_SH_R), and another variable stated in the table, as explanatory variables. The forecast horizons are 4-

Table 2.9: **Out-of-Sample *pseudo* R^2: RHS are Constant, US_SH_R, and the Variable as Stated**

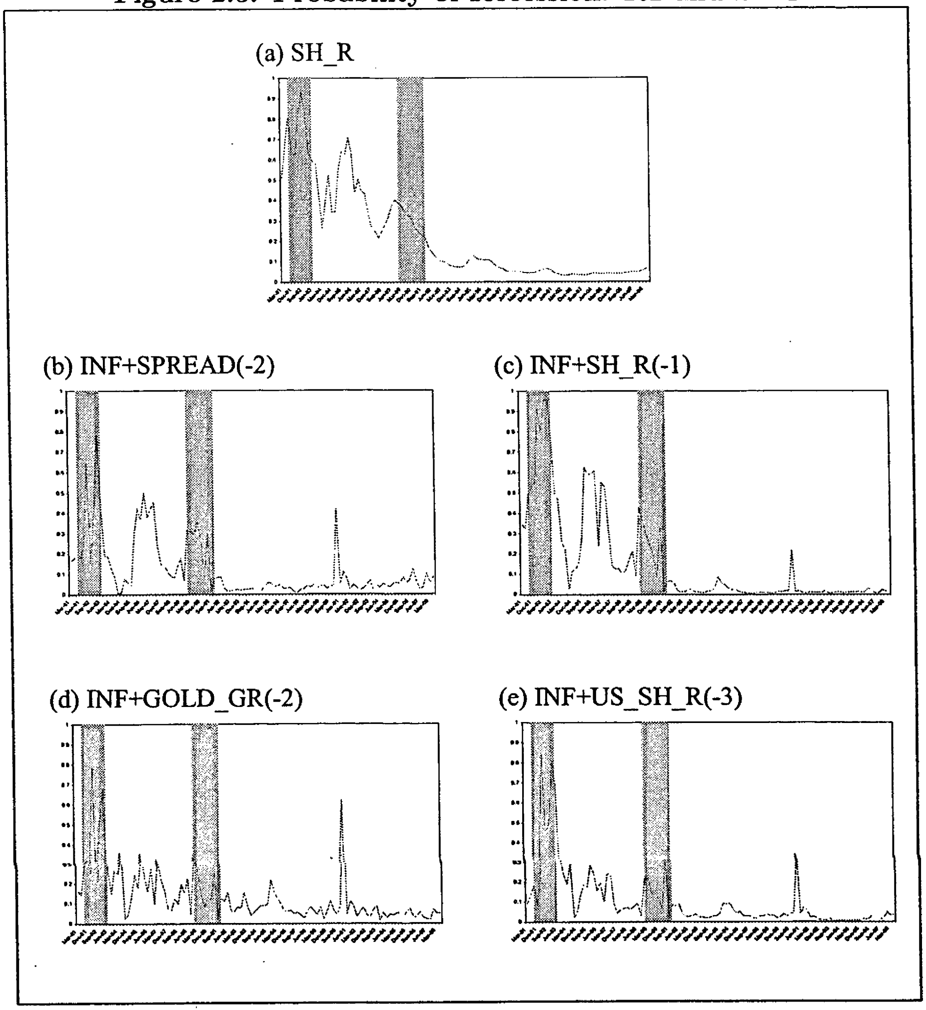
Dep Var = R1	lag	h = Quarters Ahead 4	5	6	7	8	Dep Var = R2	lag	h = Quarters Ahead 4	5	6	7	8
US_SH_R		0.234	0.198	0.167	0.131	0.090	US_SH_R		0.095	0.118	0.133	0.110	0.073
		Financial Variables							*Financial Variables*				
M3_GR	0	**0.373**	0.333	0.298	0.244	0.170	M3_GR	0	0.137	0.117	0.161	0.154	0.122
	1	0.363	0.322	0.272	0.203	0.124		1	0.092	0.137	0.173	0.150	0.110
	2	0.350	0.296	0.234	0.161	0.083		2	0.117	0.151	0.175	0.136	0.072
	3	0.327	0.261	0.195	0.126	0.051		3	0.133	0.159	0.172	0.104	0.038
	4	0.292	0.223	0.162	0.106	0.070		4	0.144	0.161	0.136	0.085	0.063
	5	0.259	0.194	0.147	0.121	0.084		5	0.148	0.124	0.114	0.097	0.075
CUR_GR	0	0.175	0.149	0.158	0.115	0.198	CUR_GR	0	0.109	0.130	0.116	0.138	0.125
	1	0.173	0.187	0.182	0.229	0.205		1	0.112	0.098	0.156	0.156	0.135
	2	0.209	0.210	0.261	0.235	**0.209**		2	0.087	0.138	0.175	0.159	**0.167**
	3	0.252	0.304	0.283	**0.259**	0.195		3	0.137	0.171	0.201	**0.204**	0.121
	4	0.350	0.320	**0.304**	0.244	0.161		4	0.166	0.193	**0.250**	0.157	0.070
	5	0.357	**0.335**	0.281	0.204	0.165		5	**0.181**	**0.230**	0.189	0.094	0.056
		Non-Financial Variables							*Non-Financial Variables*				
DWELL_GR	0	0.232	0.165	0.134	0.128	0.108	DWELL_GR	0	0.107	0.120	0.129	0.126	0.106
	1	0.201	0.159	0.157	0.142	0.101		1	0.103	0.118	0.141	0.143	0.150
	2	0.191	0.184	0.172	0.136	0.063		2	0.101	0.117	0.157	0.179	0.137
	3	0.220	0.201	0.170	0.105	0.045		3	0.097	0.136	0.209	0.167	0.032
	4	0.233	0.197	0.139	0.086	0.050		4	0.119	0.174	0.193	0.066	0.053
	5	0.231	0.166	0.121	0.094	0.089		5	0.178	0.191	0.093	0.083	0.067
		U.S. Financial Variables							*U.S. Financial Variables*				
US_SPR	0	0.239	0.200	0.170	0.132	0.094	US_SPR	0	0.044	0.089	0.109	0.087	0.047
	1	0.237	0.203	0.174	0.135	0.084		1	0.065	0.097	0.116	0.056	0.068
	2	0.249	0.212	0.167	0.121	0.077		2	0.092	0.102	0.119	0.105	0.072
	3	0.248	0.201	0.152	0.114	0.069		3	0.091	0.107	0.152	0.125	0.063
	4	0.236	0.184	0.145	0.102	0.043		4	0.086	0.136	0.163	0.106	0.064
	5	0.218	0.175	0.129	0.070	...		5	0.114	0.149	0.142	0.096	0.073

quarters to 8-quarters ahead. Full results are in Table A.8. It is obvious in Table 2.9 that the combination of U.S. short-term interest rate and growth rate of currency (CUR_GR) performs best. The best combination for 5-quarters ahead forecasting is the U.S. short-term interest rate at time t and the growth rate of currency at time $t-5$. The best combination for h-quarters ahead forecasting, $h = 6, 7, 8$, is the U.S. short-term interest rate at time t and the growth rate of currency at time $t-(10-h)$. It is also obvious here that in-sample statistics may not be good indicators for out-of-sample performances.

We can see that models with combinations of two predictors perform better than models with individual indicators out-of-sample. This can be seen from comparing the associated statistics to the ones in the first row of each table. Moreover, allowing variables with different lags into the model helps further improve its forecasting performances.

Figures 2.5 and 2.6 confirm our above finding. They plot the predicted proba-

Figure 2.5: Probability of Recession: R1 and $h = 1$



bility of R1 recessions from our out-of-sample exercises for some selected combinations in Tables 2.8 (with forecast horizon $h = 1$) and 2.9 (with forecast horizon $h = 4$), respectively. In Figure 2.5, even the combination of inflation rate and the third lag of U.S. short-term interest rate (panel e) still provides weak signal to the 1990 – 1991 recession, it does not give false signal during 1986 as in the case of using the short-term interest rate alone (panel a). In Figure 2.6, the combination of U.S. short-term interest rate and growth rate of M3 (panel b) correctly predict both recession periods.

Figure 2.6: **Probability of Recession: R1 and $h = 4$**

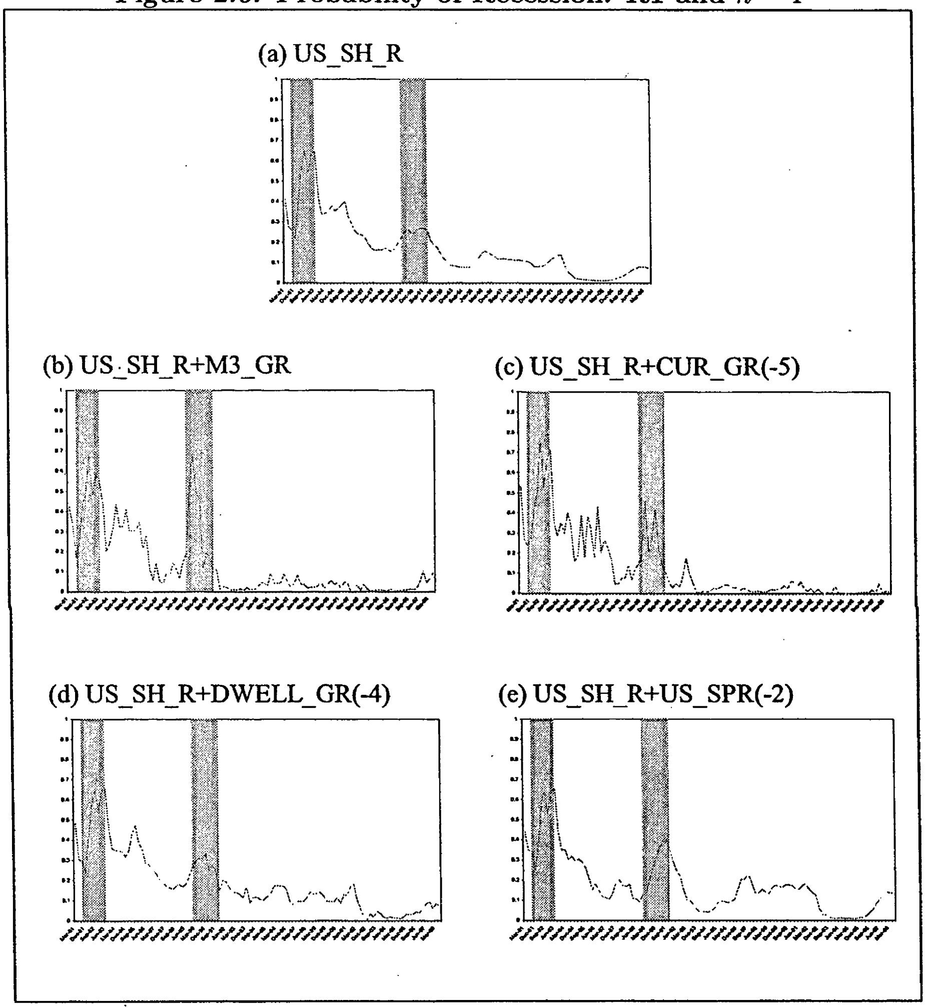


Table 2.10: **Re-Sampling *pseudo* R^2: LHS is R1, RHS are Constant and the Variable as Stated**

Dep Var = R1	h = Quarters Ahead							
	1	2	3	4	5	6	7	8
	Financial Variables							
SPREAD	0.074	0.139	0.151	0.097	0.060	...	...	...
SH_R	0.064	0.126	0.133	0.074	0.035	0.020	...	...
M3_GR	...	0.028	0.104	**0.175**	...	...	...	0.013
CUR_GR	0.018	...	...	...	0.002	0.043	0.098	**0.162**
INF	**0.149**	**0.232**	**0.165**	0.139	**0.109**	0.111	0.123	0.078
	Non-Financial Variables							
RTT_GR	0.038	0.030	...	...	...	0.000	...	...
DWELL_GR	0.086	0.075	0.028	...	...	...	...	...
	U.S. Financial Variables							
US_SPR	...	...	0.043	0.097	0.090	**0.120**	**0.125**	0.100
US_SH_R	0.012	0.044	0.064	...	...	0.035	0.039	0.017
US_LG_R	...	0.023	0.024	0.007	...	...	...	...

2.5 Re-Sampling Results

As discussed in section 2.3, the out-of-sample results may not provide leading indicators that work well in general. This section provides the results from our re-sampling exercise.

2.5.1 Individual Assessment

Table 2.10 and Table 2.11 show the values of *pseudo* R^2 from our re-sampling exercise. Table 2.10 presents the case that the LHS variable is R1 and the RHS variables are a constant and an individual variable, while Table 2.11 shows the case of LHS variable is R2. The missing numbers are *pseudo* R^2 with negative values. Full results are in Table A.9 and Table A.10.

When comparing the results in these two tables to the out-of-sample results in Table 2.4 and Table 2.5 of the previous section, we see that the best performer for short forecast horizons (1-quarter ahead to 3-quarters ahead) switches from short-term interest rate (SH_R) to inflation rate (INF) for R1, and from inflation rate to short-term interest rate for R2. However, in the overall picture

Table 2.11: **Re-Sampling *pseudo* R^2: LHS is R2, RHS are Constant and the Variable as Stated**

Dep Var = R2	h = Quarters Ahead							
	1	2	3	4	5	6	7	8
	Financial Variables							
SPREAD	0.038	0.062	0.060	0.052	0.088	0.067	...	...
SH_R	**0.113**	**0.133**	**0.120**	**0.094**	0.087	0.070	0.030	0.003
M3_GR	0.002	0.016	0.034	0.066	...	0.006	0.031	0.032
CUR_GR	0.004	0.013	0.022	0.014	0.030	0.034	**0.086**	**0.108**
INF	0.104	0.118	0.116	0.064	0.002	...	...	...
	Non-Financial Variables							
RTT_GR	0.046	0.063	...	...	...	0.015	...	0.004
DWELL_GR	0.008	0.014	0.030	0.015	...	...	0.007	0.040
	U.S. Financial Variables							
US_SPR	...	...	...	...	0.017	0.037	0.028	...
US_SH_R	0.020	0.026	0.041	0.065	0.078	**0.086**	0.058	0.032
US_LG_R	0.020	0.037	0.051	0.063	0.063	0.061	0.045	0.032

we see that our re-sampling exercise confirms our out-of-sample results: that these two are the best predictors for short forecast horizons.

This is not the same for longer forecast horizons. We see two major differences from the out-of-sample results here. First, we do not observe the outstanding performances of the U.S. short-term interest rate (US_SH_R) in our re-sampling exercise. Second, the U.S. interest rate spread (US_SPR) performs a lot better for long forecast horizons with R1.

Our re-sampling exercise still confirms that financial variables outperform non-financial variables in predicting Australian recessions. This finding stays regardless of the definition of recession or the forecast horizon used.

2.5.2 Combination of Two Predictors

Forecasting 1-quarter ahead to 3-quarters ahead, we choose inflation rate (INF) to be combined with each of the other variables. Table 2.12 reports our re-sampling *pseudo* R^2. The probit model here uses a constant, inflation rate, and another variable as stated in the table, as independent variables. The LHS variables are R1 and R2, respectively, in the left and right panels of the table.

Table 2.12: **Re-Sampling *pseudo* R^2: RHS are Constant, INF and the Variable as Stated**

Dep Var = R1	lag	h = Quarters Ahead 1	2	3	Dep Var = R2	lag	h = Quarters Ahead 1	2	3
INF		0.149	0.232	0.165	SH_R		0.113	0.133	0.120
		Financial Variables					*Financial Variables*		
SPREAD	0	0.154	0.282	**0.228**	SPREAD	0	0.110	0.137	0.141
	1	0.191	**0.294**	0.122		1	0.111	0.111	0.115
	2	**0.197**	0.237	0.092		2	0.100	0.097	0.149
	3	0.131	0.146	...		3	0.101	0.144	0.130
	4	0.117	...	0.053		4	0.171	0.154	0.075
	5	0.005	0.104	0.123		5	0.157	0.079	0.051
SH_R	0	0.122	0.272	0.191	SH_R	0	0.164	**0.178**	**0.172**
	1	0.180	0.274	...		1	**0.177**	0.153	0.142
	2	0.169	0.191	...		2	0.166	0.139	0.150
	3	0.123	0.200	0.076		3	0.155	0.157	0.145
	4	0.113	0.200	0.107		4	0.175	0.164	0.120
	5	0.134	0.197	0.119		5	0.171	0.129	0.105
		Non-Financial Variables					*Non-Financial Variables*		
RTT_GR	0	0.111	0.192	0.153	RTT_GR	0	0.069	0.111	0.101
	1	0.180	0.148	0.134		1	0.158	0.019	0.047
	2	0.111	0.190	0.134		2	0.073	0.094	0.114
	3	0.126	0.209	0.157		3	0.075	0.102	0.120
	4	0.109	0.207	0.145		4	0.088	0.123	0.107
	5	0.122	0.203	0.137		5	0.097	0.101	0.107
		U.S. Financial Variables					*U.S. Financial Variables*		
US_SH_R	0	0.101	0.198	0.144	US_SH_R	0	0.101	0.105	0.122
	1	0.106	0.199	...		1	0.104	0.114	0.138
	2	0.102	...	...		2	0.102	0.121	0.135
	3	...	...	0.111		3	0.121	0.132	0.140
	4	...	0.099	0.129		4	0.131	0.131	0.118
	5	0.069	...	0.110		5	0.123	0.097	0.099

Full results are in Table A.11. As in Table 2.8, the first row of Table 2.12 shows the performances of the best individual predictors. Each bold-faced number indicates the highest value of *pseudo* R^2 in each column.

Our re-sampling results in Table 2.12 are quite different from the out-of-sample results in Table 2.8. Firstly, we no longer observe the strong potential of the combination of inflation rate with the U.S. short-term interest rate (US_SH_R). Secondly, the results are more obvious. The combination of inflation rate and the interest rate spread (SPREAD) is the best for R1, while the combination of inflation rate and short-term interest rate (SH_R) is the best for R2.

It is not clear which variable is the best for longer forecast horizons (4-quarters ahead to 8-quarters ahead). From Table 2.10 and Table 2.11, U.S. short-term interest rate (US_SH_R), U.S. interest rate spread (US_SPR), growth rate of the currency (CUR_GR), and inflation rate (INF) seem to be close candidates.

Table 2.13: **Re-Sampling *pseudo* R^2: LHS is R1, RHS are Constant, US_SPR and the Variable as Stated**

Dep Var = R1	lag	h = Quarters Ahead				
		4	5	6	7	8
US_SPR		0.097	0.090	0.120	0.125	0.100
		Financial Variables				
M3_GR	0	0.283	0.083	0.090	0.127	0.099
	1	0.088	0.047	0.123	0.126	0.088
	2	0.062	0.095	0.121	0.115	0.076
	3	0.078	0.084	0.109	0.103	0.072
	4	0.047	0.067	0.100	0.100	0.078
	5	0.045	0.072	0.096	0.104	0.069
CUR_GR	0	0.056	0.096	0.148	0.202	0.231
	1	0.108	0.138	0.209	0.254	0.239
	2	0.140	0.186	0.239	0.253	**0.252**
	3	0.217	0.261	0.271	**0.300**	0.232
	4	0.294	0.237	**0.328**	0.284	0.171
	5	**0.297**	**0.316**	0.304	0.212	0.154
		Non-Financial Variables				
GOLD_GR	0	0.145	0.100	0.116	0.109	0.042
	1	0.140	0.113	0.112	0.070	0.071
	2	0.120	0.070	0.058	0.094	0.106
	3	0.104	0.035	0.083	0.115	0.134
	4	0.041	0.048	0.100	0.146	0.160
	5	0.051	0.051	0.129	0.167	0.144
		U.S. Financial Variables				
US_STK_GR	0	0.082	0.071	0.100	0.087	0.066
	1	0.080	0.072	0.077	0.088	0.105
	2	0.078	0.047	0.080	0.128	0.103
	3	0.039	0.040	0.122	0.125	0.082
	4	0.054	0.100	0.126	0.109	0.073
	5	0.109	0.103	0.113	0.102	0.059

We combine each of these variables to each of the other variables and compare their forecasting performances. Table 2.13 and Table 2.14 report the best results from this practice.

In Table 2.13, we use the U.S. interest rate spread (US_SPR) to be combined with each of the other variables. The LHS variable is R1. The table reports the values of *pseudo* R^2. The table shows that the combination of the U.S. interest rate spread and the growth rate of currency (CUR_GR) performs best in predicting R1 4-quarters ahead to 8-quarters ahead. This combination also obviously outperforms the best combination when we combine the U.S. short-term interest rate (US_SH_R) with other variables, which is not shown here. This is obviously different from the out-of-sample results shown in the left panel of Table 2.9.

Table 2.14 reports the *pseudo* R^2 from the case that the LHS variable is

Table 2.14: **Re-Sampling *pseudo* R^2: LHS is R2, RHS are Constant, US_SH_R and the Variable as Stated**

Dep Var = R2	lag	h = Quarters Ahead 4	5	6	7	8
US_SH_R		0.065	0.078	0.086	0.058	0.032
		Financial Variables				
M3_GR	0	0.116	0.088	0.112	0.121	0.092
	1	0.070	0.115	0.149	0.127	0.076
	2	0.088	0.135	0.148	0.078	0.032
	3	0.114	0.144	0.044	0.046	0.020
	4	0.122	0.101	0.031	0.050	...
	5	0.077	0.044	0.009	...	...
CUR_GR	0	0.045	0.067	0.077	0.111	0.119
	1	0.059	0.072	0.138	0.146	0.134
	2	0.064	0.129	**0.166**	0.128	**0.162**
	3	0.128	0.176	0.092	0.125	0.091
	4	0.162	0.179	0.123	0.104	0.046
	5	0.158	**0.193**	0.064	0.064	0.050
		Non-Financial Variables				
DWELL_GR	0	0.073	0.061	0.044	0.066	0.072
	1	0.057	0.053	0.091	0.118	0.148
	2	0.042	0.075	0.135	**0.172**	0.075
	3	0.064	0.128	0.155	0.065	...
	4	0.104	0.176	0.025	...	...
	5	**0.170**	0.039	...	0.014	0.049
		U.S. Financial Variables				
US_STK_GR	0	0.055	0.055	0.076	0.055	...
	1	0.040	0.071	0.087	0.019	0.021
	2	0.055	0.072	0.055	0.038	0.052
	3	0.061	0.047	...	0.067	0.054
	4	0.028	0.048	0.066	0.079	0.086
	5	0.038	0.089	0.057	0.100	0.055

R2, and the RHS variables are: a constant, the U.S. short-term interest rate (US_SH_R), and each of the other variables. The table shows that the combination of the U.S. short-term interest rate and the growth rate of currency (CUR_GR) performs best in predicting R2 4-quarters ahead to 8-quarters ahead. This is rather similar to the out-of-sample results shown in the right panel of Table 2.9.

2.6 Conclusion

When developing forecasting models to predict recessions in Australia, one must realize that there have been few recessions since 1970. While the most convincing evidence in support of a forecasting model is how it performs in an "out-of-sample period" (A period other than the sample period on which the model is estimated), the sparseness of recessions may make such evaluation inaccurate.

In this study, we design a re-sampling exercise that intends to avoid this problem. We show that the results of this exercise may turn out to be very different from the usual out-of-sample evaluation that focuses only on one specific period. Here, the difference is particularly remarkable with our R1 definition of recession, for long forecast horizons from 4-quarters ahead to 8-quarters ahead.

When we use R1, Australian recessions seem to follow ones of the U.S. with some lags. Moreover, there are only 2 recessionary sessions post 1980, and one of them, in 1990 – 1991, was led by the Gulf War. The U.S. recession in 1990 – 1991 is shown by some studies, including Estrella and Mishkin (1998), as one that the U.S. interest rate spread cannot predict well. This can help explain why we do not see this variable as a good predictor in our out-of-sample assessment in Section 2.4.

In contrast, for our re-sampling assessment, the U.S. interest rate spread turns out to be the best predictor for the 6-quarters to 7-quarters ahead forecast horizons (Table 2.10). Moreover, its combination with the growth rate of currency obviously outperforms the combination of the U.S. short-term interest rate and the growth rate of currency, which performs best in the out-of-sample assessment.

We believe that the results of our re-sampling exercise are more robust than ones of the out-of-sample assessment that focus on a specific evaluation period. We pick the combination of the U.S. interest rate spread and the growth rate of currency to be the best predictor for R1 recessions 4-quarters ahead to 8-quarters ahead (Table 2.13). The combination of the U.S. short-term interest rate and the growth rate of currency is the best for predicting R2 recessions 4-quarters ahead to 8-quarters ahead (Table 2.14).

Our re-sampling assessment coincides with the out-of-sample assessment in the best individual predictors for short forecast horizons (1-quarter ahead to 3-quarters ahead). They are the short-term interest rate and the inflation rate. However, we see the combination of inflation rate and interest rate spread to perform best for the case of combining two predictors under the R1 recessions, and the combination of inflation rate and short-term interest rate for R2 recessions (Table 2.12).

Our results still confirm that financial variables have strong predictive power in predicting real economic activities. They are obviously superior to non-financial leading indicators. We also see strong forecasting performances of the U.S. financial variables in predicting Australian recessions. This is similar to the finding of Smith (2005).

Appendix A

Full Results

Table A.1: In-Sample Statistics: LHS is R1, RHS are Constant and the Variable as Stated

Dependent Variable is R1		h = Quarters Ahead							
		1	2	3	4	5	6	7	8
SPREAD	pseudo R^2	0.079	0.128	0.146	0.127	0.101	0.057	0.035	0.020
	t-stat	-3.99*	-3.61*	-2.97*	-2.48*	-2.16**	-1.56	-1.28	-1.31
SH_R	pseudo R^2	0.183	0.210	0.210	0.177	0.133	0.082	0.050	0.029
	t-stat	6.14*	4.92*	4.19*	3.45*	2.80*	1.98**	1.52	1.20
LG_R	pseudo R^2	0.158	0.154	0.142	0.118	0.085	0.057	0.035	0.020
	t-stat	6.15*	4.41*	3.55*	2.85*	2.23**	1.72	1.31	0.96
D_SH_R	pseudo R^2	0.115	0.100	0.065	0.030	0.015	0.002	0.000	0.003
	t-stat	4.39*	3.15*	2.15**	1.20	0.79	0.28	-0.07	-0.34
D_LG_R	pseudo R^2	0.072	0.055	0.041	0.021	0.009	0.002	0.000	0.003
	t-stat	3.41*	2.16**	1.55	0.98	0.59	0.30	-0.05	-0.33
M3_GR	pseudo R^2	0.011	0.031	0.101	0.183	0.160	0.153	0.127	0.089
	t-stat	1.14	1.68	3.23*	4.21*	2.19**	2.35*	2.42*	2.38*
CUR_GR	pseudo R^2	0.098	0.091	0.080	0.072	0.113	0.135	0.168	0.158
	t-stat	3.61*	2.85*	2.34*	2.00**	2.55*	3.07*	4.37*	5.43*
R_M3_GR	pseudo R^2	0.029	0.024	0.000	0.018	0.024	0.025	0.012	0.008
	t-stat	-1.19	-1.05	0.12	1.39	1.33	1.34	0.84	0.72
R_CUR_GR	pseudo R^2	0.021	0.060	0.037	0.029	0.004	0.001	0.000	0.003
	t-stat	-1.50	-2.87*	-2.25*	-1.92	-0.76	-0.32	-0.06	0.65
STOCK_GR	pseudo R^2	0.063	0.046	0.071	0.044	0.006	0.000	0.006	0.001
	t-stat	-2.37*	-2.12**	-2.26*	-2.11**	-0.99	0.00	0.76	0.23
EXCH_GR	pseudo R^2	0.026	0.024	0.029	0.013	0.000	0.008	0.015	0.008
	t-stat	-2.18**	-2.08**	-2.27**	-1.56	0.11	1.02	1.28	0.94
INF	pseudo R^2	0.181	0.263	0.197	0.166	0.142	0.135	0.143	0.099
	t-stat	4.24*	5.71*	4.61*	4.49*	4.54*	4.77*	4.31*	2.98*
RTT_GR	pseudo R^2	0.049	0.054	0.000	0.006	0.016	0.031	0.004	0.004
	t-stat	-1.85	-2.17**	-0.21	1.08	1.70	2.47*	1.12	1.04
DWELL_GR	pseudo R^2	0.111	0.103	0.041	0.008	0.001	0.009	0.008	0.015
	t-stat	-3.55*	-2.95*	-2.04**	-0.91	0.36	1.00	0.84	0.99
NONRES_GR	pseudo R^2	0.020	0.013	0.000	0.000	0.001	0.000	0.001	0.003
	t-stat	-1.71	-1.57	-0.38	-0.01	0.46	-0.14	-0.51	0.94
SV_EX_GR	pseudo R^2	0.000	0.000	0.004	0.004	0.016	0.026	0.001	0.000
	t-stat	0.18	0.08	0.78	-0.83	-2.06**	-2.26**	-0.53	0.03
TOT_GR	pseudo R^2	0.042	0.030	0.013	0.002	0.019	0.007	0.023	0.015
	t-stat	-2.19**	-2.16**	-1.70	-0.76	1.69	0.80	1.17	0.77
GOLD_GR	pseudo R^2	0.005	0.001	0.019	0.040	0.014	0.006	0.002	0.002
	t-stat	-0.79	-0.3	-1.05	-1.28	-0.72	-0.48	-0.46	0.34
US_SPR	pseudo R^2	0.000	0.010	0.055	0.099	0.094	0.097	0.094	0.081
	t-stat	0.28	-1	-1.95	-2.19**	-2.09**	-2.13**	-2.67*	-3.28*
US_SH_R	pseudo R^2	0.068	0.106	0.150	0.179	0.157	0.134	0.116	0.095
	t-stat	3.55*	3.32*	3.50*	3.51*	3.43*	3.38*	3.24*	2.71*
US_LG_R	pseudo R^2	0.093	0.102	0.103	0.102	0.087	0.069	0.057	0.047
	t-stat	4.71*	3.73*	3.29*	2.95*	2.55*	2.18**	1.87	1.62
US_STK_GR	pseudo R^2	0.003	0.002	0.001	0.000	0.000	0.001	0.002	0.000
	t-stat	-0.71	-0.54	0.45	-0.21	-0.12	0.55	-0.45	-0.08

Note: * significant at 1% ** significant at 5%

Table A.2: In-Sample Statistics: LHS is R2, RHS are Constant and the Variable as Stated

Dependent Variable is R2		h = Quarters Ahead							
		1	2	3	4	5	6	7	8
SPREAD	pseudo R^2	0.029	0.044	0.050	0.051	0.089	0.076	0.024	0.003
	t-stat	-2.21**	-2.25*	-2.21**	-2.52*	-2.54*	-2.35*	-1.41	-0.5
SH_R	pseudo R^2	0.087	0.099	0.090	0.077	0.083	0.072	0.037	0.021
	t-stat	4.08*	3.50*	3.08*	2.75*	2.50*	2.36*	1.8	1.42
LG_R	pseudo R^2	0.085	0.086	0.070	0.054	0.040	0.034	0.026	0.025
	t-stat	3.80*	3.04*	2.55*	2.17**	1.88	1.72	1.49	1.54
D_SH_R	pseudo R^2	0.036	0.042	0.033	0.013	0.008	0.002	0.001	0.003
	t-stat	2.03**	1.93	1.68	1.03	0.68	0.36	-0.28	-0.46
D_LG_R	pseudo R^2	0.037	0.029	0.015	0.001	0.003	0.008	0.011	0.006
	t-stat	2.18**	1.74	1.15	0.28	-0.37	-0.55	-0.69	-0.69
M3_GR	pseudo R^2	0.023	0.023	0.034	0.049	0.022	0.044	0.052	0.055
	t-stat	1.92	1.94	1.51	2.40*	1.47	1.76	2.12**	2.31*
CUR_GR	pseudo R^2	0.035	0.076	0.069	0.055	0.050	0.034	0.069	0.084
	t-stat	2.52*	2.85*	2.40*	2.20**	2.20**	2.46*	3.58*	3.21*
R_M3_GR	pseudo R^2	0.005	0.007	0.001	0.004	0.003	0.013	0.029	0.028
	t-stat	-1.06	-1.18	-0.19	0.49	0.59	0.89	1.77	1.98**
R_CUR_GR	pseudo R^2	0.027	0.007	0.009	0.001	0.007	0.001	0.023	0.026
	t-stat	-2.11**	-0.91	-0.91	-0.21	0.74	0.24	1.88	2.48*
STOCK_GR	pseudo R^2	0.001	0.008	0.023	0.059	0.018	0.004	0.001	0.005
	t-stat	-0.34	-1.03	-1.68	-2.25*	-1.73	-1.10	0.45	0.76
EXCH_GR	pseudo R^2	0.000	0.005	0.015	0.026	0.008	0.005	0.001	0.001
	t-stat	0.06	-0.99	-1.45	-1.43	-1.01	-1.24	-0.55	0.64
INF	pseudo R^2	0.105	0.112	0.109	0.057	0.017	0.020	0.010	0.013
	t-stat	3.92*	3.97*	4.47*	3.19*	1.91	1.55	1.21	1.46
RTT_GR	pseudo R^2	0.037	0.048	0.006	0.004	0.006	0.041	0.022	0.013
	t-stat	-2.77*	-2.93*	-1.22	-1.09	1.41	3.11*	2.54*	1.89
DWELL_GR	pseudo R^2	0.011	0.014	0.030	0.018	0.009	0.006	0.014	0.047
	t-stat	-1.43	-1.79	-1.87	-1.55	-1.11	-0.94	1.6	2.64*
NONRES_GR	pseudo R^2	0.002	0.017	0.002	0.004	0.004	0.010	0.000	0.002
	t-stat	-0.49	-1.80	-0.71	1.02	1.29	1.39	0.30	0.49
SV_EX_GR	pseudo R^2	0.000	0.002	0.003	0.009	0.000	0.010	0.000	0.002
	t-stat	0.10	-0.63	-0.50	1.02	0.27	-1.25	0.11	-0.65
TOT_GR	pseudo R^2	0.039	0.024	0.003	0.002	0.000	0.011	0.026	0.000
	t-stat	-2.43*	-1.87	-0.49	-0.41	-0.10	1.91	1.89	-0.10
GOLD_GR	pseudo R^2	0.017	0.001	0.021	0.001	0.021	0.002	0.010	0.034
	t-stat	1.42	0.40	1.48	0.22	-1.73	-0.38	-0.77	-1.78
US_SPR	pseudo R^2	0.001	0.001	0.003	0.012	0.030	0.041	0.037	0.014
	t-stat	-0.44	-0.31	-0.54	-1.05	-1.67	-1.71	-1.63	-1.20
US_SH_R	pseudo R^2	0.050	0.057	0.061	0.081	0.109	0.116	0.100	0.071
	t-stat	3.17*	2.59*	2.47*	3.06*	3.67*	3.61*	3.32*	2.87*
US_LG_R	pseudo R^2	0.054	0.064	0.062	0.071	0.082	0.082	0.071	0.062
	t-stat	3.36*	2.90*	2.73*	3.1*	3.54*	3.35*	2.87*	2.60*
US_STK_GR	pseudo R^2	0.036	0.018	0.033	0.006	0.005	0.012	0.011	0.001
	t-stat	-2.13**	-1.73	-2.39*	-1.05	-0.72	1.72	2.16**	0.35

Note: * significant at 1% ** significant at 5%

Table A.3: Out-of-Sample *pseudo* R^2: LHS is R1, RHS are Constant and the Variable as Stated

Dep Var = R1	h = Quarters Ahead							
	1	2	3	4	5	6	7	8
SPREAD	0.079	0.150	0.180	0.159	0.131	0.012	-0.058	-0.001
SH_R	**0.148**	**0.198**	**0.190**	0.164	0.147	0.085	0.025	0.008
LG_R	0.123	0.118	0.092	0.079	0.069	0.047	0.024	0.006
D_SH_R	-0.220	-0.072	-0.043	-0.019	-0.002	-0.030	-0.062	-0.054
D_LG_R	-0.210	-0.180	-0.129	-0.069	-0.030	-0.019	-0.016	-0.018
M3_GR	0.005	0.033	0.092	0.177	0.164	0.165	0.144	0.108
CUR_GR	0.028	0.026	0.002	0.015	0.030	0.071	0.102	**0.177**
R_M3_GR	0.034	0.019	-0.005	0.001	0.014	0.019	0.009	0.005
R_CUR_GR	0.012	0.039	0.032	0.031	-0.002	-0.007	-0.010	-0.006
STOCK_GR	0.015	0.027	0.024	0.011	-0.002	-0.017	-0.001	-0.011
EXCH_GR	-0.059	-0.057	-0.060	-0.031	-0.009	-0.022	0.001	0.000
INF	0.147	0.183	0.151	0.169	0.148	0.145	**0.156**	0.121
RTT_GR	0.058	0.062	-0.008	-0.019	-0.006	0.001	-0.001	0.000
DWELL_GR	0.047	0.034	0.029	0.003	-0.032	-0.037	-0.009	0.013
NONRES_GR	0.003	-0.001	-0.008	-0.012	-0.007	-0.011	-0.014	-0.007
SV_EX_GR	-0.008	-0.006	-0.015	-0.004	-0.006	-0.001	-0.005	-0.005
TOT_GR	0.046	0.030	0.010	0.000	-0.006	0.000	0.012	0.011
GOLD_GR	-0.018	-0.004	0.013	0.027	0.008	0.003	-0.012	-0.007
US_SPR	-0.035	-0.007	0.067	0.130	0.120	0.123	0.115	0.090
US_SH_R	0.049	0.113	0.190	**0.234**	**0.198**	**0.167**	0.131	0.090
US_LG_R	0.097	0.109	0.115	0.113	0.087	0.064	0.031	0.015
US_STK_GR	-0.035	-0.065	-0.034	-0.017	-0.015	-0.010	-0.039	-0.029

Table A.4: Out-of-Sample *pseudo* R^2: LHS is R2, RHS are Constant and the Variable as Stated

Dep Var = R2	h = Quarters Ahead							
	1	2	3	4	5	6	7	8
SPREAD	0.026	0.049	0.057	0.062	0.091	0.081	0.020	-0.018
SH_R	0.104	0.119	0.106	0.088	0.092	0.079	0.025	0.000
LG_R	0.093	0.097	0.076	0.052	0.028	0.018	0.006	0.010
D_SH_R	0.032	0.045	0.032	0.002	-0.002	-0.011	-0.028	-0.034
D_LG_R	0.028	0.009	-0.008	-0.013	-0.034	-0.031	-0.019	-0.012
M3_GR	0.018	0.026	0.039	0.062	0.019	0.052	0.067	0.070
CUR_GR	0.040	0.050	0.063	0.056	0.057	0.029	0.084	**0.108**
R_M3_GR	0.003	0.005	-0.001	-0.003	-0.001	0.010	0.032	0.028
R_CUR_GR	0.019	0.001	0.001	-0.017	-0.019	-0.013	0.021	0.025
STOCK_GR	-0.004	0.003	0.011	-0.034	-0.004	-0.007	-0.021	-0.028
EXCH_GR	-0.012	-0.017	-0.010	0.016	-0.007	-0.003	-0.011	-0.008
INF	**0.130**	**0.123**	**0.126**	0.075	0.009	0.021	-0.001	0.005
RTT_GR	0.036	0.049	-0.049	-0.021	0.000	0.000	0.006	0.013
DWELL_GR	0.007	0.013	0.026	0.018	0.005	0.000	0.006	0.027
NONRES_GR	-0.010	-0.022	-0.009	-0.004	-0.004	-0.016	-0.008	-0.007
SV_EX_GR	-0.004	-0.010	-0.013	-0.015	-0.016	0.009	-0.006	-0.008
TOT_GR	0.051	0.029	-0.003	0.001	-0.002	-0.010	0.013	-0.008
GOLD_GR	0.016	-0.003	-0.004	-0.011	0.008	-0.010	0.008	0.021
US_SPR	-0.015	-0.009	-0.005	0.002	0.020	0.041	0.036	0.004
US_SH_R	0.055	0.071	0.075	**0.095**	**0.118**	**0.133**	**0.110**	0.073
US_LG_R	0.059	0.067	0.047	0.059	0.072	0.074	0.058	0.049
US_STK_GR	0.032	-0.012	0.020	-0.001	-0.007	-0.016	-0.021	-0.019

Table A.5: In-Sample Statistics: RHS are Constant, INF, and the Variable as Stated

Dependent Variable is R1	lag		h = Quarters Ahead 1	2	3
SPREAD		pseudo R^2	0.216	0.330	0.281
	0	t-stat	-2.66	-2.82	-2.42
		t-stat INF	3.86	5.12	4.15
		pseudo R^2	0.245	0.336	0.257
	1	t-stat	-2.98	-2.58	-1.92
		t-stat INF	3.60	5.04	3.85
		pseudo R^2	0.260	0.322	0.239
	2	t-stat	-3.11	-2.12	-1.56
		t-stat INF	3.60	5.06	3.71
		pseudo R^2	0.241	0.295	0.208
	3	t-stat	-2.57	-1.40	-0.73
		t-stat INF	3.59	4.66	3.58
		pseudo R^2	0.246	0.282	0.207
	4	t-stat	-2.60	-0.96	-0.75
		t-stat INF	3.97	5.02	3.99
		pseudo R^2	0.217	0.277	0.204
	5	t-stat	-1.57	-0.87	-0.80
		t-stat INF	4.03	5.23	4.21
SH_R		pseudo R^2	0.277	0.374	0.314
	0	t-stat	4.29	3.45	3.05
		t-stat INF	3.12	4.16	3.32
		pseudo R^2	0.295	0.365	0.281
	1	t-stat	4.37	3.12	2.46
		t-stat INF	3.08	4.43	3.32
		pseudo R^2	0.300	0.344	0.252
	2	t-stat	4.31	2.55	1.92
		t-stat INF	3.19	4.44	3.33
		pseudo R^2	0.283	0.319	0.225
	3	t-stat	3.77	2.12	1.32
		t-stat INF	3.53	4.68	3.52
		pseudo R^2	0.267	0.300	0.215
	4	t-stat	3.52	1.68	1.14
		t-stat INF	3.90	4.94	3.84
		pseudo R^2	0.238	0.288	0.209
	5	t-stat	2.78	1.46	1.01
		t-stat INF	4.04	5.32	4.15
LG_R		pseudo R^2	0.268	0.338	0.269
	0	t-stat	4.27	2.91	2.49
		t-stat INF	3.28	4.48	3.47
		pseudo R^2	0.268	0.333	0.253
	1	t-stat	4.25	2.64	2.13
		t-stat INF	3.39	4.56	3.59
		pseudo R^2	0.266	0.320	0.233
	2	t-stat	4.04	2.36	1.74
		t-stat INF	3.57	4.69	3.73
		pseudo R^2	0.259	0.308	0.222
	3	t-stat	3.75	2.14	1.46
		t-stat INF	3.88	5.00	3.98
		pseudo R^2	0.238	0.292	0.211
	4	t-stat	3.27	1.81	1.14
		t-stat INF	4.02	5.30	4.15
		pseudo R^2	0.222	0.283	0.206
	5	t-stat	2.83	1.50	0.92
		t-stat INF	4.11	5.58	4.34
D_SH_R		pseudo R^2	0.253	0.320	0.230
	0	t-stat	3.45	2.30	1.50
		t-stat INF	3.79	5.31	4.25
		pseudo R^2	0.239	0.291	0.206
	1	t-stat	2.96	1.54	0.71
		t-stat INF	3.90	5.41	4.23
		pseudo R^2	0.226	0.280	0.204
	2	t-stat	2.62	1.13	0.66
		t-stat INF	4.07	5.73	4.44
		pseudo R^2	0.202	0.272	0.197
	3	t-stat	1.61	0.80	0.19
		t-stat INF	4.19	5.89	4.54
		pseudo R^2	0.200	0.267	0.196
	4	t-stat	1.60	0.58	0.05
		t-stat INF	4.35	5.87	4.66
		pseudo R^2	0.193	0.266	0.197
	5	t-stat	1.26	0.46	0.02
		t-stat INF	4.47	6.13	4.85
D_LG_R		pseudo R^2	0.226	0.293	0.219
	0	t-stat	2.60	1.48	1.07
		t-stat INF	3.89	5.60	4.33
		pseudo R^2	0.219	0.288	0.208
	1	t-stat	2.21	1.29	0.77
		t-stat INF	4.10	5.78	4.45
		pseudo R^2	0.223	0.286	0.206
	2	t-stat	2.33	1.26	0.72
		t-stat INF	4.30	5.99	4.62
		pseudo R^2	0.214	0.282	0.203
	3	t-stat	2.10	1.18	0.65
		t-stat INF	4.50	6.29	4.81
		pseudo R^2	0.204	0.276	0.198
	4	t-stat	1.73	1.02	0.39
		t-stat INF	4.46	6.43	4.91
		pseudo R^2	0.202	0.274	0.198
	5	t-stat	1.73	0.96	0.35
		t-stat INF	4.49	6.58	4.98

Dependent Variable is R2	lag		h = Quarters Ahead 1	2	3
SPREAD		pseudo R^2	0.114	0.130	0.132
	0	t-stat	-1.20	-1.45	-1.45
		t-stat INF	3.58	3.49	3.90
		pseudo R^2	0.117	0.127	0.124
	1	t-stat	-1.17	-1.13	-1.11
		t-stat INF	3.30	3.12	3.52
		pseudo R^2	0.120	0.125	0.149
	2	t-stat	-1.19	-1.08	-1.67
		t-stat INF	3.22	3.05	3.39
		pseudo R^2	0.122	0.154	0.141
	3	t-stat	-1.33	-1.83	-1.52
		t-stat INF	3.23	2.93	3.27
		pseudo R^2	0.164	0.161	0.118
	4	t-stat	-2.79	-2.04	-0.75
		t-stat INF	3.76	3.46	3.87
		pseudo R^2	0.161	0.123	0.110
	5	t-stat	-2.62	-0.99	0.02
		t-stat INF	3.90	3.61	4.29
SH_R		pseudo R^2	0.144	0.158	0.149
	0	t-stat	2.38	2.23	1.91
		t-stat INF	2.94	2.75	3.12
		pseudo R^2	0.150	0.150	0.136
	1	t-stat	2.50	1.90	1.49
		t-stat INF	2.74	2.59	3.08
		pseudo R^2	0.145	0.141	0.141
	2	t-stat	2.34	1.63	1.52
		t-stat INF	2.79	2.71	3.15
		pseudo R^2	0.141	0.151	0.138
	3	t-stat	2.14	1.78	1.51
		t-stat INF	3.08	2.84	3.33
		pseudo R^2	0.155	0.153	0.123
	4	t-stat	2.55	1.96	1.12
		t-stat INF	3.48	3.35	3.78
		pseudo R^2	0.154	0.132	0.118
	5	t-stat	2.68	1.43	0.91
		t-stat INF	3.69	3.57	4.04
LG_R		pseudo R^2	0.145	0.153	0.138
	0	t-stat	2.37	1.92	1.53
		t-stat INF	3.03	2.94	3.39
		pseudo R^2	0.149	0.143	0.128
	1	t-stat	2.46	1.70	1.26
		t-stat INF	3.09	3.08	3.52
		pseudo R^2	0.141	0.135	0.122
	2	t-stat	2.21	1.46	1.04
		t-stat INF	3.23	3.22	3.69
		pseudo R^2	0.134	0.130	0.122
	3	t-stat	2.04	1.32	1.08
		t-stat INF	3.45	3.44	3.91
		pseudo R^2	0.126	0.128	0.119
	4	t-stat	1.84	1.32	0.93
		t-stat INF	3.56	3.56	4.03
		pseudo R^2	0.126	0.126	0.124
	5	t-stat	1.84	1.18	1.14
		t-stat INF	3.68	3.70	4.17
D_SH_R		pseudo R^2	0.125	0.136	0.125
	0	t-stat	1.54	1.49	1.20
		t-stat INF	3.55	3.50	4.01
		pseudo R^2	0.127	0.128	0.111
	1	t-stat	1.69	1.27	0.45
		t-stat INF	3.44	3.40	4.02
		pseudo R^2	0.126	0.118	0.110
	2	t-stat	1.75	0.79	0.36
		t-stat INF	3.56	3.69	4.29
		pseudo R^2	0.113	0.115	0.108
	3	t-stat	1.06	0.54	0.11
		t-stat INF	3.78	3.85	4.32
		pseudo R^2	0.113	0.114	0.109
	4	t-stat	1.05	0.52	-0.27
		t-stat INF	3.99	4.06	4.42
		pseudo R^2	0.111	0.111	0.111
	5	t-stat	0.91	-0.01	-0.24
		t-stat INF	4.09	4.08	4.44
D_LG_R		pseudo R^2	0.128	0.127	0.114
	0	t-stat	1.66	1.22	0.82
		t-stat INF	3.57	3.63	4.13
		pseudo R^2	0.124	0.120	0.108
	1	t-stat	1.64	0.89	-0.05
		t-stat INF	3.68	3.77	4.29
		pseudo R^2	0.119	0.112	0.113
	2	t-stat	1.48	0.32	-0.48
		t-stat INF	3.89	3.98	4.47
		pseudo R^2	0.107	0.112	0.115
	3	t-stat	0.62	-0.29	-0.55
		t-stat INF	3.95	4.03	4.42
		pseudo R^2	0.104	0.113	0.114
	4	t-stat	-0.22	-0.41	-0.57
		t-stat INF	3.92	3.99	4.34
		pseudo R^2	0.105	0.113	0.110
	5	t-stat	-0.29	-0.35	-0.22
		t-stat INF	3.88	3.91	4.36

Table A.5 (Continued)

Dependent Variable is R1	lag		h = Quarters Ahead 1	2	3
M3_GR		pseudo R^2	0.182	0.263	0.233
	0	t-stat	-0.36	0.14	2.16
		t-stat INF	4.35	5.81	3.77
		pseudo R^2	0.180	0.280	0.275
	1	t-stat	0.10	1.54	2.95
		t-stat INF	3.81	5.05	3.32
		pseudo R^2	0.220	0.348	0.280
	2	t-stat	2.55	3.07	1.84
		t-stat INF	3.38	4.87	4.06
		pseudo R^2	0.259	0.321	0.254
	3	t-stat	2.88	1.71	1.66
		t-stat INF	2.70	4.67	3.54
		pseudo R^2	0.269	0.346	0.258
	4	t-stat	2.37	2.07	1.84
		t-stat INF	3.69	5.48	4.23
		pseudo R^2	0.258	0.314	0.229
	5	t-stat	2.05	1.68	1.53
		t-stat INF	3.53	4.91	3.94
CUR_GR		pseudo R^2	0.193	0.263	0.200
	0	t-stat	1.21	0.22	0.49
		t-stat INF	2.99	4.70	3.74
		pseudo R^2	0.185	0.265	0.196
	1	t-stat	0.70	-0.51	-0.06
		t-stat INF	2.87	4.87	4.35
		pseudo R^2	0.184	0.264	0.213
	2	t-stat	0.68	-0.45	1.14
		t-stat INF	3.13	5.92	3.96
		pseudo R^2	0.180	0.264	0.218
	3	t-stat	0.11	0.43	1.40
		t-stat INF	3.81	5.73	3.58
		pseudo R^2	0.189	0.269	0.232
	4	t-stat	1.28	0.94	2.13
		t-stat INF	3.13	5.27	3.04
		pseudo R^2	0.200	0.286	0.232
	5	t-stat	1.76	1.72	2.09
		t-stat INF	3.00	4.61	2.84
R_M3_GR		pseudo R^2	0.182	0.263	0.233
	0	t-stat	-0.36	0.14	2.16
		t-stat INF	3.76	4.99	4.29
		pseudo R^2	0.218	0.262	0.213
	1	t-stat	-2.52	-0.02	1.32
		t-stat INF	4.72	5.69	4.50
		pseudo R^2	0.184	0.299	0.234
	2	t-stat	0.66	1.94	1.72
		t-stat INF	4.12	5.52	5.36
		pseudo R^2	0.201	0.290	0.215
	3	t-stat	1.65	1.57	1.17
		t-stat INF	3.94	5.78	4.67
		pseudo R^2	0.223	0.319	0.216
	4	t-stat	2.08	2.19	1.09
		t-stat INF	4.61	6.47	4.77
		pseudo R^2	0.200	0.273	0.202
	5	t-stat	1.27	0.82	0.64
		t-stat INF	4.45	5.91	4.66
R_CUR_GR		pseudo R^2	0.193	0.263	0.200
	0	t-stat	1.21	0.22	0.49
		t-stat INF	4.83	4.72	3.79
		pseudo R^2	0.248	0.318	0.238
	1	t-stat	-2.82	-2.50	-2.03
		t-stat INF	4.08	5.51	4.90
		pseudo R^2	0.206	0.288	0.198
	2	t-stat	-1.70	-1.75	-0.37
		t-stat INF	3.81	6.03	4.51
		pseudo R^2	0.211	0.267	0.197
	3	t-stat	-2.04	-0.69	-0.19
		t-stat INF	4.51	5.74	4.54
		pseudo R^2	0.185	0.263	0.196
	4	t-stat	-0.86	-0.31	-0.07
		t-stat INF	4.31	5.75	4.58
		pseudo R^2	0.192	0.271	0.197
	5	t-stat	-1.26	-1.08	-0.16
		t-stat INF	4.40	5.98	4.66
STOCK_GR		pseudo R^2	0.235	0.307	0.258
	0	t-stat	-2.19	-2.09	-2.12
		t-stat INF	4.02	5.65	4.61
		pseudo R^2	0.200	0.299	0.214
	1	t-stat	-1.65	-1.95	-1.73
		t-stat INF	4.02	5.42	4.42
		pseudo R^2	0.242	0.293	0.198
	2	t-stat	-2.19	-1.93	-0.53
		t-stat INF	4.45	5.58	4.52
		pseudo R^2	0.205	0.262	0.202
	3	t-stat	-1.82	0.00	0.80
		t-stat INF	4.11	5.53	4.87
		pseudo R^2	0.181	0.267	0.212
	4	t-stat	-0.48	0.86	1.56
		t-stat INF	4.20	5.62	4.53
		pseudo R^2	0.181	0.277	0.197
	5	t-stat	0.35	1.41	0.22
		t-stat INF	4.23	5.65	4.56

Dependent Variable is R2	lag		h = Quarters Ahead 1	2	3
M3_GR		pseudo R^2	0.107	0.113	0.117
	0	t-stat	0.51	0.59	0.79
		t-stat INF	3.76	3.72	3.71
		pseudo R^2	0.105	0.114	0.115
	1	t-stat	0.10	0.49	0.96
		t-stat INF	3.58	3.19	3.58
		pseudo R^2	0.116	0.127	0.110
	2	t-stat	1.12	1.51	0.46
		t-stat INF	3.29	3.44	3.90
		pseudo R^2	0.117	0.111	0.116
	3	t-stat	1.21	0.20	0.82
		t-stat INF	3.06	3.48	3.27
		pseudo R^2	0.108	0.125	0.126
	4	t-stat	0.73	1.21	1.36
		t-stat INF	3.47	3.35	4.00
		pseudo R^2	0.117	0.127	0.127
	5	t-stat	1.29	1.24	1.40
		t-stat INF	3.22	3.17	3.76
CUR_GR		pseudo R^2	0.106	0.125	0.120
	0	t-stat	0.18	1.25	0.92
		t-stat INF	3.04	2.68	2.65
		pseudo R^2	0.115	0.118	0.111
	1	t-stat	1.28	0.80	0.49
		t-stat INF	2.60	2.57	2.77
		pseudo R^2	0.114	0.114	0.111
	2	t-stat	1.23	0.62	0.51
		t-stat INF	2.76	3.09	3.03
		pseudo R^2	0.107	0.112	0.108
	3	t-stat	0.64	0.34	-0.16
		t-stat INF	2.91	3.26	3.46
		pseudo R^2	0.105	0.111	0.115
	4	t-stat	0.43	-0.36	1.08
		t-stat INF	2.94	3.55	2.87
		pseudo R^2	0.104	0.118	0.124
	5	t-stat	-0.11	1.09	1.37
		t-stat INF	3.09	2.61	2.51
R_M3_GR		pseudo R^2	0.107	0.113	0.117
	0	t-stat	0.51	0.59	0.79
		t-stat INF	3.72	4.07	4.17
		pseudo R^2	0.114	0.114	0.110
	1	t-stat	-1.23	-0.62	0.39
		t-stat INF	4.26	3.97	4.32
		pseudo R^2	0.105	0.120	0.116
	2	t-stat	0.10	0.96	1.02
		t-stat INF	3.85	3.78	4.34
		pseudo R^2	0.110	0.115	0.122
	3	t-stat	0.74	0.75	1.09
		t-stat INF	3.83	3.79	4.06
		pseudo R^2	0.114	0.136	0.150
	4	t-stat	1.23	1.61	2.09
		t-stat INF	4.30	4.16	4.78
		pseudo R^2	0.114	0.132	0.129
	5	t-stat	1.09	1.50	1.50
		t-stat INF	3.80	4.00	4.88
R_CUR_GR		pseudo R^2	0.106	0.125	0.120
	0	t-stat	0.18	1.25	0.92
		t-stat INF	3.66	3.66	3.71
		pseudo R^2	0.113	0.120	0.109
	1	t-stat	-1.01	-0.94	-0.16
		t-stat INF	3.99	3.99	4.49
		pseudo R^2	0.110	0.111	0.121
	2	t-stat	-0.82	-0.01	0.98
		t-stat INF	3.81	3.87	4.71
		pseudo R^2	0.105	0.120	0.110
	3	t-stat	-0.17	0.85	0.38
		t-stat INF	3.89	3.81	4.44
		pseudo R^2	0.111	0.112	0.139
	4	t-stat	0.93	0.38	1.99
		t-stat INF	3.89	3.86	4.13
		pseudo R^2	0.105	0.123	0.125
	5	t-stat	-0.20	1.14	1.54
		t-stat INF	3.91	3.75	4.29
STOCK_GR		pseudo R^2	0.106	0.117	0.125
	0	t-stat	-0.26	-0.78	-1.33
		t-stat INF	3.89	3.87	4.50
		pseudo R^2	0.105	0.118	0.140
	1	t-stat	-0.33	-0.97	-1.84
		t-stat INF	3.75	3.64	4.23
		pseudo R^2	0.121	0.158	0.119
	2	t-stat	-1.50	-2.13	-1.58
		t-stat INF	3.81	3.79	4.43
		pseudo R^2	0.143	0.119	0.109
	3	t-stat	-2.05	-1.21	-0.32
		t-stat INF	3.75	3.79	4.28
		pseudo R^2	0.113	0.112	0.111
	4	t-stat	-1.27	-0.51	0.73
		t-stat INF	3.91	3.87	4.14
		pseudo R^2	0.106	0.113	0.117
	5	t-stat	-0.60	0.59	1.05
		t-stat INF	3.88	4.02	4.62

Table A.5 (Continued)

Dependent Variable is R1	lag		h = Quarters Ahead 1	2	3
EXCH_GR		pseudo R^2	0.197	0.275	0.217
	0	t-stat	-1.80	-1.58	-2.11
		t-stat INF	4.18	5.58	4.45
		pseudo R^2	0.186	0.269	0.198
	1	t-stat	-1.12	-1.08	-0.62
		t-stat INF	4.12	5.82	4.44
		pseudo R^2	0.195	0.265	0.201
	2	t-stat	-1.53	-0.80	1.08
		t-stat INF	4.08	5.63	4.63
		pseudo R^2	0.185	0.265	0.213
	3	t-stat	-0.89	0.58	1.65
		t-stat INF	4.15	5.72	4.94
		pseudo R^2	0.180	0.267	0.205
	4	t-stat	-0.19	0.85	0.99
		t-stat INF	4.20	5.68	4.52
		pseudo R^2	0.188	0.280	0.205
	5	t-stat	1.03	1.35	0.98
		t-stat INF	4.30	5.57	4.59
RTT_GR		pseudo R^2	0.189	0.269	0.223
	0	t-stat	-0.92	-0.87	1.82
		t-stat INF	3.75	4.83	4.57
		pseudo R^2	0.249	0.262	0.202
	1	t-stat	-1.87	-0.09	1.12
		t-stat INF	5.10	5.65	4.59
		pseudo R^2	0.181	0.281	0.227
	2	t-stat	0.32	1.60	2.27
		t-stat INF	4.17	5.63	4.77
		pseudo R^2	0.181	0.269	0.212
	3	t-stat	0.52	0.98	1.85
		t-stat INF	4.17	5.80	4.45
		pseudo R^2	0.193	0.292	0.199
	4	t-stat	1.26	2.31	1.02
		t-stat INF	4.28	5.68	4.64
		pseudo R^2	0.194	0.263	0.197
	5	t-stat	1.46	-0.18	-0.10
		t-stat INF	4.13	5.69	4.67
DWELL_GR		pseudo R^2	0.245	0.315	0.208
	0	t-stat	-2.40	-1.88	-1.17
		t-stat INF	3.84	5.21	4.47
		pseudo R^2	0.247	0.276	0.196
	1	t-stat	-3.01	-1.36	-0.81
		t-stat INF	4.01	5.40	4.51
		pseudo R^2	0.205	0.263	0.211
	2	t-stat	-2.14	-0.13	1.01
		t-stat INF	4.34	5.38	4.45
		pseudo R^2	0.188	0.265	0.207
	3	t-stat	-1.05	0.50	0.94
		t-stat INF	4.36	5.53	4.60
		pseudo R^2	0.180	0.266	0.198
	4	t-stat	0.11	0.53	0.32
		t-stat INF	4.21	5.84	4.97
		pseudo R^2	0.197	0.285	0.226
	5	t-stat	1.33	1.34	1.62
		t-stat INF	4.30	5.15	4.65
NONRES_GR		pseudo R^2	0.190	0.267	0.199
	0	t-stat	-1.22	-0.71	0.53
		t-stat INF	4.15	5.53	4.54
		pseudo R^2	0.205	0.264	0.197
	1	t-stat	-1.70	-0.54	-0.24
		t-stat INF	4.87	5.74	4.58
		pseudo R^2	0.187	0.266	0.198
	2	t-stat	-1.19	-0.98	-0.62
		t-stat INF	4.50	5.58	4.65
		pseudo R^2	0.182	0.272	0.198
	3	t-stat	0.46	1.31	0.62
		t-stat INF	4.14	5.25	4.47
		pseudo R^2	0.180	0.263	0.199
	4	t-stat	0.21	-0.41	-0.88
		t-stat INF	4.21	5.73	4.65
		pseudo R^2	0.181	0.265	0.199
	5	t-stat	-0.27	-0.60	0.72
		t-stat INF	4.20	5.61	4.81
SV_EX_GR		pseudo R^2	0.181	0.263	0.199
	0	t-stat	0.07	-0.10	0.63
		t-stat INF	4.26	5.72	4.86
		pseudo R^2	0.182	0.264	0.207
	1	t-stat	-0.45	0.44	-1.22
		t-stat INF	4.30	5.80	4.62
		pseudo R^2	0.188	0.264	0.206
	2	t-stat	1.23	-0.57	-1.60
		t-stat INF	4.59	5.60	4.48
		pseudo R^2	0.183	0.281	0.225
	3	t-stat	-0.69	-1.76	-1.95
		t-stat INF	4.18	6.00	5.22
		pseudo R^2	0.182	0.272	0.199
	4	t-stat	-0.59	-1.29	0.63
		t-stat INF	4.18	5.54	4.78
		pseudo R^2	0.229	0.273	0.198
	5	t-stat	-2.89	-1.64	-0.54
		t-stat INF	4.66	5.66	4.52

Dependent Variable is R2	lag		h = Quarters Ahead 1	2	3
EXCH_GR		pseudo R^2	0.107	0.114	0.119
	0	t-stat	0.56	-0.59	-1.13
		t-stat INF	3.89	3.90	4.46
		pseudo R^2	0.105	0.116	0.121
	1	t-stat	-0.12	-0.76	-1.02
		t-stat INF	3.91	3.76	4.18
		pseudo R^2	0.112	0.127	0.111
	2	t-stat	-0.94	-1.28	-0.73
		t-stat INF	3.75	3.62	4.21
		pseudo R^2	0.125	0.116	0.111
	3	t-stat	-1.57	-0.95	-0.71
		t-stat INF	3.76	3.87	4.33
		pseudo R^2	0.117	0.120	0.111
	4	t-stat	-1.38	-1.34	-0.86
		t-stat INF	3.93	4.01	4.51
		pseudo R^2	0.110	0.112	0.112
	5	t-stat	-1.09	-0.34	0.58
		t-stat INF	3.76	3.88	4.56
RTT_GR		pseudo R^2	0.113	0.124	0.111
	0	t-stat	-1.09	-1.54	0.82
		t-stat INF	3.27	3.18	4.04
		pseudo R^2	0.152	0.116	0.112
	1	t-stat	-2.75	-0.77	-0.68
		t-stat INF	4.16	3.99	4.50
		pseudo R^2	0.106	0.111	0.123
	2	t-stat	-0.44	-0.28	1.93
		t-stat INF	3.81	3.84	4.34
		pseudo R^2	0.113	0.112	0.134
	3	t-stat	-1.18	0.59	2.10
		t-stat INF	4.19	3.88	4.14
		pseudo R^2	0.110	0.151	0.130
	4	t-stat	0.86	2.59	2.05
		t-stat INF	3.96	3.91	4.64
		pseudo R^2	0.129	0.120	0.114
	5	t-stat	1.84	1.28	1.03
		t-stat INF	3.53	3.79	4.32
DWELL_GR		pseudo R^2	0.107	0.114	0.118
	0	t-stat	-0.49	-0.61	-1.07
		t-stat INF	3.68	3.71	4.21
		pseudo R^2	0.109	0.125	0.114
	1	t-stat	-0.89	-1.35	-0.88
		t-stat INF	3.71	3.64	4.21
		pseudo R^2	0.121	0.120	0.111
	2	t-stat	-1.75	-1.18	-0.53
		t-stat INF	3.77	3.76	4.25
		pseudo R^2	0.122	0.118	0.113
	3	t-stat	-1.82	-1.03	-0.83
		t-stat INF	3.98	4.03	4.48
		pseudo R^2	0.120	0.123	0.120
	4	t-stat	-1.57	-1.30	1.55
		t-stat INF	3.88	3.85	4.59
		pseudo R^2	0.108	0.131	0.168
	5	t-stat	-0.65	1.85	3.02
		t-stat INF	3.78	3.91	5.02
NONRES_GR		pseudo R^2	0.105	0.122	0.109
	0	t-stat	-0.11	-1.40	-0.03
		t-stat INF	3.83	3.85	4.39
		pseudo R^2	0.129	0.115	0.111
	1	t-stat	-2.22	-0.89	0.71
		t-stat INF	4.27	4.09	4.36
		pseudo R^2	0.110	0.112	0.109
	2	t-stat	-1.06	0.34	0.53
		t-stat INF	3.96	3.91	4.35
		pseudo R^2	0.114	0.124	0.132
	3	t-stat	1.29	1.50	1.97
		t-stat INF	3.93	3.90	4.83
		pseudo R^2	0.108	0.119	0.108
	4	t-stat	0.78	1.28	0.02
		t-stat INF	3.91	3.94	4.37
		pseudo R^2	0.113	0.111	0.111
	5	t-stat	1.20	0.03	0.46
		t-stat INF	3.80	3.91	4.49
SV_EX_GR		pseudo R^2	0.105	0.114	0.115
	0	t-stat	0.08	-0.58	-0.76
		t-stat INF	3.91	4.01	4.27
		pseudo R^2	0.110	0.119	0.112
	1	t-stat	-0.76	-0.81	0.74
		t-stat INF	3.94	3.98	4.37
		pseudo R^2	0.107	0.122	0.110
	2	t-stat	-0.48	1.24	0.46
		t-stat INF	3.80	4.21	4.71
		pseudo R^2	0.113	0.112	0.116
	3	t-stat	1.06	0.37	-1.17
		t-stat INF	3.92	3.93	4.48
		pseudo R^2	0.110	0.113	0.112
	4	t-stat	1.03	-0.69	0.82
		t-stat INF	3.91	3.80	4.19
		pseudo R^2	0.122	0.113	0.118
	5	t-stat	-1.82	-0.58	-1.16
		t-stat INF	4.06	4.11	4.88

Table A.5 (Continued)

Dependent Variable is R1	lag		h = Quarters Ahead 1	2	3
TOT_GR		pseudo R^2	0.194	0.266	0.197
	0	t-stat	-1.42	-0.78	-0.08
		t-stat INF	3.95	5.30	4.31
		pseudo R^2	0.201	0.287	0.196
	1	t-stat	-1.98	-0.90	0.00
		t-stat INF	4.03	5.54	4.59
		pseudo R^2	0.198	0.288	0.211
	2	t-stat	-1.92	-0.98	1.62
		t-stat INF	4.23	5.65	4.84
		pseudo R^2	0.181	0.282	0.202
	3	t-stat	-0.48	1.70	0.88
		t-stat INF	4.20	5.65	4.73
		pseudo R^2	0.198	0.269	0.218
	4	t-stat	1.74	0.98	1.72
		t-stat INF	4.38	5.75	4.79
		pseudo R^2	0.183	0.272	0.200
	5	t-stat	0.62	1.07	0.48
		t-stat INF	4.16	5.53	4.59
GOLD_GR		pseudo R^2	0.192	0.271	0.239
	0	t-stat	-1.41	-1.07	-1.91
		t-stat INF	4.16	5.52	4.47
		pseudo R^2	0.188	0.308	0.265
	1	t-stat	-0.89	-1.65	-1.79
		t-stat INF	4.13	5.51	4.56
		pseudo R^2	0.225	0.358	0.240
	2	t-stat	-2.12	-3.00	-2.09
		t-stat INF	3.85	5.10	4.50
		pseudo R^2	0.236	0.299	0.218
	3	t-stat	-2.34	-1.99	-1.39
		t-stat INF	3.70	4.97	4.39
		pseudo R^2	0.202	0.274	0.199
	4	t-stat	-1.34	-1.08	-0.52
		t-stat INF	4.28	5.44	4.47
		pseudo R^2	0.214	0.293	0.201
	5	t-stat	-1.86	-1.88	-0.82
		t-stat INF	4.35	5.73	5.01
US_SPR		pseudo R^2	0.190	0.263	0.227
	0	t-stat	1.21	-0.18	-1.85
		t-stat INF	4.07	5.59	4.61
		pseudo R^2	0.181	0.287	0.261
	1	t-stat	-0.39	-1.81	-2.16
		t-stat INF	4.10	5.54	4.41
		pseudo R^2	0.200	0.305	0.238
	2	t-stat	-1.78	-1.97	-1.76
		t-stat INF	3.94	5.18	4.05
		pseudo R^2	0.240	0.309	0.248
	3	t-stat	-2.98	-2.18	-1.99
		t-stat INF	4.09	5.23	4.00
		pseudo R^2	0.239	0.311	0.247
	4	t-stat	-3.07	-2.22	-2.28
		t-stat INF	4.02	5.12	3.92
		pseudo R^2	0.232	0.302	0.233
	5	t-stat	-2.86	-2.08	-1.98
		t-stat INF	3.78	4.93	3.85
US_SH_R		pseudo R^2	0.197	0.293	0.267
	0	t-stat	1.59	1.80	2.47
		t-stat INF	3.82	4.78	3.60
		pseudo R^2	0.218	0.320	0.287
	1	t-stat	2.42	2.48	2.78
		t-stat INF	3.44	4.86	3.41
		pseudo R^2	0.240	0.325	0.258
	2	t-stat	3.00	2.46	2.24
		t-stat INF	3.14	4.16	3.17
		pseudo R^2	0.269	0.319	0.247
	3	t-stat	3.78	2.39	2.18
		t-stat INF	3.25	4.34	3.39
		pseudo R^2	0.255	0.308	0.239
	4	t-stat	3.57	2.25	2.08
		t-stat INF	3.34	4.44	3.60
		pseudo R^2	0.236	0.293	0.224
	5	t-stat	3.10	1.77	1.60
		t-stat INF	3.29	4.46	3.48
US_LG_R		pseudo R^2	0.218	0.298	0.241
	0	t-stat	2.77	2.12	2.17
		t-stat INF	3.80	4.75	3.72
		pseudo R^2	0.223	0.298	0.239
	1	t-stat	3.07	2.11	2.08
		t-stat INF	3.53	4.89	3.67
		pseudo R^2	0.220	0.292	0.223
	2	t-stat	2.85	1.78	1.56
		t-stat INF	3.47	4.84	3.67
		pseudo R^2	0.223	0.286	0.215
	3	t-stat	2.98	1.59	1.33
		t-stat INF	3.53	4.78	3.84
		pseudo R^2	0.214	0.277	0.209
	4	t-stat	2.60	1.31	1.13
		t-stat INF	3.61	4.89	3.94
		pseudo R^2	0.203	0.273	0.205
	5	t-stat	2.11	1.05	0.88
		t-stat INF	3.67	4.99	3.98
US_STK_GR		pseudo R^2	0.181	0.263	0.203
	0	t-stat	-0.12	0.07	1.00
		t-stat INF	4.30	5.85	4.84
		pseudo R^2	0.182	0.281	0.201
	1	t-stat	0.56	1.79	0.68
		t-stat INF	4.20	6.21	4.39
		pseudo R^2	0.192	0.265	0.199
	2	t-stat	1.23	0.63	0.70
		t-stat INF	4.45	5.71	4.59
		pseudo R^2	0.183	0.273	0.214
	3	t-stat	0.51	1.22	2.12
		t-stat INF	4.03	5.63	4.48
		pseudo R^2	0.180	0.264	0.198
	4	t-stat	-0.16	0.77	-0.80
		t-stat INF	4.20	5.81	4.57
		pseudo R^2	0.203	0.269	0.205
	5	t-stat	2.13	1.08	1.21
		t-stat INF	4.40	5.38	4.81

Dependent Variable is R2	lag		h = Quarters Ahead 1	2	3
TOT_GR		pseudo R^2	0.120	0.117	0.109
	0	t-stat	-1.54	-1.08	0.15
		t-stat INF	3.42	3.64	3.97
		pseudo R^2	0.117	0.112	0.109
	1	t-stat	-1.44	-0.37	-0.23
		t-stat INF	3.60	3.80	4.35
		pseudo R^2	0.111	0.115	0.109
	2	t-stat	-0.80	-0.71	-0.29
		t-stat INF	4.01	4.02	4.43
		pseudo R^2	0.107	0.111	0.117
	3	t-stat	-0.55	-0.12	1.17
		t-stat INF	3.88	3.93	4.37
		pseudo R^2	0.104	0.121	0.136
	4	t-stat	-0.08	1.22	1.73
		t-stat INF	3.88	4.00	4.73
		pseudo R^2	0.109	0.127	0.112
	5	t-stat	0.79	1.31	-0.44
		t-stat INF	3.81	3.63	4.19
GOLD_GR		pseudo R^2	0.114	0.112	0.120
	0	t-stat	1.13	0.07	1.12
		t-stat INF	3.81	3.91	4.35
		pseudo R^2	0.105	0.119	0.109
	1	t-stat	-0.05	1.00	-0.28
		t-stat INF	3.87	3.97	4.29
		pseudo R^2	0.112	0.112	0.140
	2	t-stat	0.93	-0.34	-2.38
		t-stat INF	3.88	4.07	3.94
		pseudo R^2	0.105	0.138	0.118
	3	t-stat	-0.18	-2.14	-0.84
		t-stat INF	3.88	3.77	4.35
		pseudo R^2	0.125	0.115	0.122
	4	t-stat	-1.79	-0.84	-1.08
		t-stat INF	3.60	3.88	4.45
		pseudo R^2	0.120	0.144	0.171
	5	t-stat	-1.20	-1.88	-2.48
		t-stat INF	3.79	4.68	5.14
US_SPR		pseudo R^2	0.105	0.112	0.109
	0	t-stat	0.13	0.22	0.61
		t-stat INF	3.81	3.92	4.34
		pseudo R^2	0.105	0.111	0.111
	1	t-stat	0.28	0.07	-0.47
		t-stat INF	3.77	3.85	4.28
		pseudo R^2	0.105	0.112	0.117
	2	t-stat	0.33	-0.24	-0.91
		t-stat INF	3.64	3.61	3.95
		pseudo R^2	0.108	0.122	0.125
	3	t-stat	-0.47	-1.08	-1.19
		t-stat INF	3.68	3.55	3.98
		pseudo R^2	0.118	0.129	0.122
	4	t-stat	-1.28	-1.32	-1.08
		t-stat INF	3.61	3.38	3.88
		pseudo R^2	0.122	0.125	0.111
	5	t-stat	-1.47	-1.13	-0.33
		t-stat INF	3.44	3.27	4.13
US_SH_R		pseudo R^2	0.119	0.129	0.130
	0	t-stat	1.54	1.49	1.49
		t-stat INF	3.27	3.17	3.52
		pseudo R^2	0.122	0.130	0.142
	1	t-stat	1.69	1.46	1.96
		t-stat INF	3.16	3.05	3.32
		pseudo R^2	0.120	0.138	0.157
	2	t-stat	1.48	1.78	2.36
		t-stat INF	2.99	2.70	2.76
		pseudo R^2	0.138	0.163	0.168
	3	t-stat	2.12	2.46	2.45
		t-stat INF	2.99	2.64	2.88
		pseudo R^2	0.158	0.172	0.167
	4	t-stat	2.83	2.63	2.14
		t-stat INF	2.88	2.89	3.02
		pseudo R^2	0.164	0.159	0.137
	5	t-stat	2.93	2.19	1.82
		t-stat INF	2.89	2.80	3.22
US_LG_R		pseudo R^2	0.124	0.138	0.135
	0	t-stat	1.84	1.82	1.72
		t-stat INF	3.29	3.19	3.82
		pseudo R^2	0.131	0.138	0.141
	1	t-stat	2.13	1.74	1.92
		t-stat INF	3.21	3.16	3.52
		pseudo R^2	0.127	0.139	0.148
	2	t-stat	1.89	1.83	2.02
		t-stat INF	3.14	3.03	3.32
		pseudo R^2	0.138	0.151	0.150
	3	t-stat	2.25	2.16	1.99
		t-stat INF	3.16	3.04	3.37
		pseudo R^2	0.145	0.153	0.143
	4	t-stat	2.63	2.16	1.77
		t-stat INF	3.13	3.05	3.47
		pseudo R^2	0.147	0.145	0.138
	5	t-stat	2.84	1.90	1.80
		t-stat INF	3.10	3.09	3.57
US_STK_GR		pseudo R^2	0.134	0.123	0.138
	0	t-stat	-1.71	-1.28	-1.84
		t-stat INF	3.84	3.97	4.41
		pseudo R^2	0.111	0.125	0.109
	1	t-stat	-0.95	-1.47	-0.18
		t-stat INF	3.89	3.80	4.46
		pseudo R^2	0.125	0.112	0.108
	2	t-stat	-1.83	-0.45	-0.24
		t-stat INF	3.79	3.93	4.41
		pseudo R^2	0.105	0.111	0.138
	3	t-stat	-0.39	-0.13	2.67
		t-stat INF	3.83	3.88	4.45
		pseudo R^2	0.108	0.123	0.120
	4	t-stat	-0.88	1.59	2.01
		t-stat INF	3.89	4.02	4.49
		pseudo R^2	0.140	0.143	0.120
	5	t-stat	2.42	2.81	1.04
		t-stat INF	3.97	3.92	4.23

Table A.6: Table A6: In-Sample Statistics: RHS are Constant, US_SH_R, and the Variable as Stated

Dependent Variable is R1	lag		h = Quarters Ahead 4	5	6	7	8
SPREAD		pseudo R^2	0.232	0.196	0.149	0.122	0.097
	0	t-stat	-1.70	-1.43	-0.84	-0.53	-0.36
		t-stat US_SH_R	2.99	2.97	3.10	3.05	2.57
		pseudo R^2	0.221	0.174	0.141	0.119	0.106
	1	t-stat	-1.51	-0.89	-0.58	-0.37	-1.09
		t-stat US_SH_R	3.14	3.21	3.32	3.19	2.38
		pseudo R^2	0.202	0.168	0.139	0.131	0.111
	2	t-stat	-1.07	-0.73	-0.57	-1.16	-1.12
		t-stat US_SH_R	3.33	3.38	3.32	2.95	2.45
		pseudo R^2	0.194	0.165	0.153	0.136	0.115
	3	t-stat	-0.87	-0.68	-1.24	-1.20	-1.13
		t-stat US_SH_R	3.42	3.38	3.16	2.99	2.49
		pseudo R^2	0.188	0.178	0.155	0.137	0.119
	4	t-stat	-0.69	-1.22	-1.23	-1.16	-1.21
		t-stat US_SH_R	3.43	3.22	3.13	2.98	2.50
		pseudo R^2	0.200	0.179	0.155	0.140	0.132
	5	t-stat	-1.22	-1.28	-1.22	-1.25	-1.66
		t-stat US_SH_R	3.28	3.20	3.13	3.00	2.49
SH_R		pseudo R^2	0.220	0.179	0.139	0.116	0.097
	0	t-stat	1.71	1.11	0.44	0.03	-0.28
		t-stat US_SH_R	1.79	1.86	2.08	2.43	2.42
		pseudo R^2	0.203	0.163	0.135	0.120	0.097
	1	t-stat	1.18	0.46	0.05	-0.31	-0.19
		t-stat US_SH_R	2.25	2.38	2.63	3.01	2.62
		pseudo R^2	0.188	0.160	0.137	0.119	0.098
	2	t-stat	0.58	0.11	-0.22	-0.14	-0.26
		t-stat US_SH_R	2.67	2.79	2.94	2.91	2.73
		pseudo R^2	0.183	0.161	0.137	0.119	0.099
	3	t-stat	0.21	-0.14	-0.03	-0.18	-0.24
		t-stat US_SH_R	2.97	2.98	2.82	2.99	2.74
		pseudo R^2	0.183	0.160	0.137	0.120	0.098
	4	t-stat	-0.08	0.02	-0.12	-0.22	-0.19
		t-stat US_SH_R	3.09	2.91	2.91	3.04	2.74
		pseudo R^2	0.182	0.160	0.137	0.119	0.097
	5	t-stat	0.07	-0.09	-0.19	-0.20	-0.06
		t-stat US_SH_R	3.02	2.93	2.98	3.09	2.69
LG_R		pseudo R^2	0.187	0.159	0.134	0.119	0.103
	0	t-stat	0.73	0.30	-0.04	-0.39	-0.60
		t-stat US_SH_R	2.04	2.13	2.33	2.86	2.80
		pseudo R^2	0.183	0.159	0.138	0.126	0.109
	1	t-stat	0.36	-0.04	-0.36	-0.65	-0.77
		t-stat US_SH_R	2.39	2.48	2.71	3.27	3.06
		pseudo R^2	0.182	0.162	0.144	0.132	0.116
	2	t-stat	0.00	-0.36	-0.61	-0.82	-0.94
		t-stat US_SH_R	2.64	2.75	3.01	3.56	3.27
		pseudo R^2	0.184	0.167	0.148	0.137	0.119
	3	t-stat	-0.29	-0.58	-0.78	-0.99	-1.03
		t-stat US_SH_R	2.86	2.97	3.27	3.84	3.34
		pseudo R^2	0.187	0.170	0.153	0.140	0.119
	4	t-stat	-0.51	-0.74	-0.96	-1.08	-1.00
		t-stat US_SH_R	3.03	3.16	3.51	3.98	3.30
		pseudo R^2	0.189	0.174	0.155	0.139	0.118
	5	t-stat	-0.69	-0.93	-1.06	-1.07	-0.95
		t-stat US_SH_R	3.17	3.33	3.68	4.02	3.21
D_SH_R		pseudo R^2	0.220	0.179	0.139	0.116	0.097
	0	t-stat	1.71	1.11	0.44	0.03	-0.28
		t-stat US_SH_R	3.56	3.39	3.30	3.14	2.65
		pseudo R^2	0.206	0.165	0.135	0.119	0.097
	1	t-stat	1.28	0.54	0.08	-0.24	0.08
		t-stat US_SH_R	3.55	3.41	3.27	3.12	2.65
		pseudo R^2	0.190	0.160	0.137	0.118	0.098
	2	t-stat	0.67	0.15	-0.19	0.16	0.27
		t-stat US_SH_R	3.52	3.36	3.24	3.19	2.69
		pseudo R^2	0.184	0.160	0.139	0.122	0.099
	3	t-stat	0.33	-0.04	0.34	0.46	0.28
		t-stat US_SH_R	3.43	3.29	3.36	3.26	2.69
		pseudo R^2	0.182	0.163	0.140	0.120	0.098
	4	t-stat	0.03	0.44	0.53	0.36	0.20
		t-stat US_SH_R	3.41	3.43	3.41	3.24	2.65
		pseudo R^2	0.185	0.164	0.138	0.119	0.103
	5	t-stat	0.46	0.58	0.37	0.20	0.62
		t-stat US_SH_R	3.47	3.46	3.36	3.17	2.74
D_LG_R		pseudo R^2	0.224	0.178	0.142	0.117	0.096
	0	t-stat	1.80	1.13	0.64	0.18	-0.18
		t-stat US_SH_R	3.56	3.41	3.27	3.07	2.56
		pseudo R^2	0.205	0.169	0.136	0.118	0.098
	1	t-stat	1.32	0.77	0.24	-0.13	-0.28
		t-stat US_SH_R	3.51	3.41	3.26	3.04	2.54
		pseudo R^2	0.192	0.161	0.137	0.120	0.100
	2	t-stat	0.82	0.27	-0.13	-0.28	-0.34
		t-stat US_SH_R	3.50	3.38	3.24	3.05	2.56
		pseudo R^2	0.185	0.160	0.138	0.120	0.099
	3	t-stat	0.43	-0.02	-0.20	-0.26	-0.22
		t-stat US_SH_R	3.46	3.32	3.20	3.02	2.56
		pseudo R^2	0.183	0.161	0.138	0.119	0.098
	4	t-stat	0.07	-0.13	-0.21	-0.17	-0.10
		t-stat US_SH_R	3.44	3.30	3.18	3.04	2.55
		pseudo R^2	0.182	0.160	0.137	0.118	0.097
	5	t-stat	-0.12	-0.20	-0.17	-0.10	0.16
		t-stat US_SH_R	3.42	3.29	3.19	3.01	2.58

Dependent Variable is R2	lag		h = Quarters Ahead 4	5	6	7	8
SPREAD		pseudo R^2	0.101	0.152	0.147	0.103	0.072
	0	t-stat	-1.63	-1.77	-1.60	-0.55	0.28
		t-stat US_SH_R	2.54	2.81	2.82	3.03	2.90
		pseudo R^2	0.133	0.146	0.121	0.100	0.078
	1	t-stat	-2.00	-1.73	-0.59	0.33	0.86
		t-stat US_SH_R	2.33	3.01	3.39	3.41	3.07
		pseudo R^2	0.130	0.117	0.119	0.107	0.079
	2	t-stat	-2.02	-0.82	0.22	0.86	1.26
		t-stat US_SH_R	2.54	3.44	3.70	3.53	3.05
		pseudo R^2	0.093	0.114	0.128	0.109	0.076
	3	t-stat	-1.03	0.09	0.77	1.27	1.02
		t-stat US_SH_R	2.87	3.75	3.81	3.53	2.92
		pseudo R^2	0.089	0.128	0.131	0.107	0.072
	4	t-stat	-0.01	0.70	1.25	1.10	0.38
		t-stat US_SH_R	3.17	3.84	3.78	3.35	2.82
		pseudo R^2	0.107	0.130	0.128	0.102	0.077
	5	t-stat	0.64	1.16	1.17	0.38	-0.65
		t-stat US_SH_R	3.43	3.85	3.60	3.18	2.70
SH_R		pseudo R^2	0.096	0.120	0.120	0.100	0.073
	0	t-stat	1.17	0.82	0.52	-0.17	-0.36
		t-stat US_SH_R	1.50	2.02	2.37	2.77	2.57
		pseudo R^2	0.104	0.118	0.117	0.103	0.073
	1	t-stat	1.25	0.77	-0.15	-0.54	-0.57
		t-stat US_SH_R	1.54	2.44	3.00	3.23	2.87
		pseudo R^2	0.101	0.111	0.121	0.105	0.075
	2	t-stat	1.22	0.14	-0.47	-0.74	-0.82
		t-stat US_SH_R	1.91	3.01	3.38	3.56	3.10
		pseudo R^2	0.086	0.114	0.126	0.107	0.073
	3	t-stat	0.57	-0.21	-0.65	-0.98	-0.57
		t-stat US_SH_R	2.44	3.32	3.56	3.69	3.00
		pseudo R^2	0.089	0.124	0.128	0.104	0.071
	4	t-stat	0.13	-0.42	-0.90	-0.75	-0.28
		t-stat US_SH_R	2.79	3.46	3.64	3.52	2.92
		pseudo R^2	0.103	0.124	0.124	0.101	0.073
	5	t-stat	-0.13	-0.66	-0.72	-0.38	0.54
		t-stat US_SH_R	3.06	3.58	3.60	3.29	2.56
LG_R		pseudo R^2	0.084	0.110	0.120	0.106	0.072
	0	t-stat	0.54	-0.27	-0.58	-0.71	-0.33
		t-stat US_SH_R	1.83	2.75	3.00	3.19	2.75
		pseudo R^2	0.083	0.111	0.122	0.102	0.070
	1	t-stat	0.27	-0.27	-0.65	-0.58	-0.12
		t-stat US_SH_R	2.18	2.88	3.15	3.41	2.80
		pseudo R^2	0.083	0.112	0.122	0.102	0.071
	2	t-stat	0.20	-0.37	-0.54	-0.40	-0.28
		t-stat US_SH_R	2.31	3.05	3.33	3.51	2.91
		pseudo R^2	0.083	0.114	0.123	0.102	0.071
	3	t-stat	0.07	-0.23	-0.36	-0.49	-0.11
		t-stat US_SH_R	2.50	3.20	3.45	3.54	2.89
		pseudo R^2	0.089	0.122	0.123	0.101	0.071
	4	t-stat	0.17	-0.08	-0.42	-0.27	-0.04
		t-stat US_SH_R	2.67	3.34	3.52	3.47	2.93
		pseudo R^2	0.103	0.121	0.122	0.100	0.071
	5	t-stat	0.25	-0.15	-0.23	-0.18	0.17
		t-stat US_SH_R	2.89	3.44	3.53	3.46	2.89
D_SH_R		pseudo R^2	0.096	0.120	0.120	0.100	0.073
	0	t-stat	1.17	0.82	0.52	-0.17	-0.36
		t-stat US_SH_R	3.31	3.81	3.61	3.21	2.71
		pseudo R^2	0.083	0.115	0.117	0.100	0.070
	1	t-stat	0.87	0.59	-0.12	-0.32	-0.16
		t-stat US_SH_R	3.16	3.74	3.50	3.16	2.77
		pseudo R^2	0.087	0.111	0.119	0.100	0.071
	2	t-stat	0.62	-0.08	-0.25	-0.08	-0.05
		t-stat US_SH_R	3.06	3.56	3.48	3.18	2.80
		pseudo R^2	0.083	0.114	0.122	0.101	0.071
	3	t-stat	-0.05	-0.15	0.10	0.18	0.22
		t-stat US_SH_R	2.95	3.58	3.50	3.19	2.81
		pseudo R^2	0.089	0.122	0.123	0.102	0.074
	4	t-stat	-0.21	0.15	0.27	0.37	0.64
		t-stat US_SH_R	3.10	3.78	3.52	3.21	2.87
		pseudo R^2	0.102	0.122	0.123	0.104	0.093
	5	t-stat	0.07	0.27	0.43	0.80	1.46
		t-stat US_SH_R	3.41	3.75	3.52	3.31	2.99
D_LG_R		pseudo R^2	0.085	0.110	0.119	0.106	0.075
	0	t-stat	0.52	-0.18	-0.40	-0.56	-0.54
		t-stat US_SH_R	3.07	3.45	3.33	3.04	2.65
		pseudo R^2	0.083	0.113	0.122	0.101	0.070
	1	t-stat	-0.22	-0.38	-0.52	-0.46	-0.05
		t-stat US_SH_R	2.91	3.40	3.33	3.08	2.75
		pseudo R^2	0.087	0.116	0.121	0.100	0.071
	2	t-stat	-0.46	-0.55	-0.46	-0.02	0.08
		t-stat US_SH_R	2.93	3.43	3.44	3.16	2.80
		pseudo R^2	0.088	0.115	0.122	0.101	0.073
	3	t-stat	-0.55	-0.37	0.15	0.24	0.40
		t-stat US_SH_R	2.86	3.52	3.45	3.17	2.80
		pseudo R^2	0.091	0.122	0.123	0.103	0.076
	4	t-stat	-0.41	0.22	0.34	0.51	0.68
		t-stat US_SH_R	3.05	3.70	3.49	3.17	2.82
		pseudo R^2	0.103	0.122	0.125	0.106	0.078
	5	t-stat	0.15	0.35	0.58	0.75	0.83
		t-stat US_SH_R	3.38	3.72	3.50	3.17	2.79

Table A.6 (Continued)

Dependent Variable is R1	lag		h = Quarters Ahead 4	5	6	7	8
M3_GR		pseudo R^2	0.331	0.293	0.259	0.218	0.163
	0	t-stat	4.46	2.24	2.49	2.54	2.50
		t-stat US_SH_R	3.16	3.30	3.18	3.11	2.50
		pseudo R^2	0.311	0.277	0.230	0.179	0.125
	1	t-stat	2.20	2.39	2.31	2.36	2.01
		t-stat US_SH_R	3.24	3.12	3.00	2.93	2.46
		pseudo R^2	0.294	0.249	0.193	0.144	0.100
	2	t-stat	2.31	2.17	2.17	1.94	0.97
		t-stat US_SH_R	3.18	3.11	3.08	3.00	2.56
		pseudo R^2	0.269	0.214	0.160	0.121	0.099
	3	t-stat	2.02	1.98	1.73	0.88	-0.50
		t-stat US_SH_R	3.20	3.15	3.16	3.12	2.64
		pseudo R^2	0.235	0.182	0.139	0.121	0.098
	4	t-stat	1.79	1.55	0.77	-0.50	-0.07
		t-stat US_SH_R	3.19	3.21	3.26	3.16	2.59
		pseudo R^2	0.205	0.162	0.138	0.118	0.098
	5	t-stat	1.48	0.80	-0.40	-0.03	0.34
		t-stat US_SH_R	3.29	3.33	3.26	3.02	2.51
CUR_GR		pseudo R^2	0.203	0.216	0.214	0.229	0.200
	0	t-stat	1.06	1.69	2.19	3.26	4.26
		t-stat US_SH_R	2.92	2.81	2.65	2.40	1.80
		pseudo R^2	0.237	0.235	0.244	0.217	0.195
	1	t-stat	1.67	2.13	3.17	4.01	3.68
		t-stat US_SH_R	2.95	2.82	2.67	2.40	1.88
		pseudo R^2	0.251	0.261	0.230	0.210	0.283
	2	t-stat	2.05	3.01	3.56	3.25	2.97
		t-stat US_SH_R	2.87	2.78	2.59	2.38	1.82
		pseudo R^2	0.295	0.263	0.237	0.232	0.185
	3	t-stat	2.97	3.55	3.25	2.82	2.76
		t-stat US_SH_R	3.22	3.11	2.93	2.64	2.08
		pseudo R^2	0.295	0.269	0.259	0.212	0.163
	4	t-stat	3.45	3.31	2.85	2.67	2.49
		t-stat US_SH_R	3.39	3.33	3.21	2.85	2.29
		pseudo R^2	0.291	0.283	0.230	0.183	0.166
	5	t-stat	3.22	2.63	2.54	2.24	2.80
		t-stat US_SH_R	3.45	3.44	3.17	2.86	2.21
R_M3_GR		pseudo R^2	0.215	0.196	0.168	0.135	0.110
	0	t-stat	2.00	1.52	1.47	1.00	0.93
		t-stat US_SH_R	3.56	3.39	3.39	3.32	2.85
		pseudo R^2	0.218	0.191	0.151	0.129	0.098
	1	t-stat	1.47	1.44	0.87	0.75	0.24
		t-stat US_SH_R	3.42	3.43	3.35	3.24	2.74
		pseudo R^2	0.210	0.173	0.145	0.119	0.107
	2	t-stat	1.44	0.81	0.65	0.15	-0.94
		t-stat US_SH_R	3.51	3.44	3.35	3.25	2.66
		pseudo R^2	0.195	0.170	0.137	0.128	0.138
	3	t-stat	0.79	0.63	0.14	-0.92	-2.10
		t-stat US_SH_R	3.51	3.39	3.39	3.24	2.66
		pseudo R^2	0.192	0.161	0.146	0.158	0.129
	4	t-stat	0.59	0.13	-0.86	-1.97	-3.02
		t-stat US_SH_R	3.45	3.43	3.44	3.23	2.63
		pseudo R^2	0.182	0.169	0.174	0.146	0.106
	5	t-stat	0.13	-0.82	-1.84	-2.73	-1.18
		t-stat US_SH_R	3.49	3.53	3.48	3.20	2.66
R_CUR_GR		pseudo R^2	0.201	0.159	0.134	0.117	0.102
	0	t-stat	-1.67	-0.43	-0.03	0.24	0.96
		t-stat US_SH_R	3.39	3.41	3.37	3.26	2.77
		pseudo R^2	0.182	0.159	0.136	0.124	0.103
	1	t-stat	-0.43	-0.04	0.21	0.91	0.81
		t-stat US_SH_R	3.49	3.41	3.37	3.28	2.77
		pseudo R^2	0.182	0.160	0.140	0.122	0.102
	2	t-stat	-0.24	0.02	0.71	0.66	0.83
		t-stat US_SH_R	3.51	3.43	3.36	3.23	2.72
		pseudo R^2	0.183	0.168	0.144	0.127	0.099
	3	t-stat	0.25	0.96	0.88	1.11	0.49
		t-stat US_SH_R	3.49	3.40	3.33	3.27	2.72
		pseudo R^2	0.194	0.172	0.149	0.122	0.098
	4	t-stat	1.21	1.10	1.38	0.76	-0.11
		t-stat US_SH_R	3.47	3.35	3.41	3.25	2.66
		pseudo R^2	0.191	0.171	0.139	0.118	0.099
	5	t-stat	1.09	1.33	0.73	-0.12	0.76
		t-stat US_SH_R	3.42	3.40	3.35	3.18	2.75
STOCK_GR		pseudo R^2	0.212	0.159	0.135	0.127	0.097
	0	t-stat	-1.95	-0.60	0.31	0.92	0.42
		t-stat US_SH_R	3.56	3.44	3.45	3.50	2.77
		pseudo R^2	0.185	0.159	0.142	0.118	0.104
	1	t-stat	-0.82	0.13	0.80	0.26	-0.85
		t-stat US_SH_R	3.54	3.48	3.48	3.28	2.60
		pseudo R^2	0.182	0.164	0.136	0.129	0.097
	2	t-stat	-0.04	0.67	0.05	-0.99	-0.12
		t-stat US_SH_R	3.52	3.46	3.40	3.13	2.68
		pseudo R^2	0.184	0.161	0.154	0.120	0.103
	3	t-stat	0.43	-0.24	-1.44	-0.37	-0.81
		t-stat US_SH_R	3.47	3.44	3.42	3.21	2.68
		pseudo R^2	0.183	0.180	0.139	0.126	0.106
	4	t-stat	-0.33	-1.70	-0.50	-1.02	-0.91
		t-stat US_SH_R	3.51	3.52	3.43	3.31	2.76
		pseudo R^2	0.204	0.162	0.145	0.129	0.104
	5	t-stat	-1.68	-0.61	-1.11	-0.99	-0.79
		t-stat US_SH_R	3.61	3.54	3.57	3.44	2.78

Dependent Variable is R2	lag		h = Quarters Ahead 4	5	6	7	8
M3_GR		pseudo R^2	0.116	0.118	0.142	0.134	0.109
	0	t-stat	1.94	1.08	1.60	2.06	2.29
		t-stat US_SH_R	2.79	3.63	3.57	3.22	2.67
		pseudo R^2	0.091	0.134	0.148	0.134	0.107
	1	t-stat	1.07	1.54	1.97	2.22	1.99
		t-stat US_SH_R	2.95	3.57	3.48	3.08	2.59
		pseudo R^2	0.105	0.139	0.150	0.134	0.073
	2	t-stat	1.54	1.83	2.03	1.82	0.65
		t-stat US_SH_R	2.73	3.41	3.35	2.95	2.74
		pseudo R^2	0.113	0.145	0.155	0.102	0.078
	3	t-stat	1.76	1.90	1.66	0.47	-0.76
		t-stat US_SH_R	2.66	3.42	3.29	3.18	2.82
		pseudo R^2	0.123	0.156	0.123	0.111	0.072
	4	t-stat	1.89	1.68	0.39	-0.86	-0.34
		t-stat US_SH_R	2.89	3.55	3.50	3.42	2.76
		pseudo R^2	0.140	0.123	0.133	0.102	0.079
	5	t-stat	1.81	0.49	-0.90	-0.41	0.71
		t-stat US_SH_R	3.23	3.64	4.02	3.24	2.62
CUR_GR		pseudo R^2	0.106	0.126	0.122	0.130	0.118
	0	t-stat	1.42	1.29	1.00	2.77	2.78
		t-stat US_SH_R	2.37	3.09	3.22	2.91	2.23
		pseudo R^2	0.103	0.116	0.145	0.141	0.121
	1	t-stat	1.40	1.01	2.37	2.69	3.59
		t-stat US_SH_R	2.43	3.25	3.20	2.74	2.25
		pseudo R^2	0.091	0.139	0.155	0.144	0.151
	2	t-stat	1.09	2.20	2.56	3.68	3.76
		t-stat US_SH_R	2.50	3.09	3.08	2.70	2.11
		pseudo R^2	0.121	0.157	0.170	0.185	0.120
	3	t-stat	2.40	2.72	3.49	3.62	2.85
		t-stat US_SH_R	2.48	3.27	3.17	2.72	2.43
		pseudo R^2	0.140	0.175	0.212	0.152	0.082
	4	t-stat	2.91	3.65	3.51	2.97	1.26
		t-stat US_SH_R	2.72	3.39	3.15	2.98	2.62
		pseudo R^2	0.158	0.212	0.172	0.110	0.079
	5	t-stat	3.80	3.56	2.86	1.15	1.15
		t-stat US_SH_R	2.99	3.37	3.39	3.00	2.55
R_M3_GR		pseudo R^2	0.090	0.117	0.134	0.138	0.107
	0	t-stat	0.76	0.88	1.02	1.96	2.13
		t-stat US_SH_R	3.12	3.74	3.71	3.61	3.06
		pseudo R^2	0.088	0.127	0.155	0.137	0.076
	1	t-stat	0.76	0.99	1.94	2.08	0.90
		t-stat US_SH_R	3.05	3.72	3.79	3.42	2.95
		pseudo R^2	0.095	0.143	0.153	0.106	0.073
	2	t-stat	0.94	1.88	2.03	0.82	-0.53
		t-stat US_SH_R	3.04	3.74	3.68	3.32	2.87
		pseudo R^2	0.114	0.149	0.128	0.103	0.100
	3	t-stat	1.81	2.06	0.86	-0.50	-1.46
		t-stat US_SH_R	3.06	3.84	3.63	3.32	2.87
		pseudo R^2	0.125	0.130	0.124	0.130	0.119
	4	t-stat	2.02	1.01	-0.45	-1.37	-2.73
		t-stat US_SH_R	3.46	3.90	3.62	3.37	2.79
		pseudo R^2	0.111	0.123	0.151	0.150	0.071
	5	t-stat	1.12	-0.39	-1.35	-2.58	-0.13
		t-stat US_SH_R	3.58	3.75	3.88	3.43	2.83
R_CUR_GR		pseudo R^2	0.081	0.121	0.117	0.127	0.101
	0	t-stat	0.05	1.06	0.33	2.22	2.52
		t-stat US_SH_R	3.15	4.00	3.65	3.55	3.08
		pseudo R^2	0.084	0.111	0.147	0.131	0.071
	1	t-stat	1.03	0.34	2.17	2.37	0.26
		t-stat US_SH_R	3.22	3.70	3.83	3.48	2.88
		pseudo R^2	0.083	0.137	0.147	0.100	0.085
	2	t-stat	0.27	1.97	2.21	0.13	1.36
		t-stat US_SH_R	3.08	3.79	3.66	3.31	2.93
		pseudo R^2	0.114	0.150	0.123	0.122	0.075
	3	t-stat	1.96	2.34	0.37	1.69	0.64
		t-stat US_SH_R	3.25	3.81	3.54	3.38	3.00
		pseudo R^2	0.130	0.124	0.151	0.108	0.097
	4	t-stat	2.39	0.58	1.96	0.91	-1.55
		t-stat US_SH_R	3.45	3.67	3.64	3.51	2.55
		pseudo R^2	0.104	0.149	0.128	0.126	0.079
	5	t-stat	0.50	1.97	0.94	-1.56	-0.95
		t-stat US_SH_R	3.36	3.83	3.71	3.08	2.72
STOCK_GR		pseudo R^2	0.132	0.122	0.117	0.103	0.061
	0	t-stat	-2.37	-1.56	-0.51	0.65	0.99
		t-stat US_SH_R	3.18	3.75	3.59	3.46	3.25
		pseudo R^2	0.099	0.113	0.119	0.106	0.079
	1	t-stat	-1.74	-0.85	0.44	0.82	1.34
		t-stat US_SH_R	3.31	3.77	3.61	3.59	2.91
		pseudo R^2	0.087	0.111	0.124	0.107	0.081
	2	t-stat	-1.12	0.29	0.70	1.18	0.87
		t-stat US_SH_R	3.21	3.64	3.75	3.33	2.82
		pseudo R^2	0.083	0.115	0.125	0.106	0.072
	3	t-stat	0.03	0.44	0.77	0.69	-0.29
		t-stat US_SH_R	3.07	3.58	3.53	3.24	2.80
		pseudo R^2	0.090	0.123	0.125	0.102	0.083
	4	t-stat	0.36	0.58	0.60	-0.47	-1.30
		t-stat US_SH_R	3.04	3.65	3.51	3.25	3.04
		pseudo R^2	0.104	0.125	0.123	0.115	0.125
	5	t-stat	0.63	0.63	-0.51	-1.45	-2.46
		t-stat US_SH_R	3.30	3.59	3.61	3.63	3.39

Table A.6 (Continued)

Dependent Variable is R1	lag		h = Quarters Ahead 4	5	6	7	8
EXCH_GR		pseudo R^2	0.182	0.164	0.162	0.156	0.120
	0	t-stat	-0.68	1.13	1.63	1.76	1.50
		t-stat US_SH_R	3.38	3.41	3.51	3.46	3.02
		pseudo R^2	0.182	0.176	0.161	0.133	0.113
	1	t-stat	0.66	1.39	1.48	1.14	1.31
		t-stat US_SH_R	3.51	3.49	3.43	3.36	2.88
		pseudo R^2	0.189	0.173	0.143	0.126	0.117
	2	t-stat	1.05	1.19	0.78	0.92	1.54
		t-stat US_SH_R	3.52	3.46	3.42	3.25	2.75
		pseudo R^2	0.197	0.168	0.145	0.139	0.115
	3	t-stat	1.21	0.80	0.92	1.51	1.43
		t-stat US_SH_R	3.55	3.47	3.40	3.23	2.71
		pseudo R^2	0.190	0.169	0.157	0.137	0.107
	4	t-stat	0.81	0.90	1.45	1.43	1.08
		t-stat US_SH_R	3.53	3.43	3.33	3.20	2.70
		pseudo R^2	0.187	0.176	0.151	0.125	0.130
	5	t-stat	0.72	1.27	1.25	0.91	2.23
		t-stat US_SH_R	3.50	3.33	3.29	3.18	2.65
INF		pseudo R^2	0.265	0.228	0.203	0.195	0.143
	0	t-stat	3.85	3.50	3.70	3.49	2.33
		t-stat US_SH_R	2.69	2.60	2.45	2.22	1.94
		pseudo R^2	0.248	0.225	0.212	0.162	0.141
	1	t-stat	3.29	3.56	3.24	2.05	1.85
		t-stat US_SH_R	2.73	2.61	2.42	2.42	2.02
		pseudo R^2	0.250	0.238	0.183	0.163	0.150
	2	t-stat	3.47	3.35	1.96	1.69	2.01
		t-stat US_SH_R	2.89	2.76	2.71	2.60	2.06
		pseudo R^2	0.259	0.204	0.180	0.169	0.152
	3	t-stat	3.06	1.85	1.57	1.81	2.13
		t-stat US_SH_R	2.98	2.87	2.73	2.55	2.02
		pseudo R^2	0.223	0.201	0.184	0.170	0.156
	4	t-stat	1.76	1.50	1.70	1.94	2.15
		t-stat US_SH_R	2.96	2.85	2.72	2.49	1.94
		pseudo R^2	0.224	0.208	0.188	0.176	0.136
	5	t-stat	1.57	1.68	1.87	2.05	1.95
		t-stat US_SH_R	3.13	2.99	2.81	2.53	2.07
RTT_GR		pseudo R^2	0.194	0.184	0.180	0.126	0.104
	0	t-stat	1.80	1.95	2.86	1.55	1.51
		t-stat US_SH_R	3.63	3.62	3.72	3.41	2.85
		pseudo R^2	0.202	0.200	0.142	0.124	0.102
	1	t-stat	1.73	2.56	1.29	1.20	1.06
		t-stat US_SH_R	3.60	3.67	3.45	3.30	2.75
		pseudo R^2	0.223	0.166	0.142	0.123	0.097
	2	t-stat	2.55	1.28	1.14	1.07	0.19
		t-stat US_SH_R	3.79	3.49	3.41	3.23	2.70
		pseudo R^2	0.191	0.168	0.144	0.119	0.098
	3	t-stat	1.37	1.24	1.21	0.38	-0.23
		t-stat US_SH_R	3.56	3.45	3.37	3.24	2.71
		pseudo R^2	0.193	0.170	0.138	0.119	0.096
	4	t-stat	1.32	1.33	0.60	-0.05	-0.15
		t-stat US_SH_R	3.51	3.39	3.39	3.25	2.70
		pseudo R^2	0.189	0.161	0.136	0.118	0.087
	5	t-stat	1.22	0.46	-0.06	-0.17	0.26
		t-stat US_SH_R	3.44	3.40	3.37	3.23	2.68
DWELL_GR		pseudo R^2	0.182	0.163	0.152	0.132	0.118
	0	t-stat	-0.53	0.65	1.24	1.04	1.13
		t-stat US_SH_R	3.50	3.62	3.60	3.41	2.87
		pseudo R^2	0.183	0.170	0.143	0.132	0.110
	1	t-stat	0.37	0.93	0.71	0.87	0.70
		t-stat US_SH_R	3.54	3.50	3.40	3.30	2.76
		pseudo R^2	0.189	0.164	0.146	0.127	0.098
	2	t-stat	0.71	0.50	0.72	0.58	0.19
		t-stat US_SH_R	3.51	3.44	3.44	3.30	2.69
		pseudo R^2	0.187	0.169	0.144	0.119	0.098
	3	t-stat	0.45	0.67	0.53	0.14	0.08
		t-stat US_SH_R	3.52	3.51	3.53	3.36	2.69
		pseudo R^2	0.188	0.165	0.137	0.119	0.098
	4	t-stat	0.51	0.41	0.02	-0.02	-0.10
		t-stat US_SH_R	3.57	3.63	3.73	3.39	2.69
		pseudo R^2	0.187	0.160	0.136	0.118	0.102
	5	t-stat	0.40	0.00	-0.04	-0.13	0.77
		t-stat US_SH_R	3.67	3.76	3.67	3.39	2.61
NONRES_GR		pseudo R^2	0.180	0.157	0.136	0.120	0.097
	0	t-stat	-0.47	-0.13	-0.70	-1.07	0.56
		t-stat US_SH_R	3.47	3.42	3.42	3.29	2.67
		pseudo R^2	0.181	0.159	0.137	0.120	0.096
	1	t-stat	0.21	-0.44	-0.83	0.74	0.07
		t-stat US_SH_R	3.51	3.44	3.39	3.24	2.71
		pseudo R^2	0.183	0.162	0.138	0.118	0.101
	2	t-stat	-0.54	-0.82	0.65	-0.01	-0.74
		t-stat US_SH_R	3.53	3.48	3.37	3.29	2.78
		pseudo R^2	0.187	0.161	0.137	0.125	0.098
	3	t-stat	-1.05	0.43	-0.24	-0.93	0.00
		t-stat US_SH_R	3.53	3.44	3.49	3.44	2.72
		pseudo R^2	0.184	0.161	0.143	0.119	0.099
	4	t-stat	0.42	-0.23	-0.95	-0.03	0.58
		t-stat US_SH_R	3.52	3.52	3.57	3.26	2.68
		pseudo R^2	0.182	0.167	0.136	0.119	0.098
	5	t-stat	-0.31	-1.09	-0.10	0.52	-0.87
		t-stat US_SH_R	3.63	3.64	3.41	3.20	2.70

Dependent Variable is R2	lag		h = Quarters Ahead 4	5	6	7	8
EXCH_GR		pseudo R^2	0.095	0.110	0.116	0.101	0.079
	0	t-stat	-1.01	-0.29	-0.17	0.61	1.44
		t-stat US_SH_R	2.95	3.73	3.62	3.42	3.02
		pseudo R^2	0.086	0.112	0.117	0.103	0.128
	1	t-stat	-0.69	-0.69	0.05	1.04	2.62
		t-stat US_SH_R	3.05	3.65	3.63	3.40	3.12
		pseudo R^2	0.088	0.112	0.120	0.144	0.150
	2	t-stat	-1.14	-0.63	0.51	2.36	3.23
		t-stat US_SH_R	3.11	3.65	3.62	3.34	3.08
		pseudo R^2	0.084	0.115	0.166	0.183	0.083
	3	t-stat	-0.55	0.54	2.28	3.22	1.42
		t-stat US_SH_R	3.06	3.73	3.61	3.51	2.88
		pseudo R^2	0.090	0.167	0.211	0.114	0.072
	4	t-stat	0.56	2.24	3.27	1.37	-0.34
		t-stat US_SH_R	3.22	3.73	3.74	3.29	2.86
		pseudo R^2	0.143	0.204	0.132	0.102	0.081
	5	t-stat	2.20	3.21	1.21	-0.48	-1.29
		t-stat US_SH_R	3.26	3.67	3.51	3.32	2.91
INF		pseudo R^2	0.102	0.109	0.117	0.100	0.072
	0	t-stat	2.20	0.10	0.31	-0.13	0.29
		t-stat US_SH_R	2.48	3.48	3.09	3.13	2.63
		pseudo R^2	0.082	0.111	0.118	0.099	0.106
	1	t-stat	0.34	0.32	-0.25	0.03	3.18
		t-stat US_SH_R	2.78	3.12	3.35	3.08	2.12
		pseudo R^2	0.085	0.111	0.119	0.133	0.092
	2	t-stat	0.49	-0.18	-0.05	3.04	2.19
		t-stat US_SH_R	2.60	3.38	3.38	2.71	2.31
		pseudo R^2	0.083	0.114	0.152	0.118	0.092
	3	t-stat	0.00	-0.08	2.66	1.96	2.11
		t-stat US_SH_R	2.79	3.50	3.08	2.80	2.34
		pseudo R^2	0.089	0.151	0.136	0.118	0.134
	4	t-stat	0.05	2.49	1.70	1.77	3.21
		t-stat US_SH_R	2.96	3.17	3.07	2.81	1.99
		pseudo R^2	0.136	0.137	0.138	0.160	0.102
	5	t-stat	2.62	1.72	1.60	2.95	2.28
		t-stat US_SH_R	2.77	3.21	3.14	2.58	2.24
RTT_GR		pseudo R^2	0.082	0.122	0.173	0.132	0.089
	0	t-stat	-0.49	1.71	3.61	2.82	1.94
		t-stat US_SH_R	3.03	3.73	3.90	3.57	2.89
		pseudo R^2	0.090	0.159	0.146	0.115	0.070
	1	t-stat	1.40	3.31	2.73	1.75	-0.14
		t-stat US_SH_R	3.04	3.76	3.78	3.31	2.88
		pseudo R^2	0.130	0.139	0.134	0.100	0.071
	2	t-stat	3.08	2.55	1.73	-0.21	-0.17
		t-stat US_SH_R	3.11	3.78	3.61	3.31	2.88
		pseudo R^2	0.113	0.133	0.122	0.101	0.071
	3	t-stat	2.47	1.88	-0.03	0.01	0.19
		t-stat US_SH_R	3.26	3.78	3.60	3.29	2.87
		pseudo R^2	0.110	0.122	0.122	0.101	0.080
	4	t-stat	2.00	0.16	0.19	0.35	-1.22
		t-stat US_SH_R	3.39	3.75	3.56	3.28	2.81
		pseudo R^2	0.102	0.121	0.122	0.109	0.102
	5	t-stat	0.07	0.08	0.25	-1.19	-2.34
		t-stat US_SH_R	3.39	3.74	3.55	3.25	2.78
DWELL_GR		pseudo R^2	0.095	0.115	0.119	0.122	0.133
	0	t-stat	-1.32	-0.88	-0.66	1.86	2.89
		t-stat US_SH_R	3.06	3.64	3.56	3.57	3.12
		pseudo R^2	0.093	0.117	0.133	0.153	0.174
	1	t-stat	-1.22	-0.95	1.64	2.89	3.17
		t-stat US_SH_R	3.17	3.60	3.73	3.56	3.09
		pseudo R^2	0.093	0.122	0.165	0.197	0.129
	2	t-stat	-1.15	1.41	2.84	3.28	1.94
		t-stat US_SH_R	3.13	3.63	3.75	3.44	2.85
		pseudo R^2	0.094	0.160	0.229	0.155	0.071
	3	t-stat	1.40	2.93	3.37	1.87	-0.08
		t-stat US_SH_R	3.00	3.80	3.79	3.29	2.89
		pseudo R^2	0.134	0.217	0.173	0.102	0.073
	4	t-stat	2.93	3.53	1.80	-0.22	-0.34
		t-stat US_SH_R	3.13	3.63	3.45	3.42	2.81
		pseudo R^2	0.208	0.179	0.123	0.102	0.076
	5	t-stat	3.73	1.86	-0.23	-0.41	-0.82
		t-stat US_SH_R	3.37	3.47	3.84	3.30	2.82
NONRES_GR		pseudo R^2	0.083	0.111	0.123	0.100	0.072
	0	t-stat	0.71	0.85	1.06	-0.16	0.19
		t-stat US_SH_R	3.01	3.63	3.53	3.28	2.90
		pseudo R^2	0.085	0.119	0.117	0.100	0.070
	1	t-stat	1.07	1.27	0.08	0.27	-0.11
		t-stat US_SH_R	3.05	3.66	3.59	3.32	2.87
		pseudo R^2	0.091	0.111	0.119	0.100	0.073
	2	t-stat	1.23	0.04	0.24	-0.19	0.51
		t-stat US_SH_R	3.09	3.64	3.61	3.30	2.85
		pseudo R^2	0.083	0.114	0.123	0.102	0.075
	3	t-stat	-0.02	0.20	-0.44	0.33	-1.24
		t-stat US_SH_R	3.03	3.70	3.61	3.29	3.01
		pseudo R^2	0.089	0.122	0.123	0.105	0.071
	4	t-stat	0.33	-0.31	0.36	-1.30	0.15
		t-stat US_SH_R	3.19	3.79	3.60	3.44	2.83
		pseudo R^2	0.103	0.122	0.128	0.100	0.072
	5	t-stat	-0.34	0.29	-1.57	0.07	-0.66
		t-stat US_SH_R	3.43	3.77	3.84	3.25	2.87

Table A.6 (Continued)

Dependent Variable is R1	lag		h = Quarters Ahead 4	5	6	7	8
SV_EX_GR	0	pseudo R^2	0.188	0.185	0.175	0.120	0.096
		t-stat	-1.45	-3.24	-3.41	-0.97	-0.32
		t-stat US_SH_R	3.60	3.56	3.59	3.26	2.71
	1	pseudo R^2	0.201	0.193	0.137	0.118	0.097
		t-stat	-2.56	-3.10	-0.48	0.00	0.32
		t-stat US_SH_R	3.51	3.50	3.38	3.24	2.71
	2	pseudo R^2	0.220	0.161	0.136	0.119	0.099
		t-stat	-3.36	-0.42	0.03	0.31	0.58
		t-stat US_SH_R	3.59	3.44	3.39	3.24	2.71
	3	pseudo R^2	0.187	0.161	0.137	0.120	0.102
		t-stat	-0.73	-0.22	0.01	0.31	0.75
		t-stat US_SH_R	3.59	3.43	3.37	3.21	2.68
	4	pseudo R^2	0.183	0.160	0.138	0.124	0.138
		t-stat	-0.11	0.06	0.43	0.86	2.72
		t-stat US_SH_R	3.52	3.42	3.36	3.24	2.79
	5	pseudo R^2	0.182	0.161	0.141	0.159	0.098
		t-stat	0.03	0.37	0.84	2.63	0.61
		t-stat US_SH_R	3.49	3.39	3.33	3.19	2.68
TOT_GR	0	pseudo R^2	0.180	0.183	0.143	0.144	0.116
		t-stat	-0.64	1.86	0.85	1.19	0.84
		t-stat US_SH_R	3.51	3.35	3.33	3.26	2.85
	1	pseudo R^2	0.198	0.163	0.154	0.131	0.151
		t-stat	1.59	0.64	1.03	0.68	1.24
		t-stat US_SH_R	3.42	3.45	3.31	3.29	2.79
	2	pseudo R^2	0.185	0.175	0.146	0.165	0.099
		t-stat	0.51	0.96	0.58	1.17	0.36
		t-stat US_SH_R	3.58	3.38	3.44	3.18	2.72
	3	pseudo R^2	0.197	0.168	0.182	0.120	0.101
		t-stat	0.97	0.55	1.17	0.34	0.50
		t-stat US_SH_R	3.49	3.54	3.28	3.32	2.71
	4	pseudo R^2	0.191	0.208	0.139	0.122	0.099
		t-stat	0.59	1.20	0.39	0.55	0.38
		t-stat US_SH_R	3.61	3.34	3.46	3.27	2.69
	5	pseudo R^2	0.218	0.160	0.139	0.118	0.098
		t-stat	1.16	0.26	0.43	0.23	0.45
		t-stat US_SH_R	3.44	3.50	3.43	3.22	2.64
GOLD_GR	0	pseudo R^2	0.204	0.165	0.136	0.117	0.099
		t-stat	-1.24	-0.58	-0.28	-0.20	0.57
		t-stat US_SH_R	3.75	3.55	3.34	3.19	2.96
	1	pseudo R^2	0.196	0.164	0.137	0.119	0.134
		t-stat	-0.87	-0.51	-0.47	0.39	1.77
		t-stat US_SH_R	4.03	3.59	3.38	3.38	3.29
	2	pseudo R^2	0.189	0.162	0.137	0.153	0.173
		t-stat	-0.62	-0.51	0.31	1.66	2.82
		t-stat US_SH_R	3.87	3.53	3.39	3.51	3.36
	3	pseudo R^2	0.190	0.160	0.163	0.183	0.165
		t-stat	-0.78	0.02	1.50	2.67	2.98
		t-stat US_SH_R	3.79	3.48	3.24	3.40	2.95
	4	pseudo R^2	0.184	0.178	0.169	0.176	0.185
		t-stat	-0.34	1.34	2.51	2.65	3.16
		t-stat US_SH_R	3.66	3.25	3.10	3.22	2.38
	5	pseudo R^2	0.191	0.200	0.180	0.192	0.167
		t-stat	1.05	2.19	2.49	2.96	4.51
		t-stat US_SH_R	3.32	3.02	3.01	2.71	2.01
US_SPR	0	pseudo R^2	0.195	0.174	0.166	0.139	0.115
		t-stat	-0.96	-0.92	-1.03	-1.26	-1.38
		t-stat US_SH_R	3.24	3.09	2.68	2.00	1.56
	1	pseudo R^2	0.196	0.177	0.157	0.136	0.103
		t-stat	-0.94	-1.02	-1.37	-1.46	-0.73
		t-stat US_SH_R	3.44	2.97	2.80	2.16	1.90
	2	pseudo R^2	0.201	0.182	0.166	0.125	0.099
		t-stat	-1.11	-1.43	-1.59	-0.75	-0.32
		t-stat US_SH_R	3.27	2.83	2.61	2.40	2.12
	3	pseudo R^2	0.201	0.176	0.142	0.120	0.106
		t-stat	-1.36	-1.47	-0.68	-0.24	-0.56
		t-stat US_SH_R	3.06	2.71	2.61	2.41	1.87
	4	pseudo R^2	0.197	0.165	0.138	0.127	0.121
		t-stat	-1.38	-0.64	-0.25	-0.64	-0.86
		t-stat US_SH_R	2.96	2.70	2.46	2.21	1.70
	5	pseudo R^2	0.187	0.161	0.146	0.142	0.114
		t-stat	-0.88	-0.28	-0.57	-0.84	-0.76
		t-stat US_SH_R	2.92	2.66	2.36	2.06	1.84
US_LG_R	0	pseudo R^2	0.195	0.174	0.165	0.139	0.115
		t-stat	-0.96	-0.82	-1.03	-1.26	-1.38
		t-stat US_SH_R	1.92	1.83	1.94	2.55	3.22
	1	pseudo R^2	0.193	0.176	0.153	0.138	0.120
		t-stat	-0.87	-1.03	-0.96	-1.23	-1.56
		t-stat US_SH_R	2.08	2.13	2.09	2.98	3.96
	2	pseudo R^2	0.193	0.173	0.151	0.140	0.128
		t-stat	-0.92	-0.82	-0.90	-1.33	-1.78
		t-stat US_SH_R	2.30	2.28	2.38	3.61	4.49
	3	pseudo R^2	0.195	0.175	0.157	0.153	0.141
		t-stat	-0.98	-1.00	-1.14	-1.70	-2.06
		t-stat US_SH_R	2.44	2.47	2.72	4.09	4.54
	4	pseudo R^2	0.194	0.178	0.165	0.161	0.146
		t-stat	-0.97	-1.16	-1.42	-1.96	-2.39
		t-stat US_SH_R	2.63	2.78	3.11	4.49	4.68
	5	pseudo R^2	0.194	0.183	0.170	0.182	0.185
		t-stat	-1.05	-1.38	-1.63	-2.19	-2.94
		t-stat US_SH_R	2.83	3.01	3.39	4.72	5.47
US_STK_GR	0	pseudo R^2	0.179	0.157	0.137	0.118	0.098
		t-stat	-0.13	0.08	0.75	-0.54	-0.14
		t-stat US_SH_R	3.51	3.44	3.41	3.20	2.72
	1	pseudo R^2	0.181	0.163	0.137	0.118	0.103
		t-stat	0.31	0.92	-0.44	-0.03	-1.63
		t-stat US_SH_R	3.56	3.49	3.34	3.27	2.66
	2	pseudo R^2	0.186	0.161	0.138	0.125	0.107
		t-stat	0.97	-0.48	-0.03	-1.42	-1.52
		t-stat US_SH_R	3.52	3.42	3.41	3.19	2.72
	3	pseudo R^2	0.183	0.161	0.140	0.126	0.102
		t-stat	-0.17	0.25	-1.17	-1.25	-0.95
		t-stat US_SH_R	3.50	3.47	3.33	3.21	2.69
	4	pseudo R^2	0.182	0.157	0.148	0.128	0.104
		t-stat	0.02	-1.42	-1.42	-1.13	-0.96
		t-stat US_SH_R	3.52	3.40	3.37	3.27	2.72
	5	pseudo R^2	0.187	0.158	0.142	0.123	0.108
		t-stat	-1.24	-1.36	-1.12	-0.86	-0.93
		t-stat US_SH_R	3.45	3.41	3.37	3.25	2.88

Dependent Variable is R2	lag		h = Quarters Ahead 4	5	6	7	8
SV_EX_GR	0	pseudo R^2	0.088	0.109	0.135	0.100	0.075
		t-stat	0.79	0.00	-1.81	-0.28	-1.01
		t-stat US_SH_R	2.97	3.66	3.77	3.29	2.87
	1	pseudo R^2	0.083	0.123	0.117	0.101	0.077
		t-stat	0.36	-1.36	0.09	-0.76	-1.08
		t-stat US_SH_R	3.08	3.67	3.63	3.33	2.97
	2	pseudo R^2	0.095	0.111	0.122	0.108	0.071
		t-stat	-1.35	0.02	-0.89	-1.22	0.26
		t-stat US_SH_R	2.99	3.68	3.63	3.49	2.88
	3	pseudo R^2	0.083	0.120	0.135	0.101	0.072
		t-stat	-0.28	-1.20	-1.61	0.06	0.26
		t-stat US_SH_R	3.07	3.81	3.94	3.30	2.86
	4	pseudo R^2	0.093	0.132	0.122	0.102	0.076
		t-stat	-0.95	-1.31	0.15	0.35	-0.38
		t-stat US_SH_R	3.33	4.05	3.59	3.27	2.97
	5	pseudo R^2	0.113	0.121	0.122	0.105	0.072
		t-stat	-1.28	0.10	0.34	-0.37	-0.61
		t-stat US_SH_R	3.49	3.75	3.55	3.36	2.88
TOT_GR	0	pseudo R^2	0.082	0.109	0.133	0.136	0.072
		t-stat	-0.30	0.11	2.18	1.98	-0.15
		t-stat US_SH_R	3.07	3.69	3.76	3.56	2.90
	1	pseudo R^2	0.082	0.121	0.144	0.100	0.070
		t-stat	-0.17	1.61	1.67	-0.34	0.05
		t-stat US_SH_R	3.07	3.71	3.84	3.44	3.02
	2	pseudo R^2	0.091	0.136	0.120	0.100	0.076
		t-stat	1.35	1.61	-0.47	-0.03	0.68
		t-stat US_SH_R	3.03	3.59	3.84	3.51	2.90
	3	pseudo R^2	0.107	0.115	0.122	0.105	0.073
		t-stat	1.61	-0.40	0.00	0.64	0.56
		t-stat US_SH_R	3.02	3.87	3.82	3.34	2.87
	4	pseudo R^2	0.089	0.122	0.127	0.102	0.071
		t-stat	-0.28	0.08	0.66	0.58	-0.10
		t-stat US_SH_R	3.26	3.88	3.63	3.31	2.86
	5	pseudo R^2	0.103	0.125	0.122	0.101	0.070
		t-stat	0.07	0.59	0.40	-0.27	0.05
		t-stat US_SH_R	3.42	3.76	3.63	3.36	2.84
GOLD_GR	0	pseudo R^2	0.084	0.118	0.118	0.103	0.089
		t-stat	0.59	-1.23	-0.05	-0.44	-1.39
		t-stat US_SH_R	3.08	3.57	3.51	3.25	2.52
	1	pseudo R^2	0.096	0.111	0.123	0.122	0.070
		t-stat	-1.58	-0.28	-0.65	-1.39	0.15
		t-stat US_SH_R	3.01	3.71	3.61	3.00	2.93
	2	pseudo R^2	0.084	0.118	0.148	0.100	0.079
		t-stat	-0.37	-0.75	-1.42	0.11	0.96
		t-stat US_SH_R	3.12	3.77	3.65	3.32	2.93
	3	pseudo R^2	0.095	0.152	0.122	0.105	0.079
		t-stat	-0.69	-1.69	-0.16	0.65	1.25
		t-stat US_SH_R	3.41	3.95	3.61	3.20	2.74
	4	pseudo R^2	0.141	0.125	0.123	0.104	0.100
		t-stat	-2.03	-0.46	0.30	0.84	1.70
		t-stat US_SH_R	4.15	3.89	3.40	3.03	2.22
	5	pseudo R^2	0.106	0.121	0.122	0.120	0.124
		t-stat	-0.64	0.03	0.39	1.58	2.22
		t-stat US_SH_R	3.99	3.58	3.17	2.66	1.79
US_SPR	0	pseudo R^2	0.082	0.109	0.118	0.101	0.072
		t-stat	0.28	-0.21	-0.41	-0.37	0.21
		t-stat US_SH_R	2.95	3.24	2.93	2.50	2.39
	1	pseudo R^2	0.084	0.113	0.118	0.100	0.080
		t-stat	-0.48	-0.51	-0.32	0.36	1.16
		t-stat US_SH_R	2.50	3.01	3.09	3.00	2.87
	2	pseudo R^2	0.090	0.113	0.120	0.112	0.085
		t-stat	-0.81	-0.49	0.38	1.22	1.59
		t-stat US_SH_R	2.44	3.09	3.67	3.42	2.96
	3	pseudo R^2	0.088	0.114	0.136	0.123	0.081
		t-stat	-0.68	0.29	1.37	1.78	0.59
		t-stat US_SH_R	2.44	3.73	4.07	3.43	2.76
	4	pseudo R^2	0.089	0.137	0.148	0.115	0.081
		t-stat	0.02	1.32	1.90	1.31	0.80
		t-stat US_SH_R	3.05	4.41	3.76	3.23	2.83
	5	pseudo R^2	0.113	0.147	0.138	0.114	0.086
		t-stat	1.03	1.88	1.38	0.96	0.91
		t-stat US_SH_R	3.80	3.99	3.34	2.86	2.81
US_LG_R	0	pseudo R^2	0.082	0.109	0.118	0.101	0.072
		t-stat	0.25	-0.21	-0.41	-0.37	0.21
		t-stat US_SH_R	1.01	1.64	1.80	1.76	1.17
	1	pseudo R^2	0.089	0.110	0.118	0.100	0.070
		t-stat	0.84	0.16	-0.38	-0.36	-0.18
		t-stat US_SH_R	0.84	1.71	2.14	2.25	2.15
	2	pseudo R^2	0.093	0.111	0.120	0.103	0.073
		t-stat	0.97	0.14	-0.35	-0.62	-0.57
		t-stat US_SH_R	1.03	2.06	2.48	2.87	2.58
	3	pseudo R^2	0.090	0.114	0.126	0.109	0.075
		t-stat	0.79	0.02	-0.67	-1.08	-0.77
		t-stat US_SH_R	1.34	2.40	2.84	3.24	2.79
	4	pseudo R^2	0.093	0.123	0.130	0.108	0.077
		t-stat	0.63	-0.30	-0.98	-1.20	-1.06
		t-stat US_SH_R	1.80	2.84	3.17	3.49	3.10
	5	pseudo R^2	0.103	0.124	0.128	0.109	0.074
		t-stat	0.26	-0.54	-1.00	-1.33	-0.80
		t-stat US_SH_R	2.36	3.07	3.34	3.80	3.22
US_STK_GR	0	pseudo R^2	0.085	0.113	0.130	0.114	0.072
		t-stat	-1.05	-0.60	1.79	2.82	0.29
		t-stat US_SH_R	3.13	3.70	3.82	3.37	2.87
	1	pseudo R^2	0.088	0.127	0.134	0.109	0.085
		t-stat	-0.61	1.88	2.80	0.35	-1.09
		t-stat US_SH_R	3.09	3.55	3.66	3.32	2.83
	2	pseudo R^2	0.097	0.128	0.119	0.118	0.094
		t-stat	1.90	2.61	0.32	-1.78	-2.31
		t-stat US_SH_R	3.00	3.63	3.41	3.26	2.92
	3	pseudo R^2	0.101	0.116	0.133	0.129	0.098
		t-stat	2.80	0.51	-1.54	-2.08	-2.62
		t-stat US_SH_R	3.22	3.74	3.82	3.25	2.85
	4	pseudo R^2	0.089	0.136	0.146	0.128	0.090
		t-stat	0.32	-1.81	-2.41	-2.09	-2.35
		t-stat US_SH_R	3.23	3.68	3.66	3.41	2.91
	5	pseudo R^2	0.116	0.142	0.148	0.117	0.079
		t-stat	-1.70	-2.37	-2.96	-2.22	-1.03
		t-stat US_SH_R	3.17	3.66	3.86	3.36	2.83

Table A.7: Out-of-Sample *pseudo* R^2: RHS are Constant, INF, and the Variable as Stated

Dep Var = R1	lag	h = Quarters Ahead			Dep Var = R2	lag	h = Quarters Ahead		
		1	2	3			1	2	3
SPREAD	0	0.156	0.232	0.246	SPREAD	0	0.128	0.137	0.146
	1	0.216	0.270	0.227		1	0.132	0.129	0.137
	2	0.244	0.252	0.201		2	0.137	0.134	0.159
	3	0.230	0.220	0.042		3	0.145	0.162	0.156
	4	0.241	0.138	0.088		4	0.187	0.176	0.128
	5	0.144	0.057	0.160		5	0.189	0.134	0.104
SH_R	0	0.200	0.271	0.259	SH_R	0	0.157	0.160	0.149
	1	0.258	0.269	0.231		1	0.165	0.138	0.129
	2	0.250	0.220	0.205		2	0.161	0.140	0.152
	3	0.246	0.244	0.134		3	0.160	0.163	0.153
	4	0.251	0.207	0.139		4	0.183	0.170	0.122
	5	0.217	0.189	0.162		5	0.185	0.133	0.108
LG_R	0	0.217	0.242	0.201	LG_R	0	0.162	0.155	0.128
	1	0.218	0.193	0.181		1	0.170	0.141	0.104
	2	0.196	0.167	0.172		2	0.157	0.122	0.098
	3	0.197	0.191	0.165		3	0.145	0.111	0.103
	4	0.197	0.181	0.157		4	0.135	0.117	0.108
	5	0.186	0.171	0.155		5	0.134	0.117	0.117
D_SH_R	0	-0.319	0.036	0.103	D_SH_R	0	0.132	0.136	0.130
	1	0.064	0.132	0.128		1	0.146	0.127	0.102
	2	0.084	0.149	0.145		2	0.148	0.112	0.115
	3	0.134	0.182	0.105		3	0.125	0.117	0.113
	4	0.158	0.171	0.102		4	0.131	0.118	0.102
	5	0.158	0.166	0.134		5	0.127	0.105	0.096
D_LG_R	0	-0.024	0.025	0.061	D_LG_R	0	0.141	0.121	0.110
	1	-0.008	0.081	0.106		1	0.132	0.114	0.095
	2	0.047	0.139	0.139		2	0.127	0.103	0.060
	3	0.112	0.180	0.149		3	0.118	0.063	0.072
	4	0.152	0.188	0.144		4	0.085	0.071	0.092
	5	0.166	0.192	0.144		5	0.090	0.093	0.103
M3_GR	0	0.146	0.179	0.175	M3_GR	0	0.125	0.122	0.133
	1	0.144	0.198	0.226		1	0.129	0.125	0.134
	2	0.168	0.240	0.218		2	0.141	0.145	0.111
	3	0.227	0.250	0.230		3	0.146	0.084	0.116
	4	0.229	0.268	0.232		4	0.121	0.135	0.151
	5	0.234	0.256	0.205		5	0.139	0.147	0.146
CUR_GR	0	0.109	0.155	0.106	CUR_GR	0	0.120	0.104	0.120
	1	0.111	0.166	0.134		1	0.099	0.115	0.118
	2	0.095	0.166	0.100		2	0.121	0.118	0.118
	3	0.123	0.148	0.131		3	0.122	0.111	0.070
	4	0.107	0.168	0.144		4	0.124	0.077	0.111
	5	0.126	0.175	0.202		5	0.105	0.114	0.125
R_M3_GR	0	0.146	0.179	0.175	R_M3_GR	0	0.125	0.122	0.133
	1	0.141	0.173	0.164		1	0.138	0.122	0.124
	2	0.149	0.200	0.179		2	0.129	0.128	0.134
	3	0.170	0.219	0.181		3	0.133	0.123	0.136
	4	0.181	0.231	0.183		4	0.142	0.157	0.182
	5	0.177	0.207	0.169		5	0.136	0.154	0.144
R_CUR_GR	0	0.109	0.155	0.106	R_CUR_GR	0	0.120	0.104	0.120
	1	0.132	0.183	0.183		1	0.133	0.121	0.109
	2	0.171	0.215	0.144		2	0.129	0.100	0.108
	3	0.189	0.189	0.152		3	0.112	0.103	0.114
	4	0.154	0.182	0.149		4	0.109	0.115	0.160
	5	0.169	0.197	0.153		5	0.113	0.137	0.139
STOCK_GR	0	0.140	0.208	0.171	STOCK_GR	0	0.125	0.122	0.132
	1	0.157	0.193	0.156		1	0.120	0.124	0.069
	2	0.169	0.193	0.149		2	0.137	0.072	0.117
	3	0.160	0.131	0.095		3	0.077	0.117	0.114
	4	0.152	0.167	0.176		4	0.118	0.120	0.113
	5	0.143	0.204	0.136		5	0.125	0.112	0.099
EXCH_GR	0	0.116	0.161	0.123	EXCH_GR	0	0.118	0.079	0.115
	1	0.100	0.095	0.130		1	0.095	0.103	0.127
	2	0.097	0.153	0.155		2	0.111	0.128	0.111
	3	0.123	0.185	0.128		3	0.142	0.111	0.121
	4	0.140	0.184	0.163		4	0.135	0.131	0.122
	5	0.122	0.187	0.166		5	0.131	0.117	0.117

Table A.7 (Continued)

Dep Var = R1	lag	h = Quarters Ahead			Dep Var = R2	lag	h = Quarters Ahead		
		1	2	3			1	2	3
RTT_GR	0	0.131	0.161	0.144	RTT_GR	0	0.108	0.132	0.101
	1	0.225	0.154	0.140		1	0.177	0.024	0.082
	2	0.139	0.140	0.145		2	0.081	0.102	0.137
	3	0.148	0.194	0.171		3	0.099	0.124	0.119
	4	0.148	0.184	0.162		4	0.132	0.125	0.136
	5	0.154	0.183	0.156		5	0.113	0.124	0.128
DWELL_GR	0	0.105	0.051	0.145	DWELL_GR	0	0.122	0.119	0.126
	1	0.143	0.188	0.128		1	0.132	0.132	0.129
	2	0.171	0.152	0.076		2	0.145	0.132	0.126
	3	0.159	0.146	0.136		3	0.150	0.131	0.127
	4	0.110	0.097	0.154		4	0.147	0.134	0.136
	5	0.105	0.163	0.193		5	0.132	0.142	0.168
NONRES_GR	0	0.141	0.170	0.134	NONRES_GR	0	0.113	0.094	0.117
	1	0.166	0.165	0.125		1	0.121	0.119	0.121
	2	0.153	0.181	0.142		2	0.126	0.118	0.121
	3	0.124	0.190	0.150		3	0.137	0.134	0.120
	4	0.149	0.184	0.153		4	0.127	0.097	0.120
	5	0.151	0.181	0.156		5	0.115	0.120	0.117
SV_EX_GR	0	0.138	0.173	0.126	SV_EX_GR	0	0.125	0.111	0.106
	1	0.145	0.176	0.130		1	0.118	0.100	0.111
	2	0.131	0.181	0.147		2	0.116	0.109	0.110
	3	0.144	0.161	0.135		3	0.116	0.111	0.137
	4	0.144	0.183	0.153		4	0.110	0.126	0.123
	5	0.148	0.179	0.158		5	0.153	0.128	0.119
TOT_GR	0	0.161	0.182	0.147	TOT_GR	0	0.148	0.127	0.109
	1	0.170	0.185	0.151		1	0.145	0.115	0.124
	2	0.152	0.178	0.158		2	0.134	0.129	0.125
	3	0.151	0.194	0.163		3	0.133	0.124	0.120
	4	0.148	0.196	0.177		4	0.130	0.118	0.145
	5	0.159	0.205	0.163		5	0.119	0.136	0.105
GOLD_GR	0	0.141	0.189	0.193	GOLD_GR	0	0.135	0.119	0.107
	1	0.155	0.231	0.217		1	0.126	0.099	0.115
	2	0.197	0.247	0.192		2	0.110	0.115	0.151
	3	0.193	0.209	0.178		3	0.122	0.143	0.127
	4	0.181	0.206	0.156		4	0.141	0.119	0.145
	5	0.191	0.127	0.145		5	0.144	0.170	0.171
US_SPR	0	0.121	0.145	0.177	US_SPR	0	0.114	0.112	0.116
	1	0.114	0.186	0.235		1	0.124	0.116	0.118
	2	0.151	0.217	0.196		2	0.125	0.115	0.123
	3	0.231	0.226	0.216		3	0.122	0.126	0.139
	4	0.231	0.221	0.220		4	0.131	0.142	0.137
	5	0.220	0.209	0.201		5	0.143	0.136	0.114
US_SH_R	0	0.084	0.115	0.205	US_SH_R	0	0.137	0.141	0.145
	1	0.112	0.175	0.254		1	0.147	0.139	0.156
	2	0.171	0.216	0.212		2	0.140	0.150	0.168
	3	0.259	0.237	0.206		3	0.157	0.174	0.187
	4	0.247	0.224	0.187		4	0.177	0.192	0.171
	5	0.221	0.176	0.165		5	0.188	0.168	0.141
US_LG_R	0	0.131	0.141	0.157	US_LG_R	0	0.141	0.135	0.106
	1	0.146	0.161	0.173		1	0.140	0.109	0.123
	2	0.154	0.173	0.156		2	0.111	0.123	0.141
	3	0.180	0.170	0.147		3	0.130	0.144	0.148
	4	0.172	0.161	0.118		4	0.148	0.149	0.138
	5	0.161	0.109	0.111		5	0.152	0.136	0.128
US_STK_GR	0	0.120	0.131	0.134	US_STK_GR	0	0.141	0.050	0.115
	1	0.029	0.117	0.139		1	0.096	0.119	0.118
	2	0.114	0.162	0.141		2	0.132	0.113	0.112
	3	0.133	0.164	0.159		3	0.122	0.113	0.143
	4	0.148	0.188	0.118		4	0.123	0.109	0.108
	5	0.181	0.166	0.156		5	0.143	0.130	0.117

Table A.8: Out-of-Sample *pseudo* R^2: RHS are Constant, US_SH_R, and the Variable as Stated

Dep Var = R1	lag	h = Quarters Ahead 4	5	6	7	8	Dep Var = R2	lag	h = Quarters Ahead 4	5	6	7	8
SPREAD	0	0.289	0.245	0.085	-0.022	0.022	SPREAD	0	0.113	0.156	0.154	0.098	0.040
	1	0.284	0.103	0.013	0.082	0.088		1	0.139	0.152	0.117	0.081	0.039
	2	0.170	0.060	0.133	0.134	0.084		2	0.139	0.107	0.104	0.080	0.055
	3	0.127	0.171	0.173	0.126	0.061		3	0.096	0.089	0.107	0.092	0.067
	4	0.208	0.206	0.161	0.104	0.036		4	0.074	0.087	0.125	0.100	0.066
	5	0.241	0.193	0.143	0.084	0.023		5	0.073	0.110	0.133	0.095	0.046
SH_R	0	0.202	0.199	0.139	0.070	0.036	SH_R	0	0.098	0.117	0.124	0.081	0.035
	1	0.234	0.169	0.110	0.090	0.076		1	0.107	0.113	0.098	0.081	0.043
	2	0.213	0.146	0.131	0.120	0.080		2	0.105	0.083	0.103	0.083	0.057
	3	0.188	0.164	0.154	0.120	0.073		3	0.074	0.086	0.112	0.092	0.064
	4	0.199	0.184	0.154	0.116	0.066		4	0.070	0.095	0.125	0.097	0.063
	5	0.219	0.185	0.152	0.111	0.066		5	0.077	0.108	0.130	0.093	0.050
LG_R	0	0.164	0.159	0.143	0.119	0.083	LG_R	0	0.080	0.094	0.115	0.094	0.054
	1	0.195	0.174	0.154	0.126	0.076		1	0.071	0.090	0.115	0.092	0.055
	2	0.210	0.184	0.160	0.123	0.075		2	0.068	0.092	0.119	0.089	0.057
	3	0.219	0.190	0.157	0.120	0.078		3	0.067	0.098	0.122	0.092	0.057
	4	0.222	0.185	0.154	0.123	0.070		4	0.077	0.105	0.124	0.091	0.056
	5	0.216	0.182	0.155	0.113	0.056		5	0.089	0.108	0.122	0.085	0.048
D_SH_R	0	0.202	0.199	0.139	0.070	0.036	D_SH_R	0	0.098	0.117	0.124	0.081	0.035
	1	0.246	0.174	0.098	0.074	0.076		1	0.096	0.109	0.104	0.060	0.041
	2	0.211	0.126	0.097	0.119	0.077		2	0.087	0.090	0.094	0.066	0.048
	3	0.163	0.130	0.156	0.123	0.069		3	0.067	0.080	0.104	0.084	0.057
	4	0.167	0.189	0.159	0.116	0.068		4	0.057	0.092	0.121	0.093	0.052
	5	0.223	0.190	0.153	0.112	0.074		5	0.075	0.107	0.125	0.081	0.043
D_LG_R	0	0.159	0.176	0.152	0.116	0.070	D_LG_R	0	0.083	0.080	0.095	0.085	0.056
	1	0.221	0.185	0.151	0.110	0.060		1	0.056	0.082	0.108	0.086	0.059
	2	0.224	0.182	0.143	0.100	0.058		2	0.060	0.094	0.117	0.090	0.057
	3	0.222	0.175	0.135	0.099	0.064		3	0.071	0.100	0.125	0.091	0.057
	4	0.211	0.165	0.135	0.107	0.035		4	0.081	0.111	0.125	0.090	0.058
	5	0.199	0.164	0.140	0.077	0.024		5	0.096	0.113	0.125	0.088	0.064
M3_GR	0	0.373	0.333	0.298	0.244	0.170	M3_GR	0	0.137	0.117	0.161	0.154	0.122
	1	0.363	0.322	0.272	0.203	0.124		1	0.092	0.137	0.173	0.150	0.110
	2	0.350	0.296	0.234	0.161	0.083		2	0.117	0.151	0.175	0.136	0.072
	3	0.327	0.261	0.195	0.126	0.051		3	0.133	0.159	0.172	0.104	0.038
	4	0.292	0.223	0.162	0.106	0.070		4	0.144	0.161	0.136	0.085	0.063
	5	0.259	0.194	0.147	0.121	0.084		5	0.148	0.124	0.114	0.097	0.075
CUR_GR	0	0.175	0.149	0.158	0.115	0.198	CUR_GR	0	0.109	0.130	0.116	0.138	0.125
	1	0.173	0.187	0.182	0.229	0.205		1	0.112	0.098	0.156	0.156	0.135
	2	0.209	0.210	0.261	0.235	**0.209**		2	0.087	0.138	0.175	0.159	**0.167**
	3	0.252	0.304	0.283	**0.259**	0.195		3	0.137	0.171	0.201	**0.204**	0.121
	4	0.350	0.320	**0.304**	0.244	0.161		4	0.166	0.193	**0.250**	0.157	0.070
	5	0.357	**0.335**	0.281	0.204	0.165		5	**0.181**	**0.230**	0.189	0.094	0.056
R_M3_GR	0	0.261	0.233	0.200	0.149	0.103	R_M3_GR	0	0.098	0.124	0.151	0.157	0.112
	1	0.268	0.229	0.181	0.141	0.085		1	0.096	0.128	0.181	0.146	0.073
	2	0.260	0.208	0.173	0.126	0.084		2	0.103	0.156	0.172	0.103	0.072
	3	0.244	0.204	0.161	0.128	0.082		3	0.131	0.158	0.140	0.105	0.083
	4	0.240	0.191	0.164	0.135	0.097		4	0.140	0.129	0.137	0.122	0.113
	5	0.227	0.195	0.177	0.149	0.093		5	0.114	0.123	0.151	0.152	0.062
R_CUR_GR	0	0.256	0.192	0.158	0.121	0.090	R_CUR_GR	0	0.075	0.102	0.117	0.139	0.104
	1	0.228	0.189	0.157	0.130	0.089		1	0.079	0.098	0.166	0.138	0.058
	2	0.227	0.188	0.162	0.128	0.085		2	0.078	0.145	0.164	0.091	0.068
	3	0.224	0.198	0.168	0.133	0.078		3	0.128	0.159	0.124	0.107	0.051
	4	0.240	0.204	0.173	0.126	0.076		4	0.146	0.113	0.154	0.088	0.073
	5	0.236	0.203	0.160	0.118	0.079		5	0.097	0.139	0.118	0.101	0.068
STOCK_GR	0	0.239	0.191	0.144	0.135	0.082	STOCK_GR	0	0.055	0.112	0.119	0.084	0.028
	1	0.232	0.176	0.166	0.120	0.080		1	0.090	0.104	0.103	0.076	0.076
	2	0.216	0.195	0.153	0.124	0.063		2	0.087	0.091	0.102	0.106	0.076
	3	0.228	0.184	0.170	0.110	0.078		3	0.071	0.081	0.134	0.103	0.038
	4	0.221	0.206	0.147	0.123	0.069		4	0.062	0.118	0.133	0.075	0.045
	5	0.244	0.181	0.162	0.113	0.065		5	0.104	0.121	0.110	0.080	0.084
EXCH_GR	0	0.190	0.198	0.169	0.166	0.115	EXCH_GR	0	0.095	0.097	0.123	0.099	0.074
	1	0.227	0.190	0.185	0.142	0.101		1	0.081	0.109	0.120	0.102	0.141
	2	0.213	0.201	0.164	0.129	0.104		2	0.091	0.104	0.124	0.153	-0.029
	3	0.237	0.196	0.167	0.148	0.099		3	0.083	0.108	0.192	-0.048	0.063
	4	0.233	0.198	0.185	0.144	0.088		4	0.091	0.180	0.053	0.084	0.049
	5	0.227	0.209	0.174	0.128	0.115		5	0.157	0.057	0.118	0.081	0.033

Table A.8 (Continued)

Dep Var = R1	lag	h = Quarters Ahead 4	5	6	7	8	Dep Var = R2	lag	h = Quarters Ahead 4	5	6	7	8
INF	0	0.289	0.246	0.221	0.202	0.146	INF	0	0.120	0.096	0.128	0.096	0.058
	1	0.278	0.248	0.233	0.181	0.142		1	0.079	0.112	0.119	0.094	0.112
	2	0.287	0.266	0.218	0.183	0.149		2	0.092	0.102	0.121	0.141	0.095
	3	0.302	0.245	0.215	0.188	0.148		3	0.081	0.105	0.172	0.123	0.095
	4	0.277	0.242	0.220	0.186	0.151		4	0.087	0.153	0.153	0.122	0.146
	5	0.279	0.250	0.223	0.192	0.134		5	0.141	0.142	0.154	0.170	0.108
RTT_GR	0	0.232	0.210	0.192	0.138	0.097	RTT_GR	0	0.066	0.126	0.153	0.131	0.093
	1	0.237	0.216	0.169	0.134	0.086		1	0.097	0.130	0.152	0.122	0.050
	2	0.252	0.200	0.169	0.128	0.071		2	0.101	0.135	0.152	0.086	0.061
	3	0.240	0.203	0.169	0.116	0.060		3	0.113	0.142	0.115	0.095	0.064
	4	0.243	0.203	0.153	0.106	0.069		4	0.124	0.102	0.128	0.096	0.017
	5	0.235	0.183	0.141	0.113	0.081		5	0.087	0.114	0.128	0.044	0.101
DWELL_GR	0	0.232	0.165	0.134	0.128	0.108	DWELL_GR	0	0.107	0.120	0.129	0.126	0.106
	1	0.201	0.159	0.157	0.142	0.101		1	0.103	0.118	0.141	0.143	0.150
	2	0.191	0.184	0.172	0.136	0.063		2	0.101	0.117	0.157	0.179	0.137
	3	0.220	0.201	0.170	0.105	0.045		3	0.097	0.136	0.209	0.167	0.032
	4	0.233	0.197	0.139	0.086	0.050		4	0.119	0.174	0.193	0.066	0.053
	5	0.231	0.166	0.121	0.094	0.089		5	0.178	0.191	0.093	0.083	0.067
NONRES_GR	0	0.224	0.192	0.159	0.120	0.081	NONRES_GR	0	0.088	0.111	0.108	0.101	0.063
	1	0.228	0.188	0.153	0.123	0.072		1	0.089	0.094	0.124	0.097	0.062
	2	0.227	0.186	0.156	0.114	0.048		2	0.074	0.107	0.125	0.094	0.060
	3	0.226	0.188	0.148	0.090	0.073		3	0.086	0.107	0.129	0.093	0.047
	4	0.225	0.178	0.121	0.116	0.066		4	0.089	0.114	0.128	0.084	0.059
	5	0.217	0.160	0.153	0.114	0.061		5	0.100	0.115	0.118	0.088	0.061
SV_EX_GR	0	0.238	0.203	0.178	0.128	0.082	SV_EX_GR	0	0.071	0.101	0.153	0.105	0.066
	1	0.235	0.202	0.162	0.126	0.085		1	0.077	0.128	0.128	0.099	0.050
	2	0.250	0.193	0.162	0.126	0.072		2	0.107	0.111	0.128	0.085	0.067
	3	0.234	0.193	0.161	0.118	0.079		3	0.090	0.115	0.124	0.099	0.064
	4	0.229	0.192	0.155	0.126	0.086		4	0.094	0.096	0.133	0.096	0.072
	5	0.226	0.184	0.160	0.128	0.083		5	0.080	0.119	0.128	0.101	0.068
TOT_GR	0	0.233	0.203	0.170	0.151	0.108	TOT_GR	0	0.094	0.115	0.130	0.133	0.064
	1	0.222	0.194	0.174	0.138	0.138		1	0.092	0.111	0.146	0.099	0.066
	2	0.227	0.200	0.169	0.169	0.080		2	0.081	0.128	0.126	0.100	0.073
	3	0.234	0.199	0.199	0.125	0.082		3	0.105	0.106	0.133	0.104	0.071
	4	0.233	0.230	0.161	0.129	0.080		4	0.090	0.116	0.140	0.102	0.063
	5	0.247	0.190	0.162	0.123	0.084		5	0.102	0.126	0.133	0.095	0.063
GOLD_GR	0	0.240	0.197	0.166	0.111	0.088	GOLD_GR	0	0.086	0.101	0.113	0.109	0.074
	1	0.241	0.200	0.152	0.124	0.132		1	0.084	0.101	0.136	0.110	0.057
	2	0.242	0.180	0.156	0.166	0.166		2	0.077	0.119	0.139	0.079	0.051
	3	0.227	0.187	0.185	0.184	0.167		3	0.104	0.125	0.122	0.082	0.071
	4	0.227	0.206	0.200	0.188	0.183		4	0.128	0.115	0.119	0.099	0.105
	5	0.228	0.216	0.205	0.195	0.144		5	0.102	0.110	0.128	0.123	0.126
US_SPR	0	0.239	0.200	0.170	0.132	0.094	US_SPR	0	0.044	0.089	0.109	0.087	0.047
	1	0.237	0.203	0.174	0.135	0.084		1	0.065	0.097	0.116	0.056	0.068
	2	0.249	0.212	0.167	0.121	0.077		2	0.092	0.102	0.119	0.105	0.072
	3	0.248	0.201	0.152	0.114	0.069		3	0.091	0.107	0.152	0.125	0.063
	4	0.236	0.184	0.145	0.102	0.043		4	0.086	0.136	0.163	0.106	0.064
	5	0.218	0.175	0.129	0.070	-0.006		5	0.114	0.149	0.142	0.096	0.073
US_LG_R	0	0.239	0.200	0.170	0.132	0.094	US_LG_R	0	0.044	0.089	0.109	0.087	0.047
	1	0.232	0.204	0.142	0.125	0.087		1	0.069	0.087	0.108	0.090	0.050
	2	0.233	0.178	0.145	0.131	0.082		2	0.077	0.082	0.115	0.093	0.002
	3	0.219	0.181	0.165	0.140	0.077		3	0.068	0.092	0.126	0.056	0.017
	4	0.214	0.194	0.170	0.121	0.096		4	0.073	0.098	0.088	0.054	0.038
	5	0.221	0.193	0.143	0.120	0.100		5	0.082	0.068	0.085	0.058	0.024
US_STK_GR	0	0.218	0.188	0.162	0.101	0.069	US_STK_GR	0	0.092	0.110	0.113	0.090	0.059
	1	0.226	0.197	0.132	0.110	0.067		1	0.086	0.107	0.121	0.093	0.056
	2	0.233	0.162	0.145	0.118	0.088		2	0.083	0.106	0.119	0.092	0.061
	3	0.196	0.174	0.156	0.128	0.079		3	0.082	0.101	0.116	0.086	0.076
	4	0.209	0.191	0.168	0.126	0.081		4	0.080	0.107	0.127	0.114	0.068
	5	0.225	0.198	0.161	0.124	0.073		5	0.093	0.110	0.148	0.098	0.048

Table A.9: Re-Sampling *pseudo* R^2: LHS is R1, RHS are Constant and the Variable as Stated

Dep Var = R1	h = Quarters Ahead							
	1	2	3	4	5	6	7	8
SPREAD	0.074	0.139	0.151	0.097	0.060	-0.028	-0.028	-0.009
SH_R	0.064	0.126	0.133	0.074	0.035	0.020	-0.003	-0.025
LG_R	0.070	0.072	0.041	-0.010	-0.012	-0.024	-0.041	-0.050
D_SH_R	-0.104	-0.062	-0.074	-0.108	-0.110	-0.136	-0.163	-0.138
D_LG_R	-0.106	-0.128	-0.133	-0.130	-0.112	-0.094	-0.093	-0.098
M3_GR	-0.009	0.028	0.104	**0.175**	-0.052	-0.054	-0.001	0.013
CUR_GR	0.018	-0.004	-0.047	-0.045	0.002	0.043	0.098	**0.162**
R_M3_GR	-0.024	-0.047	-0.044	-0.019	-0.035	-0.118	-0.132	-0.067
R_CUR_GR	0.006	0.053	0.026	0.001	-0.021	-0.020	-0.017	-0.013
STOCK_GR	0.030	0.022	0.019	0.004	-0.007	-0.014	-0.004	-0.019
EXCH_GR	-0.016	-0.019	-0.003	-0.007	-0.008	-0.012	0.003	-0.003
INF	**0.149**	**0.232**	**0.165**	0.139	**0.109**	0.111	0.123	0.078
RTT_GR	0.038	0.030	-0.031	-0.028	-0.020	0.000	-0.008	-0.008
DWELL_GR	0.086	0.075	0.028	-0.013	-0.035	-0.030	-0.017	-0.014
NONRES_GR	-0.008	-0.009	-0.013	-0.015	-0.010	-0.014	-0.013	-0.009
SV_EX_GR	-0.015	-0.013	-0.017	-0.005	0.004	0.017	-0.016	-0.016
TOT_GR	0.033	0.018	-0.002	-0.012	0.006	-0.016	-0.003	-0.064
GOLD_GR	-0.017	-0.035	-0.021	0.009	-0.034	-0.030	-0.034	-0.073
US_SPR	-0.033	-0.021	0.043	0.097	0.090	**0.120**	**0.125**	0.100
US_SH_R	0.012	0.044	0.064	-0.039	-3.607	0.035	0.039	0.017
US_LG_R	-0.009	0.023	0.024	0.007	-0.016	-0.034	-0.064	-0.089
US_STK_GR	-0.021	-0.015	-0.023	-0.029	-0.032	-0.030	-0.060	-0.056

Table A.10: Re-Sampling *pseudo* R^2: LHS is R2, RHS are Constant and the Variable as Stated

Dep Var = R2	h = Quarters Ahead							
	1	2	3	4	5	6	7	8
SPREAD	0.038	0.062	0.060	0.052	**0.088**	0.067	-0.014	-0.057
SH_R	**0.113**	**0.133**	**0.120**	**0.094**	0.087	0.070	0.030	0.003
LG_R	0.084	0.079	0.070	0.052	0.033	0.024	0.016	0.013
D_SH_R	0.020	0.032	0.014	-0.027	-0.041	-0.054	-0.069	-0.060
D_LG_R	0.000	-0.005	-0.020	-0.048	-0.087	-0.089	-0.061	-0.034
M3_GR	0.002	0.016	0.034	0.066	-0.007	0.006	0.031	0.032
CUR_GR	0.004	0.013	0.022	0.014	0.030	0.034	**0.086**	**0.108**
R_M3_GR	-0.004	-0.007	-0.045	-0.025	-0.035	-0.055	0.000	0.004
R_CUR_GR	0.025	-0.003	-0.004	-0.030	-0.039	-0.031	0.003	0.014
STOCK_GR	-0.021	-0.010	-0.002	-0.068	-0.023	-0.007	-0.012	-0.023
EXCH_GR	-0.016	-0.012	-0.015	0.012	-0.009	-0.002	-0.010	-0.009
INF	0.104	0.118	0.116	0.064	0.002	-0.016	-0.021	-0.015
RTT_GR	0.046	0.063	-0.028	-0.018	-0.015	0.015	-0.003	0.004
DWELL_GR	0.008	0.014	0.030	0.015	-0.012	-0.016	0.007	0.040
NONRES_GR	-0.018	-0.017	-0.011	-0.004	-0.003	-0.009	-0.012	-0.008
SV_EX_GR	-0.016	-0.012	-0.036	-0.013	-0.015	0.006	-0.015	-0.018
TOT_GR	0.028	0.015	-0.032	-0.022	-0.014	-0.001	0.005	-0.067
GOLD_GR	-0.007	-0.019	-0.022	-0.041	0.004	-0.048	-0.048	-0.008
US_SPR	-0.021	-0.026	-0.022	-0.009	0.017	0.037	0.028	-0.004
US_SH_R	0.020	0.026	0.041	0.065	0.078	**0.086**	0.058	0.032
US_LG_R	0.020	0.037	0.051	0.063	0.063	0.061	0.045	0.032
US_STK_GR	0.028	-0.009	0.026	-0.009	-0.018	-0.016	-0.014	-0.058

Table A.11: Re-Sampling *pseudo* R^2: RHS are Constant, INF, and the Variable as Stated

Dep Var = R1	lag	h = Quarters Ahead 1	2	3	Dep Var = R2	lag	h = Quarters Ahead 1	2	3
SPREAD	0	0.154	0.282	0.228	SPREAD	0	0.110	0.137	0.141
	1	0.191	0.294	0.122		1	0.111	0.111	0.115
	2	0.197	0.237	0.092		2	0.100	0.097	0.149
	3	0.131	0.146	-0.050		3	0.101	0.144	0.130
	4	0.117	-0.107	0.053		4	0.171	0.154	0.075
	5	0.005	0.104	0.123		5	0.157	0.079	0.051
SH_R	0	0.122	0.272	0.191	SH_R	0	0.164	0.178	0.172
	1	0.180	0.274	-0.196		1	0.177	0.153	0.142
	2	0.169	0.191	-0.017		2	0.166	0.139	0.150
	3	0.123	0.200	0.076		3	0.155	0.157	0.145
	4	0.113	0.200	0.107		4	0.175	0.164	0.120
	5	0.134	0.197	0.119		5	0.171	0.129	0.105
LG_R	0	0.161	0.252	0.133	LG_R	0	0.153	0.144	0.139
	1	0.152	0.194	0.029		1	0.149	0.141	0.123
	2	0.094	0.135	0.088		2	0.147	0.129	0.113
	3	0.086	0.178	0.115		3	0.140	0.123	0.115
	4	0.110	0.168	0.104		4	0.128	0.122	0.113
	5	0.105	0.152	0.098		5	0.128	0.121	0.122
D_SH_R	0	-0.054	0.140	0.066	D_SH_R	0	0.118	0.127	0.117
	1	0.060	0.149	-0.039		1	0.127	0.107	0.074
	2	0.046	0.123	0.027		2	0.124	0.086	0.074
	3	0.024	0.126	0.007		3	0.083	0.072	0.060
	4	0.043	0.122	0.008		4	0.078	0.071	0.050
	5	0.054	0.120	0.058		5	0.069	0.058	0.066
D_LG_R	0	0.046	0.095	0.005	D_LG_R	0	0.110	0.109	0.092
	1	-0.005	0.091	-0.021		1	0.109	0.097	0.048
	2	0.011	0.115	0.049		2	0.103	0.066	0.000
	3	0.039	0.149	0.083		3	0.066	0.000	-0.009
	4	0.058	0.155	0.067		4	0.004	-0.011	0.021
	5	0.089	0.162	0.076		5	0.011	0.037	0.074
M3_GR	0	0.121	0.208	0.191	M3_GR	0	0.089	0.110	0.109
	1	0.126	0.238	0.221		1	0.091	0.106	0.114
	2	0.166	0.299	0.095		2	0.103	0.132	0.079
	3	0.193	0.147	0.056		3	0.112	0.046	0.058
	4	0.077	0.121	0.113		4	0.068	0.075	0.114
	5	0.019	0.165	0.125		5	0.067	0.100	0.114
CUR_GR	0	0.121	0.188	0.095	CUR_GR	0	0.084	0.084	0.083
	1	0.107	0.179	0.118		1	0.061	0.087	0.082
	2	0.074	0.165	0.103		2	0.066	0.088	0.085
	3	0.089	0.166	0.124		3	0.068	0.092	0.090
	4	0.097	0.188	0.154		4	0.081	0.096	0.112
	5	0.110	0.207	0.189		5	0.090	0.110	0.123
R_M3_GR	0	0.121	0.208	0.191	R_M3_GR	0	0.089	0.110	0.109
	1	0.127	0.195	0.154		1	0.102	0.094	0.099
	2	0.128	0.248	0.155		2	0.079	0.113	0.103
	3	0.145	0.222	0.073		3	0.095	0.089	0.069
	4	0.141	0.186	0.060		4	0.097	0.100	0.151
	5	0.070	0.140	0.104		5	0.054	0.114	0.118
R_CUR_GR	0	0.121	0.188	0.095	R_CUR_GR	0	0.084	0.084	0.083
	1	0.156	0.246	0.161		1	0.105	0.109	0.084
	2	0.140	0.192	0.124		2	0.085	0.072	0.072
	3	0.136	0.188	0.132		3	0.067	0.065	0.076
	4	0.113	0.194	0.139		4	0.058	0.081	0.130
	5	0.137	0.214	0.126		5	0.063	0.105	0.117
STOCK_GR	0	0.160	0.260	0.190	STOCK_GR	0	0.081	0.106	0.117
	1	0.154	0.233	0.164		1	0.084	0.111	0.066
	2	0.167	0.232	0.146		2	0.106	0.053	0.098
	3	0.148	0.190	0.113		3	0.051	0.101	0.100
	4	0.129	0.193	0.161		4	0.086	0.105	0.103
	5	0.119	0.222	0.128		5	0.097	0.104	0.090
EXCH_GR	0	0.143	0.220	0.168	EXCH_GR	0	0.090	0.083	0.093
	1	0.118	0.198	0.145		1	0.077	0.088	0.105
	2	0.124	0.205	0.155		2	0.080	0.115	0.097
	3	0.124	0.208	0.125		3	0.115	0.098	0.101
	4	0.123	0.204	0.148		4	0.115	0.114	0.104
	5	0.097	0.204	0.148		5	0.096	0.098	0.099

Table A.11 (Continued)

Dep Var = R1	lag	h = Quarters Ahead			Dep Var = R2	lag	h = Quarters Ahead		
		1	2	3			1	2	3
RTT_GR	0	0.111	0.192	0.153	RTT_GR	0	0.069	0.111	0.101
	1	0.180	0.148	0.134		1	0.158	0.019	0.047
	2	0.111	0.190	0.134		2	0.073	0.094	0.114
	3	0.126	0.209	0.157		3	0.075	0.102	0.120
	4	0.109	0.207	0.145		4	0.088	0.123	0.107
	5	0.122	0.203	0.137		5	0.097	0.101	0.107
DWELL_GR	0	0.138	0.196	0.147	DWELL_GR	0	0.080	0.105	0.111
	1	0.165	0.217	0.109		1	0.099	0.121	0.109
	2	0.157	0.172	0.080		2	0.119	0.117	0.092
	3	0.126	0.165	0.124		3	0.119	0.097	0.090
	4	0.080	0.127	0.126		4	0.085	0.096	0.116
	5	0.068	0.181	0.172		5	0.065	0.123	0.168
NONRES_GR	0	0.131	0.203	0.134	NONRES_GR	0	0.079	0.096	0.103
	1	0.151	0.203	0.129		1	0.090	0.104	0.105
	2	0.146	0.222	0.145		2	0.101	0.107	0.105
	3	0.088	0.207	0.148		3	0.107	0.119	0.105
	4	0.125	0.209	0.145		4	0.098	0.090	0.101
	5	0.126	0.204	0.139		5	0.092	0.101	0.106
SV_EX_GR	0	0.138	0.217	0.148	SV_EX_GR	0	0.089	0.103	0.075
	1	0.140	0.218	0.152		1	0.089	0.079	0.104
	2	0.131	0.217	0.154		2	0.076	0.109	0.096
	3	0.134	0.223	0.164		3	0.097	0.097	0.118
	4	0.128	0.221	0.128		4	0.081	0.105	0.097
	5	0.151	0.212	0.143		5	0.127	0.106	0.101
TOT_GR	0	0.155	0.227	0.157	TOT_GR	0	0.098	0.113	0.083
	1	0.162	0.223	0.155		1	0.106	0.091	0.102
	2	0.148	0.216	0.164		2	0.091	0.105	0.101
	3	0.133	0.234	0.145		3	0.095	0.102	0.112
	4	0.139	0.218	0.168		4	0.093	0.112	0.127
	5	0.119	0.209	0.093		5	0.100	0.109	0.044
GOLD_GR	0	0.149	0.229	0.196	GOLD_GR	0	0.091	0.098	0.086
	1	0.138	0.258	0.217		1	0.081	0.091	0.090
	2	0.175	0.292	0.186		2	0.081	0.094	0.134
	3	0.172	0.240	0.166		3	0.081	0.128	0.097
	4	0.143	0.222	0.131		4	0.112	0.091	0.090
	5	0.159	0.196	0.095		5	0.100	0.118	0.132
US_SPR	0	0.128	0.198	0.179	US_SPR	0	0.085	0.099	0.096
	1	0.116	0.232	0.222		1	0.088	0.098	0.100
	2	0.138	0.245	0.177		2	0.087	0.094	0.103
	3	0.194	0.241	0.206		3	0.085	0.104	0.113
	4	0.188	0.248	0.212		4	0.090	0.111	0.103
	5	0.184	0.242	0.185		5	0.086	0.094	0.091
US_SH_R	0	0.101	0.198	0.144	US_SH_R	0	0.101	0.105	0.122
	1	0.106	0.199	-0.387		1	0.104	0.114	0.138
	2	0.102	-0.131	-0.450		2	0.102	0.121	0.135
	3	-1.545	-4.690	0.111		3	0.121	0.132	0.140
	4	-3.268	0.099	0.129		4	0.131	0.131	0.118
	5	0.069	-0.193	0.110		5	0.123	0.097	0.099
US_LG_R	0	0.098	0.201	0.148	US_LG_R	0	0.102	0.122	0.133
	1	0.118	0.191	0.129		1	0.113	0.129	0.142
	2	0.113	0.166	0.108		2	0.116	0.134	0.140
	3	0.097	0.146	0.096		3	0.130	0.138	0.142
	4	0.073	0.130	0.062		4	0.133	0.135	0.128
	5	0.060	0.078	0.028		5	0.130	0.117	0.114
US_STK_GR	0	0.107	0.199	0.146	US_STK_GR	0	0.101	0.071	0.116
	1	0.118	0.207	0.144		1	0.083	0.121	0.102
	2	0.132	0.204	0.140		2	0.113	0.095	0.088
	3	0.114	0.212	0.160		3	0.081	0.090	0.125
	4	0.126	0.207	0.120		4	0.092	0.102	0.102
	5	0.148	0.179	0.124		5	0.119	0.123	0.079

Table A.12: Re-Sampling *pseudo* R^2: LHS is R1, RHS are Constant, US_SPR, and the Variable as Stated

Dep Var = R1	lag	h = Quarters Ahead 4	5	6	7	8
SPREAD	0	0.045	0.023	-0.043	0.005	0.026
	1	-0.001	-0.155	0.022	0.072	0.079
	2	-3.030	-0.029	0.083	0.113	0.071
	3	-0.120	0.063	0.127	0.105	0.063
	4	0.041	0.107	0.122	0.104	0.063
	5	0.113	0.114	0.125	0.114	0.076
SH_R	0	-0.371	0.006	0.121	0.114	0.069
	1	-2.386	0.058	0.118	0.109	0.076
	2	-9.238	0.058	0.111	0.112	0.075
	3	-37.729	0.058	0.117	0.112	0.072
	4	-0.125	0.059	0.114	0.109	0.073
	5	-0.198	-0.003	0.104	0.112	0.091
LG_R	0	-22.520	-0.017	0.113	0.113	0.077
	1	-20.984	0.000	0.105	0.110	0.073
	2	-4.529	-0.039	0.091	0.099	0.067
	3	-28.610	-0.097	0.070	0.093	0.065
	4	-46.871	-1.224	0.063	0.089	0.059
	5	-1.919	-0.276	0.056	0.077	0.051
D_SH_R	0	-0.686	0.001	0.069	0.051	0.028
	1	-0.131	-0.044	0.014	0.042	0.069
	2	-0.257	-0.117	0.002	0.080	0.070
	3	-6.109	-0.111	0.064	0.100	0.068
	4	-0.226	-0.030	0.090	0.097	0.066
	5	-0.149	-0.093	0.059	0.072	0.092
D_LG_R	0	-2.492	-0.057	0.087	0.089	0.049
	1	-3.972	-0.028	0.065	0.068	0.044
	2	-0.545	-0.072	0.031	0.053	0.042
	3	-17.821	-0.135	0.022	0.064	0.051
	4	-15.629	-0.205	0.032	0.069	0.002
	5	-13.950	-0.185	0.018	-0.016	-0.028
M3_GR	0	0.283	0.083	0.090	0.127	0.099
	1	0.088	0.047	0.123	0.126	0.088
	2	0.062	0.095	0.121	0.115	0.076
	3	0.078	0.084	0.109	0.103	0.072
	4	0.047	0.067	0.100	0.100	0.078
	5	0.045	0.072	0.096	0.104	0.069
CUR_GR	0	0.056	0.096	0.148	0.202	0.231
	1	0.108	0.138	0.209	0.254	0.239
	2	0.140	0.186	0.239	0.253	**0.252**
	3	0.217	0.261	0.271	**0.300**	0.232
	4	0.294	0.237	**0.328**	0.284	0.171
	5	**0.297**	**0.316**	0.304	0.212	0.154
R_M3_GR	0	0.089	0.059	-0.006	0.030	0.052
	1	0.077	-0.066	0.004	0.070	0.045
	2	-0.045	-0.077	0.060	0.072	0.092
	3	-0.098	0.020	0.071	0.126	0.124
	4	0.003	0.033	0.130	0.172	0.145
	5	0.023	0.101	0.162	0.149	0.095
R_CUR_GR	0	0.066	0.064	0.099	0.107	0.098
	1	0.062	0.067	0.103	0.116	0.083
	2	0.069	0.079	0.103	0.094	0.084
	3	0.080	0.079	0.095	0.114	0.082
	4	0.087	0.066	0.113	0.111	0.092
	5	0.066	0.082	0.109	0.120	0.090
STOCK_GR	0	0.109	0.059	0.100	0.136	0.093
	1	0.082	0.069	0.126	0.112	0.064
	2	0.080	0.093	0.091	0.095	0.042
	3	0.094	0.048	0.103	0.085	0.072
	4	0.070	0.071	0.087	0.101	0.053
	5	0.077	0.064	0.107	0.080	0.031
EXCH_GR	0	0.084	0.078	0.110	0.136	0.111
	1	0.089	0.080	0.109	0.120	0.106
	2	0.087	0.055	0.098	0.114	0.106
	3	0.088	0.077	0.112	0.128	0.092
	4	0.091	0.089	0.125	0.115	0.079
	5	0.094	0.095	0.105	0.097	0.110

Table A.12 (Continued)

Dep Var = R1	lag	h = Quarters Ahead 4	5	6	7	8
INF	0	0.211	0.175	0.196	0.201	0.132
	1	0.191	0.190	0.218	0.174	0.149
	2	0.209	0.209	0.194	0.203	0.167
	3	0.200	0.150	0.203	0.201	0.148
	4	0.144	0.172	0.209	0.186	0.154
	5	0.211	0.177	0.184	0.182	0.125
RTT_GR	0	0.063	0.069	0.126	0.117	0.093
	1	0.070	0.088	0.111	0.116	0.094
	2	0.088	0.081	0.111	0.117	0.088
	3	0.090	0.084	0.113	0.115	0.083
	4	0.088	0.083	0.111	0.108	0.087
	5	0.089	0.085	0.105	0.111	0.074
DWELL_GR	0	0.081	0.074	0.118	0.133	0.118
	1	0.075	0.074	0.097	0.082	0.060
	2	0.068	0.063	0.073	0.059	0.026
	3	0.071	0.065	0.057	0.052	0.018
	4	0.068	0.014	0.048	0.041	0.043
	5	0.024	0.013	0.046	0.077	0.086
NONRES_GR	0	0.075	0.073	0.102	0.111	0.090
	1	0.077	0.073	0.108	0.114	0.082
	2	0.082	0.078	0.111	0.109	0.061
	3	0.088	0.082	0.105	0.085	0.082
	4	0.092	0.079	0.088	0.107	0.090
	5	0.087	0.068	0.107	0.115	0.086
SV_EX_GR	0	0.097	0.100	0.144	0.116	0.086
	1	0.106	0.118	0.101	0.108	0.083
	2	0.122	0.052	0.100	0.107	0.092
	3	0.060	0.069	0.095	0.110	0.093
	4	0.061	0.051	0.102	0.115	0.128
	5	0.073	0.073	0.114	0.159	0.089
TOT_GR	0	0.093	0.083	0.089	0.100	0.035
	1	0.093	0.068	0.093	0.056	0.000
	2	0.086	0.078	0.058	0.025	-0.012
	3	0.091	0.029	0.016	0.030	-0.008
	4	0.055	0.009	0.045	0.043	0.053
	5	0.004	0.029	0.060	0.090	0.084
GOLD_GR	0	0.145	0.100	0.116	0.109	0.042
	1	0.140	0.113	0.112	0.070	0.071
	2	0.120	0.070	0.058	0.094	0.106
	3	0.104	0.035	0.083	0.115	0.134
	4	0.041	0.048	0.100	0.146	0.160
	5	0.051	0.051	0.129	0.167	0.144
US_SH_R	0	-0.241	-3.677	-0.017	0.011	-0.014
	1	-9.123	-0.022	0.031	0.041	0.027
	2	-8.875	0.053	0.072	0.079	0.051
	3	-0.013	0.052	0.077	0.080	0.053
	4	-0.008	0.040	0.078	0.083	0.035
	5	-0.040	0.033	0.084	0.063	-0.003
US_LG_R	0	-0.241	-3.570	-0.017	0.011	-0.014
	1	-0.056	-0.143	-0.003	0.017	0.003
	2	-0.034	-0.072	0.008	0.032	0.018
	3	-0.041	-0.077	0.006	0.038	0.032
	4	-0.063	-0.108	0.020	0.057	0.033
	5	-0.388	-0.137	0.041	0.057	0.026
US_STK_GR	0	0.082	0.071	0.100	0.087	0.066
	1	0.080	0.072	0.077	0.088	0.105
	2	0.078	0.047	0.080	0.128	0.103
	3	0.039	0.040	0.122	0.125	0.082
	4	0.054	0.100	0.126	0.109	0.073
	5	0.109	0.103	0.113	0.102	0.059

Table A.13: Re-Sampling *pseudo* R^2: LHS is R2, RHS are Constant, US_SH_R, and the Variable as Stated

Dep Var = R2	lag	h = Quarters Ahead 4	5	6	7	8
SPREAD	0	0.090	0.143	0.111	0.020	-0.025
	1	0.140	0.134	0.054	0.011	-0.002
	2	0.130	0.065	0.040	-0.003	0.020
	3	0.056	0.039	-0.078	0.035	0.034
	4	0.019	0.004	-0.083	0.052	0.032
	5	-0.014	0.039	-0.059	0.050	0.011
SH_R	0	0.075	0.068	0.039	0.006	-0.014
	1	0.085	0.070	0.018	0.022	0.013
	2	0.081	0.040	0.043	0.031	0.025
	3	0.051	0.049	-0.038	0.053	0.031
	4	0.044	0.019	0.022	0.059	0.032
	5	0.015	0.056	0.023	0.061	0.030
LG_R	0	0.051	0.044	0.041	-0.004	0.006
	1	0.044	0.039	0.020	0.025	0.025
	2	0.042	0.033	0.048	0.040	0.027
	3	0.039	0.054	0.048	0.029	0.032
	4	0.050	0.052	-0.097	0.053	0.037
	5	0.022	0.053	-0.184	0.035	0.040
D_SH_R	0	0.075	0.068	0.039	0.006	-0.014
	1	0.066	0.062	0.029	0.023	0.015
	2	0.048	0.029	0.050	-0.002	0.023
	3	0.019	0.047	-0.047	0.028	0.030
	4	0.024	0.049	-0.043	0.002	0.036
	5	0.021	0.060	-0.088	0.006	0.054
D_LG_R	0	0.047	0.004	-0.017	-0.018	-0.012
	1	-0.003	-0.006	0.008	0.026	0.012
	2	-0.016	0.002	0.051	-0.033	0.014
	3	-0.007	0.040	-0.123	0.006	0.017
	4	0.023	0.049	0.004	0.006	0.032
	5	0.022	0.050	0.012	0.027	0.036
M3_GR	0	0.116	0.088	0.112	0.121	0.092
	1	0.070	0.115	0.149	0.127	0.076
	2	0.088	0.135	0.148	0.078	0.032
	3	0.114	0.144	0.044	0.046	0.020
	4	0.122	0.101	0.031	0.050	-0.020
	5	0.077	0.044	0.009	-2.210	-0.085
CUR_GR	0	0.045	0.067	0.077	0.111	0.119
	1	0.059	0.072	0.138	0.146	0.134
	2	0.064	0.129	0.166	0.128	0.162
	3	0.128	0.176	0.092	0.125	0.091
	4	0.162	0.179	0.123	0.104	0.046
	5	0.158	0.193	0.064	0.064	0.050
R_M3_GR	0	0.050	0.077	-0.053	0.116	0.083
	1	0.053	0.005	0.146	0.130	0.034
	2	0.031	0.122	0.145	0.040	0.029
	3	0.101	0.141	0.031	0.049	0.050
	4	0.122	0.070	0.045	0.073	0.058
	5	0.041	0.042	0.021	0.048	-0.044
R_CUR_GR	0	0.031	0.040	0.060	0.075	0.065
	1	0.029	0.049	0.107	0.108	0.031
	2	0.032	0.091	0.124	0.026	0.052
	3	0.081	0.127	-0.007	0.058	0.018
	4	0.113	0.059	-0.056	0.036	0.044
	5	0.024	0.088	-0.016	0.062	0.039
STOCK_GR	0	0.020	0.066	0.066	0.024	-0.018
	1	0.052	0.063	0.055	0.024	0.027
	2	0.057	0.056	0.057	0.009	0.048
	3	0.047	0.050	0.018	0.061	-0.010
	4	0.039	0.063	0.030	0.003	0.031
	5	0.035	-0.004	-0.001	0.053	0.075
EXCH_GR	0	0.051	0.047	0.075	0.046	0.027
	1	0.045	0.066	0.074	0.065	0.099
	2	0.056	0.062	0.085	0.063	0.116
	3	0.050	0.074	0.102	0.127	0.055
	4	0.056	0.139	0.122	0.079	0.034
	5	0.096	0.091	0.029	0.049	0.034

Table A.13 (Continued)

Dep Var = R2	lag	h = Quarters Ahead 4	5	6	7	8
INF	0	0.097	0.058	0.047	0.017	0.011
	1	0.048	0.040	0.040	0.037	0.054
	2	0.032	0.032	0.054	-0.025	0.073
	3	0.026	0.056	-0.140	0.077	0.081
	4	0.042	0.090	0.033	0.083	0.142
	5	0.057	0.085	0.037	0.143	0.111
RTT_GR	0	0.038	0.064	0.123	0.068	0.034
	1	0.045	0.102	0.090	0.019	0.022
	2	0.084	0.078	0.016	0.015	0.031
	3	0.067	0.080	-0.010	0.045	0.041
	4	0.079	0.040	0.013	0.032	0.019
	5	0.017	0.044	-0.002	0.037	0.090
DWELL_GR	0	0.073	0.061	0.044	0.066	0.072
	1	0.057	0.053	0.091	0.118	0.148
	2	0.042	0.075	0.135	0.172	0.075
	3	0.064	0.128	0.155	0.065	-0.051
	4	0.104	0.176	0.025	-0.043	-0.002
	5	0.170	0.039	-0.147	0.014	0.049
NONRES_GR	0	0.057	0.070	0.062	0.037	0.015
	1	0.059	0.061	0.070	0.052	0.029
	2	0.048	0.063	0.080	0.014	0.035
	3	0.052	0.074	-0.078	0.016	0.035
	4	0.058	0.055	-0.142	0.042	0.047
	5	0.024	0.059	0.010	0.066	0.049
SV_EX_GR	0	0.048	0.063	0.107	0.045	0.018
	1	0.046	0.085	0.071	0.052	0.008
	2	0.075	0.061	0.068	0.011	0.003
	3	0.049	0.071	0.001	-0.128	0.023
	4	0.047	0.044	-0.032	0.030	-0.001
	5	0.013	0.041	0.002	0.002	0.046
TOT_GR	0	0.039	0.061	0.089	0.055	-0.102
	1	0.042	0.078	0.080	-0.063	-0.072
	2	0.064	0.086	-0.001	-0.063	0.027
	3	0.079	0.013	-0.063	0.035	0.036
	4	0.011	-0.019	-2.816	-0.003	0.028
	5	-0.021	0.050	-1.048	0.035	0.040
GOLD_GR	0	0.023	0.045	0.045	0.018	0.011
	1	0.044	0.048	0.053	0.033	0.016
	2	0.039	0.052	0.039	0.024	0.021
	3	0.045	0.072	-0.022	-0.010	0.047
	4	0.077	0.056	-0.026	0.047	0.081
	5	0.033	0.004	-0.002	0.080	0.097
US_SPR	0	0.037	0.041	0.024	-0.037	-0.012
	1	0.038	0.031	0.017	0.019	0.037
	2	0.042	0.034	0.058	0.052	0.056
	3	0.034	0.063	-0.168	0.096	0.042
	4	0.042	0.081	0.058	0.084	0.035
	5	0.049	0.116	0.073	0.066	0.046
US_LG_R	0	0.037	0.041	0.024	-0.037	-0.012
	1	0.034	0.046	0.044	0.007	0.007
	2	0.039	0.047	0.061	-0.046	0.016
	3	0.033	0.058	-0.044	-0.216	0.029
	4	0.037	0.043	-0.299	-0.077	0.040
	5	0.015	0.056	-0.138	-3.256	0.037
US_STK_GR	0	0.055	0.055	0.076	0.055	-0.012
	1	0.040	0.071	0.087	0.019	0.021
	2	0.055	0.072	0.055	0.038	0.052
	3	0.061	0.047	-0.727	0.067	0.054
	4	0.028	0.048	0.066	0.079	0.086
	5	0.038	0.089	0.057	0.100	0.055

Chapter 3

Forecasting with Bayesian VARs: Does Larger Mean Better?

Conceptually, the impressive forecasting performance of the Bayesian VARs may be further improved by expanding the number of variables into the models. This chapter compares the forecasting performance of a large Bayesian VAR with 131 variables to much smaller models. Since the performance of a Bayesian regression can be affected by a hyperparameter governing the overall tightness of the prior distribution, we perform our investigation with careful consideration to this effect. Our results support the idea that larger Bayesian VARs perform better than smaller ones. However, when the hyperparameter of the prior of a smaller model is carefully chosen, the improvement in performances of larger models is not as impressive as previously thought. Even a 3-variable model with appropriately chosen shrinkage parameter will produce much better forecasts than those reported in the literature.

3.1 Introduction

In forecasting macroeconomic variables, there is an excellent record of the Bayesian VARs in the literature. For example, Robertson and Tallman (1999) report that various Bayesian VAR specifications outperform unrestricted VARs, while Litterman (1986) shows that a Bayesian VAR outperforms an ARIMA, a univariate AR, and the best known commercial forecasting services in out-of-sample forecasting.

According to Litterman (1986), there are at least two advantages of the Bayesian VARs over other nonstructural econometric models. First, since there are many relationships among macroeconomic variables not fully understood by economists, Bayesian VARs, which allow some uncertainty over the true structure of the economy, perform better in forecasting than other models that are fully based on just a single economic structure. Second, under the situation of a limited amount of observations, Bayesian VARs allow the incorporation of more information into account. A larger amount of parameters can be fitted into the model through assigning appropriate weight to the prior information.

With these advantages, one may argue that larger Bayesian VARs can outperform smaller models in forecasting. Since the exact structure of the economy is not known and the problem about the degrees of freedom is ameliorated, larger Bayesian VARs seem to have an advantage over smaller ones.

Recent forecasting literature is also supportive for the practice of incorporating a large number of variables into the models. Many methods are proposed or extended to allow this practice. These include, for example, the dynamic factor models of Stock and Watson (2002a) and Forni et al. (2000), and the factor-augmented VAR of Bernanke et al. (2005). There is a lot of evidence to show this practice improves the forecasting performances of the models. See, for example, Bernanke and Boivin (2003), Stock and Watson (2002b), D'Agostino and Giannone (2007), and Forni et al. (2003).

Bańbura et al. (2008) (henceforth BGR) demonstrate that Bayesian VARs admit a large number of endogenous variables. They investigate empirically whether this practice is desirable. According to the authors, a large Bayesian VAR with 131 variables performs better than smaller models with 3, 7, and 20 variables in out-of-sample forecasting. The largest model clearly outperforms the two smallest ones, but its forecasting performance, however, can be matched by the model with 20 variables.

The Bayesian VAR estimator, however, depends on a hyperparameter determining the relative weight given to the prior information, and as a consequence the out-of-sample forecasting performance of a model is influenced by this hyperparameter as well. BGR's findings therefore are based on the particular way that they determine the value of this hyperparameter. We do not find the BGR's method the most natural way of setting this value, and there is no reason to believe that their results will be robust if this parameter value is chosen in a different way. In section 3.3, we show that if we assign different values to this hyperparameter, larger VARs of BGR may not outperform smaller ones.

We, then, determine a suitable hyperparameter value for each model, which makes the most out of each model given our pre-evaluation period. Given a model and a forecast horizon, we find the hyperparameter value that minimizes the magnitude of out-of-sample forecast errors in a part of the pre-evaluation period. After that, we assign this suitable value to that model during our out-of-sample assessment in an evaluation period. This is shown in section 3.4. Our result in this section supports BGR's finding that larger models perform better in the overall picture. However, the performances of the larger models are not dramatically different from that of the smallest model.

We realize that the suitable hyperparameter value can vary over time. The time-varying hyperparameter may affect different models in different magnitudes. To make our study more robust, we make two additional experiments. First, we allow the suitable hyperparameter value of each model to change every 10 years. For each additional 10 years of observations, we re-calculate a suitable hyperparameter value for each model. After that, we use this hyperparameter value in making forecasts for the next 10 years, until we re-calculate a new hyperparameter value again. We assess the performances of Bayesian VARs under this practice. The result of this experiment is shown in section 3.5. Contrary to our expectation, this practice does not improve the forecasting performances of any model specifications.

Second, we apply an updating scheme for the hyperparameter value. With additional data, we calculate the effect of a small change in value of the hyperparameter. If the change signals an improvement in the forecasting performance of a model, a new hyperparameter value will be applied to the model in making the forecast for the next period. Section 3.6 reports the result from this experiment. It shows that our updating scheme can just marginally improve the forecasting performance of each model specification as well.

Apart from these, section 3.2 shows the details of the model and the estimation method used in this paper, and section 3.7 concludes the paper.

3.2 Estimated Model

We estimate the same Bayesian VARs as BGR. Let $Y_t = (y_{1,t}\ y_{2,t}\ \dots\ y_{m,t})'$ be an $m \times 1$ column vector of m endogenous variables in period t. The VAR has its reduced form as:

$$\underset{m\times 1}{Y_t} = \underset{m\times m}{A_1}\ \underset{m\times 1}{Y_{t-1}} + \underset{m\times m}{A_2}\ \underset{m\times 1}{Y_{t-2}} + \dots + \underset{m\times m}{A_p}\ \underset{m\times 1}{Y_{t-p}} + \underset{m\times 1}{\mathbf{c}} + \underset{m\times 1}{U_t}, \tag{3.1}$$

where $\mathbf{c} = (c_1, \dots, c_m)'$ is the vector of constants, and $U_t = (u_{1,t}\ \dots\ u_{m,t})'$ is the vector of unknown disturbances. We assume that:

$$\underset{m\times 1}{U_t} \sim N(\underset{m\times 1}{\mathbf{0}}, \underset{m\times m}{\Psi}),$$

where the time-invariant matrix Ψ is a positive definite matrix.

Let $X_t = (Y'_{t-1}, \dots, Y'_{t-p}, 1)'$ be a column vector containing p lags of Y_t and a constant 1. With observations $t = 1, \dots, T$, we can rearrange the VAR from (3.1) into:

$$\underset{T\times m}{Y} = \underset{T\times k}{X}\ \underset{k\times m}{B} + \underset{T\times m}{U}, \tag{3.2}$$

where $Y = (Y_1, \dots, Y_T)'$ is the matrix of dependent variables, $X = (X_1, \dots, X_T)'$

is the matrix of independent variables, $B = (A_1, ..., A_p, \mathbf{c})'$ is the matrix of unknown coefficients, $U = (U_1, ..., U_T)'$ is the matrix of disturbances, and $k = mp + 1$ is the total number of independent variables. Let $\mathbf{u}$ be the column vector obtained by stacking the columns of the disturbance matrix U from (3.2). The above assumption on U_t is equivalent to:

$$\underset{Tm\times 1}{\mathbf{u}} \sim N(\underset{Tm\times 1}{\mathbf{0}}, \underset{m\times m}{\Psi} \otimes \underset{T\times T}{I}),$$

where $\otimes$ represents the Kronecker product, and I is an identity matrix.

With the seemingly-unrelated-regressions (SUR) structure, the efficient estimator for B is the same as an unrestricted OLS estimator, which is:

$$\hat{B} = (X'X)^{-1}(X'Y). \tag{3.3}$$

A major problem with this estimator is that when the model becomes larger, through increasing the number of endogenous variables m or the number of lags p, while the number of observations T is still finite, the estimator becomes more unreliable or even uncomputable. Bayesian VARs help avoid this problem.

According to the Bayesian VAR approach, the coefficients in the model are treated as random variables, with given means and variances. The prior information about these means and variances is imposed, and we update this information with the sample observations, using Bayes' law. The end result is the posterior distribution of the coefficients with estimated means and variances. With suitable adjustment on a parameter of the model, there is no requirement on the total number of observations. This is because these observations are only used to update the prior distribution.

The main issue of implementing Bayesian VARs is about the specification of prior distribution. Litterman (1986) suggests imposing a form of prior distributions, generally referred to as Minnesota prior. The prior puts the means of the coefficients at the point that makes Y_t be a vector of univariate random walks, i.e. the means are at $A_1 = I_{m\times m}$ and $A_2, ..., A_p = \mathbf{0}_{m\times m}$. It may or may not allow for drift. The coefficients are also uncorrelated with each other, with prior variances given by:

$$Var[(A_l)_{ij}] = \begin{cases} \frac{\lambda^2}{l^2}, & i = j, \\ \pi\frac{\lambda^2}{l^2}\frac{\sigma_i^2}{\sigma_j^2}, & otherwise, \end{cases}$$

where $(A_l)_{ij}$ is the ij-th element of the l-th lag coefficient matrix A_l, $\lambda \geq 0$ is the hyperparameter determining the overall tightness of the distribution around the random walk, σ_i^2, $i = 1, ..., m$, is the variance of disturbance term of the variable $y_{i,t}$ in the VAR, and $\pi \in (0, 1]$ is another hyperparameter, reflecting the relative importance of other endogenous variables $j \neq i$ in accounting for the variation of variable i. The prior on the intercept $\mathbf{c}$ is diffuse, i.e. the variance is very high.

Recall that the variance close to zero means the distribution is very tight around the mean value. Lowering the value of λ toward zero means tightening the prior distribution toward the random walk. The term l^2 is added to reflect that the longer lagged variables should have less effects on the current variation of each variable. That is the coefficients in front of these variables should be tightened more toward zero. The hyperparameter π has the same function as l^2, but for other endogenous variables $j \neq i$. It captures the idea that in explaining the variation of a variable, own lags are more important than lags of the other variables. At last, the ratio σ_i^2/σ_j^2 is used to account for the difference in the units of measurement of different variables i and j.

For more detailed discussion on the prior variances, see Litterman (1986) or Robertson and Tallman (1999).

The original versions of Bayesian VARs assume the covariance matrix Ψ to be diagonal, fixed, and known. This is considered to be very restrictive. The prior distribution imposed in this model, as recommended by Kadiyala and Karlsson (1997), is assumed to be a Normal-(Inverted)-Wishart, which has the form:

$$\underset{km\times 1}{\mathbf{b}} \mid \underset{m\times m}{\Psi} \sim N(\underset{km\times 1}{\tilde{\mathbf{b}}}, \underset{m\times m}{\Psi} \otimes \underset{k\times k}{\tilde{\Omega}}) \quad \text{and} \quad \underset{m\times m}{\Psi} \sim iW(\underset{m\times m}{\tilde{\Psi}}, \alpha), \tag{3.4}$$

where $\mathbf{b}$ is the column vector obtained by stacking columns of the matrix B from (3.2). The degree of freedom of the inverted-Wishart distribution is set at $\alpha = m + 2$. This makes the prior mean and variance of the coefficients to be $E(\mathbf{b}) = \tilde{\mathbf{b}}$ and $Var(\mathbf{b}) = \tilde{\Psi} \otimes \tilde{\Omega}$.

Following Kadiyala and Karlsson (1997) and BGR, the parameters of the distribution in (3.4), $\tilde{\mathbf{b}}$, $\tilde{\Omega}$, and $\tilde{\Psi}$, are chosen to match the Minnesota prior. The parameter $\tilde{\mathbf{b}}$ is obtained by stacking columns of the matrix $\tilde{B}$, given by:

$$\underset{k\times m}{\tilde{B}} = \begin{bmatrix} diag(\delta_1, ..., \delta_m) \\ \cdots\cdots\cdots\cdots\cdots \\ \underset{(k-m-1)\times m}{\mathbf{0}} \\ \cdots\cdots\cdots\cdots\cdots \\ b_1 \ \cdots \ b_m \end{bmatrix},$$

where $diag(\delta_1, ..., \delta_m)$ is an $m \times m$ diagonal matrix with values $\delta_1, ..., \delta_m$ along its main diagonal, δ_i, $i = 1, ..., m$, can be either 0 or 1, and b_i, $i = 1, ..., m$, is a constant or zero. Originally, Litterman sets each δ_i, $i = 1, ..., m$, equal to 1. However, following BGR, it is more appropriate to set this value at $\delta_j = 0$ for any mean-reverting variable j.

The parameters $\tilde{\Psi}$ and $\tilde{\Omega}$ are set to be:

$$\underset{m\times m}{\tilde{\Psi}} = diag(\sigma_1^2, ..., \sigma_m^2),$$

and

$$\underset{k\times k}{\tilde{\Omega}} = \lambda^2 \cdot diag\left(\frac{1}{\sigma_1^2}, ..., \frac{1}{\sigma_m^2}; \frac{1}{2^2 \cdot \sigma_1^2}, ..., \frac{1}{2^2 \cdot \sigma_m^2}; ...; \frac{1}{p^2 \cdot \sigma_1^2}, ..., \frac{1}{p^2 \cdot \sigma_m^2}; \frac{1}{\lambda^2 \cdot \epsilon}\right), \tag{3.5}$$

where ϵ is a very small number. These parameters make the prior variance of the coefficients, $\tilde{\Psi} \otimes \tilde{\Omega}$, follow the Minnesota prior, with one exception that the hyperparameter π must be equal to 1 (See Kadiyala and Karlsson, 1997, or Robertson and Tallman, 1999, for more details). In practice, each parameter σ_i^2 is set to be the variance of the OLS residual from a univariate autoregressive model of order p of the variable $y_{i,t}$.

The posterior distribution of this model is also Normal-(Inverted)-Wishart, given by:

$$\underset{km\times 1}{\mathbf{b}} \mid \underset{m\times m}{\Psi}, \underset{T\times m}{Y} \sim N(\underset{km\times 1}{\bar{\mathbf{b}}}, \underset{m\times m}{\Psi} \otimes \underset{k\times k}{\bar{\Omega}}) \quad \text{and} \quad \underset{m\times m}{\Psi} \mid \underset{T\times m}{Y} \sim iW(\underset{m\times m}{\bar{\Psi}}, T+\alpha), \tag{3.6}$$

where $\bar{\Omega} = (\tilde{\Omega}^{-1} + X'X)^{-1}$, $\bar{\mathbf{b}}$ is obtained from stacking columns of the matrix $\bar{B}$, given by:

$$\bar{B} = (\tilde{\Omega}^{-1} + X'X)^{-1}(\tilde{\Omega}^{-1}\tilde{B} + X'Y), \tag{3.7}$$

and $\bar{\Psi}$ is given by:

$$\bar{\Psi} = Y'Y - \bar{B}'(\tilde{\Omega}^{-1} + X'X)\bar{B} + \tilde{B}'\tilde{\Omega}^{-1}\tilde{B} + \tilde{\Psi}. \tag{3.8}$$

Normally, the posterior mean $\bar{\mathbf{b}}$ is used as the point estimate of the model.

With the OLS estimator in (3.3), the estimator of the posterior mean from (3.7) can be rewritten as:

$$\bar{B} = (\tilde{\Omega}^{-1} + X'X)^{-1}(\tilde{\Omega}^{-1}\tilde{B} + (X'X)\hat{B}). \tag{3.9}$$

The estimator in (3.9) looks similar to a weighted average between the prior mean $\tilde{B}$ and the OLS estimator $\hat{B}$ of the model. It is actually a shrinkage estimator that shrinks the OLS estimate toward the prior mean, which is the random walk in this case. Since λ determines the magnitude of the matrix $\tilde{\Omega}$, setting different values of λ is equivalent to assigning different relative weights to the prior information. In one extreme, if $\lambda = 0$, we give the whole weight toward the prior information. If $\lambda = +\infty$, we give the whole weight toward the OLS estimator.

Mathematically, the main problem with the OLS estimator (3.3) is the singularity of the matrix $X'X$. The posterior mean of the Bayesian VARs as in (3.9) avoids this problem by summing the diagonal matrix $\tilde{\Omega}^{-1}$ into the matrix $X'X$. This technique produces a feasible and more reliable (less variance) estimator when the number of parameters is too large relative to the number of observations.

3.3 Performances with Different Hyperparameter Values

The main method we use in evaluating the performance of each VAR specification is the out-of-sample assessment. We follow BGR's practice closely. The data set is of Stock and Watson (2005), which have 132 monthly macroeconomic indicators starting from January 1959 to December 2003.

Let $\hat{Y}_{t+h|t}^{(\mu,\lambda)} = (\hat{y}_{1,t+h|t}^{(\mu,\lambda)} \ldots \hat{y}_{m,t+h|t}^{(\mu,\lambda)})'$ denote the point estimate of the h-steps ahead forecast obtained from the model μ with the hyperparameter value λ. The point estimate of the one-step ahead forecast is computed from:

$$\hat{Y}_{t+1|t}^{(\mu,\lambda)\prime} = X_{t+1}'\bar{B}^{(\mu,\lambda)}, \tag{3.10}$$

where $\bar{B}^{(\mu,\lambda)}$ is the posterior mean of the coefficients from the model μ with the hyperparameter value λ. For the case of $p > h > 1$ that we consider, we can recursively construct a matrix of independent variables $X_{t+h|t}^{(\mu,\lambda)}$, given by:

$$X_{t+h|t}^{(\mu,\lambda)} = (\hat{Y}_{t+h-1|t}^{(\mu,\lambda)\prime}, \ldots, \hat{Y}_{t+1|t}^{(\mu,\lambda)\prime}, Y_t', \ldots, Y_{t+h-p}', 1)', \tag{3.11}$$

using the forecasts $\hat{Y}_{t+h-1|t}^{(\mu,\lambda)}, \ldots, \hat{Y}_{t+1|t}^{(\mu,\lambda)}$ and the sample observations $Y_t, \ldots, Y_{t+h-p}$. The point estimate of the h-steps forecast, then, is computed from:

$$\hat{Y}_{t+h|t}^{(\mu,\lambda)\prime} = X_{t+h|t}^{(\mu,\lambda)\prime}\bar{B}^{(\mu,\lambda)}. \tag{3.12}$$

The random walk is used as our benchmark model. The estimator can be obtained by setting λ equal to 0, which makes the h-steps ahead forecast from this model to be the same across all model specifications μ. We use $\hat{Y}_{t+h|t}^{(0)}$ to denote the h-steps ahead forecast from this benchmark model. Most of the parameters δ_i are set to be 1, except for some stationary variables specified by BGR, of which δ_i are set to be 0 (See the last column of the Appendix D).

The out-of-sample assessment is conducted for forecast horizons h equal to $1, 3, 6$, and 12. Let t_0 and t_1 denote the position of *January 1971* and *December 2003* in the data set. For each forecast horizon h, we compute $\hat{Y}^{(\mu,\lambda)}_{t+h|t}$ in each period $t = t_0 - h, ..., t_1 - h$ (396 times). The order of the VAR is $p = 13$. The parameters and posterior mean in each model for each t are computed from the most recent 10 years of sample observations up to time t (Rolling scheme, 120 observations). We set the small number ϵ, the parameter governing prior variances of the constant terms in the matrix $\tilde{\Omega}$ in (3.5), to be 10^{-10}.

The forecasting performance is measured in terms of out-of-sample Mean Squared Forecast Error (MSFE). For the model μ, the value λ, the forecast horizon h, and the variable i, we have:

$$MSFE^{(\mu,\lambda)}_{i,h} = \frac{1}{t_1 - t_0 + 1} \sum_{t=t_0-h}^{t_1-h} \left(y_{i,t+h} - \hat{y}^{(\mu,\lambda)}_{i,t+h|t} \right)^2. \tag{3.13}$$

The results are reported for MSFE in relative to one of the benchmark model (Random walk with drift), given by:

$$RMSFE^{(\mu,\lambda)}_{i,h} = \frac{MSFE^{(\mu,\lambda)}_{i,h}}{MSFE^{(0)}_{i,h}}. \tag{3.14}$$

A number smaller than 1 for $RMSFE^{(\mu,\lambda)}_{i,h}$ implies that the model μ with value λ performs better than the random walk.

The variables of interest i are 1) employment (EMPL), measured by the number of employees on non-farm payrolls, 2) consumer price index (CPI) representing the price level, and 3) the Federal Fund Rate (FFR) representing the monetary instrument.

Following BGR, there are 4 VAR specifications μ which are:

1. $SMALL$. There are only 3 variables of interest; 1) EMPL, 2) FFR, and 3) CPI.

2. CEE. This is the model of Christiano et al. (1999). There are 7 variables, 3 as in $SMALL$, and 4) index of sensitive material prices, 5) non-borrowed reserves, 6) total reserves, and 7) M2 money stock.

3. $MEDIUM$. There are 20 variables, 7 as in CEE, and 8) Personal Income, 9) Real Consumption, 10) Industrial Production, 11) Capacity Utilization, 12) Unemployment Rate, 13) Housing Starts, 14) Producer Price Index, 15) Personal Consumption Expenditures Price Deflator, 16) Average Hourly Earnings, 17) M1 money stock, 18) Standard and Poor's Price Index, 19) Yields on 10 year U.S. Treasury Bond, and 20) effective exchange rate.

4. $LARGE$ This specification includes all indicators in the data set, except spot market price index of all commodities (PSCCOM).

We report our first out-of-sample assessment result in Table 3.1, using the same hyperparameter values as in BGR. That is $\lambda = \infty$ for $\mu = SMALL$, $\lambda = 0.262$ for $\mu = CEE$, $\lambda = 0.108$ for $\mu = MEDIUM$, and $\lambda = 0.035$ for $\mu = LARGE$. This result is qualitatively similar to Table 1 of BGR. It can be seen clearly that larger models perform better than smaller ones.

BGR assign hyperparameter values to keep the in-sample fit of all models in the pre-evaluation period to be the same, for the forecast horizon $h = 1$. Specifically, let T_0 denote the position of *December 1969* in the data set. Define the in-sample 1-step ahead mean squared forecast errors ($msfe$) for a model

Table 3.1: **BVARs different λ, Out-of-Sample Relative MSFE, 1971 - 2003**

		SMALL	*CEE*	*MEDIUM*	*LARGE*
$h = 1$	EMPL	1.02	0.65	0.54	0.45
	FFR	1.65	0.90	0.79	0.75
	CPI	0.81	0.55	0.51	0.51
$h = 3$	EMPL	0.85	0.63	0.50	0.37
	FFR	1.57	1.12	0.96	0.92
	CPI	0.60	0.43	0.40	0.40
$h = 6$	EMPL	0.90	0.79	0.66	0.51
	FFR	1.84	1.30	1.31	1.24
	CPI	0.59	0.44	0.37	0.40
$h = 12$	EMPL	0.84	0.96	0.87	0.81
	FFR	2.48	1.49	1.56	1.80
	CPI	0.74	0.60	0.44	0.45
λ		∞	0.262	0.108	0.035

μ, a hyperparameter value λ, and a variable i as:

$$msfe_i^{(\mu,\lambda)} = \frac{1}{T_0 - p - 1} \sum_{t=p}^{T_0 - 1} (\hat{y}_{i,t+1|t}^{(\mu,\lambda)} - y_{i,t+1})^2.$$

Note that $\hat{y}_{i,t+1|t}^{(\mu,\lambda)}$ is the in-sample forecast (Estimated value) for $y_{i,t+1}$ within the period from *January 1960* $(t = 1)$ to *December 1969* $(t = T_0)$.

Next, estimate the unrestricted OLS VAR of the *SMALL* model using the data from *January 1960* to *December 1969*, and figure out the in-sample fit (Fit), given by:

$$Fit = \frac{1}{3} \sum_{i \in \mathcal{I}} \frac{msfe_i^{(\mu,\lambda)}}{msfe_i^{(0)}} \Bigg|_{\mu = SMALL,\ \lambda = +\infty},$$

where $\mathcal{I} = \{EMPL, FFR, CPI\}$ is the set of variables of interest.

At last, for each model $\mu \neq SMALL$, determine from grid search the hyperparameter $\lambda^{(\mu,Fit)}$ that gives the in-sample fit of the model closest to the in-sample fit of the unrestricted OLS VAR. Specifically, the hyperparameter

$\lambda^{(\mu,Fit)}$ can be defined as:

$$\lambda^{(\mu,Fit)} = \underset{\lambda}{\text{argmin}} \left| Fit - \frac{1}{3} \sum_{i \in \mathcal{I}} \frac{msfe_i^{(\mu,\lambda)}}{msfe_i^{(0)}} \right| .$$

We see that the way BGR set the hyperparameter values biasses against small models. First, note that the *SMALL* Bayesian VAR of BGR is actually the unrestricted OLS VAR. This is because the hyperparameter of the model is set at $\lambda = +\infty$. The *SMALL* model does not benefit from shrinkage estimation at all. Next, observe that larger models will be assigned with lower values of the hyperparameter λ. This is a usual result as a larger OLS model provides a better in-sample fit to the sample observations. To set the in-sample fit at a given level, this model must be pulled away more from its OLS estimate. However, since the shrinkage estimator improves the forecasting performance of a model by avoiding the problem of overfitting into the sample observations[1], this way of assigning the hyperparameter values provides more benefits to larger models.

To show this empirically, we set up a new out-of-sample assessment that assigns the same hyperparameter value across all model specifications. Each hyperparameter value $\lambda = 0.035$, 0.108, and 0.262 is applied to all specifications each time in this assessment. Everything else stays the same. Table 3.2 reports the relative MSFE under this new assessment.

Comparing Table 3.2 to Table 3.1, we can see an obvious improvement of the performances of small models. Even the smallest model can benefit from shrinkage estimation. The *SMALL* model is a 3-variable VAR with 13 lags, which results in 40 coefficients to be estimated per equation including the

[1] Zha (1998) provides a good discussion on this point.

Table 3.2: **BVARs same λ, Out-of-Sample Relative MSFE, 1971 - 2003**

$\lambda = 0.035$		*SMALL*	*CEE*	*MEDIUM*	*LARGE*
$h = 1$	EMPL	0.64	0.61	0.52	0.45
	FFR	1.00	0.95	0.87	0.75
	CPI	0.57	0.51	0.51	0.51
$h = 3$	EMPL	0.63	0.56	0.48	0.37
	FFR	1.06	1.03	1.02	0.92
	CPI	0.47	0.38	0.38	0.40
$h = 6$	EMPL	0.73	0.63	0.58	0.51
	FFR	1.11	1.11	1.28	1.24
	CPI	0.46	0.34	0.34	0.40
$h = 12$	EMPL	0.93	0.71	0.72	0.81
	FFR	1.22	1.27	1.56	1.80
	CPI	0.51	0.40	0.38	0.45
$\lambda = 0.108$		*SMALL*	*CEE*	*MEDIUM*	*LARGE*
$h = 1$	EMPL	0.54	0.59	0.54	0.51
	FFR	0.96	0.86	0.79	0.75
	CPI	0.55	0.51	0.51	0.55
$h = 3$	EMPL	0.49	0.55	0.50	0.40
	FFR	1.08	0.94	0.96	0.94
	CPI	0.47	0.39	0.40	0.46
$h = 6$	EMPL	0.55	0.66	0.66	0.54
	FFR	1.18	1.03	1.31	1.33
	CPI	0.48	0.37	0.37	0.47
$h = 12$	EMPL	0.65	0.76	0.87	0.96
	FFR	1.29	1.21	1.56	1.86
	CPI	0.55	0.47	0.44	0.59
$\lambda = 0.262$		*SMALL*	*CEE*	*MEDIUM*	*LARGE*
$h = 1$	EMPL	0.57	0.65	0.66	0.65
	FFR	0.92	0.90	0.85	0.82
	CPI	0.55	0.55	0.54	0.62
$h = 3$	EMPL	0.53	0.63	0.63	0.47
	FFR	1.14	1.12	1.04	1.05
	CPI	0.48	0.43	0.44	0.51
$h = 6$	EMPL	0.60	0.79	0.82	0.62
	FFR	1.29	1.30	1.51	1.49
	CPI	0.50	0.44	0.39	0.52
$h = 12$	EMPL	0.65	0.96	1.10	1.09
	FFR	1.50	1.49	1.85	1.96
	CPI	0.59	0.60	0.49	0.69

constant term. It is estimated with 120 observations each time. The smallest amount of shrinkage in this experiment, $\lambda = 0.262$, can remarkably improve the forecasting performances of this model.

The obvious improvement of the forecasting performances for larger models also disappears. This point is apparent for the hyperparameter values $\lambda = 0.108$ and $\lambda = 0.262$. Especially, for the case of fixing the value at $\lambda = 0.262$, the $SMALL$ model outperforms other larger models in the overall picture. This suggests that the value of hyperparameter λ must be chosen more carefully.

3.4 Performances with Suitable Hyperparameter Values

In this section, we consider a procedure for choosing a "suitable" hyperparameter value for each model based on a training sample. We choose the value that leads to the minimum relative MSFE of the variables of interest in an out-of-sample assessment. The observations up to *December 1980* are employed to figure out each suitable hyperparameter value for a given VAR specification μ and a given forecasting span h. After that we will assess each model with this hyperparameter value, using our evaluation period from *January 1981* to *December 2003*.

In searching for a suitable hyperparameter value, let τ_0 and τ_1 denote the position of *January 1971* and *December 1980* in the data set, respectively. For a forecasting span $h \in \{1, 3, 6, 12\}$ and an arbitrary hyperparameter value $\tilde{\lambda}$, we can compute a forecast $\hat{Y}^{(\mu,\tilde{\lambda})}_{\tau+h|\tau}$ for each period $\tau = \tau_0 - h, ..., \tau_1 - h$ (120 times), using the same setting as in the previous section ($p = 13$, rolling scheme with 120 observations, $\epsilon = 10^{-10}$). These point forecasts can be used to compute $MSFE^{(\mu,\tilde{\lambda})}_{i,h}$ and $RMSFE^{(\mu,\tilde{\lambda})}_{i,h}$ from (3.13) and (3.14), for each variable of interest i. Let $TV_h^{(\mu,\lambda)} \equiv RMSFE_{EMPL,h}^{(\mu,\lambda)} + RMSFE_{CPI,h}^{(\mu,\lambda)} + RMSFE_{FFR,h}^{(\mu,\lambda)}$ denote our target variable, which is the sum of relative MSFE of the variables of interest. Note that each relative MSFE does not depend on the unit of measurement of each variable, since it is in relative term. We find the suitable hyperparameter value λ_h^μ for each specification μ and each forecast horizon h from grid search such that:

$$\lambda_h^\mu = \underset{\tilde{\lambda}}{\text{argmin}} \; TV_h^{(\mu,\tilde{\lambda})}. \tag{3.15}$$

Since there is no natural upper bound for the hyperparameter λ, a good grid search for λ_h^μ should cover a wide range of possible values between 0 and $+\infty$. To avoid this, we calculate the derivative $\partial \, TV_h^{(\mu,\lambda)}/\partial\lambda$ to help search for each hyperparameter value λ_h^μ. This allows us to search for the value λ_h^μ in steps, and helps reduce the task.

From (3.13) and (3.14), we have:

$$\frac{\partial \, TV_h^{(\mu,\lambda)}}{\partial\lambda} = \frac{\partial \, MSFE_{EMPL,h}^{(\mu,\lambda)}/\partial\lambda}{MSFE_{EMPL,h}^{(0)}} + \frac{\partial \, MSFE_{CPI,h}^{(\mu,\lambda)}/\partial\lambda}{MSFE_{CPI,h}^{(0)}} + \frac{\partial \, MSFE_{FFR,h}^{(\mu,\lambda)}/\partial\lambda}{MSFE_{FFR,h}^{(0)}}, \tag{3.16}$$

where for each variable $i \in \{EMPL, CPI, FFR\}$,

$$\begin{aligned} \frac{\partial \, MSFE_{i,h}^{(\mu,\lambda)}}{\partial\lambda} &= \frac{1}{\tau_1 - \tau_0 + 1} \sum_{t=\tau_0-h}^{\tau_1-h} \frac{\partial}{\partial\lambda} \left(y_{i,t+h} - \hat{y}_{i,t+h|t}^{(\mu,\lambda)} \right)^2, \\ &= \frac{1}{\tau_1 - \tau_0 + 1} \sum_{t=\tau_0-h}^{\tau_1-h} -2 \left(y_{i,t+h} - \hat{y}_{i,t+h|t}^{(\mu,\lambda)} \right) \frac{\partial \, \hat{y}_{i,t+h|t}^{(\mu,\lambda)}}{\partial\lambda}. \end{aligned} \tag{3.17}$$

Given an $m \times n$ matrix Z, we use $\partial Z/\partial\lambda$ to denote the gradient matrix of Z with respect to λ. The gradient matrix $\partial Z/\partial\lambda$ has the same dimension as Z

with $\partial z_{ij}/\partial\lambda$ as its ij-th element. This is the same for a gradient vector $\partial \mathbf{z}/\partial\lambda$ of an $m \times 1$ vector $\mathbf{z}$. The value of $\partial\, \hat{y}^{(\mu,\lambda)}_{i,t+h|t}/\partial\lambda$ in (3.17) can be taken from the gradient vector $\partial\, \hat{Y}^{(\mu,\lambda)}_{t+h|t}/\partial\lambda$. From (3.12), the gradient vector $\partial\, \hat{Y}^{(\mu,\lambda)}_{t+h|t}/\partial\lambda$ can be written as:

$$\left(\frac{\partial\, \hat{Y}^{(\mu,\lambda)}_{t+h|t}}{\partial\lambda}\right)' = \left(\frac{\partial\, X^{(\mu,\lambda)}_{t+h|t}}{\partial\lambda}\right)' \bar{B}^{(\mu,\lambda)} + X^{(\mu,\lambda)\prime}_{t+h|t}\left(\frac{\partial\, \bar{B}^{(\mu,\lambda)}}{\partial\lambda}\right), \qquad (3.18)$$

where $\partial\bar{B}^{(\mu,\lambda)}/\partial\lambda$ is given by[2]:

$$\frac{\partial\bar{B}^{(\mu,\lambda)}}{\partial\lambda} = \left(\tilde{\Omega}^{(\mu,\lambda)\,-1} + X^{(\mu)\prime}X^{(\mu)}\right)^{-1}\left(\frac{\partial\, \tilde{\Omega}^{(\mu,\lambda)\,-1}}{\partial\lambda}\right)\left(\tilde{B}^{(\mu)} - \bar{B}^{(\mu,\lambda)}\right). \qquad (3.19)$$

We can compute the gradient vector $\partial\, \hat{Y}^{(\mu,\lambda)}_{t+h|t}/\partial\lambda$ recursively, using (3.18) and, according to (3.10) and (3.11), the following equations:

$$\left(\frac{\partial\, \hat{Y}^{(\mu,\lambda)}_{t+1|t}}{\partial\lambda}\right)' = X'_{t+1}\left(\frac{\partial\, \bar{B}^{(\mu,\lambda)}}{\partial\lambda}\right), \qquad (3.20)$$

and

$$\frac{\partial\, X^{(\mu,\lambda)}_{t+h|t}}{\partial\lambda} = \left(\left(\frac{\partial\, \hat{Y}^{(\mu,\lambda)}_{t+h-1|t}}{\partial\lambda}\right)' \cdots \left(\frac{\partial\, \hat{Y}^{(\mu,\lambda)}_{t+1|t}}{\partial\lambda}\right)' \mathbf{0}_{1\times(k-mh+m)}\right)'. \qquad (3.21)$$

We perform grid search in finding the values of λ^{μ}_{h} for each forecasting horizon h and each model specification μ. We search for the suitable hyperparameter λ^{μ}_{h} with 3 decimal places, which makes our grid search composed of 4 steps as follow[3]:

[2] See the Appendix B for the derivation of $\partial\bar{B}^{(\mu,\lambda)}/\partial\lambda$ in (3.19).

[3] It is possible that our grid search may not return the optimal hyperparameter λ^{μ}_{h} as defined in (3.15), if the function $TV^{(\mu,\lambda)}_{h}$ is not smooth. Regarding this problem, we have tried minimizing the function with respect to the value of λ for each forecast horizon $h = 1, 3, 6, 12$ of the *SMALL* model, using the add-on application OPTMUM in GAUSS. It

1. Calculate the values of target variable $TV_h^{(\mu,\lambda)}$ and gradient $\partial TV_h^{(\mu,\lambda)}/\partial\lambda$ for each of 11 values of λ, which are 0.001 and $1, 2, ..., 10$. Figure out the possible region of the hyperparameter value λ_h^μ [4]. Let λ_{s1} denote the lower bound of this region.

2. Calculate the values of target variable $TV_h^{(\mu,\lambda)}$ and gradient $\partial TV_h^{(\mu,\lambda)}/\partial\lambda$ for each of 9 values of λ, which are $\lambda_{s1}+0.1, \lambda_{s1}+0.2, ..., \lambda_{s1}+0.9$. Figure out the possible region of the hyperparameter value λ_h^μ. Let λ_{s2} denote the lower bound of this region.

3. Calculate the values of target variable $TV_h^{(\mu,\lambda)}$ and gradient $\partial TV_h^{(\mu,\lambda)}/\partial\lambda$ for each of 9 values of λ, which are $\lambda_{s2} + 0.01, \lambda_{s2} + 0.02, ..., \lambda_{s2} + 0.09$. Figure out the possible region of the hyperparameter value λ_h^μ. Let λ_{s3} denote the lower bound of this region.

4. Calculate the values of target variable $TV_h^{(\mu,\lambda)}$ for 9 values of λ, which are $\lambda_{s3} + 0.001, \lambda_{s3} + 0.002, ..., \lambda_{s3} + 0.009$. The suitable hyperparameter value λ_h^μ is the one associated with the minimum value of $TV_h^{(\mu,\lambda)}$ from these 4 steps.

Table 3.3 reports the optimal hyperparameter λ_h^μ with the associated values of target variable $TV_h^{(\mu,\lambda)}$ and gradient, for each forecast horizon h and each model specification μ. The details of grid search can be found in Appendix C.

Using the suitable hyperparameter values λ_h^μ from Table 3.3, we perform the out-of-sample assessment. Let t_0 and t_1 denote the position of *January 1981* and *December 2003*, respectively, in the data set. We compute the forecast

returns a similar result to our grid search. The use of this program, however, is not practical for larger models, as it consumes a lot of time even for the smallest model.

[4] This should be the region that has a negative gradient at its lower bound. This tells that the values in the region will generate smaller values of $TV_h^{(\mu,\lambda)}$ than one at the lower bound.

Table 3.3: λ_h^μ from grid search and $TV_h^{(\mu,\lambda)}$, 1971 - 1980

		$SMALL$	CEE	$MEDIUM$	$LARGE$
$h=1$	λ_h^μ	0.130	0.129	0.096	0.053
	$TV_h^{(\mu,\lambda)}$	1.950	1.849	1.665	1.598
	gradient	0.008	0.001	-0.000	0.029
$h=3$	λ_h^μ	0.111	0.143	0.117	0.072
	$TV_h^{(\mu,\lambda)}$	1.876	1.739	1.695	1.614
	gradient	0.001	0.004	0.024	-0.008
$h=6$	λ_h^μ	0.130	0.134	0.017	0.059
	$TV_h^{(\mu,\lambda)}$	1.887	1.852	2.191	2.288
	gradient	0.007	0.004	-0.350	-0.033
$h=12$	λ_h^μ	0.102	0.049	0.020	0.006
	$TV_h^{(\mu,\lambda)}$	1.912	1.896	2.254	2.483
	gradient	-0.000	-0.020	-0.672	-1.646

$\hat{Y}_{t+h|t}^{(\mu,\lambda_h^\mu)}$ in each period $t = t_0 - h, ..., t_1 - h$ (276 times) with VAR of order $p = 13$, using the most recent 10 years of observations (Rolling scheme, 120 observations), and the parameter ϵ at 10^{-10}. These forecasts are used to calculate the relative MSFE in (3.14), for 3 variables of interest $i =$ EMPL, FFR, and CPI. Table 3.4 reports the result of this assessment, with the associated values of $TV_h^{(\mu,\lambda_h^\mu)} = \sum_{i\in\mathcal{I}} RMSFE_i^{(\mu,\lambda_h^\mu)}$ and λ_h^μ.

Since our evaluation period has been changed from the previous section, we also construct Table 3.5 for the purpose of comparison. In this table, we use the same setting as in Table 3.1 of the previous section, but the evaluation period has been changed to one from *January 1981* to *December 2003*.

According to Table 3.4, the $LARGE$ model performs best in the overall picture. This supports the finding of BGR that larger Bayesian VARs perform better than smaller models in forecasting the three key macroeconomic variables. However, we see a significant difference between Table 3.4 and Table 3.5.

With results similar to Table 3.5, it looks like adding more variables into the

Table 3.4: **BVARs with λ_h^μ, Out-of-Sample Relative MSFE, 1981 - 2003**

		SMALL	*CEE*	*MEDIUM*	*LARGE*
$h=1$	EMPL	0.53	0.62	0.53	0.49
	FFR	0.96	0.85	0.93	0.80
	CPI	0.62	0.60	0.57	0.53
	$TV_h^{(\mu,\lambda_h^\mu)}$	2.104	2.067	2.025	1.822
	λ_h^μ	0.130	0.129	0.096	0.053
$h=3$	EMPL	0.42	0.56	0.44	0.37
	FFR	1.23	1.06	1.13	0.95
	CPI	0.59	0.52	0.54	0.51
	$TV_h^{(\mu,\lambda_h^\mu)}$	2.234	2.145	2.107	1.830
	λ_h^μ	0.111	0.143	0.117	0.072
$h=6$	EMPL	0.53	0.77	0.63	0.49
	FFR	1.47	1.12	1.17	1.05
	CPI	0.62	0.51	0.43	0.50
	$TV_h^{(\mu,\lambda_h^\mu)}$	2.612	2.391	2.225	2.045
	λ_h^μ	0.130	0.134	0.017	0.059
$h=12$	EMPL	0.72	0.91	0.82	0.69
	FFR	1.47	1.24	1.75	1.75
	CPI	0.78	0.54	0.47	0.52
	$TV_h^{(\mu,\lambda_h^\mu)}$	2.966	2.696	3.045	2.970
	λ_h^μ	0.102	0.049	0.020	0.006

Table 3.5: **BVARs different λ, Out-of-Sample Relative MSFE, 1981 - 2003**

		SMALL	*CEE*	*MEDIUM*	*LARGE*
$h=1$	EMPL	0.81	0.68	0.54	0.47
	FFR	1.71	0.99	0.94	0.78
	CPI	0.83	0.64	0.57	0.54
	$TV_h^{(\mu,\lambda)}$	3.360	2.313	2.049	1.794
$h=3$	EMPL	0.67	0.65	0.43	0.32
	FFR	1.74	1.43	1.12	0.89
	CPI	0.73	0.58	0.53	0.48
	$TV_h^{(\mu,\lambda)}$	3.132	2.649	2.077	1.689
$h=6$	EMPL	0.89	0.95	0.62	0.45
	FFR	2.44	1.58	1.36	1.00
	CPI	0.77	0.59	0.48	0.47
	$TV_h^{(\mu,\lambda)}$	4.105	3.117	2.461	1.925
$h=12$	EMPL	1.04	1.33	0.93	0.82
	FFR	3.18	1.63	1.40	1.64
	CPI	1.04	0.80	0.51	0.56
	$TV_h^{(\mu,\lambda)}$	5.266	3.761	2.840	3.014
λ		∞	0.262	0.108	0.035

VAR helps to significantly improve its forecasting performances. The models with 7 and 20 variables perform much better than the 3-variable model. Since Bayesian or shrinkage estimation allows us to use all available information in making forecasts, adding as many data as possible like the *LARGE* model helps further improve the forecasting performances.

However, Table 3.4 shows that this impression is false. This is a result of allowing no shrinkage at all for the *SMALL* model. If we use Bayesian or shrinkage estimation with the *SMALL* model, the improvement of larger VARs over the 3-variable VAR becomes minimal. Specifically, the 7-variable and 20-variable models do not seem to have a clear edge over the 3-variable model, and the improvement of the 131-variable VAR is much less pronounced than what Table 3.5 implies. This also features after we have tried to make the most out of each model given our pre-evaluation period.

Table 3.6: λ_h^μ **from grid search and** $TV_h^{(\mu,\lambda)}$**, 1971 - 1990**

		SMALL	*CEE*	*MEDIUM*	*LARGE*
$h=1$	λ_h^μ	0.164	0.102	0.078	0.044
	$TV_h^{(\mu,\lambda)}$	2.003	1.920	1.793	1.666
	gradient	-0.002	0.007	0.010	-0.069
$h=3$	λ_h^μ	0.089	0.085	0.066	0.048
	$TV_h^{(\mu,\lambda)}$	2.044	1.857	1.849	1.691
	gradient	-0.005	0.003	-0.036	-0.063
$h=6$	λ_h^μ	0.059	0.066	0.022	0.043
	$TV_h^{(\mu,\lambda)}$	2.195	1.986	2.246	2.199
	gradient	-0.030	0.014	0.033	-0.002
$h=12$	λ_h^μ	0.073	0.055	0.047	0.005
	$TV_h^{(\mu,\lambda)}$	2.407	2.227	2.691	2.772
	gradient	0.018	0.064	0.126	12.717

3.5 Repeated Calculations of Hyperparameter Values

It can be the case that the optimal hyperparameter value λ_h^μ varies with time. Allowing some changes for the value may improve the forecasting performance of each Bayesian VAR. In this section, we allow this change every 10 years. We repeat our practice in the previous section; finding the suitable hyperparameter value after we have an additional 10 years of observations.

Table 3.6 reports the suitable hyperparameter value λ_h^μ with the associated values of target variable and gradient for each forecast horizon h and each model μ, using the observations from *January 1971* to *December 1990.* Table 3.7 reports the same values, using the observations from *January 1971* to *December 2000.* The suitable hyperparameter values reported in Table 3.6 and Table 3.7 look different from ones in Table 3.3 of the previous section. However, the values are relatively similar in these two tables.

Next, we use the hyperparameter values from Table 3.3, Table 3.6, and Table 3.7 in assessing the out-of-sample forecasting performances of our Bayesian

Table 3.7: λ_h^μ **from grid search and** $TV_h^{(\mu,\lambda)}$**, 1971 - 2000**

		SMALL	*CEE*	*MEDIUM*	*LARGE*
$h = 1$	λ_h^μ	0.168	0.101	0.077	0.043
	$TV_h^{(\mu,\lambda)}$	2.030	1.963	1.828	1.704
	gradient	-0.001	-0.007	0.015	-0.008
$h = 3$	λ_h^μ	0.098	0.083	0.062	0.046
	$TV_h^{(\mu,\lambda)}$	2.045	1.877	1.852	1.699
	gradient	-0.002	-0.018	-0.018	0.096
$h = 6$	λ_h^μ	0.071	0.064	0.022	0.041
	$TV_h^{(\mu,\lambda)}$	2.199	2.004	2.212	2.176
	gradient	0.000	-0.010	-0.225	-0.028
$h = 12$	λ_h^μ	0.086	0.054	0.043	0.005
	$TV_h^{(\mu,\lambda)}$	2.434	2.276	2.690	2.724
	gradient	-0.017	-0.004	0.134	4.146

VARs. The values from Table 3.3 are used to make forecasts from *January 1981* to *December 1990*. Ones from Table 3.6 are used for the forecasts from *January 1991* to *December 2000*, and ones from Table 3.7 for *January 2001* to *December 2003*. Table 3.8 reports the values of out-of-sample relative MSFE from this exercise. The variable TV_h^μ represents the sum of relative MSFE of our variables of interest for model μ and forecast horizon h. Comparing the results in Table 3.8 and Table 3.4, we can see that our exercise can just marginally improve the forecasting performances of the models.

3.6 An Updating Scheme for the Hyperparameter

Another way to allow changes in hyperparameter values is to use an updating scheme that is sensitive to previous forecasting performances of the model. In this section, we apply an updating scheme that makes use of each additional observation in determining whether to change the hyperparameter value of a model. Such adaptive schemes will only improve forecasting performance if the underlying data generating process (DGP) is changing through time. What

Table 3.8: **BVARs with Varied λ, Out-of-Sample Relative MSFE, 1981 - 2003**

		SMALL	*CEE*	*MEDIUM*	*LARGE*
$h=1$	EMPL	0.53	0.62	0.53	0.49
	FFR	0.96	0.85	0.93	0.80
	CPI	0.62	0.60	0.57	0.53
	TV_h^μ	2.102	2.063	2.025	1.821
$h=3$	EMPL	0.43	0.57	0.45	0.37
	FFR	1.23	1.05	1.12	0.95
	CPI	0.59	0.51	0.52	0.50
	TV_h^μ	2.241	2.133	2.090	1.819
$h=6$	EMPL	0.58	0.79	0.61	0.48
	FFR	1.46	1.08	1.16	1.05
	CPI	0.60	0.49	0.44	0.50
	TV_h^μ	2.650	2.362	2.202	2.036
$h=12$	EMPL	0.75	0.91	0.78	0.71
	FFR	1.46	1.24	1.73	1.76
	CPI	0.77	0.54	0.51	0.52
	TV_h^μ	2.987	2.693	3.018	2.987

the "optimal" adaptive scheme will be depends on how the underlying DGP is changing over time.

Although there is a strong belief that there has been structural change in the economic system during our sample period, there is no precise information about how the parameters have changed. Therefore, instead of making an arbitrary assumption about mechanism that governs such changes and then deriving the optimal adaptive scheme for that mechanism, we consider an adaptive scheme that makes good sense to us. From the practice of forecasting, we know that adaptive schemes that give very high weight to new information often chase noises and do not perform very well. Hence, we consider the following scheme.

Let t_0 and t_1 represent the positions of *January 1981* and *December 2003*, respectively. We start from the suitable hyperparameter value of each model and each forecast horizon from Table 3.3. We use λ^{μ}_{h,t_0} to denote this initial hyperparameter value. Let $\lambda^{\mu}_{h,T}$ denote the value used in a given period T. At the start of each period $T \in [t_0, t_1]$, we use a model μ in making a forecast $\hat{Y}^{\mu,\lambda^{\mu}_{h,T}}_{T|T-h}$. At the end of the period T, after realizing the actual data Y_T, we calculate the square forecast error from the Bayesian VAR μ:

$$SFE^{(\mu,\lambda^{\mu}_{h,T})}_{i,h;T} \equiv (y_{i,T} - \hat{y}^{\mu,\lambda^{\mu}_{h,T}}_{T|T-h})^2, \tag{3.22}$$

as well as the square forecast error from the benchmark model $SFE^{(0)}_{i,h;T}$ for each variable of interest $i \in \mathcal{I}$.

We also calculate at this point in time the indicators:

$$INDC^{(\mu,\lambda)}_{h;T} \equiv \sum_{i\in\mathcal{I}} \frac{SFE^{(\mu,\lambda)}_{i,h;T}}{\sum_{t=t_0}^{T} SFE^{(0)}_{i,h;t} + (276 + t_0 - T)SFE^{(0)}_{i,h;T}}, \tag{3.23}$$

for 3 values of the hyperparameter λ, which are $\lambda^{\mu}_{h,T}$, $\lambda^{\mu}_{h,T} + 0.001$, and $\lambda^{\mu}_{h,T} - 0.001$[5]. We use the indicator $INDC^{(\mu,\lambda)}_{h;T}$ to approximate the marginal increase in the sum of relative MSFE from using different values of λ at time T. Observe that the term $\sum_{t=t_0}^{T} SFE^{(0)}_{i,h;t}$ in the denominator increases as T increases. We put the term $(276+t_0-T)SFE^{(0)}_{i,h;T}$[6] into the denominator as well to make the value of $INDC^{(\mu,\lambda)}_{h;T}$ relatively stable along the time T. Otherwise, the value $\lambda^{\mu}_{h,T}$ will be more fluctuated for small T and very stable for larger T, if we fix a constant threshold as in the following.

Among these 3 hyperparameter values, we first choose the one that gives the minimum value of $INDC^{(\mu,\lambda)}_{h;T}$. If it is $\lambda = \lambda^{\mu}_{h,T}$, we also use this value as $\lambda^{\mu}_{h,T+1}$ in the next period. Otherwise, for $\lambda \in \{\lambda^{\mu}_{h,T} + 0.001, \lambda^{\mu}_{h,T} - 0.001\}$, if

[5] We use stepsize equal to 0.001 in every case, except for the case of *LARGE* model with $h = 12$ that we use step size at 0.0005.

[6] Recall that $276 = t_1 - t_0 - 1$ is the total number of repetitions in our out-of-sample exercise.

Table 3.9: **BVARs with Varied λ, Out-of-Sample Relative MSFE,** 1981 – 2003

		$SMALL$	CEE	$MEDIUM$	$LARGE$
$h=1$	EMPL	0.53	0.62	0.52	0.48
	FFR	0.95	0.84	0.91	0.78
	CPI	0.61	0.60	0.57	0.53
	TV_h^{μ}	2.098	2.053	2.003	1.798
	$\lambda_{h,0}^{\mu}$	0.130	0.129	0.096	0.053
$h=3$	EMPL	0.42	0.56	0.43	0.36
	FFR	1.23	1.04	1.12	0.91
	CPI	0.59	0.52	0.53	0.48
	TV_h^{μ}	2.237	2.124	2.079	1.747
	$\lambda_{h,0}^{\mu}$	0.111	0.143	0.117	0.072
$h=6$	EMPL	0.53	0.77	0.44	0.47
	FFR	1.47	1.11	1.04	1.02
	CPI	0.62	0.50	0.50	0.49
	TV_h^{μ}	2.613	2.383	1.978	1.975
	$\lambda_{h,0}^{\mu}$	0.130	0.134	0.017	0.059
$h=12$	EMPL	0.72	0.92	0.81	0.57
	FFR	1.47	1.24	1.63	1.54
	CPI	0.78	0.54	0.46	0.44
	TV_h^{μ}	2.966	2.698	2.908	2.548
	$\lambda_{h,0}^{\mu}$	0.102	0.049	0.020	0.006

$(INDC_{h;T}^{(\mu,\lambda_{h,T}^{\mu})} - INDC_{h;T}^{(\mu,\lambda)})$ is higher than 0.0001 we use this new value as $\lambda_{h,T+1}^{\mu}$ in the next period. Observe that we can increase the fluctuation of the hyperparameter value $\lambda_{h,T}^{\mu}$ by increasing the stepsize (Currently at 0.001) and lowering the threshold value (Currently at 0.0001). Actually, we have made some experiments with a range of threshold values and stepsizes. The setting reported here yields the best results. Note also that in this process we use the information up to period T to figure out the hyperparameter value $\lambda_{h,T+1}^{\mu}$ that will be applied in the next time period $T+1$.

At the end of the exercise, we calculate relative MSFE for each variable of interest $i \in \mathcal{I}$ from the square forecast errors $SFE(\mu, \lambda_{h,T}^{\mu})_{i,h;T}$ calculated at the start of each period T. The relative MSFE can be written as:

$$RMSFE_{i,h}^{(\mu)} = \frac{\sum_{t=t_0}^{t_1} SFE_{i,h;t}^{(\mu,\lambda_{h,t}^{\mu})}}{\sum_{t=t_0}^{t_1} SFE_{i,h;t}^{(0)}} \tag{3.24}$$

Table 3.9 reports the relative MSFE from this exercise.

Comparing Table 3.9 to Table 3.4, there is just a small improvement of the forecasting performance of each Bayesian VAR from this exercise. This improvement, however, does not affect our finding in Section 3.4 that the forecasting performances of the larger models are not impressively better than that of the smallest model.

3.7 Conclusion

Bayesian or shrinkage estimation allows us to use all available information to forecast key economic indicators. BGR show us this point using the U.S. data. The results of BGR, similar to our Table 3.1 or Table 3.5, implicitly imply that a 3-variable VAR with only target variables is grossly inadequate.

However, this impression is false and is a result of their practice of not allowing shrinkage at all for this 3-variable model. This 3-variable VAR has 13 lags, estimated using 120 observations. We have shown that if we use a shrinkage estimator for this 3-variable model with an appropriate hyperparameter value, the improvement of larger models will be minimal. Specifically, the 7-variable and 20-variable models considered in BGR do not seem to have a clear edge over the 3-variable model, and the improvement of the 131-variable model is much less pronounced than what BGR implies.

We try allowing for time-varying hyperparameter values as well, but the result we have found so far is that the time-varying scheme just marginally improves the performance of each model. It does not change our previous conclusion either.

In this study, we also demonstrate a way to figure out the suitable hyperparameter value for each model specification with a given forecast horizon. The value is chosen based on the out-of-sample forecasting performances in the test period, which is a part of the pre-evaluation period. This process takes time for the *LARGE* model. The estimation of the *LARGE* model involves calculating for the inverse of matrix of dimension $(1,704 \times 1,704)$. Since we have to estimate the model 120 times for each value of λ and each forecast horizon h in the grid search shown in the Appendix C, it costs us about 3 days for each of the 4 steps in the search under the computation of a Pentium Core 2 processor. The whole process of the grid search, which is composed of 4 steps, requires about two weeks.

However, in a real forecasting practice we need to figure this suitable hyperparameter value just once. We think that this is the process that should be taken rather than depending on an arbitrarily chosen value. Moreover, as can be observed from Table 3.6 and Table 3.7, the value tends to be stable for a long enough series as in the case of U.S. data.

It would be convenient if we can figure out some patterns of changes in the suitable hyperparameter values of the Bayesian VARs. For example, the values may decrease for longer forecast horizons or bigger model specifications. Our results so far have not shown any obvious pattern. However, a more thorough investigation in this direction is still interesting.

Another interesting way to deal with the hyperparameter value is to figure out a good updating scheme. The scheme that can tell a suitable hyperparameter value when there is an additional actual realization looks very attractive for a real forecasting practice. Unfortunately, this hyperparameter has a non-linear relationship with the forecasting performances of the model. One might have to depend on a relatively complicated framework to figure out an optimal updating scheme for the model. However, our results with a simple sensible scheme in Section 3.6 are encouraging.

Appendix B

Gradient Matrix of the Coefficients

To simplify the notation, let $Z \equiv \tilde{\Omega}^{-1} + X'X$. The posterior mean of the matrix $\bar{B}$ can be written as:

$$\bar{B} = Z^{-1}(\tilde{\Omega}^{-1}\tilde{B} + X'Y). \tag{B.1}$$

The derivative of $\bar{B}$ with respect to λ can be computed from:

$$\begin{aligned} \frac{\partial \bar{B}}{\partial \lambda} &= Z^{-1}\left(\frac{\partial}{\partial \lambda}(\tilde{\Omega}^{-1}\tilde{B} + X'Y)\right) + \left(\frac{\partial Z^{-1}}{\partial \lambda}\right)(\tilde{\Omega}^{-1}\tilde{B} + X'Y), \\ &= Z^{-1}\left(\frac{\partial}{\partial \lambda}\tilde{\Omega}^{-1}\tilde{B}\right) + \left(\frac{\partial Z^{-1}}{\partial \lambda}\right)(\tilde{\Omega}^{-1}\tilde{B} + X'Y). \end{aligned} \tag{B.2}$$

From Magnus and Neudecker (1999), the derivative of an inverse matrix of functions Z can be written as:

$$\frac{\partial Z^{-1}}{\partial \lambda} = -Z^{-1}\left(\frac{\partial Z}{\partial \lambda}\right)Z^{-1}. \tag{B.3}$$

Since the matrix $X'X$ is not a function of λ, the derivative $\partial Z/\partial \lambda$ is:

$$\frac{\partial Z}{\partial \lambda} = \frac{\partial \tilde{\Omega}^{-1}}{\partial \lambda} = -\frac{2}{\lambda^3} diag(\sigma_1^2, ..., \sigma_m^2; 2^2 \cdot \sigma_1^2, ..., 2^2 \cdot \sigma_m^2; ...; p^2 \cdot \sigma_1^2, ..., p^2 \cdot \sigma_m^2; 0). \tag{B.4}$$

The derivative $\partial(\tilde{\Omega}^{-1}\tilde{B})/\partial \lambda$ can also be written as:

$$\frac{\partial}{\partial \lambda}\tilde{\Omega}^{-1}\tilde{B} = \left(\frac{\partial \tilde{\Omega}^{-1}}{\partial \lambda}\right)\tilde{B}, \tag{B.5}$$

and the value of $\partial\tilde{\Omega}^{-1}/\partial \lambda$ is as in (B.4).

Totally, from (B.1) - (B.5), we have:

$$\begin{aligned}\frac{\partial \bar{B}}{\partial \lambda} &= Z^{-1}\left(\frac{\partial \tilde{\Omega}^{-1}}{\partial \lambda}\right)\tilde{B} - Z^{-1}\left(\frac{\partial \tilde{\Omega}^{-1}}{\partial \lambda}\right)Z^{-1}(\tilde{\Omega}^{-1}\tilde{B} + X'Y), \\ &= Z^{-1}\left(\frac{\partial \tilde{\Omega}^{-1}}{\partial \lambda}\right)(\tilde{B} - \bar{B}). \qquad \text{(B.6)}\end{aligned}$$

Appendix C

Grid Search, 1971 - 1980

Table C.1: SMALL, 1971 – 1980

h	step1 λ	step1 $TV_h^{(\mu,\lambda)}$	step1 grad	step2 λ	step2 $TV_h^{(\mu,\lambda)}$	step2 grad	step3 λ	step3 $TV_h^{(\mu,\lambda)}$	step3 grad	step4 λ	step4 $TV_h^{(\mu,\lambda)}$	step4 grad
	0.001	2.711	- 134.132	0.1	1.956	- 0.463	0.11	1.952	- 0.278	0.121	1.950	- 0.110
	1	2.473	0.472	0.2	1.972	0.542	0.12	1.951	- 0.124	0.122	1.950	- 0.096
	2	2.831	0.274	0.3	2.040	0.741	**0.13**	**1.950**	**0.008**	0.123	1.950	- 0.082
	3	3.052	0.177	0.4	2.114	0.740	0.14	1.951	0.121	0.124	1.950	- 0.069
	4	3.197	0.119	0.5	2.186	0.694	0.15	1.952	0.220	0.125	1.950	- 0.055
1	5	3.297	0.083	0.6	2.253	0.641	0.16	1.955	0.305	0.126	1.950	- 0.042
	6	3.368	0.060	0.7	2.315	0.591	0.17	1.958	0.378	0.127	1.950	- 0.029
	7	3.419	0.045	0.8	2.371	0.546	0.18	1.962	0.441	0.128	1.950	- 0.017
	8	3.458	0.034	0.9	2.424	0.506	0.19	1.967	0.496	0.129	1.950	- 0.004
	9	3.488	0.026									
	10	3.511	0.021									

h	step1 λ	step1 $TV_h^{(\mu,\lambda)}$	step1 grad	step2 λ	step2 $TV_h^{(\mu,\lambda)}$	step2 grad	step3 λ	step3 $TV_h^{(\mu,\lambda)}$	step3 grad	step4 λ	step4 $TV_h^{(\mu,\lambda)}$	step4 grad
	0.001	2.701	- 100.100	0.1	1.877	- 0.216	0.11	1.876	- 0.017	**0.111**	**1.876**	**0.001**
	1	2.460	0.384	0.2	1.919	0.797	0.12	1.876	0.142	0.112	1.876	0.018
	2	2.701	0.148	0.3	2.010	0.958	0.13	1.878	0.274	0.113	1.876	0.034
	3	2.805	0.071	0.4	2.103	0.888	0.14	1.882	0.386	0.114	1.876	0.051
	4	2.857	0.038	0.5	2.186	0.774	0.15	1.886	0.483	0.115	1.876	0.067
3	5	2.886	0.022	0.6	2.258	0.664	0.16	1.891	0.566	0.116	1.876	0.082
	6	2.903	0.014	0.7	2.320	0.572	0.17	1.897	0.638	0.117	1.876	0.098
	7	2.915	0.009	0.8	2.373	0.496	0.18	1.904	0.700	0.118	1.876	0.113
	8	2.922	0.006	0.9	2.420	0.435	0.19	1.911	0.752	0.119	1.876	0.128
	9	2.927	0.005									
	10	2.931	0.003									

h	step1 λ	step1 $TV_h^{(\mu,\lambda)}$	step1 grad	step2 λ	step2 $TV_h^{(\mu,\lambda)}$	step2 grad	step3 λ	step3 $TV_h^{(\mu,\lambda)}$	step3 grad	step4 λ	step4 $TV_h^{(\mu,\lambda)}$	step4 grad
	0.001	2.775	5.383	0.1	1.893	- 0.477	0.11	1.889	- 0.278	0.131	1.887	0.004
	1	2.234	0.268	0.2	1.905	0.433	0.12	1.887	- 0.127	0.132	1.887	0.015
	2	2.416	0.119	0.3	1.957	0.556	**0.13**	**1.887**	**- 0.007**	0.133	1.887	0.025
	3	2.499	0.056	0.4	2.011	0.518	0.14	1.887	0.092	0.134	1.887	0.035
	4	2.540	0.028	0.5	2.060	0.458	0.15	1.889	0.175	0.135	1.887	0.045
6	5	2.561	0.016	0.6	2.103	0.404	0.16	1.891	0.245	0.136	1.887	0.055
	6	2.573	0.009	0.7	2.141	0.360	0.17	1.893	0.305	0.137	1.887	0.064
	7	2.581	0.006	0.8	2.175	0.324	0.18	1.897	0.355	0.138	1.887	0.074
	8	2.585	0.004	0.9	2.206	0.294	0.19	1.901	0.398	0.139	1.887	0.083
	9	2.589	0.003									
	10	2.591	0.002									

h	step1 λ	step1 $TV_h^{(\mu,\lambda)}$	step1 grad	step2 λ	step2 $TV_h^{(\mu,\lambda)}$	step2 grad	step3 λ	step3 $TV_h^{(\mu,\lambda)}$	step3 grad	step4 λ	step4 $TV_h^{(\mu,\lambda)}$	step4 grad
	0.001	2.935	248.436	0.1	1.912	- 0.099	0.11	1.913	0.313	0.101	1.912	- 0.049
	1	2.255	0.177	0.2	1.981	0.781	0.12	1.918	0.563	**0.102**	**1.912**	**- 0.000**
	2	2.364	0.067	0.3	2.047	0.540	0.13	1.924	0.711	0.103	1.912	0.046
	3	2.411	0.032	0.4	2.093	0.414	0.14	1.932	0.793	0.104	1.912	0.090
	4	2.434	0.017	0.5	2.131	0.346	0.15	1.940	0.833	0.105	1.912	0.132
12	5	2.448	0.010	0.6	2.163	0.298	0.16	1.948	0.846	0.106	1.912	0.172
	6	2.455	0.006	0.7	2.191	0.260	0.17	1.957	0.842	0.107	1.912	0.209
	7	2.460	0.004	0.8	2.215	0.228	0.18	1.965	0.827	0.108	1.912	0.246
	8	2.463	0.003	0.9	2.237	0.200	0.19	1.973	0.805	0.109	1.913	0.280
	9	2.466	0.002									
	10	2.467	0.001									

Table C.2: CEE, 1971 – 1980

h	step1			step2			step3			step4		
h	λ	$TV_h^{(\mu,\lambda)}$	grad	λ	$TV_h^{(\mu,\lambda)}$	grad	λ	$TV_h^{(\mu,\lambda)}$	grad	λ	$TV_h^{(\mu,\lambda)}$	grad
1	0.001	2.661	- 51.314	0.1	1.858	- 0.632	0.11	1.853	- 0.366	0.121	1.850	- 0.134
	1	2.485	0.676	0.2	1.875	0.580	0.12	1.850	- 0.153	0.122	1.850	- 0.116
	2	3.086	0.555	0.3	1.945	0.787	0.13	1.849	0.017	0.123	1.850	- 0.098
	3	3.617	0.508	0.4	2.027	0.834	0.14	1.850	0.153	0.124	1.850	- 0.081
	4	4.103	0.462	0.5	2.110	0.826	0.15	1.852	0.263	0.125	1.849	- 0.063
	5	4.541	0.415	0.6	2.192	0.797	0.16	1.855	0.353	0.126	1.849	- 0.047
	6	4.932	0.369	0.7	2.270	0.764	0.17	1.859	0.426	0.127	1.849	- 0.030
	7	5.280	0.326	0.8	2.344	0.731	0.18	1.864	0.487	0.128	1.849	- 0.014
	8	5.585	0.287	0.9	2.416	0.702	0.19	1.869	0.538	**0.129**	**1.849**	**0.001**
	9	5.855	0.253									
	10	6.092	0.222									

h	step1			step2			step3			step4		
h	λ	$TV_h^{(\mu,\lambda)}$	grad	λ	$TV_h^{(\mu,\lambda)}$	grad	λ	$TV_h^{(\mu,\lambda)}$	grad	λ	$TV_h^{(\mu,\lambda)}$	grad
3	0.001	2.654	- 20.586	0.1	1.760	- 1.124	0.11	1.750	- 0.775	0.141	1.739	- 0.030
	1	2.405	0.654	0.2	1.759	0.591	0.12	1.744	- 0.485	0.142	1.739	- 0.013
	2	2.931	0.445	0.3	1.836	0.873	0.13	1.741	- 0.246	**0.143**	**1.739**	**0.004**
	3	3.332	0.362	0.4	1.926	0.919	0.14	1.739	- 0.048	0.144	1.739	0.021
	4	3.662	0.300	0.5	2.018	0.899	0.15	1.739	0.114	0.145	1.739	0.037
	5	3.935	0.249	0.6	2.105	0.853	0.16	1.741	0.248	0.146	1.739	0.053
	6	4.163	0.209	0.7	2.188	0.799	0.17	1.744	0.359	0.147	1.739	0.069
	7	4.355	0.176	0.8	2.265	0.746	0.18	1.748	0.451	0.148	1.739	0.084
	8	4.518	0.150	0.9	2.337	0.697	0.19	1.753	0.527	0.149	1.739	0.100
	9	4.656	0.128									
	10	4.775	0.110									

h	step1			step2			step3			step4		
h	λ	$TV_h^{(\mu,\lambda)}$	grad	λ	$TV_h^{(\mu,\lambda)}$	grad	λ	$TV_h^{(\mu,\lambda)}$	grad	λ	$TV_h^{(\mu,\lambda)}$	grad
6	0.001	2.763	50.819	0.1	1.865	- 0.839	0.11	1.858	- 0.547	0.131	1.852	- 0.054
	1	2.791	1.044	0.2	1.882	0.777	0.12	1.854	- 0.293	0.132	1.852	- 0.034
	2	3.665	0.730	0.3	1.981	1.130	0.13	1.852	- 0.074	0.133	1.852	- 0.015
	3	4.290	0.534	0.4	2.099	1.212	0.14	1.852	0.113	**0.134**	**1.852**	**0.004**
	4	4.754	0.401	0.5	2.221	1.217	0.15	1.854	0.272	0.135	1.852	0.023
	5	5.105	0.308	0.6	2.342	1.196	0.16	1.857	0.407	0.136	1.852	0.041
	6	5.379	0.243	0.7	2.460	1.162	0.17	1.862	0.522	0.137	1.852	0.060
	7	5.597	0.196	0.8	2.574	1.124	0.18	1.868	0.621	0.138	1.852	0.078
	8	5.774	0.160	0.9	2.685	1.084	0.19	1.874	0.705	0.139	1.852	0.095
	9	5.919	0.132									
	10	6.041	0.111									

h	step1			step2			step3			step4		
h	λ	$TV_h^{(\mu,\lambda)}$	grad	λ	$TV_h^{(\mu,\lambda)}$	grad	λ	$TV_h^{(\mu,\lambda)}$	grad	λ	$TV_h^{(\mu,\lambda)}$	grad
12	0.001	3.016	220.535	0.1	1.945	1.373	0.01	2.273	- 56.549	0.041	1.899	- 0.807
	1	3.318	1.279	0.2	2.084	1.429	0.02	1.984	- 12.212	0.042	1.898	- 0.686
	2	4.256	0.658	0.3	2.238	1.644	0.03	1.918	- 3.228	0.043	1.897	- 0.573
	3	4.746	0.365	0.4	2.407	1.716	0.04	1.899	- 0.937	0.044	1.897	- 0.467
	4	5.043	0.245	0.5	2.578	1.680	0.05	1.896	0.056	0.045	1.896	- 0.367
	5	5.259	0.193	0.6	2.742	1.605	0.06	1.899	0.647	0.046	1.896	- 0.273
	6	5.437	0.166	0.7	2.898	1.522	0.07	1.908	1.008	0.047	1.896	- 0.184
	7	5.593	0.147	0.8	3.046	1.439	0.08	1.919	1.215	0.048	1.896	- 0.100
	8	5.731	0.131	0.9	3.186	1.358	0.09	1.932	1.324	**0.049**	**1.896**	**- 0.020**
	9	5.854	0.116									
	10	5.964	0.103									

Table C.3: MEDIUM, 1971 – 1980

h	step1			step2			step3			step4		
	λ	$TV_h^{(\mu,\lambda)}$	grad	λ	$TV_h^{(\mu,\lambda)}$	grad	λ	$TV_h^{(\mu,\lambda)}$	grad	λ	$TV_h^{(\mu,\lambda)}$	grad
	0.001	2.659	- 44.169	0.1	1.665	0.122	0.01	2.101	- 28.281	0.091	1.665	- 0.176
	1	2.479	0.714	0.2	1.743	1.047	0.02	1.926	- 11.827	0.092	1.665	- 0.139
	2	3.009	0.383	0.3	1.851	1.082	0.03	1.831	- 7.657	0.093	1.665	- 0.102
	3	3.299	0.211	0.4	1.957	1.038	0.04	1.767	- 5.185	0.094	1.665	- 0.067
	4	3.453	0.122	0.5	2.058	0.982	0.05	1.725	- 3.451	0.095	1.665	- 0.033
1	5	3.556	0.075	0.6	2.153	0.924	0.06	1.697	- 2.208	**0.096**	**1.665**	**- 0.000**
	6	3.605	0.049	0.7	2.243	0.867	0.07	1.679	- 1.317	0.097	1.665	0.032
	7	3.651	0.033	0.8	2.327	0.812	0.08	1.670	- 0.677	0.098	1.665	0.063
	8	3.669	0.024	0.9	2.405	0.761	0.09	1.665	- 0.215	0.099	1.665	0.093
	9	3.690	0.017									
	10	3.743	0.013									

h	step1			step2			step3			step4		
	λ	$TV_h^{(\mu,\lambda)}$	grad	λ	$TV_h^{(\mu,\lambda)}$	grad	λ	$TV_h^{(\mu,\lambda)}$	grad	λ	$TV_h^{(\mu,\lambda)}$	grad
	0.001	2.677	- 9.505	0.1	1.698	- 0.360	0.11	1.696	- 0.113	0.111	1.696	- 0.092
	1	2.282	0.488	0.2	1.731	0.656	0.12	1.696	0.074	0.112	1.696	- 0.071
	2	2.599	0.199	0.3	1.806	0.796	0.13	1.697	0.218	0.113	1.695	- 0.051
	3	2.739	0.096	0.4	1.887	0.814	0.14	1.700	0.330	0.114	1.695	- 0.032
	4	2.806	0.052	0.5	1.967	0.780	0.15	1.704	0.417	0.115	1.695	- 0.013
3	5	2.851	0.031	0.6	2.042	0.722	0.16	1.708	0.487	0.116	1.695	0.006
	6	2.871	0.020	0.7	2.111	0.658	0.17	1.713	0.542	**0.117**	**1.695**	**0.023**
	7	2.892	0.013	0.8	2.174	0.596	0.18	1.719	0.588	0.118	1.696	0.041
	8	2.897	0.009	0.9	2.230	0.539	0.19	1.725	0.625	0.119	1.696	0.058
	9	2.911	0.007									
	10	2.927	0.005									

h	step1			step2			step3			step4		
	λ	$TV_h^{(\mu,\lambda)}$	grad	λ	$TV_h^{(\mu,\lambda)}$	grad	λ	$TV_h^{(\mu,\lambda)}$	grad	λ	$TV_h^{(\mu,\lambda)}$	grad
	0.001	2.824	46.860	0.1	2.231	0.548	0.01	2.255	- 25.316	0.011	2.234	- 18.655
	1	2.895	0.379	0.2	2.323	1.095	0.02	2.193	1.846	0.012	2.218	- 13.367
	2	3.122	0.132	0.3	2.433	1.061	0.03	2.215	1.604	0.013	2.206	- 9.227
	3	3.212	0.059	0.4	2.532	0.933	0.04	2.223	0.294	0.014	2.199	- 6.025
	4	3.251	0.031	0.5	2.619	0.797	0.05	2.224	- 0.143	0.015	2.194	- 3.575
6	5	3.279	0.018	0.6	2.692	0.677	0.06	2.222	- 0.118	0.016	2.192	- 1.726
	6	3.292	0.011	0.7	2.755	0.577	0.07	2.222	0.051	**0.017**	**2.191**	**- 0.350**
	7	3.308	0.008	0.8	2.808	0.497	0.08	2.223	0.237	0.018	2.191	0.652
	8	3.310	0.005	0.9	2.855	0.432	0.09	2.226	0.405	0.019	2.192	1.363
	9	3.321	0.004									
	10	3.327	0.003									

h	step1			step2			step3			step4		
	λ	$TV_h^{(\mu,\lambda)}$	grad	λ	$TV_h^{(\mu,\lambda)}$	grad	λ	$TV_h^{(\mu,\lambda)}$	grad	λ	$TV_h^{(\mu,\lambda)}$	grad
	0.001	3.174	206.889	0.1	2.937	5.071	0.01	2.477	- 63.264	0.021	2.254	1.166
	1	4.683	1.130	0.2	3.253	2.440	**0.02**	**2.254**	**- 0.672**	0.022	2.256	2.712
	2	5.355	0.379	0.3	3.487	2.273	0.03	2.309	9.260	0.023	2.260	4.021
	3	5.607	0.162	0.4	3.708	2.149	0.04	2.416	11.555	0.024	2.264	5.136
	4	5.719	0.082	0.5	3.915	1.985	0.05	2.532	11.425	0.025	2.270	6.092
12	5	5.782	0.046	0.6	4.104	1.800	0.06	2.641	10.228	0.026	2.276	6.916
	6	5.827	0.028	0.7	4.275	1.613	0.07	2.736	8.710	0.027	2.284	7.629
	7	5.849	0.018	0.8	4.428	1.437	0.08	2.816	7.262	0.028	2.292	8.249
	8	5.858	0.012	0.9	4.563	1.275	0.09	2.882	6.038	0.029	2.300	8.789
	9	5.866	0.008									
	10	5.875	0.006									

* Step 4 is also applied for the values of λ between 0.06 and 0.07 in this case. The one shown in the table yields better results.

Table C.4: LARGE, 1971 – 1980

h	step1 λ	step1 $TV_h^{(\mu,\lambda)}$	step1 grad	step2 λ	step2 $TV_h^{(\mu,\lambda)}$	step2 grad	step3 λ	step3 $TV_h^{(\mu,\lambda)}$	step3 grad	step4 λ	step4 $TV_h^{(\mu,\lambda)}$	step4 grad
1	0.001	2.545	- 244.283	0.1	1.642	1.353	0.01	1.863	-22.827	0.051	1.598	- 0.136
	1	2.063	0.062	0.2	1.777	1.193	0.02	1.704	-10.355	0.052	1.598	- 0.051
	2	2.127	0.008	0.3	1.876	0.783	0.03	1.635	- 4.328	**0.053**	**1.598**	**0.029**
	3	2.524	0.003	0.4	1.940	0.503	0.04	1.607	- 1.560	0.054	1.598	0.104
	4	2.988	0.001	0.5	1.979	0.328	0.05	1.599	- 0.227	0.055	1.599	0.175
	5	3.529	- 0.001	0.6	2.005	0.221	0.06	1.600	0.473	0.056	1.599	0.242
	6	6.178	- 0.000	0.7	2.028	0.153	0.07	1.607	0.873	0.057	1.599	0.304
	7	8.458	- 0.002	0.8	2.036	0.110	0.08	1.617	1.114	0.058	1.599	0.364
	8	16.511	- 0.003	0.9	2.052	0.082	0.09	1.629	1.263	0.059	1.600	0.420
	9	21.860	0.000									
	10	30.033	0.000									

h	step1 λ	step1 $TV_h^{(\mu,\lambda)}$	step1 grad	step2 λ	step2 $TV_h^{(\mu,\lambda)}$	step2 grad	step3 λ	step3 $TV_h^{(\mu,\lambda)}$	step3 grad	step4 λ	step4 $TV_h^{(\mu,\lambda)}$	step4 grad
3	0.001	2.581	- 236.990	0.1	1.626	0.673	0.01	1.978	-13.317	0.071	1.614	- 0.059
	1	1.788	0.018	0.2	1.690	0.489	0.02	1.834	-13.480	**0.072**	**1.614**	**- 0.008**
	2	1.801	0.002	0.3	1.727	0.272	0.03	1.726	- 8.171	0.073	1.614	0.040
	3	2.003	0.001	0.4	1.749	0.162	0.04	1.665	- 4.363	0.074	1.614	0.087
	4	2.270	0.003	0.5	1.760	0.102	0.05	1.633	- 2.136	0.075	1.614	0.130
	5	2.481	0.005	0.6	1.767	0.066	0.06	1.619	- 0.855	0.076	1.614	0.171
	6	3.540	0.013	0.7	1.775	0.046	0.07	1.614	- 0.113	0.077	1.615	0.210
	7	4.673	0.023	0.8	1.778	0.032	0.08	1.615	0.314	0.078	1.615	0.246
	8	7.618	0.047	0.9	1.780	0.024	0.09	1.620	0.550	0.079	1.615	0.281
	9	10.733	0.073									
	10	15.450	0.108									

h	step1 λ	step1 $TV_h^{(\mu,\lambda)}$	step1 grad	step2 λ	step2 $TV_h^{(\mu,\lambda)}$	step2 grad	step3 λ	step3 $TV_h^{(\mu,\lambda)}$	step3 grad	step4 λ	step4 $TV_h^{(\mu,\lambda)}$	step4 grad
6	0.001	2.740	- 216.779	0.1	2.340	1.585	0.01	2.413	22.720	0.051	2.294	- 1.519
	1	2.579	0.023	0.2	2.448	0.700	0.02	2.471	- 5.454	0.052	2.292	- 1.295
	2	2.608	0.004	0.3	2.500	0.373	0.03	2.393	- 8.107	0.053	2.291	- 1.081
	3	2.815	0.002	0.4	2.529	0.217	0.04	2.327	- 4.777	0.054	2.290	- 0.880
	4	3.207	0.003	0.5	2.544	0.134	0.05	2.296	- 1.756	0.055	2.290	- 0.689
	5	3.189	0.007	0.6	2.554	0.087	0.06	2.288	0.106	0.056	2.289	- 0.510
	6	4.683	0.016	0.7	2.558	0.059	0.07	2.295	1.075	0.057	2.289	- 0.341
	7	5.786	0.026	0.8	2.568	0.042	0.08	2.308	1.496	0.058	2.288	- 0.182
	8	8.262	0.060	0.9	2.575	0.030	0.09	2.324	1.615	**0.059**	**2.288**	**- 0.033**
	9	12.477	0.099									
	10	17.466	0.143									

h	step1 λ	step1 $TV_h^{(\mu,\lambda)}$	step1 grad	step2 λ	step2 $TV_h^{(\mu,\lambda)}$	step2 grad	step3 λ	step3 $TV_h^{(\mu,\lambda)}$	step3 grad	step4 λ	step4 $TV_h^{(\mu,\lambda)}$	step4 grad
12	0.001	3.102	- 103.754	0.1	3.671	5.650	0.01	2.605	47.772	0.001	3.102	-103.754
	1	4.385	0.065	0.2	4.006	2.074	0.02	3.008	24.394	0.002	2.901	-237.304
	2	4.365	0.009	0.3	4.159	1.113	0.03	3.148	7.992	0.003	2.691	-171.138
	3	4.564	0.004	0.4	4.248	0.640	0.04	3.218	6.958	0.004	2.561	- 90.969
	4	5.129	0.003	0.5	4.293	0.390	0.05	3.294	8.189	0.005	2.500	- 36.068
	5	4.882	0.007	0.6	4.320	0.251	0.06	3.379	8.593	**0.006**	**2.483**	**- 1.646**
	6	6.462	0.014	0.7	4.339	0.169	0.07	3.463	8.180	0.007	2.493	20.192
	7	7.717	0.020	0.8	4.348	0.119	0.08	3.541	7.377	0.008	2.520	34.204
	8	8.988	0.045	0.9	4.376	0.086	0.09	3.610	6.484	0.009	2.559	42.915
	9	13.824	0.094									
	10	17.860	0.114									

* Step 4 is also applied for the values of λ between 0.001 and 0.01 in this case. The one shown in the table yields better results.

Appendix D

Description of The Data

Table F.1: Description of the Data

No	Symbol	Long Description	SMALL	CEE	MEDIUM	log	$\delta_i = 0$
1	A0M052	Personal Income (Ar, Bil. Chain 2000 $)				x	
2	A0M051	Personal Income Less Transfer Payments (Ar, Bil. Chain 2000 $)			x	x	
3	A0M224_R	Real Consumption (Ac) A0M224/Gmdc			x	x	
4	A0M057	Manufacturing And Trade Sales (Mil. Chain 1996 $)				x	
5	A0M059	Sales Of Retail Stores (Mil. Chain 2000 $)				x	
6	IPS10	Industrial Production Index - Total Index			x	x	
7	IPS11	Industrial Production Index - Products, Total				x	
8	IPS299	Industrial Production Index - Final Products				x	
9	IPS12	Industrial Production Index - Consumer Goods				x	
10	IPS13	Industrial Production Index - Durable Consumer Goods				x	
11	IPS18	Industrial Production Index - Nondurable Consumer Goods				x	
12	IPS25	Industrial Production Index - Business Equipment				x	
13	IPS32	Industrial Production Index - Materials				x	
14	IPS34	Industrial Production Index - Durable Goods Materials				x	
15	IPS38	Industrial Production Index - Nondurable Goods Materials				x	
16	IPS43	Industrial Production Index - Manufacturing (Sic)				x	
17	IPS307	Industrial Production Index - Residential Utilities				x	
18	IPS306	Industrial Production Index - Fuels				x	
19	PMP	Napm Production Index (Percent)					x
20	A0M082	Capacity Utilization (Mfg)			x		
21	LHEL	Index Of Help-Wanted Advertising In Newspapers (1967=100;Sa)					
22	LHELX	Employment: Ratio; Help-Wanted Ads:No. Unemployed Clf					
23	LHEM	Civilian Labor Force: Employed, Total (Thous.,Sa)				x	
24	LHNAG	Civilian Labor Force: Employed, Nonagric.Industries (Thous.,Sa)				x	
25	LHUR	Unemployment Rate: All Workers, 16 Years & Over (%,Sa)			x		
26	LHU680	Unemploy.By Duration: Average(Mean)Duration In Weeks (Sa)					
27	LHU5	Unemploy.By Duration: Persons Unempl.Less Than 5 Wks (Thous.,Sa)				x	
28	LHU14	Unemploy.By Duration: Persons Unempl.5 To 14 Wks (Thous.,Sa)				x	
29	LHU15	Unemploy.By Duration: Persons Unempl.15 Wks + (Thous.,Sa)				x	
30	LHU26	Unemploy.By Duration: Persons Unempl.15 To 26 Wks (Thous.,Sa)				x	
31	LHU27	Unemploy.By Duration: Persons Unempl.27 Wks + (Thous,Sa)				x	
32	A0M005	Average Weekly Initial Claims, Unemploy. Insurance (Thous.)				x	
33	CES002	Employees On Nonfarm Payrolls - Total Private	x	x	x	x	
34	CES003	Employees On Nonfarm Payrolls - Goods-Producing				x	
35	CES006	Employees On Nonfarm Payrolls - Mining				x	
36	CES011	Employees On Nonfarm Payrolls - Construction				x	
37	CES015	Employees On Nonfarm Payrolls - Manufacturing				x	
38	CES017	Employees On Nonfarm Payrolls - Durable Goods				x	
39	CES033	Employees On Nonfarm Payrolls - Nondurable Goods				x	
40	CES046	Employees On Nonfarm Payrolls - Service-Providing				x	
41	CES048	Employees On Nonfarm Payrolls - Trade, Transportation, And Utilities				x	
42	CES049	Employees On Nonfarm Payrolls - Wholesale Trade				x	
43	CES053	Employees On Nonfarm Payrolls - Retail Trade				x	
44	CES088	Employees On Nonfarm Payrolls - Financial Activities				x	
45	CES140	Employees On Nonfarm Payrolls - Government				x	
46	A0M048	Employee Hours In Nonag. Establishments (Ar, Bil. Hours)				x	
47	CES151	Average Weekly Hours Of Production Or Nonsupervisory Workers On Private Nonfar					x
48	CES155	Average Weekly Hours Of Production Or Nonsupervisory Workers On Private Nonfar					
49	AOM001	Average Weekly Hours, Mfg. (Hours)					x
50	PMEMP	Napm Employment Index (Percent)					x
51	HSFR	Housing Starts:Nonfarm(1947-58);Total Farm&Nonfarm(1959-)(Thous.,Sa)			x	x	x
52	HSNE	Housing Starts:Northeast (Thous.U.)S.A.				x	x
53	HSMW	Housing Starts:Midwest(Thous.U.)S.A.				x	x

Description of the Data (Cont.)

No	Symbol	Long Description	SMALL	CEE	MEDIUM	log	$\delta_i = 0$
54	HSSOU	Housing Starts:South (Thous.U.)S.A.				x	x
55	HSWST	Housing Starts:West (Thous.U.)S.A.				x	x
56	HSBR	Housing Authorized: Total New Priv Housing Units (Thous.,Saar)				x	x
57	HSBNE	Houses Authorized By Build. Permits:Northeast(Thou.U.)S.A				x	x
58	HSBMW	Houses Authorized By Build. Permits:Midwest(Thou.U.)S.A.				x	x
59	HSBSOU	Houses Authorized By Build. Permits:South(Thou.U.)S.A.				x	x
60	HSBWST	Houses Authorized By Build. Permits:West(Thou.U.)S.A.				x	x
61	PMI	Purchasing Managers' Index (Sa)					x
62	PMNO	Napm New Orders Index (Percent)					x
63	PMDEL	Napm Vendor Deliveries Index (Percent)					x
64	PMNV	Napm Inventories Index (Percent)					x
65	A0M008	Mfrs' New Orders, Consumer Goods And Materials (Bil. Chain 1982 $)				x	
66	A0M007	Mfrs' New Orders, Durable Goods Industries (Bil. Chain 2000 $)				x	
67	A0M027	Mfrs' New Orders, Nondefense Capital Goods (Mil. Chain 1982 $)				x	
68	A1M092	Mfrs' Unfilled Orders, Durable Goods Indus. (Bil. Chain 2000 $)				x	
69	A0M070	Manufacturing And Trade Inventories (Bil. Chain 2000 $)				x	
70	A0M077	Ratio, Mfg. And Trade Inventories To Sales (Based On Chain 2000 $)					
71	FM1	Money Stock: M1(Curr,Trav.Cks,Dem Dep,Other Ck'Able Dep)(Bil$,Sa)			x	x	
72	FM2	Money Stock:M2(M1+O'Nite Rps,Euro$,G/P&B/D Mmmfs&Sav&Sm Time Dep(Bil$,Sa),		x	x	x	
73	FM3	Money Stock: M3(M2+Lg Time Dep,Term Rp'S&Inst Only Mmmfs)(Bil$,Sa)				x	
74	FM2DQ	Money Supply - M2 In 1996 Dollars (Bci)				x	
75	FMFBA	Monetary Base, Adj For Reserve Requirement Changes(Mil$,Sa)				x	
76	FMRRA	Depository Inst Reserves:Total,Adj For Reserve Req Chgs(Mil$,Sa)		x	x	x	
77	FMRNBA	Depository Inst Reserves:Nonborrowed,Adj Res Req Chgs(Mil$,Sa)		x	x	x	
78	FCLNQ	Commercial & Industrial Loans Oustanding In 1996 Dollars (Bci)				x	
79	FCLBMC	Wkly Rp Lg Com'L Banks:Net Change Com'L & Indus Loans(Bil$,Saar)					x
80	CCINRV	Consumer Credit Outstanding - Nonrevolving(G19)				x	
81	A0M095	Ratio, Consumer Installment Credit To Personal Income (Pct.)					
82	FSPCOM	S&P'S Common Stock Price Index: Composite (1941-43=10)			x	x	
83	FSPIN	S&P'S Common Stock Price Index: Industrials (1941-43=10)				x	
84	FSDXP	S&P'S Composite Common Stock: Dividend Yield (% Per Annum)					
85	FSPXE	S&P'S Composite Common Stock: Price-Earnings Ratio (%,Nsa)				x	
86	FYFF	Interest Rate: Federal Funds (Effective) (% Per Annum,Nsa)	x	x	x		
87	CP90	Cmmercial Paper Rate (Ac)					
88	FYGM3	Interest Rate: U.S.Treasury Bills,Sec Mkt,3-Mo.(% Per Ann,Nsa)					
89	FYGM6	Interest Rate: U.S.Treasury Bills,Sec Mkt,6-Mo.(% Per Ann,Nsa)					
90	FYGT1	Interest Rate: U.S.Treasury Const Maturities,1-Yr.(% Per Ann,Nsa)					
91	FYGT5	Interest Rate: U.S.Treasury Const Maturities,5-Yr.(% Per Ann,Nsa)					
92	FYGT10	Interest Rate: U.S.Treasury Const Maturities,10-Yr.(% Per Ann,Nsa)			x		
93	FYAAAC	Bond Yield: Moody'S Aaa Corporate (% Per Annum)					
94	FYBAAC	Bond Yield: Moody'S Baa Corporate (% Per Annum)					
95	SCP90	Cp90-Fyff					x
96	SFYGM3	Fygm3-Fyff					x
97	SFYGM6	Fygm6-Fyff					x
98	SFYGT1	Fygt1-Fyff					x
99	SFYGT5	Fygt5-Fyff					x
100	SFYGT10	Fygt10-Fyff					x
101	SFYAAAC	Fyaaac-Fyff					x
102	SFYBAAC	Fybaac-Fyff					x
103	EXRUS	United States;Effective Exchange Rate(Merm)(Index No.)			x	x	
104	EXRSW	Foreign Exchange Rate: Switzerland (Swiss Franc Per U.S.$)				x	
105	EXRJAN	Foreign Exchange Rate: Japan (Yen Per U.S.$)				x	
106	EXRUK	Foreign Exchange Rate: United Kingdom (Cents Per Pound)				x	

Description of the Data (Cont.)

No	Symbol	Long Description	SMALL	CEE	MEDIUM	log	$\delta_1 = 0$
107	EXRCAN	Foreign Exchange Rate: Canada (Canadian $ Per U.S.$)				x	
108	PWFSA	Producer Price Index: Finished Goods (82=100,Sa)			x	x	
109	PWFCSA	Producer Price Index:Finished Consumer Goods (82=100,Sa)				x	
110	PWIMSA	Producer Price Index:Intermed Mat.Supplies & Components(82=100,Sa)				x	
111	PWCMSA	Producer Price Index:Crude Materials (82=100,Sa)				x	
112	PSM99Q	Index Of Sensitive Materials Prices (1990=100)(Bci-99A)		x	x	x	
113	PMCP	Napm Commodity Prices Index (Percent)					x
114	PUNEW	Cpi-U: All Items (82-84=100,Sa)	x	x	x	x	
115	PU83	Cpi-U: Apparel & Upkeep (82-84=100,Sa)				x	
116	PU84	Cpi-U: Transportation (82-84=100,Sa)				x	
117	PU85	Cpi-U: Medical Care (82-84=100,Sa)				x	
118	PUC	Cpi-U: Commodities (82-84=100,Sa)				x	
119	PUCD	Cpi-U: Durables (82-84=100,Sa)				x	
120	PUS	Cpi-U: Services (82-84=100,Sa)				x	
121	PUXF	Cpi-U: All Items Less Food (82-84=100,Sa)				x	
122	PUXHS	Cpi-U: All Items Less Shelter (82-84=100,Sa)				x	
123	PUXM	Cpi-U: All Items Less Medical Care (82-84=100,Sa)				x	
124	GMDC	Pce,Impl Pr Defl:Pce (1987=100)			x	x	
125	GMDCD	Pce,Impl Pr Defl:Pce; Durables (1987=100)				x	
126	GMDCN	Pce,Impl Pr Defl:Pce; Nondurables (1996=100)				x	
127	GMDCS	Pce,Impl Pr Defl:Pce; Services (1987=100)				x	
128	CES275	Average Hourly Earnings Of Production Or Nonsupervisory Workers On Private No			x	x	
129	CES277	Average Hourly Earnings Of Production Or Nonsupervisory Workers On Private No				x	
130	CES278	Average Hourly Earnings Of Production Or Nonsupervisory Workers On Private No				x	
131	HHSNTN	U. Of Mich. Index Of Consumer Expectations(Bcd-83)					

Chapter 4

Shrinkage Estimation of Vector Autoregressive Models

First, we show a close relationship between the Bayesian VAR with Litterman prior and the ridge regression, which is a well-known shrinkage estimator. Next, we compare the out-of-sample forecasting performances of the Bayesian VAR with Litterman prior to other shrinkage estimators, such as the LASSO (Least Absolute Shrinkage and Selection Operator) or the elastic net, which, in the literature, have some record in outperforming ridge regression. In doing this, we modify the LASSO and the elastic net to admit the decay rate of effects of lags as in the case of the Bayesian VAR. The LASSO outperforms the Bayesian VAR in our empirical study on the U.S. macroeconomic data, which supports the idea that there is some redundancy in the VAR with 3 endogenous variables and 13 lags.

4.1 Introduction

In recent years, we have witnessed a growing interest in forecasting models that take into account a large number of predictors. This can be linked to the improvement in computer technology and the idea that people in the real world usually incorporate a lot of information into their economic decision framework. A popular method in this line of literature is the approximate factor model, popularized by Stock and Watson (2002a) and Bai and Ng (2002). Stock and Watson (2002b), De Mol et al. (2008), and Bernanke and Boivin (2003) among others show that this method performs very well in forecasting important macroeconomic indicators. The principle components technique employed to estimate common factors in the approximate factor models does not impose any limit on the number of predictors that can be considered in a model.

Since the principle components estimator can be regarded as a shrinkage estimator, other alternative shrinkage estimators are also proposed as close substitutes to the approximate factor model. These include the Bayesian regression with normal priors and the LASSO, which is a Bayesian regression with double-exponential priors, by De Mol et al. (2008), and the partial least squares regression by Groen and Kapetanios (2008). These authors show that the forecasts by these methods are asymptotically similar to ones of the approximate factor model.

Boivin and Ng (2006) issue an important caution to the practice of putting a huge number of predictors into the approximate factor model. They warn that smaller models can outperform larger ones in forecasting if many uninformative predictors are incorporated into the latter. In the principle components regression, principle components associated with higher variances are considered as true signals, while ones with lower variances are discarded as noise. Since uninformative predictors can increase variances without informing much about the true signals, we may end up throwing away the signals, while keeping the noise, in the case of larger models.

Bai and Ng (2008) use the elastic net, proposed by Zou and Hastie (2005) as an extension of the LASSO, as their automated variable selector. They show that some smaller models with the selected variables clearly outperform larger models that incorporate all the variables in the data set. They explain their results using the argument made by Boivin and Ng (2006).

We see that we can interpret the results in Bai and Ng (2008) differently. Throwing away some predictors can reduce the variances of its estimated coefficients as well as its forecasts. This is similar to fixing the coefficients in front of the thrown away predictors at zeros. If we choose correctly to throw away the uninformative predictors, this will not affect the expected value of the forecast but will reduce its variance, and may help improve the out-of-sample forecasting performances in any experiment with repeated sampling.

Our bias-variance tradeoff argument is more general as it does not depend on the mechanism specific to the principle components regression. In this way, we think that results as in Bai and Ng (2008) can be generalized to the case of VARs as well. Consider a VAR with long lags. It is hard to believe that all the lags included into the model are informative in predicting the endogenous variables. We see that this explains the impressive performances of the Bayesian VAR with Litterman prior over the unrestricted OLS VAR as recorded in the literature. See, for example, Litterman (1986) and Robertson and Tallman (1999).

Here, we want to study the forecasting performances of Bayesian VAR with Litterman prior in comparison to some selected shrinkage estimators. We are interested in Bayesian VAR with Litterman prior because it is a popular framework in macroeconomic forecasting for central banks and international organizations around the world. See, for example, Doan et al. (1986), Ciccarelli and Rebucci (2003), and Zha (1998). Recently, Bańbura et al. (2008) demonstrate that it is possible and satisfactory to incorporate a large set of endogenous variables with long lags into the model. This will expose the framework to an even wider range of applications in the future.

We choose the LASSO, the elastic net, and a procedure that uses LASSO as the variable selector as our selected alternatives. Note that Hsu et al. (2008) apply the LASSO as a subset selection method on VAR before. However, we are interested in the forecasting performances of its estimates in comparison with the Bayesian VAR here. To motivate our interest, we first note that there is a close relationship between the Bayesian VAR with Litterman prior and the ridge regression, which is a well-known shrinkage regression. We demonstrate this relationship in Section 4.2. This means some shrinkage estimators that outperform ridge regression in the literature may outperform Bayesian VAR with Litterman prior as well, when applied to the VAR.

Tibshirani (1996), while proposing the LASSO, employs a simulation study to show that if the true (unknown) coefficients of a model are composed of 1) a lot of zeros and just a small number of large values, or 2) a few zeros and a moderate number of moderate-sized values, LASSO can outperform ridge regression in out-of-sample prediction. Later, Zou and Hastie (2005) show that the elastic net can further improve the predictive performances of the LASSO. We modify the LASSO and the elastic net to make them take into account the decaying effects of lagged variables. This enables similar treatment on lagged variables as in the usual Bayesian VAR. The modification will be explained in more details in Section 4.3.

The simulation study in Tibshirani (1996) assumes i.i.d. and exogenously given predictors. This is not the case for a VAR. In Section 4.4, we perform a simulation study that uses some VARs as the true models. Our study generally confirms the results in Tibshirani (1996). However, we do not see the difference in performances between the LASSO and the elastic net. In one case study, we also use a VARMA(1,1) as the true model. Under this case, the application of Bayesian information criteria to the unrestricted VAR outperforms all of the shrinkage estimators under the range of parameters considered.

In Section 4.5, we employ all the aforementioned techniques to the real US data set obtained from Stock and Watson (2005). There are 3 endogenous variables, with 13 lags in the VAR model that we consider. Even though we do not see a big improvement of the forecasting performances, the application of LASSO on the VAR tends to dominate the Bayesian VAR with Litterman prior for most of the forecast horizons considered. It also clearly demonstrates the problem of redundancy in the VAR. Section 4.6 concludes the paper.

4.2 Bayesian VAR and Ridge Regression

The Bayesian VAR can be applied to nonstationary time series. See, for example, Litterman (1986), Robertson and Tallman (1999), or Bańbura et al. (2008). However, to make the relationship between Bayesian VAR with Litterman prior and the standard ridge regression explicit, we consider only stationary time series here. This also makes it convenient as we can demean the series without loss of generality. The results of the study, however, can be generalized to nonstationary time series.

Let $Y_t = (y_{1t}, ..., y_{nt})'$ be a $n \times 1$ vector of n stationary and demeaned endogenous variables in period t. We consider the VAR relationship given by:

$$Y_t = A_1 Y_{t-1} + A_2 Y_{t-2} + ... + A_p Y_{t-p} + U_t, \tag{4.1}$$

where $A_i, i = 1, ..., p$, is a $n \times n$ matrix of unknown coefficients, and $U_t = (u_{1t}, ..., u_{nt})'$ is a $n \times 1$ vector of unknown disturbances. We assume that:

$$\underset{n\times 1}{U_t} \sim \text{i.i.d. } N(\underset{n\times 1}{\mathbf{0}}, \underset{n\times n}{\Psi}),$$

where $\mathbf{0}$ is a vector of zeros, and the time-invariant matrix Ψ is a positive definite matrix.

Let $X_t = (Y'_{t-1}, ..., Y'_{t-p})'$ be a $np \times 1$ vector containing p lags of Y_t, and T be a scalar representing the total number of observations (Y_t, X_t). Upon setting $t = 1$ for the first observation, i.e. assuming history starts at time $1 - p$, we can rearrange the VAR from (4.1) into:

$$Y = XB + U, \tag{4.2}$$

where $Y = (Y_1, ..., Y_T)'$ is the $T \times n$ matrix of dependent variables, $X = (X_1, ..., X_T)'$ is the $T \times np$ matrix of independent variables, $B = (A_1, ..., A_p)'$ is the $np \times n$ matrix of unknown coefficients, and $U = (U_1, ..., U_T)'$ is the $T \times n$ matrix of disturbances. Let $\mathbf{u}$ be the vector obtained from stacking the columns of the disturbance matrix U from (4.2). The above assumption on U_t is equivalent to:

$$\underset{Tn\times 1}{\mathbf{u}} \sim N(\underset{Tn\times 1}{\mathbf{0}}, \underset{n\times n}{\Psi} \otimes \underset{T\times T}{I}),$$

where $\otimes$ denotes the Kronecker product, and I is an identity matrix.

With the seemingly-unrelated-regressions (SUR) structure, the efficient estimator for B is the same as an unrestricted OLS estimator, which is:

$$\hat{B}^{(ols)} = (X'X)^{-1}X'Y. \tag{4.3}$$

We use $\hat{B}_i^{(ols)}, i = 1, ..., n$, to represent the i-th column of the estimate $\hat{B}^{(ols)} = (\hat{B}_1^{(ols)}, ..., \hat{B}_n^{(ols)})$, and $Y_i = (y_{i1}, ..., y_{iT})'$ to represent the i-th column of the matrix of dependent variables Y. The estimator in (4.3) implies:

$$\hat{B}_i^{(ols)} = (X'X)^{-1}X'Y_i. \tag{4.4}$$

We know that the estimator in (4.4) is consistent:

$$\sqrt{T}(\underset{np\times 1}{\hat{B}_i^{(ols)}} - \underset{np\times 1}{B_i}) \xrightarrow{p} N(\underset{np\times 1}{\mathbf{0}}, \sigma_i^2 \operatorname{plim}[T \cdot \underset{np\times np}{(X'X)^{-1}}]),$$

where σ_i^2 is the (i,i)-th element of the covariance matrix Ψ. This implies that with large enough number of sample observations T, we can approximately treat $\hat{B}_i^{(ols)}$ as:

$$\underset{np\times 1}{\hat{B}_i^{(ols)}} \approx N(\underset{np\times 1}{B_i}, \sigma_i^2 (X'X)^{-1}). \tag{4.5}$$

The unrestricted OLS estimator (4.3) or (4.4) is not defined when the number of independent variables np is higher than the number of observations T. This is because $X'X$ becomes a singular matrix. Even when the OLS estimator is well defined, a large value of np or strong multicollinearity between variables in the matrix X will increase the variance of the estimate $\hat{B}_i^{(ols)}$. We expect shrinkage regressions to help ameliorate these problems.

4.2.1 Bayesian VAR with Litterman Prior

Bayesian VAR with Litterman prior is a popular shrinkage regression in macroeconomic forecasting. See, for example, Ciccarelli and Rebucci (2003) and Doan et al. (1986). The method assumes each unknown coefficient of the model (4.1) to admit Litterman prior distribution. The sample observations are, then, used to update this distribution according to Bayes' law. We follow Kadiyala and Karlsson (1997) and Bańbura et al. (2008) to put a modified version of Litterman prior as our prior distribution. From now on, when we refer to Bayesian VAR, we mean the method described in this section. Since only the posterior mean is used in forecasting, we will focus on this estimate only.

Let $\mathbf{b} = vec(B)$ be a $n^2p \times 1$ vector obtained from stacking columns of the matrix of coefficients B in (4.2). We assume that the vector $\mathbf{b}$ has a Normal-

(Inverted)-Wishart distribution, given by:

$$\underset{n^2p\times 1}{\mathbf{b}} \mid \underset{n\times n}{\Psi} \sim N(\underset{n^2p\times 1}{\tilde{\mathbf{b}}}, \underset{n\times n}{\Psi} \otimes (1/\lambda^{(bvar)}) \cdot \underset{np\times np}{\tilde{\Omega}}) \qquad \text{and} \qquad \underset{n\times n}{\Psi} \sim iW(\underset{n\times n}{\tilde{\Psi}}, \phi),$$

where $\tilde{\mathbf{b}}$ is the prior mean value of vector $\mathbf{b}$, $\lambda^{(bvar)}$ is a scalar parameter determining the overall tightness of prior distribution around mean, $\tilde{\Omega}$ is the matrix of parameters determining prior variance of $\mathbf{b}$, $\tilde{\Psi}$ is the prior mean value of the covariance matrix Ψ, and ϕ is the degree of freedom of the inverted-Wishart distribution.

We choose parameters of the prior $(\tilde{\mathbf{b}}, \tilde{\Omega}, \tilde{\Psi}, \phi)$ to make it match the modified Litterman prior. With stationary and demeaned variables, the prior mean of the vector of coefficients is set at $\tilde{\mathbf{b}} = \mathbf{0}_{(Nn\times 1)}$. The matrices $\tilde{\Psi}$ and $\tilde{\Omega}$ are at:

$$\tilde{\Psi} = diag(\sigma_1^2, ..., \sigma_n^2),$$

and

$$\tilde{\Omega} = diag\left(\frac{1}{1^\pi \cdot \sigma_1^2}, ..., \frac{1}{1^\pi \cdot \sigma_n^2}; \frac{1}{2^\pi \cdot \sigma_1^2}, ..., \frac{1}{2^\pi \cdot \sigma_n^2}; ...; \frac{1}{p^\pi \cdot \sigma_1^2}, ..., \frac{1}{p^\pi \cdot \sigma_n^2}\right), \tag{4.6}$$

where π is another parameter of the model determining the decay rate of effects of lags. The degree of freedom is set at $\phi = n + 2$ to make prior mean and variance be $\mathbb{E}(\mathbf{b}) = \tilde{\mathbf{b}}$ and $var(\mathbf{b}) = \tilde{\Psi} \otimes (1/\lambda^{(bvar)}) \cdot \tilde{\Omega}$, respectively.

Recall that we define the matrix of unknown coefficients in (4.2) as $B = (A_1, ..., A_p)'$. The variance of prior distribution $var(\mathbf{b}) = \tilde{\Psi} \otimes (1/\lambda^{(bvar)}) \cdot \tilde{\Omega}$

implies:

$$var[(A_l)_{ij}] = \begin{cases} \frac{(1/\lambda^{(bvar)})}{l^{\pi}}, & \text{if } i = j, \\ \frac{(1/\lambda^{(bvar)})}{l^{\pi}} \frac{\sigma_i^2}{\sigma_j^2}, & \text{if } i \neq j, \end{cases}$$

where $(A_l)_{ij}$ is the (i, j)-th element of the l-th lag coefficient matrix $A_l, l = 1, ..., p$. This is suggested by Litterman (1986). We can see that $\lambda^{(bvar)}$ determines the overall magnitude of prior variance of each coefficient $(A_l)_{ij}$. The term l^{π} is added to reflect the declining effects of lagged variables. The ratio (σ_i^2/σ_j^2) is used to adjust the difference in units of measurement between different variables i and j.

The posterior mean $\hat{\mathbf{b}}^{(bvar)}$, which we will use as the Bayesian VAR estimate, is obtained from stacking the columns of the matrix $\hat{B}^{(bvar)}$, given by:

$$\hat{B}^{(bvar)} = (\lambda^{(bvar)}\tilde{\Omega}^{-1} + X'X)^{-1}X'Y.$$

Let $\hat{B}_i^{(bvar)}$ be the i-th column of the estimate matrix $\hat{B}^{(bvar)}$. The estimator of $\hat{B}_i^{(bvar)}$ can be written as:

$$\hat{B}_i^{(bvar)} = (\lambda^{(bvar)}\tilde{\Omega}^{-1} + X'X)^{-1}X'Y_i. \tag{4.7}$$

When the OLS estimator in (4.4) is well defined, the estimator in (4.7) is the same as:

$$\hat{B}_i^{(bvar)} = (\lambda^{(bvar)}\tilde{\Omega}^{-1} + X'X)^{-1}(X'X)\hat{B}_i^{(ols)}. \tag{4.8}$$

It is clear from (4.8) that the (Euclidean) norm of this estimate is shorter than that of the OLS estimate, $\|\hat{B}_i^{(bvar)}\| < \|\hat{B}_i^{(ols)}\|$.

Bańbura et al. (2008) show that Bayesian VAR helps avoid the problem of undefined OLS estimator, when the number of independent variables np is a lot larger than the number of observations T. Even when the OLS estimator is well defined, Litterman (1986), Robertson and Tallman (1999), and our previous chapter, for example, show that Bayesian VAR clearly outperforms the unrestricted OLS VAR in out-of-sample forecasting.

4.2.2 Ridge Regression

Ridge regression was originally introduced by Hoerl and Kennard (1970) to deal with the problem of singularity of the matrix $X'X$ in an OLS regression. However, we can look at it as a regression that imposes a penalty on the size of the estimated coefficients as well. The ridge regression estimate $\hat{B}_i^{(rr)}$ can be defined as:

$$\hat{B}_i^{(rr)} = \underset{\tilde{B}_i \in \mathbb{R}^{np}}{\operatorname{argmin}} \left\{ (Y_i - X\tilde{B}_i)'(Y_i - X\tilde{B}_i) + \lambda^{(rr)}(\tilde{B}_i'\tilde{B}_i) \right\}, \tag{4.9}$$

where $\lambda^{(rr)} > 0$ is the ridge parameter that penalizes larger L_2 norm of the estimate $\hat{B}_i^{(rr)}$.

The ridge regression estimator of $\hat{B}_i^{(rr)}$ is:

$$\hat{B}_i^{(rr)} = (\lambda^{(rr)} I + X'X)^{-1} X'Y_i, \tag{4.10}$$

or:

$$\hat{B}_i^{(rr)} = (\lambda^{(rr)} I + X'X)^{-1} (X'X) \hat{B}_i^{(ols)},$$

when the OLS estimator in (4.4) is well defined. It is well known that even though the ridge regression estimator shrinks each estimated coefficient toward zero, it does not put any at exactly zero. The ridge regression estimate is not invariant under scaling of the independent variables. This can be seen from our proof of Proposition 4.2 in the next section. In practice, the independent variables are usually standardized before being applied with ridge regression.

In the static regression context, there is ample evidence that the ridge regression outperforms OLS in out-of-sample forecasting. This is explained using the well-known bias-variance tradeoff in statistical decision theory. See Chapter 2 of Hastie et al. (2001) for more details on the topic. This bias-variance tradeoff can *not* be written explicitly for the case of VAR. However, we expect that similar logic should be applicable. Accordingly, shrinkage regressions, like ridge, tend to work well under the conditions that 1) the norm of the true unknown coefficients $\|B_i\|$ is not too large, 2) there is a strong multicollinearity between independent variables, and 3) the number of observations T is not too large in comparison with the number of independent variables np. We see that these are all relevant to the VAR.

4.2.3 Bayesian VAR as An Augmented Ridge Regression

We want to show a close relationship between Bayesian VAR and ridge regression in this section. We start from our first observation that the estimator (4.7) can be seen as the solution of the optimization problem:

$$\hat{B}_i^{(bvar)} = \underset{\tilde{B}_i \in \mathbb{R}^{np}}{\operatorname{argmin}} \left\{ (Y_i - X\tilde{B}_i)'(Y_i - X\tilde{B}_i) + \lambda^{(bvar)}(\tilde{B}_i'\tilde{\Omega}^{-1}\tilde{B}_i) \right\}. \qquad (4.11)$$

We state this as our first result:

Proposition 4.1 *The posterior mean of Bayesian VAR with modified Litterman prior is the solution of the optimization problem (4.11).*

Proof. The first-order condition of the problem (4.11) is:

$$-2X'(Y_i - X\hat{B}_i^{(bvar)}) + 2\lambda^{(bvar)}\tilde{\Omega}^{-1}\hat{B}_i^{(bvar)} = 0,$$
$$(\lambda^{(bvar)}\tilde{\Omega}^{-1} + X'X)\hat{B}_i^{(bvar)} = X'Y_i,$$

which yields the estimator of $\hat{B}_i^{(bvar)}$ as in (4.7). ■

Proposition 4.1 sees Bayesian VAR as a regression that penalizes weighted sum of the squared size of the coefficients. This is the reason that Bayesian VAR, like ridge regression, usually does not put any estimated coefficient at zero.

Next, define $\hat{y}_{i,T+1|T}^{(\mu)} = \hat{B}_i^{(\mu)\prime}X_{T+1}$ as the period $T+1$ forecast from a regression μ given information up to time T, where $X_{T+1} = (Y_T', ..., Y_{T-p+1}')'$ be the $np \times 1$ vector of all endogenous variables from period $T-p+1$ to period T. Let $X^* = X\tilde{\Omega}^{1/2}$ be a matrix of scaled independent variables, and $X_{T+1}^* = \tilde{\Omega}^{1/2}X_{T+1}$ be the corresponding scaling of the vector X_{T+1}. Our next result says that the ridge regression forecast of period $T+1$ using X^* and X_{T+1}^* is exactly the same as the Bayesian VAR forecast of the period using X and X_{T+1}.

Proposition 4.2 *Provided that* $\lambda^{(rr)} = \lambda^{(bvar)}$*, the Bayesian VAR forecast* $\hat{y}_{i,T+1|T}^{(bvar)}$ *is exactly the same as the ridge regression forecast* $\hat{y}_{i,T+1|T}^{(rr)*} = \hat{B}_i^{(rr)*\prime}X_{T+1}^*$*, where:*

$$\hat{B}_i^{(rr)*} = (\lambda^{(rr)}I + X^{*\prime}X^*)^{-1}X^{*\prime}Y_i.$$

Proof. The estimator of $\hat{B}_i^{(rr)*}$ can be re-written as:

$$\begin{aligned}
\hat{B}_i^{(rr)*} &= (\lambda^{(rr)} I + X^{*\prime}X^*)^{-1}X^{*\prime}Y_i, \\
&= (\tilde{\Omega}^{1/2}\lambda^{(rr)}\tilde{\Omega}^{-1}\tilde{\Omega}^{1/2} + \tilde{\Omega}^{1/2}X'X\tilde{\Omega}^{1/2})^{-1}\tilde{\Omega}^{1/2}X'Y_i, \\
&= \tilde{\Omega}^{-1/2}(\lambda^{(bvar)}\tilde{\Omega}^{-1} + X'X)^{-1}X'Y_i, \\
&= \tilde{\Omega}^{-1/2}\hat{B}_i^{(bvar)}.
\end{aligned}$$

Since we have $X^*_{T+1} = \tilde{\Omega}^{1/2}X_{T+1}$, $\hat{B}_i^{(rr)*} = \tilde{\Omega}^{-1/2}\hat{B}_i^{(bvar)}$ implies $\hat{y}^{(rr)*}_{i,T+1|T} = \hat{B}_i^{(bvar)\prime}\tilde{\Omega}^{-1/2}\tilde{\Omega}^{1/2}X_{T+1} = \hat{y}^{(bvar)}_{i,T+1|T}$. ■

If the parameter π in the matrix $\tilde{\Omega}$ from (4.6) is set to be $\pi = 0$, the transformation X^* is actually the standardization of X. Since it is common in the literature to standardize the matrix X before applying ridge regression, Proposition 4.2 tells us that our Bayesian VAR with $\pi = 0$ will just produce the same forecasts as ridge regression under the standard practice.

However, it is common in application that π is set at $\pi = 1$ or $\pi = 2$. See, for example, the settings in Robertson and Tallman (1999), Bańbura et al. (2008), or Kadiyala and Karlsson (1997). This may cause some differences in the forecasts between the standard practice of ridge regression and Bayesian VAR. We develop a relationship between these two in the following.

Let $S = diag(1/\sigma_1, ..., 1/\sigma_n; ...; 1/\sigma_1, ..., 1/\sigma_n)$ be the standardization matrix, and $W = diag(1^\pi, ..., 1^\pi; ...; p^\pi, ..., p^\pi)$ be the weight matrix. It is easy to see that $X^s = XS$ is the matrix of standardized independent variables, and $W^{-1}S^2$ is equal to the parameter $\tilde{\Omega}$ of Bayesian VAR. Define the augmented ridge regression estimator $\hat{B}_i^{(arr)}$ as the solution of the optimization:

$$\hat{B}_i^{(arr)} = \underset{\tilde{B}_i \in \mathbb{R}^{np}}{\operatorname{argmin}} \left\{ (Y_i - X^s\tilde{B}_i)'(Y_i - X^s\tilde{B}_i) + \lambda^{(arr)}(\tilde{B}_i' W \tilde{B}_i) \right\}. \qquad (4.12)$$

We have that:

Corollary 1 *Given that* $\lambda^{(arr)} = \lambda^{(bvar)}$*, the augmented ridge regression forecast* $\hat{y}^{(arr)}_{i,T+1|T} = \hat{B}^{(arr)\prime}_i X^s_{T+1}$ *is exactly the same as the Bayesian VAR forecast* $\hat{y}^{(bvar)}_{i,T+1|T}$.

Proof. From Proposition 4.1 and the proof of Proposition 4.2, we have:

$$\begin{aligned}
\hat{B}^{(arr)}_i &= (\lambda^{(arr)} W + X^{s\prime} X^s)^{-1} X^s Y_i, \\
&= (S\lambda^{(arr)} W S^{-2} S + SX'XS)^{-1} SX'Y_i, \\
&= S^{-1}(\lambda^{(bvar)} \tilde{\Omega}^{-1} + X'X)^{-1} X'Y_i, \\
&= S^{-1} \hat{B}^{(bvar)}_i.
\end{aligned}$$

Then, the statement of the corollary follows. ■

Problem (4.12) can also be seen as applying Bayesian VAR with parameter $\tilde{\Omega} = W^{-1}$ to the standardized independent variables. Note that since we use an iterative process in generating forecasts for other forecast horizons, Proposition 4.2 and Corollary 1 can be applied to these forecasts as well.

4.3 LASSO and the Elastic Net

4.3.1 LASSO

Since ridge regression usually keeps all coefficients nonzero, it is criticized because 1) it does not create a parsimonious model in comparison to OLS, and 2) it may generate huge prediction errors if the matrix of true (unknown) coefficients is sparse. The technique called LASSO (Least Absolute Shrinkage and Selection Operator), proposed by Tibshirani (1996), can avoid these problems.

The LASSO estimate $\hat{B}_i^{(lasso)}$ can be defined as:

$$\hat{B}_i^{(lasso)} = \underset{\tilde{B}_i \in \mathbb{R}^{np}}{\operatorname{argmin}} \left\{ (Y_i - X\tilde{B}_i)'(Y_i - X\tilde{B}_i) + \lambda^{(lasso)} |\tilde{B}_i|_1 \right\}, \quad (4.13)$$

where $|\tilde{B}_i|_1 = \sum_{j=1}^{np} |\tilde{b}_{ji}|$ denotes the L_1 norm of the vector $\tilde{B}_i$, while $\tilde{b}_{ji}, j = 1, ..., np$, is the j-th element of the vector $\tilde{B}_i$. The estimate $\hat{B}_i^{(lasso)}$ sets some coefficients at zeros, creating a parsimonious model. Unfortunately, there is no closed-form solution for the problem (4.13) in general.

A closely related definition of the LASSO estimate to (4.13) is:

$$\hat{B}_i^{(lasso)} = \underset{\tilde{B}_i \in \mathbb{R}^{np}}{\operatorname{argmin}} (Y_i - X\tilde{B}_i)'(Y_i - X\tilde{B}_i), \quad (4.14)$$
$$\text{subject to } |\tilde{B}|_1 \leq \theta^{(lasso)},$$

where $\theta^{(lasso)}$ is a scalar governing the size of the L_1 norm of the LASSO estimate $\hat{B}_i^{(lasso)}$. These two definitions are equivalent; that is, for a given $\lambda^{(lasso)}$, $0 \leq \lambda^{(lasso)} < \infty$, there exists a $\theta^{(lasso)}$ such that the two definitions share the same estimate (Osborne, Presnell, and Turlach, 2000b). It is a lot more convenient for a computational algorithm to use (4.14), as it can check for the magnitude of the L_1 norm of $\hat{B}_i^{(lasso)}$ easily.

There are various methods to solve for the LASSO estimate. See Schmidt (2005) for a survey on these methods. We use the algorithm called LARS-LASSO here. To solve for the LASSO estimate, first note that the estimate $\hat{B}_i^{(lasso)}$ solves the necessary condition:

$$2X'(Y_i - X\hat{B}_i^{(lasso)}) = \lambda^{(lasso)} \Xi^{(lasso)}, \quad (4.15)$$

where $\Xi^{(lasso)} = (\xi_1, ..., \xi_{np})'$ is of the following form: $\xi_j = 1$ if $\hat{b}_{ji}^{(lasso)} > 0$,

$\xi_j = -1$ if $\hat{b}_{ji}^{(lasso)} < 0$, and $\xi_j \in [-1, 1]$ if $\hat{b}_{ji}^{(lasso)} = 0$; while $\hat{b}_{ji}^{(lasso)}$ is the j-th element of the LASSO estimate $\hat{B}_i^{(lasso)}$. Observe that according to (4.15), if X_j and X_k are the j-th and k-th variables in the matrix X (the j-th and k-th columns) with $\hat{b}_{ji}^{(lasso)}, \hat{b}_{ki}^{(lasso)} \neq 0$, we have:

$$|X_j'(Y_i - X\hat{B}_i^{(lasso)})| = |X_k'(Y_i - X\hat{B}_i^{(lasso)})| = |\lambda^{(lasso)}/2|. \qquad (4.16)$$

LARS (Least Angle Regression) is a model selection algorithm, proposed by Efron et al. (2004), which adds a variable into the "active set", denoted by $\mathcal{A}$, one-by-one. Let $\hat{Y}_{i,\mathcal{A}}$ be the current forecast of Y_i according to the active set $\mathcal{A}$, and $\hat{c}_j = X_j'(Y_i - \hat{Y}_{i,\mathcal{A}})$ be the inner product between the j-th variable X_j and the current forecast error. Starting from an empty active set $\mathcal{A} = \emptyset$, the variable that possesses the highest absolute value $|\hat{c}_j|$ will be added into the active set $\mathcal{A}$ first. After that, LARS generates forecast in the direction that equates $|\hat{c}_j|$ of variables j in the active set to the absolute value $|\hat{c}_k|$ of the newly added variable k. This means the variables j in the active set $\mathcal{A}$ share the same value of $|\hat{c}_j|$.

The LARS-LASSO algorithm is a modification of LARS that uses the condition (4.15) in determining the signs of the nonzero estimated coefficients. Accordingly, we have:

$$\text{sign}(X_j'(Y_i - X\hat{B}_i^{(lasso)})) = \text{sign}(\hat{b}_{ji}^{(lasso)}), \qquad (4.17)$$

for any nonzero estimate $\hat{b}_{ji}^{(lasso)} \neq 0$, and a positive penalty $\lambda^{(lasso)} > 0$. LARS-LASSO uses the constraint in (4.14) as its stopping condition. See Appendix E for more details about LARS and its LARS-LASSO modification.

The simulation study in Tibshirani (1996) shows that LASSO performs much better than ridge regression if the vector of true (unknown) coefficients B_i possesses 1) a lot of zeros and a small number of large values, or 2) some zeros and a moderate number of moderate-sized values. LASSO is also shown in Zou and Hastie (2005) and chapter 3 of Hastie et al. (2001) to outperform ridge regression in real applications using biological prostate cancer data.

4.3.2 The Elastic Net

Zou and Hastie (2005) provide three scenarios that make the implementation of LASSO inappropriate:

(1) If the number of independent variables is higher than the number of observations $N > T$, LASSO can select at most T variables.

(2) If there is a group of variables with very high pairwise correlation, LASSO tends to select only one variable from the group and does not care which one is selected.

(3) In an empirical assessment, if there is high multicollinearity between independent variables, LASSO is usually dominated by ridge regression.

All of these are relevant to our study.

Zou and Hastie (2005), hence, propose the elastic net that combines the LASSO and ridge penalties. The naive version of the elastic net estimate $\hat{B}_i^{(nen)}$ can be defined as:

$$\hat{B}_i^{(nen)} = \underset{\tilde{B}_i \in \mathbb{R}^{np}}{\operatorname{argmin}} \left\{ (Y_i - X\tilde{B}_i)'(Y_i - X\tilde{B}_i) + \lambda_1^{(en)} |\tilde{B}_i|_1 + \lambda_2^{(en)} (\tilde{B}_i'\tilde{B}_i) \right\}. \quad (4.18)$$

However, this naive version tends to over-shrink the estimated coefficients in many empirical examinations. The elastic net estimate $\hat{B}_i^{(en)}$ is therefore defined as:

$$\hat{B}_i^{(en)} = (1 + \lambda_2^{(en)})\hat{B}_i^{(nen)}. \tag{4.19}$$

Generally, there is no closed-form solutions for the problem (4.18).

The elastic net is shown to outperform LASSO in various simulations and empirical studies. These also include the cases when LASSO dominates ridge regression. This finding may not be surprising, however, if we allow the parameter $\lambda_2^{(en)}$ to be very close to zero. It means the elastic net considers LASSO forecast as one of its options.

Problem (4.18) can be transformed into another problem similar to (4.13) of LASSO. First, define some artificial variables as follow:

$$X^+ = (1 + \lambda_2^{(en)})^{-1/2} \begin{pmatrix} X \\ \sqrt{\lambda_2^{(en)}} I_{np} \end{pmatrix}, \; Y_i^+ = \begin{pmatrix} Y_i \\ 0_{np} \end{pmatrix}, \text{ and } \delta = \frac{\lambda_1^{(en)}}{\sqrt{1 + \lambda_2^{(en)}}}.$$

Next, let $B_i^+ = (1 + \lambda_2^{(en)})^{1/2}\tilde{B}_i$ be an artificial vector of coefficients, and $\hat{B}_i^+$ be the solution of the problem:

$$\hat{B}_i^+ = \underset{B_i^+ \in \mathbb{R}^{np}}{\operatorname{argmin}} \left\{ (Y_i^+ - X^+B_i^+)'(Y_i^+ - X^+B_i^+) + \delta |B_i^+|_1 \right\}. \tag{4.20}$$

Straightforward calculation shows that the naive elastic net estimate is $\hat{B}_i^{(nen)} = (1 + \lambda_2^{(en)})^{-1/2}\hat{B}_i^+$. Hence, the elastic net estimate $\hat{B}_i^{(en)}$ can be written as:

$$\hat{B}_i^{(en)} = (1 + \lambda_2^{(en)})^{1/2}\hat{B}_i^+. \tag{4.21}$$

Solving for $\hat{B}_i^+$ from the problem (4.20) is exactly the same as solving for the

LASSO estimate $\hat{B}_i^{(lasso)}$ from the problem (4.13). Zou and Hastie (2005) propose applying LARS to the problem (4.20) and adjust the estimate according to equation (4.21). This process is called LARS-EN.

4.3.3 The Modified Elastic Net

We are interested in applying different weights of the coefficients into the elastic net. As seen in section 4.2.3, this will allow us to standardize the independent variables as well as putting unequal effects to different-lagged variables. We define the modified naive elastic net estimate $\hat{B}_i^{(mnen)}$ as the solution of the optimization problem:

$$\hat{B}_i^{(mnen)} = \underset{\tilde{B}_i \in \mathbb{R}^{np}}{\operatorname{argmin}} \left\{ \begin{array}{l} (Y_i - X\tilde{B}_i)'(Y_i - X\tilde{B}_i) \\ \qquad +\lambda_1^{(men)}|\tilde{\Omega}^{-1/2}\tilde{B}_i|_1 + \lambda_2^{(men)}(\tilde{B}_i'\tilde{\Omega}^{-1}\tilde{B}_i) \end{array} \right\}, \tag{4.22}$$

and define the modified elastic net estimate $\hat{B}_i^{(men)}$ as:

$$\hat{B}_i^{(men)} = (1 + \lambda_2^{(men)})\hat{B}_i^{(mnen)}. \tag{4.23}$$

Note that if we have $\lambda_2^{(men)} = 0$, the modified elastic net estimate can be seen as the modified LASSO estimate as well.

To implement our modified elastic net, let $X^* = X\tilde{\Omega}^{1/2}$ be the matrix of scaled independent variables as in the previous section. Next, following Zou and Hastie (2005), define:

$$X^{++} = (1 + \lambda_2^{(men)})^{-1/2} \begin{pmatrix} X^* \\ \sqrt{\lambda_2^{(men)}} I_{np} \end{pmatrix}, \qquad Y_i^{++} = \begin{pmatrix} Y_i \\ \mathbf{0}_{np} \end{pmatrix},$$

and

$$\delta = \frac{\lambda_1^{(men)}}{\sqrt{1+\lambda_2^{(men)}}}.$$

Construct another artificial vector of coefficients $B_i^{++} = (1+\lambda_2^{(men)})^{1/2}\tilde{\Omega}^{-1/2}\tilde{B}_i$. Then, proposition 4.3 follows:

Proposition 4.3 *Let $\hat{B}_i^{++}$ be the solution of the problem:*

$$\hat{B}_i^{++} = \underset{B_i^{++}\in\mathbb{R}^{np}}{\text{argmin}} \left\{ (Y_i^{++} - X^{++}B_i^{++})'(Y_i^{++} - X^{++}B_i^{++}) + \delta|B_i^{++}|_1 \right\}.$$

The modified elastic net estimate $\hat{B}_i^{(men)}$ is given by:

$$\hat{B}_i^{(men)} = (1+\lambda_2^{(men)})^{1/2}\tilde{\Omega}^{1/2}\hat{B}_i^{++}.$$

Proof. Substitute $B_i^{++} = (1+\lambda_2^{(men)})^{1/2}\tilde{\Omega}^{-1/2}\tilde{B}_i$ into the optimization problem and multiply the matrices. The problem will turn into problem (4.22). Hence, the modified naive elastic net estimate is $\hat{B}_i^{(mnen)} = (1+\lambda_2^{(men)})^{-1/2}\tilde{\Omega}^{1/2}\hat{B}_i^{++}$ and the modified elastic net estimate as stated in the proposition. ■

Proposition 4.3 implies that we can use LARS-EN to implement our modified elastic net as well.

4.4 Simulation Study

Ideally, we should generate a large cross section of time series, apply the menus of above methods on each time series, and average the results across all time series. In that way, if a small number of series happen to possess unusual features, averaging across a large number of series will make their effect negligible. However, this practice is very costly in this study. This is because we have to determine parameters of various regressions for each of them. The process we use to determine these parameters, which will be explained more clearly below, involves grid searches that compare out-of-sample performances of various parameter values in a test period. It takes time to determine each parameter of each regression.

Here, to remedy this problem, we use the time dimension instead. With each data generating process, we simulate only one time series. However, we simulate a large number of observations for this series. We next apply on the series the rolling window procedure usually used in empirical studies with real data. In the procedure, we consider a sub-period of the series each time. The next sub-period is made different from the current sub-period by adding in some new observations outside the current sub-period and deleting out some observations inside. We then average the results across these sub-periods. Since the generated data are stationary, we see that the problems from strange features in this exercise should be minimal.

We employ 3 data generating processes, each of which will be explained later. With each data generating process, we generate $1,800$ observations of $n = 3$ endogenous variables, starting from a (13×3) matrix of zeros (Resulting in totally $1,813$ observations). The same set of errors is used in each process. To minimize the effect of the initial condition, we discard the first 560 observations. This leaves $1,253$ observations of $n = 3$ endogenous variables in each data set. The first 253 observations will be used as the pre-evaluation period, and the remaining $1,000$ as the evaluation period.

For notational convenience, we also use $Y_t = (y_{1t}, y_{2t}, y_{3t})', t = 1, ..., 1253$ to denote our simulated series. We apply all the methods explained in previous sections to each simulated data set. We consider 4 forecast horizons, $h \in$

$\{1, 3, 6, 12\}$, and 3 decay rates of effects of lags, $\pi \in \{0, 1, 2\}$. The order of VAR in each regression is fixed at $p = 13$. We use the unrestricted VAR as our benchmark regression.

In addition to the regressions we described so far, we also add two additional procedures. First, we apply variable selection using Bayesian information criteria (BIC) to the unrestricted VAR. This is for comparison purposes. We figure out the BIC of the unrestricted VARs with 1 to 13 lags of every endogenous variable. The model with the smallest BIC value will be chosen. Second, we combine LASSO and Bayesian VAR together. Since some previous simulation studies show that ridge regression applies a more suitable level of shrinkage to the case of non-sparse matrix of true (unknown) coefficients, the practice of using LASSO as the variable selector and applying the shrinkage of Bayesian VAR may outperform applying Bayesian VAR alone.

Let $\hat{Y}^{(\mu,\pi)}_{\tau+h|\tau} = (\hat{y}^{(\mu,\pi)}_{1,\tau+h|\tau} \dots \hat{y}^{(\mu,\pi)}_{3,\tau+h|\tau})'$ denote the h-steps ahead forecast from period τ, obtained from the regression μ with the decay rate of lags π. The one-step ahead forecast is computed from:

$$\hat{Y}^{(\mu,\pi)\prime}_{\tau+1|\tau} = X'_{\tau+1}\hat{B}^{(\mu,\pi)}.$$

For horizons $h > 1$, we use the iterative process to produce the forecasts. Since we have $h < p$, we can recursively construct a matrix of independent variables $X^{(\mu,\pi)}_{\tau+h|\tau}$ according to:

$$X^{(\mu,\pi)}_{\tau+h|\tau} = (\hat{Y}^{(\mu,\pi)\prime}_{\tau+h-1|\tau}, \dots, \hat{Y}^{(\mu,\pi)\prime}_{\tau+1|\tau}, Y'_{\tau}, \dots, Y'_{\tau+h-p})',$$

using the forecasts $\hat{Y}^{(\mu,\pi)}_{\tau+h-1|\tau}, \dots, \hat{Y}^{(\mu,\pi)}_{\tau+1|\tau}$ and the observations $Y_{\tau}, \dots, Y_{\tau+h-p}$. The point estimate of the h-steps forecast, then, is computed from:

$$\hat{Y}^{(\mu,\pi)\prime}_{\tau+h|\tau} = X^{(\mu,\pi)\prime}_{\tau+h|\tau}\hat{B}^{(\mu,\pi)}.$$

We use $\hat{Y}^{(0)}_{\tau+h|\tau}$ to denote the h-steps ahead forecast from the benchmark unrestricted OLS VAR.

The observations from $t = 254$ to $t = 1253$ (1,000 observations) are used as our evaluation period. Let $\tau_0 = 254$ and $\tau_1 = 1253$ denote the first and the last observations of the period. For a given forecast horizon h, we compute forecasts $\hat{Y}^{(\mu,\pi)}_{\tau+h|\tau}$ in each period $\tau = \tau_0 - h, ..., \tau_1 - h$ (1,000 times), using the most recent 120 observations (Rolling scheme, 120 observations). That is we start by using the observations from $t = \tau_0 - h - 120$ to $t = \tau_0 - h$ to estimate the regression μ and construct the forecast $\hat{Y}^{(\mu,\pi)}_{\tau_0|\tau_0-h}$. Next, we use the observations from $t = \tau_0 - h - 119$ to $t = \tau_0 - h + 1$ to again estimate the regression μ and construct the forecast $\hat{Y}^{(\mu,\pi)}_{\tau_0+1|\tau_0-h+1}$. We repeat this process until we get all 1,000 forecasts for each regression.

The forecasting performance of each forecasting method is measured in terms of Mean Squared Forecast Error (MSFE). For a regression μ with parameter π and forecast horizon h, we have:

$$MSFE^{(\mu,\pi)}_{i,h} = \frac{1}{\tau_1 - \tau_0 + 1} \sum_{\tau=\tau_0-h}^{\tau_1-h} \left(y_{i,\tau+h} - \hat{y}^{(\mu,\pi)}_{i,\tau+h|\tau} \right)^2,$$

where i indexes the variable of interest, and $\hat{y}^{(\mu,\pi)}_{i,\tau+h|\tau}$ is the forecast for the i variable, which can be taken from the vector $\hat{Y}^{(\mu,\pi)}_{\tau+h|\tau}$. In this simulation study, we are interested in forecasting all 3 endogenous variables. This means $i = 1, 2, 3$.

We report the MSFE of each regression relative to the benchmark regression. That is, we report the Relative MSFE (RMSFE), given by:

$$RMSFE_{i,h}^{(\mu,\pi)} = \frac{MSFE_{i,h}^{(\mu,\pi)}}{MSFE_{i,h}^{(0)}}.$$

A number smaller than one for $RMSFE_{i,h}^{(\mu,\pi)}$ implies that the regression μ with parameter π performs better than the unrestricted OLS VAR.

The parameters $\lambda^{(\mu)}$ of a regression μ are chosen from out-of-sample performance in the test period. This test period covers the observations from $t = 134$ to $t = 253$ (120 observations). For the Bayesian VAR, we select the parameter $\lambda^{(bvar)}$ that minimizes the sum of $RMSFE_{i,h}^{(\mu,\pi)}$ of all 3 endogenous variables in our test period. We refer the readers to our previous chapter for more details of this selection procedure. For the modified LASSO, we cannot choose the value of parameter $\lambda^{(lasso)}$ directly. This is because the LARS algorithm, which we use to implement LASSO, does not allow us to do so. We hence select the ratio $s^{(lasso)} = |\hat{B}_i^{(lasso)}|_1/|\hat{B}_i^{(ols)}|_1$, and fix it in our evaluation period instead. This means we set $\theta^{(lasso)} = s^{(lasso)}|\hat{B}_i^{(ols)}|_1$ in each date. We set up a grid of 90 values, ranging from 0.01 to 0.90, to search for the ratio $s^{(lasso)}$ that minimizes the sum of $RMSFE_{i,h}^{(\mu,\pi)}$ of all 3 endogenous variables in the test period.

Similar to the modified LASSO, we cannot directly choose the parameters $\lambda_1^{(men)}$ and $\lambda_1^{(mnen)}$. We instead choose and fix the ratios $s^{(k)} = |\hat{B}_i^{(k)}|_1/|\hat{B}_i^{(ols)}|_1$, $k = men, mnen$. In searching for the pairs of parameters $(s^{(mnen)}, \lambda_2^{(mnen)})$ and $(s^{(men)}, \lambda_2^{(men)})$ for our modified elastic nets, we set up a grid of 420 pairs of values. This results from the Cartesian product of a set of 21 values of $s^{(k)}, k = men, mnen$, ranging from 0.10 to 0.90, and another set of 20 values of $\lambda^{(k)}, k = men, mnen$, ranging from 0.001 to 500. We use the same criteria in choosing the parameters. For the combination of LASSO and Bayesian VAR, we select variables using LASSO with its associated $s^{(lasso)}$ first, and select the parameter $\lambda^{(bvar)}$ again under the model with selected variables only.

In all of the following, it turns out that the out-of-sample performances of our modified elastic net (MEN) are usually outperformed by either the LASSO or the modified naive elastic net (MNEN). Hence, we do not report its results.

4.4.1 Simulation 1

In this simulation, we generate the data according to the VAR in (4.1):

$$Y_t = B^{(s1)\prime} X_t + U_t, \quad \text{where } U_t \sim \text{i.i.d. } N(\mathbf{0}, \Psi^{(s1)}). \tag{4.24}$$

With data described in Section 4.5, we first estimate the unrestricted OLS VAR of order $p = 13$ with 3 endogenous variables, using the whole data set. Next, with a 5-percent significant level, we recursively eliminate insignificant estimates one-by-one, until there is none of them left. The resulted matrix of coefficients is used as $B^{(s1)}$, and the estimated covariance matrix as $\Psi^{(s1)}$ in (4.24). These matrices are shown in Appendix F. Our matrix $B^{(s1)}$ contains a lot of zeros, with the absolute values of non-zero elements between 0.03 and 0.55.

Table 4.1 reports the parameters from grid searches for various regressions. It also reports the associated sums of $RMSFE_{i,h}^{(\mu,\pi)}$ of the 3 endogenous variables during the test period. For each regression, the first column contains the parameter(s), while the second column the sum of $RMSFE_{i,h}^{(\mu,\pi)}$. For each regression and each forecast horizon h, the bold-faced numbers indicate the smallest sum of $RMSFE_{i,h}^{(\mu,\pi)}$ across the decay rates π. That is, for example,

Table 4.1: **Parameters from grid searches, Simulation 1**

		BVAR		LASSO		LASSO+BVAR		MNEN	
		$\lambda^{(bvar)}$	*sum*	$s^{(lasso)}$	*sum*	$s^{(lasso)}$, $\lambda^{(bvar)}$	*sum*	$s^{(mnen)}$, $\lambda_2^{(mnen)}$	*sum*
$\pi = 0$	$h = 1$	156.25	2.299	0.20	2.185	0.20, 118.15	2.284	0.22, 0.10	2.185
	$h = 3$	287.27	2.249	**0.20**	**2.192**	0.20, 173.13	2.234	**0.38, 1.00**	**2.192**
	$h = 6$	**369.82**	**2.349**	**0.21**	**2.280**	**0.21, 177.78**	**2.317**	**0.22, 0.05**	**2.280**
	$h = 12$	**516.53**	**2.656**	**0.16**	**2.559**	**0.16, 198.37**	**2.611**	**0.18, 0.10**	**2.559**
$\pi = 1$	$h = 1$	**39.06**	**2.220**	**0.19**	**2.165**	**0.19, 27.41**	**2.195**	**0.18, 0.001**	**2.166**
	$h = 3$	**45.65**	**2.246**	0.21	2.196	**0.21, 29.54**	**2.223**	0.22, 0.05	2.196
	$h = 6$	53.28	2.359	0.18	2.294	0.18, 26.85	2.340	0.18, 0.001	2.294
	$h = 12$	94.26	2.670	0.16	2.568	0.16, 24.03	2.641	0.18, 0.10	2.569
$\pi = 2$	$h = 1$	11.81	2.221	0.17	2.198	0.17, 10.61	2.222	0.14, 0.001	2.199
	$h = 3$	8.60	2.256	0.17	2.220	0.17, 7.00	2.252	0.18, 0.001	2.221
	$h = 6$	8.75	2.371	0.17	2.315	0.17, 5.86	2.358	0.18, 0.001	2.316
	$h = 12$	23.11	2.680	0.16	2.597	0.16, 9.47	2.672	0.18, 0.001	2.599

in the case of Bayesian VAR the decay rate $\pi = 1$ yields the smallest RMSFE for the forecast horizons $h = 1, 3$, and the decay rate $\pi = 0$ for the horizons $h = 6, 12$.

One interesting feature of Table 4.1 is that given a forecast horizon h, when the decay rate of lags π is higher, the grid search for each regression usually yields parameter(s) associated with smaller penalty on the magnitude of the estimate $\hat{B}_i^{(\mu)}$. For example, given the horizon $h = 1$, the parameters of Bayesian VAR $\lambda^{(bvar)}$ are at 156.25, 39.06, and 11.81, when the decay rate π are at 0,1, and 2, respectively. Our grid search yields parameters that try to offset the decay rates on the effects of lagged variables.

Using the parameters reported in Table 4.1, the out-of-sample RMSFEs of each regression in the evaluation period are shown in Table 4.2. For the LASSO, the combination of LASSO and Bayesian VAR (LASSO+BVAR), and the modified naive elastic net (MNEN), we also report the average numbers of variables selected during the evaluation period (1,000 estimations). These numbers are reported in the column named "no. var". The bold-faced numbers in each column are those that on average perform best across the values of the decay rate π for each forecast horizon h.

Table 4.2: RMSFEs of different regressions, Simulation 1

$\pi = 0$		VAR+BIC	BVAR	LASSO		LASSO+BVAR		MNEN	
		RMSFE	RMSFE	RMSFE	no. var	RMSFE	no. var	RMSFE	no. var
$h = 1$	var1	0.838	0.784	0.760	10.30	0.777	10.30	0.760	10.31
	var2	0.936	0.865	0.887	13.14	0.875	13.14	0.887	13.15
	var3	1.004	0.832	0.841	10.87	0.828	10.87	0.841	10.88
	average	0.926	0.827	0.829		0.827		0.829	
$h = 3$	var1	0.867	0.837	0.799	10.32	0.829	10.32	0.801	9.74
	var2	0.913	0.849	0.871	13.14	0.865	13.14	0.871	12.73
	var3	1.016	0.842	0.845	10.87	0.845	10.87	0.850	10.44
	average	0.932	0.843	0.838		0.846		0.841	
$h = 6$	var1	0.817	0.810	0.791	11.00	0.818	11.00	0.791	10.98
	var2	0.880	0.839	0.856	13.66	0.843	13.66	0.856	13.63
	var3	1.051	0.893	0.884	11.35	0.886	11.35	0.884	11.33
	average	0.916	**0.847**	0.844		0.849		0.844	
$h = 12$	var1	0.822	0.840	0.830	7.76	0.839	7.76	0.830	8.01
	var2	0.840	0.847	0.850	10.77	0.848	10.77	0.850	11.02
	var3	0.985	0.923	0.904	8.81	0.908	8.81	0.903	8.98
	average	0.882	**0.870**	0.861		0.865		0.861	

$\pi = 1$		VAR+BIC	BVAR	LASSO		LASSO+BVAR		MNEN	
		RMSFE	RMSFE	RMSFE	no. var	RMSFE	no. var	RMSFE	no. var
$h = 1$	var1	0.838	0.753	0.743	13.00	0.754	13.00	0.741	12.30
	var2	0.936	0.846	0.860	14.19	0.854	14.19	0.863	13.66
	var3	1.004	0.865	0.825	11.66	0.849	11.66	0.829	11.10
	average	0.926	**0.821**	**0.809**		**0.819**		**0.811**	
$h = 3$	var1	0.867	0.785	0.772	14.35	0.777	14.35	0.771	14.35
	var2	0.913	0.850	0.862	15.18	0.850	15.18	0.862	15.18
	var3	1.016	0.865	0.831	12.71	0.854	12.71	0.831	12.71
	average	0.932	**0.833**	**0.821**		**0.827**		**0.821**	
$h = 6$	var1	0.817	0.791	0.774	12.33	0.783	12.33	0.774	12.32
	var2	0.880	0.844	0.856	13.66	0.846	13.66	0.856	13.65
	var3	1.051	0.909	0.889	11.10	0.900	11.10	0.889	11.09
	average	0.916	0.848	0.840		**0.843**		0.840	
$h = 12$	var1	0.822	0.835	0.828	10.91	0.838	10.91	0.829	11.17
	var2	0.840	0.846	0.851	12.53	0.857	12.53	0.851	12.81
	var3	0.985	0.937	0.891	10.01	0.896	10.01	0.890	10.27
	average	0.882	0.873	0.857		**0.863**		**0.857**	

$\pi = 2$		VAR+BIC	BVAR	LASSO		LASSO+BVAR		MNEN	
		RMSFE	RMSFE	RMSFE	no. var	RMSFE	no. var	RMSFE	no. var
$h = 1$	var1	0.838	0.746	0.752	15.71	0.745	15.71	0.744	13.77
	var2	0.936	0.861	0.861	15.97	0.864	15.97	0.866	14.40
	var3	1.004	0.916	0.839	13.58	0.919	13.58	0.853	11.86
	average	0.926	0.841	0.817		0.843		0.821	
$h = 3$	var1	0.867	0.758	0.765	15.72	0.757	15.72	0.767	16.33
	var2	0.913	0.857	0.864	15.97	0.858	15.97	0.863	16.43
	var3	1.016	0.897	0.845	13.58	0.908	13.58	0.843	14.14
	average	0.932	0.838	0.825		0.841		0.824	
$h = 6$	var1	0.817	0.776	0.772	15.73	0.772	15.73	0.774	16.34
	var2	0.880	0.849	0.853	15.97	0.852	15.97	0.853	16.43
	var3	1.051	0.929	0.883	13.58	0.927	13.58	0.880	14.13
	average	0.916	0.852	**0.836**		0.850		**0.835**	
$h = 12$	var1	0.822	0.830	0.836	15.11	0.830	15.11	0.839	16.34
	var2	0.840	0.845	0.857	15.48	0.848	15.48	0.860	16.43
	var3	0.985	0.954	0.874	12.99	0.935	12.99	0.871	14.12
	average	0.882	0.877	**0.856**		0.871		0.857	

From Table 4.2, it is obvious that each regression outperforms the OLS VAR, which is our benchmark model. The performance gain is about 10 to 20 percent for every regression. On average, all shrinkage regressions can outperform the use of BIC with OLS VAR as well. This gain in performances presents for every decay rate π applied.

Looking at the best performances of each regression for a given forecast horizon h, the LASSO and the MNEN turn out to be the best among all regressions. These two outperform the Bayesian VAR by about 0.5 to 2 percent, and outperform the use of BIC on OLS VAR by about 10 percent. Another interesting feature is that out of 39 independent variables (13 lags of 3 endogenous variables), the LASSO and the modified naive elastic net can obtain their performances from just about one third (no. var is about 13) of these variables.

In our simulation 1, the best decay rate for forecast horizons $h = 1, 3$ is at $\pi = 1$. However, it varies for $h = 6, 12$. We can see some differences in the performances across the decay rates, even our grid search tends to obtain parameter values that try to offset the effects of the decay rates, as explained before.

4.4.2 Simulation 2

We use the estimated coefficients from the unrestricted OLS VAR on the whole data set, without setting any coefficients to zero. The resulted matrix of estimated coefficients $B^{(s2)}$ and $\Psi^{(s2)}$ will replace $B^{(s1)}$ and $\Psi^{(s1)}$ in (4.24). The matrix $B^{(s2)}$ and $\Psi^{(s2)}$ are also shown in Appendix F. All the elements in the matrix $B^{(s2)}$ have their absolute values between 0.002 and 0.53. To make

Table 4.3: Parameters from grid searches, Simulation 2

		BVAR		LASSO		LASSO+BVAR		MNEN	
		$\lambda^{(bvar)}$	*sum*	$s^{(lasso)}$	*sum*	$s^{(lasso)}$, $\lambda^{(bvar)}$	*sum*	$s^{(mnen)}$, $\lambda_2^{(mnen)}$	*sum*
$\pi = 0$	$h = 1$	120.76	2.404	0.28	2.412	0.28, 126.25	2.443	0.42, 0.50	2.411
	$h = 3$	192.90	2.357	0.25	2.378	**0.25, 168.66**	**2.376**	0.50, 1	2.377
	$h = 6$	**268.74**	**2.322**	**0.16**	**2.315**	0.16, 216.26	2.375	**0.18, 0.10**	**2.315**
	$h = 12$	**356.00**	**2.573**	**0.15**	**2.502**	**0.15, 229.57**	**2.568**	**0.22, 0.50**	**2.503**
$\pi = 1$	$h = 1$	**27.99**	**2.342**	**0.23**	**2.359**	**0.23, 30.86**	**2.369**	**0.46, 1**	**2.358**
	$h = 3$	**30.19**	**2.355**	**0.20**	**2.370**	0.20, 32.28	2.380	**0.42, 1**	**2.370**
	$h = 6$	42.72	2.339	0.13	2.318	**0.13, 31.92**	**2.359**	0.14, 0.01	2.318
	$h = 12$	57.39	2.595	0.15	2.514	0.15, 23.80	2.570	0.14, 0.001	2.515
$\pi = 2$	$h = 1$	7.04	2.367	0.15	2.369	0.15, 10.82	2.394	0.14, 0.001	2.371
	$h = 3$	5.59	2.375	0.17	2.381	0.17, 7.89	2.393	0.18, 0.05	2.381
	$h = 6$	7.89	2.358	0.14	2.341	0.14, 8.86	2.380	0.14, 0.001	2.341
	$h = 12$	11.81	2.607	0.13	2.543	0.13, 6.47	2.593	0.14, 0.001	2.545

the results comparable to simulation 1, we use the same set of simulated errors and the same procedure to generate this data set.

Table 4.3 reports the parameters from grid searches for various regressions, as well as the associated sums of $RMSFE_{i,h}^{(\mu,\pi)}$ in the test period. Observe that we still get the parameter(s) that offsets the difference in decay rate π. The bold-faced numbers indicate the smallest sum of $RMSFE_{i,h}^{(\mu,\pi)}$ across the decay rates π for each forecast horizon h.

Table 4.4 reports the out-of-sample RMSFEs of each regression in our evaluation period. Similar to results in simulation 1, shrinkage regressions outperform the unrestricted OLS VAR and the use of BIC to select variables for OLS VAR. However, the LASSO and the modified naive elastic net do not outperform Bayesian VAR in this simulated data set. It can be said that the performances of these three are the same for any given forecast horizon h.

4.4.3 Simulation 3

The data set is generated according to a VARMA(1,1) model, given by:

$$Y_t = B^{(s3)}Y_{t-1} + U_t + \Theta^{(s3)}U_{t-1}, \quad U_t \sim \text{i.i.d. } N(\mathbf{0}, \Psi^{(s3)}), \tag{4.25}$$

Table 4.4: **RMSFEs of different regressions, Simulation 2**

$\pi = 0$		**VAR+BIC**	**BVAR**	**LASSO**		**LASSO+BVAR**		**MNEN**	
		RMSFE	RMSFE	RMSFE	no. var	RMSFE	no. var	RMSFE	no. var
$h = 1$	var1	0.872	0.804	0.802	15.80	0.811	15.80	0.802	15.85
	var2	0.966	0.869	0.884	17.28	0.881	17.28	0.884	17.33
	var3	1.029	0.835	0.831	15.42	0.841	15.42	0.831	15.47
	average	0.956	0.836	0.839		0.845		0.839	
$h = 3$	var1	0.929	0.857	0.839	14.03	0.864	14.03	0.839	14.11
	var2	0.975	0.881	0.902	15.92	0.884	15.92	0.902	16.02
	var3	1.057	0.861	0.862	13.80	0.870	13.80	0.862	13.91
	average	0.987	0.866	0.868		0.873		0.867	
$h = 6$	var1	0.879	0.850	0.857	8.55	0.861	8.55	0.855	8.79
	var2	0.906	0.852	0.870	11.15	0.864	11.15	0.870	11.35
	var3	1.062	0.886	0.926	9.42	0.912	9.42	0.924	9.60
	average	0.949	**0.863**	0.884		**0.879**		0.883	
$h = 12$	var1	0.832	0.841	0.834	7.92	0.842	7.92	0.834	7.75
	var2	0.847	0.848	0.850	10.51	0.851	10.51	0.849	10.33
	var3	0.927	0.878	0.875	8.90	0.878	8.90	0.876	8.77
	average	0.869	**0.856**	0.853		0.857		0.853	

$\pi = 1$		**VAR+BIC**	**BVAR**	**LASSO**		**LASSO+BVAR**		**MNEN**	
		RMSFE	RMSFE	RMSFE	no. var	RMSFE	no. var	RMSFE	no. var
$h = 1$	var1	0.872	0.782	0.790	15.91	0.786	15.91	0.788	16.46
	var2	0.966	0.856	0.878	16.78	0.868	16.78	0.876	17.36
	var3	1.029	0.869	0.836	14.54	0.879	14.54	0.838	15.34
	average	0.956	**0.836**	**0.835**		**0.844**		**0.834**	
$h = 3$	var1	0.929	0.812	0.813	14.04	0.812	14.04	0.811	15.15
	var2	0.975	0.880	0.903	15.37	0.891	15.37	0.900	16.30
	var3	1.057	0.885	0.875	12.82	0.900	12.82	0.871	14.11
	average	0.987	**0.859**	**0.864**		**0.868**		**0.861**	
$h = 6$	var1	0.879	0.834	0.833	9.50	0.833	9.50	0.831	10.04
	var2	0.906	0.860	0.883	11.76	0.877	11.76	0.882	12.27
	var3	1.062	0.907	0.941	9.11	0.931	9.11	0.935	9.56
	average	0.949	0.867	0.886		0.880		0.882	
$h = 12$	var1	0.832	0.837	0.834	10.79	0.836	10.79	0.833	10.12
	var2	0.847	0.851	0.854	12.87	0.862	12.87	0.852	12.34
	var3	0.927	0.880	0.866	10.10	0.869	10.10	0.868	9.61
	average	0.869	0.856	**0.851**		**0.856**		**0.851**	

$\pi = 2$		**VAR+BIC**	**BVAR**	**LASSO**		**LASSO+BVAR**		**MNEN**	
		RMSFE	RMSFE	RMSFE	no. var	RMSFE	no. var	RMSFE	no. var
$h = 1$	var1	0.872	0.781	0.789	14.81	0.786	14.81	0.788	14.18
	var2	0.966	0.876	0.895	15.55	0.893	15.55	0.897	14.96
	var3	1.029	0.921	0.875	13.15	0.947	13.15	0.881	12.60
	average	0.956	0.860	0.853		0.875		0.855	
$h = 3$	var1	0.929	0.792	0.804	16.18	0.797	16.18	0.803	16.41
	var2	0.975	0.889	0.906	16.62	0.906	16.62	0.905	16.87
	var3	1.057	0.922	0.889	14.30	0.953	14.30	0.889	14.65
	average	0.987	0.868	0.866		0.885		0.866	
$h = 6$	var1	0.879	0.824	0.823	14.19	0.827	14.19	0.823	14.19
	var2	0.906	0.870	0.886	14.96	0.882	14.96	0.886	14.95
	var3	1.062	0.936	0.923	12.58	0.970	12.58	0.923	12.58
	average	0.949	0.876	**0.877**		0.893		**0.877**	
$h = 12$	var1	0.832	0.834	0.842	13.52	0.835	13.52	0.842	14.19
	var2	0.847	0.852	0.859	14.37	0.858	14.37	0.861	14.94
	var3	0.927	0.887	0.855	12.07	0.879	12.07	0.854	12.55
	average	0.869	0.858	0.852		0.857		0.852	

Table 4.5: Parameters from grid searches, Simulation 3

		BVAR		LASSO		LASSO+BVAR		MNEN	
		$\lambda^{(bvar)}$	*sum*	$s^{(lasso)}$	*sum*	$s^{(lasso)}$, $\lambda^{(bvar)}$	*sum*	$s^{(mnen)}$, $\lambda_2^{(mnen)}$	*sum*
$\pi = 0$	$h = 1$	2.85	2.630	0.38	2.362	0.38, 3.12	2.588	0.38, 0.001	2.362
	$h = 3$	29.22	2.429	0.33	2.465	0.33, 25.00	2.371	0.34, 0.05	2.466
	$h = 6$	342.94	2.376	0.12	2.386	0.12, 138.41	2.382	0.18, 0.50	2.386
	$h = 12$	10^6	2.196	0.01	2.195	0.01, 90.70	2.175	0.22, 30	2.195
$\pi = 1$	$h = 1$	2.64	2.339	0.20	2.211	0.20, 2.53	2.334	0.22, 0.10	2.214
	$h = 3$	17.22	2.236	0.19	2.386	0.19, 15.87	2.241	0.90, 5	2.347
	$h = 6$	87.34	2.273	0.06	2.318	0.06, 120.76	2.315	0.62, 10	2.320
	$h = 12$	10^6	2.195	0.01	2.195	0.01, 78.31	2.162	0.42, 50	2.195
$\pi = 2$	$h = 1$	**1.80**	**2.164**	**0.10**	**2.158**	**0.10, 1.95**	**2.163**	**0.22, 1**	**2.147**
	$h = 3$	**9.47**	**2.098**	**0.07**	**2.229**	**0.07, 9.41**	**2.076**	**0.86, 10**	**2.092**
	$h = 6$	**34.20**	**2.172**	**0.03**	**2.295**	**0.03, 40.57**	**2.190**	**0.82, 10**	**2.194**
	$h = 12$	**10^6**	**2.195**	**0.01**	**2.195**	**0.01, 66.10**	**2.156**	**0.62, 30**	**2.191**

where the matrices $B^{(s3)}$, $\Theta^{(s3)}$, and $\Psi^{(s3)}$ are the maximum likelihood estimates of VARMA(1,1) from the same data set as in simulations 1 and 2. These matrices will also be shown in Appendix F. The same set of simulated errors and the same procedure as in Simulations 1 and 2 are used to generate the data set.

Table 4.5 reports the parameters from grid searches, as well as the associated performances in the test period. Then, Table 4.6 reports the out-of-sample performances of the regressions in the evaluation period. The bold-faced numbers are still those that averagely perform best across the decay rates π for each forecast horizon h.

According to Table 4.6, the performances of LASSO turn out to be the worst, while ones of the combination of LASSO and BVAR are the best among shrinkage regressions. This is the opposite case to simulation 1, of which the results are in Table 4.2. Observe also that in this simulation, the application of BIC on OLS VAR provides the best performances of all in every forecast horizon.

Looking across all simulations, we can see that the LASSO can outperform the Bayesian VAR when the matrix of true (unknown) coefficients is a sparse matrix. Under this condition, it is desirable to work with LASSO as well,

Table 4.6: **RMSFEs of different regressions, Simulation 3**

$\pi = 0$		**VAR+BIC**	**BVAR**	**LASSO**		**LASSO+BVAR**		**MNEN**	
		RMSFE	RMSFE	RMSFE	no. var	RMSFE	no. var	RMSFE	no. var
$h = 1$	var1	0.704	0.915	0.862	11.20	0.835	11.20	0.863	11.17
	var2	0.699	0.844	0.838	28.37	0.837	28.37	0.838	28.37
	var3	0.709	0.864	0.808	23.19	0.848	23.19	0.808	23.17
	average	0.704	0.874	0.836		0.840		0.836	
$h = 3$	var1	0.687	0.813	0.838	5.35	0.797	5.35	0.840	4.76
	var2	0.696	0.776	0.846	27.12	0.775	27.12	0.844	27.05
	var3	0.687	0.787	0.838	20.12	0.782	20.12	0.837	19.69
	average	0.690	0.792	0.840		0.785		0.840	
$h = 6$	var1	0.706	0.805	0.927	1.18	0.801	1.18	0.927	1.18
	var2	0.745	0.779	0.836	17.54	0.788	17.54	0.836	17.84
	var3	0.702	0.772	0.822	2.72	0.752	2.72	0.822	2.73
	average	0.718	0.786	0.862		0.780		0.862	
$h = 12$	var1	0.749	0.830	0.831	1.00	0.772	1.00	0.831	1.00
	var2	0.845	0.846	0.847	3.21	0.850	3.21	0.846	2.97
	var3	0.772	0.839	0.840	1.00	0.787	1.00	0.840	1.00
	average	0.789	**0.838**	0.839		0.803		0.839	

$\pi = 1$		**VAR+BIC**	**BVAR**	**LASSO**		**LASSO+BVAR**		**MNEN**	
		RMSFE	RMSFE	RMSFE	no. var	RMSFE	no. var	RMSFE	no. var
$h = 1$	var1	0.704	0.822	0.857	8.66	0.793	8.66	0.858	8.80
	var2	0.699	0.773	0.795	23.99	0.773	23.99	0.793	24.84
	var3	0.709	0.800	0.783	19.15	0.797	19.15	0.783	19.54
	average	0.704	0.798	0.812		0.787		0.811	
$h = 3$	var1	0.687	0.766	0.800	7.08	0.756	7.08	0.823	3.21
	var2	0.696	0.747	0.810	23.56	0.741	23.56	0.770	37.50
	var3	0.687	0.748	0.795	18.46	0.741	18.46	0.780	22.47
	average	0.690	0.754	0.801		0.746		0.791	
$h = 6$	var1	0.706	0.773	0.938	1.18	0.786	1.18	0.953	1.14
	var2	0.745	0.768	0.809	14.01	0.762	14.01	0.800	25.48
	var3	0.702	0.752	0.833	2.67	0.737	2.67	0.840	2.76
	average	0.718	0.764	0.860		0.762		0.864	
$h = 12$	var1	0.749	0.831	0.831	1.00	0.769	1.00	0.831	1.00
	var2	0.845	0.846	0.846	3.43	0.847	3.43	0.846	4.55
	var3	0.772	0.840	0.840	1.00	0.784	1.00	0.840	1.00
	average	0.789	0.839	**0.839**		0.800		**0.839**	

$\pi = 2$		**VAR+BIC**	**BVAR**	**LASSO**		**LASSO+BVAR**		**MNEN**	
		RMSFE	RMSFE	RMSFE	no. var	RMSFE	no. var	RMSFE	no. var
$h = 1$	var1	0.704	0.766	0.752	10.17	0.755	10.17	0.746	17.25
	var2	0.699	0.740	0.768	19.95	0.736	19.95	0.751	37.36
	var3	0.709	0.767	0.772	16.75	0.758	16.75	0.768	31.16
	average	0.704	**0.758**	**0.764**		**0.749**		**0.755**	
$h = 3$	var1	0.687	0.730	0.794	3.70	0.719	3.70	0.743	11.38
	var2	0.696	0.726	0.771	17.30	0.716	17.30	0.724	39.04
	var3	0.687	0.720	0.761	13.05	0.709	13.05	0.722	37.89
	average	0.690	**0.725**	**0.775**		**0.715**		**0.730**	
$h = 6$	var1	0.706	0.747	0.895	1.48	0.767	1.48	0.776	9.07
	var2	0.745	0.756	0.785	11.77	0.752	11.77	0.767	39.04
	var3	0.702	0.731	0.816	3.04	0.730	3.04	0.760	37.28
	average	0.718	**0.745**	**0.832**		**0.750**		**0.768**	
$h = 12$	var1	0.749	0.831	0.831	1.00	0.766	1.00	0.831	1.11
	var2	0.845	0.846	0.847	6.36	0.845	6.36	0.847	37.24
	var3	0.772	0.840	0.840	1.37	0.782	1.37	0.840	2.90
	average	0.789	0.839	0.839		**0.798**		0.839	

since it outperforms the application of BIC on OLS VAR. The LASSO tends to be worse than the Bayesian VAR when the matrix of true coefficients is not sparse. Moreover, under the case that the true data generating process is a VARMA(1,1), there is no need to apply shrinkage regressions at all.

Even to a small extent, the decay rate π affects the forecasting performances of each regression as well. We see that this is another area to improve the forecasting performances of shrinkage regressions. Even our simulations show that better performances in the test period cannot insure better performances in the evaluation period, but the simulations show that the performances in the test period are not bad indicators of the performances in the evaluation period. If there is no superior substitutes, these can be used to select the decay rate to be applied with each regression.

4.5 Empirical Study

In this study, we employ the US data set of Stock and Watson (2005), which covers the period from *January 1959* to *December 2003*. There are 3 endogenous variables i, which are 1) the growth rate of employment (ΔEMPL), computed from the number of employees on non-farm payrolls, 2) the change of annual inflation (ΔINF), computed from the (seasonally adjusted) consumer price index, and 3) the change in (effective) Federal Fund Rate (ΔFFR). These variables are calculated as in De Mol et al. (2008) to obtain stationarity. That is we have ΔEMPL as the monthly growth rate of the number of employees on non-farm payrolls, ΔINF is calculated from $\Delta INF_t = 100 \times \{\log(CPI_t/CPI_{t-12}) - \log(CPI_{t-1}/CPI_{t-13})\}$, and ΔFFR is a monthly change.

Similar to the previous section, we consider 4 forecast horizons, $h \in \{1, 3, 6, 12\}$.

Table 4.7: **Parameters from grid searches, 3 variables, test period 1971-1980**

		BVAR		LASSO		LASSO+BVAR		MNEN	
		$\lambda^{(bvar)}$	*sum*	$s^{(lasso)}$	*sum*	$s^{(lasso)}$, $\lambda^{(bvar)}$	*sum*	$s^{(mnen)}$, $\lambda_2^{(mnen)}$	*sum*
$\pi = 0$	$h = 1$	173.13	1.943	**0.25**	**2.005**	**0.25,115.60**	**1.948**	**0.50, 1**	**2.004**
	$h = 3$	**216.26**	**2.062**	0.23	2.880	**0.23, 152.42**	**2.059**	**0.46, 1**	**2.088**
	$h = 6$	**318.88**	**2.021**	**0.20**	**1.969**	**0.20, 156.25**	**1.990**	**0.22, 0.1**	**1.969**
	$h = 12$	**108.51**	**2.700**	**0.26**	**2.739**	**0.26, 132.12**	**2.724**	**0.54, 1**	**2.738**
$\pi = 1$	$h = 1$	**53.28**	**1.940**	0.21	2.014	0.21, 35.86	1.965	0.42, 1	2.009
	$h = 3$	38.10	2.069	**0.17**	**2.113**	0.17, 23.80	2.083	0.38, 1	2.109
	$h = 6$	50.30	2.044	0.20	2.009	0.20, 24.03	2.032	0.30, 0.5	2.008
	$h = 12$	13.32	2.714	0.27	2.795	0.27, 17.22	2.747	0.62, 1	2.786
$\pi = 2$	$h = 1$	25.25	1.988	0.15	2.070	0.15, 24.75	2.022	0.86, 10	2.025
	$h = 3$	8.65	2.112	0.12	2.155	0.12, 4.96	2.135	0.90, 5	2.130
	$h = 6$	9.64	2.084	0.15	2.048	0.15, 4.83	2.084	0.14, 0.001	2.050
	$h = 12$	1.28	2.733	0.28	2.806	0.28, 1.57	2.755	0.86, 1	2.739

The order of the VAR in each regression is fixed at $p = 13$. We are interested in forecasting all of these 3 endogenous variables. The test period is set from *January 1971* to *December 1980* (120 times), while the evaluation period is from *January 1981* to *December 2003* (276 times). Let τ_0 and τ_1 denote the positions of the first and the last observations in each period. We produce the forecasts $\hat{Y}^{(\mu,\pi)}_{\tau+h|\tau}$ in each time $\tau = \tau_0 - h, ..., \tau_1 - h$ using the most recent 10 years of sample observations up to time τ (Rolling scheme, 120 observations).

Table 4.7 reports the parameter(s) of each regression μ from grid search, as well as its associated sum of $RMSFE^{(\mu,\pi)}_{i,h}$ of the 3 endogenous variables i during the test period. Since the out-of-sample performances of the modified elastic net (MEN) are dominated by the LASSO, we do not report its results here. As in the previous section, bold-faced numbers indicate the best performances across the decay rates π for each forecast horizon h.

Table 4.7 looks closer to Table 4.1 in Simulation 1 and Table 4.3 in Simulation 2 than Table 4.5 in Simulation 3. In most of the cases, the decay rate $\pi = 0$ performs best. Lag discounting does not seem to matter. The results from Table 4.7 do *not* imply the decaying effects of lagged variables as in the case of VARMA(1,1) in our Simulation 3.

Table 4.8: **RMSFEs of different regressions, 3 variables, evaluation period 1981-2003**

		VAR +BIC	BVAR	LASSO		LASSO + BVAR		MNEN	
		RMSFE	RMSFE	RMSFE	no. var	RMSFE	no. var	RMSFE	no. var
$h=1$	ΔEMPL	0.889	**0.778**	0.788	11.37	0.771	11.37	0.787	11.47
	ΔFFR	0.403	**0.409**	0.462	16.00	0.476	16.00	0.462	16.07
	ΔINF	0.830	**0.743**	0.681	11.72	0.712	11.72	0.681	11.79
	average	0.707	**0.643**	0.644		0.653		0.643	
$h=3$	ΔEMPL	0.852	0.862	**0.825**	10.69	0.880	10.15	0.827	10.23
	ΔFFR	0.523	0.616	**0.589**	13.79	0.639	14.93	0.637	15.01
	ΔINF	0.791	0.696	**0.652**	9.19	0.667	10.75	0.642	10.83
	average	0.722	0.725	**0.689**		0.729		0.702	
$h=6$	ΔEMPL	0.893	0.870	**0.847**	7.85	0.888	7.85	0.907	3.45
	ΔFFR	0.473	0.524	**0.547**	13.45	0.549	13.45	0.487	7.94
	ΔINF	0.889	0.779	**0.740**	9.27	0.760	9.27	0.804	5.11
	average	0.752	0.724	**0.711**		0.732		0.733	
$h=12$	ΔEMPL	0.795	0.804	**0.770**	11.99	0.770	11.99	0.772	12.77
	ΔFFR	0.397	0.490	**0.444**	16.43	0.451	16.43	0.448	17.01
	ΔINF	0.692	0.677	**0.613**	12.32	0.651	12.32	0.617	12.94
	average	0.628	0.657	**0.609**		0.624		0.612	

Using the appropriate parameter(s) reported in Table 4.7, the out-of-sample RMSFEs of each regression during our evaluation period are reported in Table 4.8. Here, instead of evaluating the performances for every value of decay rate π, we choose only one value according to the information in Table 4.7. That is, we choose a value of π for each regression μ in each forecast horizon h that makes the regression perform best during the test period. The bold-faced numbers indicate the regression that on average performs best for each forecast horizon h.

According to Table 4.8, every shrinkage estimator outperforms the OLS VAR, which is the benchmark model. Each usually outperforms the application of BIC on the OLS VAR as well. For forecast horizons $h = 3$ to $h = 12$, the LASSO is obviously the best performer. Its performance gain over the application of BIC with OLS VAR is, on average, about 2 to 4 percent, while over the Bayesian VAR is about 1 to 4 percent. For the horizon $h = 1$, it can be said that the performance of LASSO is at the same level as Bayesian VAR and the modified naive elastic net.

The results in Table 4.8 look relatively similar to that of Table 4.2 in our

Simulation 1. The LASSO on average performs best there as well. This implies that the data generating process of the three selected U.S. variables agrees more with that of our Simulation 1. Recall that the true matrix of coefficients in this simulation is a sparse matrix (See again Appendix F). This indicates the redundancy of the VAR with long lags.

Consider also the average number of variables selected (no. var) by the LASSO. We can see that the relatively superior performances of the LASSO can be obtained from only one third of the total 39 independent variables (13 lags of 3 endogenous variables). If one uses the LASSO as the variable selector before applying the Bayesian VAR, performances similar to that of applying the Bayesian VAR alone can be obtained as well. This can be seen from comparing RMSFEs of the procedure LASSO+BVAR to that of the BVAR. This also confirms the redundancy of the VAR with long lags.

At last, Table 4.9 reports the proportion of each variable selected by the LASSO out of the 276 repetitions in our evaluation period. The table reports the cases of (i) the forecast horizon $h = 3$ (One-quarter ahead) and the decay rate $\pi = 1$, and (ii) the forecast horizon $h = 12$ (One-year ahead) and the decay rate $\pi = 0$ only. Results of the other cases are rather similar. The numbers $1 - 10$ in parentheses indicate the ten most frequently chosen variables in each forecasting equation. For example, in predicting $\Delta EMPL$ 12-periods ahead, the second lag of $\Delta EMPL$ itself, which is chosen in all 276 repetitions, is the most frequently chosen variable.

According to Table 4.9, the LASSO does not pick the same lags of every endogenous variable into each forecasting equation. The recent lags of each endogenous variable turn to be among the most frequently chosen ones in predicting itself in the future. Apart from these, the chosen variables seem

Table 4.9: **Proportion of each variables selected by the LASSO**

(i) $\pi = 1$ and $h = 3$

LAGS	$\Delta EMPL$			ΔFFR			ΔINF		
	$\Delta EMPL$	ΔFFR	ΔINF	$\Delta EMPL$	ΔFFR	ΔINF	$\Delta EMPL$	ΔFFR	ΔINF
1	92.4% (3)	25.7%	50.4% (8)	98.6% (2)	100.0% (1)	42.8%	71.7% (6)	89.9% (3)	100.0% (1)
2	100.0% (1)	43.8%	78.6% (4)	44.2%	81.9% (3)	59.4% (6)	86.6% (4)	79.3% (5)	52.5% (7)
3	100.0% (1)	50.7% (7)	18.5%	51.4% (10)	71.7% (5)	27.5%	21.0%	4.0%	31.5%
4	65.6% (5)	18.1%	51.1% (6)	2.5%	73.9% (4)	52.5% (9)	4.7%	12.0%	10.5%
5	19.2%	13.8%	43.1%	40.9%	17.8%	26.8%	5.4%	8.0%	37.7% (10)
6	17.4%	6.9%	43.5%	37.0%	19.9%	12.3%	6.9%	4.3%	1.4%
7	14.1%	6.9%	7.2%	29.0%	55.1% (7)	50.4%	4.3%	0.0%	0.0%
8	26.1%	7.6%	45.7% (10)	12.7%	42.4%	41.3%	40.9% (9)	14.1%	0.0%
9	2.2%	16.7%	0.0%	34.1%	54.7% (8)	25.7%	17.0%	0.0%	47.8% (8)
10	0.0%	0.0%	8.7%	22.8%	16.3%	2.9%	12.0%	2.9%	21.4%
11	9.1%	17.8%	7.2%	5.4%	42.8%	8.3%	0.0%	4.7%	13.8%
12	5.8%	1.4%	0.0%	11.6%	7.6%	2.2%	5.4%	1.1%	100.0% (1)
13	0.7%	48.9% (9)	0.4%	11.2%	22.5%	13.8%	0.0%	2.9%	0.0%

(ii) $\pi = 0$ and $h = 12$

LAGS	$\Delta EMPL$			ΔFFR			ΔINF		
	$\Delta EMPL$	ΔFFR	ΔINF	$\Delta EMPL$	ΔFFR	ΔINF	$\Delta EMPL$	ΔFFR	ΔINF
1	87.0% (3)	1.8%	12.7%	93.8% (2)	100.0% (1)	25.4%	57.6% (7)	47.1%	84.4% (2)
2	100.0% (1)	15.2%	45.3% (8)	44.9%	71.0% (5)	46.7%	83.3% (3)	71.0% (4)	47.5% (9)
3	98.9% (2)	44.6% (9)	15.6%	50.7%	72.5% (3)	21.7%	51.8% (8)	6.2%	26.8%
4	76.1% (5)	16.3%	51.4% (6)	2.9%	68.1% (6)	41.3%	6.5%	13.0%	11.6%
5	40.2%	13.8%	43.1%	40.2%	17.0%	27.9%	14.5%	31.5%	42.4%
6	21.4%	7.2%	43.8% (10)	29.0%	17.4%	19.6%	18.5%	12.0%	3.3%
7	19.2%	7.2%	12.3%	34.8%	64.1% (8)	54.3%	7.6%	11.2%	6.9%
8	27.5%	16.3%	50.4% (7)	28.3%	42.8%	63.4% (9)	62.3% (6)	34.4%	2.2%
9	12.0%	22.8%	4.3%	62.7% (10)	59.1%	28.3%	46.0% (10)	5.4%	68.1% (5)
10	0.0%	0.4%	28.6%	20.3%	22.1%	14.9%	25.4%	18.1%	46.0% (10)
11	23.6%	40.9%	41.7%	15.2%	58.3%	13.8%	17.8%	23.9%	20.7%
12	12.3%	26.1%	1.1%	64.5% (7)	42.8%	9.8%	18.5%	25.7%	100.0% (1)
13	29.7%	79.3% (4)	5.1%	36.2%	72.5% (3)	38.4%	0.0%	26.4%	31.9%

to reflect seasonal and computational effects. For example, the 12-th lag of ΔINF, which is the most frequently chosen variable in predicting itself 3-periods or 12-periods ahead, may reflect the yearly difference transformation that we employ to make the variable stationary. It seems that we can mimic the results of the LASSO by employing a model selection procedure that allows different lags of each independent variable to enter the forecasting model. However, this can be very costly if we want to incorporate many independent variables with various lags in our forecasting practice.

4.6 Conclusion

We demonstrate a close relationship between the Bayesian VAR and ridge regression. Accordingly, the Bayesian VAR can be seen as applying ridge regression on the scaled matrix of independent variables in the VAR relationship. In this way, we see that the LASSO and the elastic net, which are shown in the literature to outperform ridge regression under some circumstances, may outperform the Bayesian VAR as well.

We develop a way to implement different decay rate of effects of lags on the LASSO, as well as on the elastic net proposed by Zou and Hastie (2005). Our simulation study confirms the study in Tibshirani (1996) that if the matrix of true (unknown) coefficients is a sparse matrix, the LASSO will outperform the Bayesian VAR. This happens even when we use VAR as our true model, which makes all regressors endogenous. However, we do not see the difference in performances between the LASSO and the elastic net.

The empirical results on the U.S. data set with 3 endogenous variables and 13 lags in the VAR relationship shows that LASSO can help forecasting in a VAR framework. Applying the LASSO on the VAR usually outperforms the Bayesian VAR. The results also suggest that there is strong evidence of redundancy of the 3-variable VAR with 13 lags.

We see that there is a relationship between our study and Boivin and Ng (2006) as well. Redundancy should become more problematic if we add more irrelevant endogenous variables into the VAR. In this way, bigger models may not always be preferable. The extension of the LASSO or some other automatic variable selectors that can take into account group effects (all lags of one endogenous variable) is an interesting topic to be pursued in the future.

Appendix E

LARS

The method of least angle regression (LARS) is a new model selection algorithm purposed by Efron et al. (2004). It adds a variable into the selected set one-by-one. A variable to be added, however, must, firstly, be considered important in explaining the dependent variable, and, secondly, not be too similar to the variables selected beforehand.

LARS is quick. In each step, it increases the length of the projection just enough to make the inner product between the residual and the newly added variable have the same magnitude as with all the previously selected variables. Since the residual is not orthogonal to the space of previously selected variables, the newly added variable can be highly correlated to the previously chosen ones. We explain the algorithm of LARS below. For more details about LARS, see Efron et al. (2004).

The algorithm of LARS begins by setting the forecast $\hat{Y}_i^{(0)} = 0$. Suppose $\hat{Y}_i^{(q-1)}$ is the current estimate, compute the vector of inner products:

$$\hat{\mathbf{c}} = X'(Y_i - \hat{Y}_i^{(q-1)}),$$

with $\hat{c}_j$ be the j-element of this vector $\hat{\mathbf{c}}$. Let $\mathcal{A} = \{j : |\hat{c}_j| = \hat{C}\}$ be the set of indices of selected variables, where $\hat{C} = \max_j |\hat{c}_j|$. We use $s_j = \text{sign}(\hat{c}_j)$ to represent the sign of the inner product $\hat{c}_j$, and define the matrix of active variables corresponding to $\mathcal{A}$ as:

$$X_{\mathcal{A}} \doteq (s_j X_j)_{j \in \mathcal{A}}.$$

Let $\alpha_{\mathcal{A}} = (1'_{\mathcal{A}}(X'_{\mathcal{A}}X_{\mathcal{A}})^{-1}1_{\mathcal{A}})^{-1/2}$ be a scalar, where $1_{\mathcal{A}}$ is a vector of ones of length equaling the number of elements in $\mathcal{A}$. We construct the unit equiangular vector with columns of the matrix $X_{\mathcal{A}}$ as:

$$u_{\mathcal{A}} = X_{\mathcal{A}} w_{\mathcal{A}},$$

where $w_{\mathcal{A}} = \alpha_{\mathcal{A}}(X'_{\mathcal{A}} X_{\mathcal{A}})^{-1} 1_{\mathcal{A}}$ is a vector of weights. Observe that:

$$X'_{\mathcal{A}} u_{\mathcal{A}} = \alpha_{\mathcal{A}} 1_{\mathcal{A}}, \tag{E.1}$$

which implies that $\alpha_{\mathcal{A}}$ is the equal magnitude of inner product between each variable in the active matrix $X_{\mathcal{A}}$ and the unit vector $u_{\mathcal{A}}$.

Define the vector of inner products between variables in X and $u_{\mathcal{A}}$ as:

$$\mathbf{a} = X' u_{\mathcal{A}},$$

with a_j be the j-th element of the vector $\mathbf{a}$. LARS updates the forecast according to:

$$\hat{Y}_i^{(q)} = \hat{Y}_i^{(q-1)} + \hat{\gamma} u_{\mathcal{A}}, \tag{E.2}$$

with:

$$\hat{\gamma} = \min_{j \in \mathcal{A}^c}{}^{+} \left(\frac{\hat{C} - \hat{c}_j}{\alpha_{\mathcal{A}} - a_j}, \frac{\hat{C} + \hat{c}_j}{\alpha_{\mathcal{A}} + a_j} \right), \tag{E.3}$$

where the minimization considers positive components only, and $\mathcal{A}^c$ is the complement of the set $\mathcal{A}$.

Equation (E.3) implies that the length $\hat{\gamma}$ applied to the unit vector $u_{\mathcal{A}}$ is chosen for a new variable to be added into the set of selected variables $\mathcal{A}$ next step.

To see this, given a scalar $\gamma > 0$, define:

$$\hat{Y}_i(\gamma) = \hat{Y}_i^{(q-1)} + \gamma u_{\mathcal{A}}. \tag{E.4}$$

Then, the next step inner product between a variable X_j and $(Y_i - \hat{Y}_i(\gamma))$ becomes:

$$c_j(\gamma) = X_j'(Y_i - \hat{Y}_i(\gamma)) = \hat{c}_j - \gamma a_j. \tag{E.5}$$

For any previously selected variable $j \in \mathcal{A}$, the definition of the set $\mathcal{A}$ and (E.1) yield:

$$|c_j(\gamma)| = \hat{C} - \gamma\alpha_{\mathcal{A}}. \tag{E.6}$$

For one of the variables not been selected yet $j \in \mathcal{A}^c$, the scalar $\hat{\gamma}$ in (E.3) equates (E.5) and (E.6), which means that this variable will be in the selected set $\mathcal{A}$ next step.

To implement the LASSO, we need a modification on LARS that enforces the condition:

$$\text{sign}(\hat{b}_{ji}^{(lasso)}) = \text{sign}(\hat{c}_j) = s_j, \tag{E.7}$$

where $\hat{b}_{ji}^{(lasso)}$ is the j-th element of the estimate $\hat{B}_i^{(lasso)}$. See the formal argument for the need of this condition in Lemma 8 of Efron et al. (2004). Consider the case that we have just completed the q-th step of LARS, while the estimate from the $(q-1)$-th step still satisfies the condition (E.7). Let $\hat{Y}_i^{(q-1)} = X\hat{B}_i^{(q-1)}$, and $\hat{b}_{ji}^{(q-1)}$ be the j-th element of the vector $\hat{B}_i^{(q-1)}$. According to (E.4), $\hat{b}_{ji}^{(q-1)}$ is equal to $\hat{b}_{ji}^{(q-2)} + \hat{\gamma} s_j w_{\mathcal{A},j}$ for $j \in \mathcal{A}$, where $w_{\mathcal{A},j}$ is the corresponding element in the vector $w_{\mathcal{A}}$, and equal to 0 for $j \in \mathcal{A}^c$.

From (E.4), we can see that:

$$\hat{Y}_i(\gamma) = X\hat{B}_i(\gamma), \qquad \text{where} \qquad \hat{b}_{ji}(\gamma) = \hat{b}_{ji}^{(q-1)} + \gamma s_j w_{\mathcal{A},j}, \text{ for } j \in \mathcal{A}.$$

The coefficient $\hat{b}_{ji}(\gamma)$ will change sign at:

$$\gamma_j = -\hat{b}_{ji}^{(q-1)}/(s_j w_{\mathcal{A},j}),$$

and the first such change occurs at:

$$\tilde{j} = \min_{j \in \mathcal{A}}{}^{+}\{\gamma_j\},$$

where the minimization considers only positive components. Define a scalar $\tilde{\gamma}$ as:

$$\tilde{\gamma} = \begin{cases} \gamma_{\tilde{j}}, & \text{if there exists } \gamma_j > 0, \\ +\infty, & \text{otherwise.} \end{cases}$$

For any $\gamma > \tilde{\gamma}$, the condition (E.7) is violated, because $\hat{b}_{\tilde{j}i}(\gamma)$ has switched signs, while $\hat{c}_{\tilde{j}}$ has not.

The *LARS-LASSO Modification* is: if $\hat{\gamma} > \tilde{\gamma}$, apply $\gamma = \tilde{\gamma}$ in the ongoing LARS step, and remove $\tilde{j}$ from the calculation of the next equiangular direction. That is:

$$\hat{Y}_i^{(q)} = \hat{Y}_i^{(q-1)} + \tilde{\gamma} u_{\mathcal{A}}, \qquad \text{and} \qquad \mathcal{A}^{(q)} = \mathcal{A}^{(q-1)} - \{\tilde{j}\}, \tag{E.8}$$

rather than (E.2). If only a single index j is added or deleted from the set

$\mathcal{A}$ in each step, it can be shown that the LARS-LASSO yields all the LASSO solutions. See Theorem 1 of Efron et al. (2004).

Appendix F

Matrices for Simulations 1 - 3

$$\underset{(39\times3)}{B^{(s1)}} = \begin{bmatrix} 0.197 & 0.541 & 0 \\ 0 & 0.334 & 0.062 \\ 0 & 0 & 0.278 \\ 0.305 & 0.266 & 0.151 \\ 0 & -0.176 & 0.049 \\ 0 & 0 & 0 \\ 0.225 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0.094 & 0 & 0 \\ -0.036 & -0.203 & 0 \\ -0.090 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0.050 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & -0.214 & 0 \\ 0 & 0.215 & 0.090 \\ 0 & 0 & 0 \\ 0 & 0.112 & 0.055 \\ 0 & 0 & 0 \\ 0 & 0.291 & 0 \\ 0 & 0.086 & 0 \\ 0 & 0 & 0.105 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & -0.124 & 0 \\ 0 & 0 & 0.160 \\ 0 & -0.373 & 0 \\ 0 & 0 & 0.046 \\ 0 & 0 & -0.465 \\ 0 & 0 & 0 \\ 0 & 0.160 & 0 \\ 0 & 0 & 0.091 \end{bmatrix} . \qquad \underset{(39\times3)}{B^{(s2)}} = \begin{bmatrix} 0.180 & 0.529 & 0.096 \\ 0.013 & 0.309 & 0.052 \\ 0.053 & 0.080 & 0.247 \\ 0.275 & 0.291 & 0.119 \\ -0.009 & -0.175 & 0.034 \\ -0.043 & -0.022 & 0.043 \\ 0.216 & 0.125 & 0.046 \\ -0.025 & -0.065 & 0.023 \\ -0.010 & 0.083 & -0.025 \\ 0.125 & 0.078 & -0.009 \\ -0.037 & -0.179 & 0.004 \\ -0.066 & -0.069 & 0.021 \\ 0.092 & -0.069 & -0.009 \\ -0.016 & -0.063 & 0.051 \\ -0.022 & 0.066 & 0.069 \\ -0.006 & -0.165 & -0.100 \\ -0.019 & -0.010 & 0.003 \\ -0.015 & 0.063 & 0.006 \\ -0.031 & 0.101 & -0.109 \\ -0.019 & -0.235 & 0.019 \\ 0.033 & 0.165 & 0.089 \\ 0.025 & 0.109 & 0.030 \\ -0.027 & 0.077 & 0.054 \\ 0.044 & 0.102 & -0.008 \\ 0.078 & 0.359 & 0.028 \\ -0.024 & 0.083 & -0.030 \\ -0.025 & 0.088 & 0.082 \\ -0.039 & -0.030 & 0.091 \\ -0.003 & -0.068 & 0.034 \\ -0.013 & 0.036 & 0.074 \\ 0.015 & -0.241 & 0.013 \\ -0.048 & -0.089 & -0.033 \\ 0.030 & 0.021 & 0.129 \\ -0.026 & -0.402 & -0.003 \\ 0.031 & -0.075 & 0.051 \\ -0.024 & 0.007 & -0.469 \\ -0.002 & 0.093 & 0.018 \\ -0.040 & 0.156 & -0.035 \\ 0.049 & 0.104 & 0.073 \end{bmatrix} .$$

$$\Psi^{(s1)} = \begin{bmatrix} 0.039 & 0.011 & -0.001 \\ 0.011 & 0.245 & 0.008 \\ -0.001 & 0.008 & 0.064 \end{bmatrix}. \qquad \Psi^{(s2)} = \begin{bmatrix} 0.037 & 0.011 & -0.001 \\ 0.011 & 0.233 & 0.008 \\ -0.001 & 0.008 & 0.061 \end{bmatrix}.$$

$$B^{(s3)} = \begin{bmatrix} 1.043 & 0.053 & -0.280 \\ 1.297 & 0.111 & -1.162 \\ 0.122 & 0.167 & 0.534 \end{bmatrix}. \qquad \Theta^{(s3)} = \begin{bmatrix} -0.806 & -0.039 & 0.308 \\ -0.788 & 0.271 & 1.234 \\ -0.163 & -0.102 & -0.388 \end{bmatrix}.$$

$$\Psi^{(s3)} = \begin{bmatrix} 0.036 & 0.013 & 0.000 \\ 0.013 & 0.274 & 0.011 \\ 0.000 & 0.011 & 0.081 \end{bmatrix}.$$

Chapter 5

Conclusion

This study considers macroeconomic forecasting with non-structural models. In Chapter 2, we assess the predictive performances of financial variables in comparison with non-financial leading indicators in predicting Australian recessions. In the context of forecasting, it is more appropriate to assess the predictive performances of these predictors using the out-of-sample method. However, there are rare recessions in Australia. Hence, we apply the re-sampling assessment scheme in figuring out robust predictors for Australian recessions. This involves repeatedly re-shuffling the data set and re-applying the out-of-sample assessment to the re-shuffled set. We find that the results from this re-sampling assessment may be very different from that of the usual out-of-sample assessment.

When defining recessions with the turning points from the Melbourne Institute (R1 definition), Australian recessions seem to follow U.S. recessions with some lags. Our out-of-sample assessment reports that the U.S. short-term interest rate becomes the best predictor in forecasting the Australian recessions 4-quarters ahead to 6-quarters ahead. We see that this is the result of the fact that there are only 2 recessionary sessions in our out-of-sample evaluation period, which starts from 1st quarter of 1981, and one of these two sessions is led by the Gulf War. Estrella and Mishkin (1998) show that the U.S. recession associated with the Gulf War cannot be predicted well by the U.S. interest rate spread, which normally performs well in predicting earlier recessions.

In our re-sampling assessment, the U.S. short-term interest rate performs poorly in predicting Australian recessions under this R1 definition. Its 4-quarter ahead to 6-quarter ahead predictive performances are outperformed by many other variables, including the inflation rate and the U.S. interest rate spread. In contrast, the U.S. interest rate spread turns to be the best predictor for the 6-quarter to 7-quarter ahead forecast horizons.

We believe that the results from our re-sampling assessment are more robust than ones from the usual out-of-sample assessment. This means the best predictors from the re-sampling assessment should perform better in predicting Australian recessions in the future. Our re-sampling assessment confirms the usual out-of-sample assessment for the findings that financial variables usually outperform non-financial leading indicators, and the U.S. financial variables have strong performances in predicting Australian recessions.

Chapter 3 re-examines the finding of Bańbura et al. (2008) that the out-of-sample forecasting performances of the Bayesian VAR with Litterman prior can be improved by expanding the number of endogenous variables into the model. In their study, the model, with 131 endogenous variables, obviously outperforms smaller models with 3 and 7 variables. However, the model with 20 endogenous variables yields performances that can match the largest model.

After showing that the forecasting performances of various specifications can be affected by the value of the parameter that governs the overall dispersion of the prior distribution, we argue that the way Bańbura et al. (2008) set the values of this parameter in various specifications is not convincing. The authors set the values to fix the in-sample fit of various specifications to be the same as the unrestricted OLS VAR with 3 endogenous variables. This implies the smallest Bayesian VAR in their study is the unrestricted OLS VAR, and the larger specifications can benefit more from shrinkages.

We see the more pertinent way is to compare the best performances of these specifications. However, we still follow the logic of the out-of-sample assessment method, and attempt to find the suitable parameter values that should lead to the best performances after considering the pre-evaluation period only.

Hence, we set the latter part of the pre-evaluation period to be our test period, and determine the suitable parameter values according to the out-of-sample performances in this test period.

We show that we can determine the optimal parameter values in this test period. We apply each of these parameter values to the associated specification in our out-of-sample assessment. Our results still confirm the finding of Bańbura et al. (2008) that larger specifications perform better than smaller ones. However, the improvement in the performances of larger models is a lot less pronounced than the one shown by Bańbura et al. (2008).

Chapter 4 considers the Bayesian VAR with Litterman prior from a different perspective. We think that the mechanism of the bias-variance tradeoff in statistics can be applied to the VAR as well. We start by showing a close relationship between the Bayesian VAR with Litterman prior and the ridge regression. With this close relationship, we suspect that the impressive performances of the Bayesian VAR with Litterman prior recorded in, for example, Litterman (1986) and Robertson and Tallman (1999) can be explained by the same mechanism that makes the ridge regression outperform the OLS regression.

This also means some alternative shrinkage regressions that have a record to outperform the ridge regression may outperform the Bayesian VAR with Litterman prior as well. We select the LASSO, the elastic net, and a procedure that uses LASSO as the variable selector before applying the Bayesian VAR with Litterman prior as our alternatives. We also modify the LASSO and the elastic net to make them take into account the declining effects of lags of endogenous variables.

Our simulation study confirms our priori expectation that the LASSO and the elastic net applied on a VAR outperform the Bayesian VAR with Litterman prior, when the matrix of true (unknown) coefficients is sparse. The reverse is true when all the elements of the matrix of true (unknown) coefficients are nonzero. However, we cannot observe the difference in performances between the LASSO and the elastic net, when applied on the VAR.

The empirical result that uses the U.S. employment, inflation rate, and the Federal Fund Rate data to predict themselves shows that the LASSO outperforms the Bayesian VAR with Litterman prior for most of the forecast horizons considered. This confirms that there is a redundancy problem in a small VAR with just 3 endogenous variables and 13 lags.

This is also an early caution to the practice of incorporating a large number of endogenous variables into the Bayesian VAR. There would be no harm to incorporate additional variables, which are informative. However, it would be wrong to think that just adding some more variables will benefit the forecasting performances of the model.

References

Alles, L. (1995). The Australian term structure as a predictor of real economic activity. *Australian Economic Review 28*(4), 71–85.

Anderson, H. M. and F. Vahid (2001). Predicting the probability of a recession with nonlinear autoregressive leading-indicator models. *Macroeconomic Dynamics 5*, 482–505.

Ang, A., M. Piazzesi, and M. Wei (2006). What does the yield curve tell us about GDP growth? *Journal of Econometrics 131*, 359–403.

Bai, J. and S. Ng (2002, January). Determining the number of factors in approximate factor models. *Econometrica 70*(1), 191–221.

Bai, J. and S. Ng (2008). Forecasting economic time series using targeted predictors. *Journal of Econometrics 146*, 304–317.

Bańbura, M., D. Giannone, and L. Reichlin (2008, November). Large Bayesian VARs. Working Paper Series 966, European Central Bank.

Bates, J. M. and C. W. Granger (1969). The combination of forecasts. *Operations Research Quarterly 20*, 451–68.

Bernanke, B., J. Boivin, and P. S. Eliasz (2005, January). Measuring the effects of monetary policy: A factor-augmented vector autoregressive (FAVAR) approach. *The Quarterly Journal of Economics 120*(1), 387–422.

Bernanke, B. S. and J. Boivin (2003, April). Monetary policy in a data-rich environment. *Journal of Monetary Economics 50*(3), 525–546.

Bernard, H. and S. Gerlach (1998, July). Does the term structure predict recessions? The international evidence. *International Journal of Finance & Economics 3*(3), 195–215.

Birchenhall, C. R., H. Jessen, D. R. Osborn, and P. Simpson (1999, July). Predicting U.S. business-cycle regimes. *Journal of Business & Economic Statistics 17*(3), 313–23.

Birchenhall, C. R., D. R. Osborn, and M. Sensier (2001, May). Predicting UK business cycle regimes. *Scottish Journal of Political Economy 48*(2), 179–95.

Boivin, J. and S. Ng (2006). Are more data always better for factor analysis? *Journal of Econometrics 132*, 169–194.

Box, G. E. and G. M. Jenkins (1970). *Time series analysis: Forecasting and control.* San Francisco, Holden-Day.

Burns, A. F. and W. C. Mitchell (1946). *Measuring Business Cycles.* New York: National Bureau of Economic Research.

Burr, T. L. and H. A. Fry (2005, August). Biased regression: The case for cautious application. *Technometrics 47*(3), 284–296.

Chauvet, M. and S. Potter (2002). Predicting a recession: Evidence from the yield curve in the presence of structural breaks. *Economics Letters 77*, 245–253.

Chauvet, M. and S. Potter (2005). Forecasting recessions using the yield curve. *Journal of Forecasting 24*, 77–103.

Christiano, L. J., M. Eichenbaum, and C. L. Evans (1999). Monetary policy shocks: What have we learned and to what end? In J. B. Taylor and M. Woodford (Eds.), *Handbook of Macroeconomics*, Volume 1, Chapter 2, pp. 65–148. Elsevier.

Ciccarelli, M. and A. Rebucci (2003). Bayesian VARs: A survey of the recent literature with an application to the European monetary system. IMF Working Paper WP/03/102, International Monetary Fund.

Cooper, R. L. (1972, Winter). The predictive performance of quarterly econometric models of the United States. In *Econometric Models of Cyclical Behavior, Vols. 1 and 2*, NBER Chapters, pp. 813–948. National Bureau of Economic Research, Inc.

D'Agostino, A. and D. Giannone (2007, November). Comparing alternative predictors based on large-panel factor models. CEPR Discussion Papers 6564, C.E.P.R. Discussion Papers.

De Mol, C., D. Giannone, and L. Reichlin (2008). Forecasting using a large number of predictors: Is Bayesian shrinkage a valid alternative to principal components? *Journal of Econometrics 146*, 318–328.

Diebold, F. X. (1998, Spring). The past, present, and future of macroeconomic forecasting. *Journal of Economic Perspective 12*(2), 175–192.

Diebold, F. X. and G. D. Rudebusch (1989, July). Scoring the leading indicators. *Journal of Business 62*(3), 369–91.

Doan, T., R. Litterman, and C. Sims (1986). Forecasting and conditional projection using realistic prior distributions. Research Department Staff Report 93, Federal Reserve Bank of Minneapolis.

Dotsey, M. (1998). The predictive content of the interest rate term spread for future economic growth. *Economic Quarterly 84*(3), 31–51. Federal Reserve Bank of Richmond.

Dueker, M. (1997). Strengthening the case for the yield curve as a predictor of U.S. recessions. *Federal Reserve Bank of St. Louise Review* (Mar), 41–51.

Dueker, M. (2005, January). Dynamic forecasts of qualitative variables: A Qual VAR model of U.S. recessions. *Journal of Business & Economic Statistics 23*, 96–104.

Dueker, M. J. and K. Wesche (2005). Forecasting macro variables with a Qual VAR business cycle turning point index. Technical report.

Efron, B., T. Hastie, I. Johnstone, and R. Tibshirani (2004, April). Least angle regression. *Annals of Statistics 32*(2), 407–451.

Elliott, G. and A. Timmermann (2004, September). Optimal forecast combinations under general loss functions and forecast error distributions. *Journal of Econometrics 122*(1), 47–79.

Engle, R. F. and C. W. J. Granger (1987, March). Co-integration and error correction: Representation, estimation, and testing. *Econometrica 55*(2), 251–76.

Estrella, A. (1998). A new measure of fit for equations with dichotomous dependent variables. *Journal of Business and Economic Statistics 16*(2), 198–205.

Estrella, A. and G. A. Hardouvelis (1991, June). The term structure as a predictor of real economic activity. *Journal of Finance 46*(2), 555–76.

Estrella, A. and F. Mishkin (1998). Predicting U.S. recessions: Financial variables as leading indicators. *The Review of Economics and Statistics 80*(1), 45–61.

Estrella, A. and F. S. Mishkin (1997, July). The predictive power of the term structure of interest rates in Europe and the United States: Implications for the European Central Bank. *European Economic Review 41*(7), 1375–1401.

Estrella, A. and A. P. Rodrigues (1998). Consistent covariance matrix estimation in probit models with autocorrelated errors. Working Paper, Federal Reserve Bank of New York.

Estrella, A., A. P. Rodrigues, and S. Schich (2003). How stable is the predictive power of the yield curve? Evidence from Germany and the United States. *The Review of Economics and Statistics 85*(3), 629–644.

Fisher, C. and B. Felmingham (1998). The Australian yield curve as a leading indicator for consumption growth. *Applied Financial Economics 8*, 627–635.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2000, November). The generalized dynamic-factor model: Identification and estimation. *The Review of Economics and Statistics 82*(4), 540–554.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2003, September). Do financial variables help forecasting inflation and real activity in the Euro area? *Journal of Monetary Economics 50*(6), 1243–1255.

Frank, I. E. and J. H. Friedman (1993, May). A statistical view of some chemometrics regression tools. *Technometrics 35*(2), 109–135.

Granger, C. W. J. (1969, July). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica 37*(3), 424–38.

Groen, J. J. J. and G. Kapetanios (2008). Revisiting useful approaches to data-rich macroeconomic forecasting. Staff Reports 327, Federal Reserve Bank of New York.

Hamilton, J. D. (1989, March). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica 57*(2), 357–84.

Hamilton, J. D. (1994). *Time Series Analysis.* Princeton: Princeton University Press.

Hamilton, J. D. and G. Perez-Quiros (1996, January). What do the leading indicators lead? *Journal of Business 69*(1), 27–49.

Harding, D. and A. Pagan (1999). Knowing the cycle. Melbourne Institute Working Paper No. 12/99, University of Melbourne.

Harding, D. and A. Pagan (2002). Dissecting the cycle: A methodological investigation. *Journal of Monetary Economics 49*, 365–381.

Harding, D. and A. Pagan (2006, May). Synchronization of cycles. *Journal of Econometrics 132*(1), 59–79.

Harding, D. and A. Pagan (2009, January). An econometric analysis of some models for constructed binary time series. NCER Working Paper Series 39, National Centre for Econometric Research.

Hastie, T., R. Tibshirani, and J. Friedman (2001). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* Springer-Verlag.

Hoerl, A. E. and R. W. Kennard (1970, February). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics 12*(1), 55–67.

Hsu, N.-J., H.-L. Hung, and Y.-M. Chang (2008). Subset selection for vector autoregressive processes using Lasso. *Computational Statistics and Data Analysis 52*, 3645–3657.

Issler, J. V. and F. Vahid (2006). The missing link: Using the NBER recession indicator to construct coincident and leading indices of economic activity. *Journal of Econometrics 132*, 281–303.

Kadiyala, K. R. and S. Karlsson (1997). Numberical methods for estimation and inference in Bayesian VAR-models. *Journal of Applied Econometrics 12*(2), 99–132.

Karunaratne, N. D. (2002). Predicting Australian growth and recession via the yield curve. *Economic Analysis & Policy 32*(2), 233–250.

Kauppi, H. and P. Saikkonen (2008, 07). Predicting U.S. recessions with dynamic binary response models. *The Review of Economics and Statistics 90*(4), 777–91.

Litterman, R. B. (1986). Forecasting with Bayesian vector autoregressions - five years of experience. *Journal of Business and Economic Statistics 4*(1), 25–38.

Lowe, P. (1992). The term structure of interest rates, real activity and inflation. Research Discussion Paper 9204, Reserve Bank of Australia.

Magnus, J. R. and H. Neudecker (1999). *Matrix Differential Calculus with Applications in Statistics and Econometrics* (2nd ed.). John Wiley & Sons.

Marcellino, M. (2006). Leading indicators. In G. Elliott, C. Granger, and A. Timmermann (Eds.), *Handbook of Economic Forecasting*, Volume 1, Chapter 16, pp. 879–960. Elsevier.

Moore, G. H. (1990). *Leading Indicators for the 1990's.* New York: Dow Jones-Erwin.

Naylor, T. H., T. G. Seaks, and D. W. Wichern (1972, August). Box-Jenkins methods: An alternative to econometric models. *International Statistical Review / Revue Internationale de Statistique 40*(2), 123–37.

Nelson, C. R. (1972, December). The prediction performance of the FRB-MIT-PENN model of the U.S. economy. *American Economic Review 62*(5), 902–17.

Osborne, M. R., B. Presnell, and B. A. Turlach (2000a). A new approach to variable selection in least squares problems. *IMA Journal of Numerical Analysis 20*, 389–404.

Osborne, M. R., B. Presnell, and B. A. Turlach (2000b). On the LASSO and its dual. *Journal of Computational and Graphical Statistics 9*(2), 319–337.

Poke, J. and G. Wells (2007, December). The term spread and GNP growth in Australia. CAMA Working Paper No. 27/2007, Center for Applied Macroeconomic Analysis, Australian National University.

Potter, S. M. (1995). A nonlinear approach to U.S. GNP. *Journal of Applied Econometrics 10*(2), 109–25.

Robertson, J. C. and E. W. Tallman (1999, Q1). Vector autoregressions: Forecasting and reality. *Federal Reserve Bank of Atlanta Economic Review 84*(1), 4–18.

Sargent, T. J. and C. A. Sims (1977). Business cycle modeling without pretending to have too much a priori economic theory. Technical report.

Schmidt, M. (2005, December). Least squares optimization with l1-norm regularization. Cs542b project report, Department of Computer Sciences, University of British Columbia.

Sims, C. A. (1980, January). Macroeconomics and reality. *Econometrica 48*(1), 1–48.

Sims, C. A. and T. Zha (1998). Bayesian methods for dynamic multivariate models. *International Economic Review 39*(4), 949–968.

Smith, P. (2005). What economic indicators have led Australian GDP over the past decade? *Quarterly Bulletin of Economic Trends 3*, 19–22.

Stock, J. H. and M. W. Watson (1989). New indexes of coincident and leading economic indicators. In *NBER Macroeconomics Annual 1989, Volume 4*, NBER Chapters, pp. 351–409. National Bureau of Economic Research.

Stock, J. H. and M. W. Watson (1992). A procedure for predicting recessions with leading indicators: Econometric issues and recent experience. NBER Working Papers 4014, National Bureau of Economic Research.

Stock, J. H. and M. W. Watson (2002a, December). Forecasting using principal components from a large number of predictors. *Journal of American Statistical Association 97*(460), 1167–1179.

Stock, J. H. and M. W. Watson (2002b, April). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics 20*(2), 147–62.

Stock, J. H. and M. W. Watson (2003). Forecasting output and inflation: The role of asset prices. *Journal of Economic Literature 41*(3), 788–829.

Stock, J. H. and M. W. Watson (2005). Implications of dynamic factor models for VAR analysis. NBER Working Papers 11467, National Bureau of Economic Research.

Teräsvirta, T. (1994). Specification, estimation and evaluation of smooth transition autoregressive models. *Journal of the American Statistical Association 89*, 208–18.

Teräsvirta, T. (2005, May). Forecasting economic variables with nonlinear models. Working Paper Series in Economics and Finance 598, Stockholm School of Economics.

Tibshirani, R. (1996). Regression shrinkage and selection via the LASSO. *Journal of the Royal Statistical Society: Series B 58*(1), 267–288.

Verbeek, M. (2004). *A Guide to Modern Econometrics* (2 ed.). West Sussex: John Wiley & Sons.

West, K. D. (2006). Forecast evaluation. In G. Elliott, C. Granger, and A. Timmermann (Eds.), *Handbook of Economic Forecasting*, Volume 1, Chapter 3, pp. 99–134. Elsevier.

Zellner, A. and C. Hong (1989). Forecasting international growth rates using Bayesian shrinkage and other procedures. *Journal of Econometrics 40*, 183–202.

Zellner, A. and C. Min (1999). Forecasting turning points in countries' output growth rates. *Journal of Econometrics 88*, 203–303.

Zha, T. (1998, Q1). A dynamic multivariate model for use in formulating policy. *Economic Review 83*(1), 16–29. Federal Reserve Bank of Atlanta.

Zou, H. and T. Hastie (2005). Regularization and variable selection via the Elastic Net. *Journal of the Royal Statistical Society: Series B 67*(2), 301–320.