A Systematic Analysis of Potential Leading Indicators in the United States through Vector Autoregression

Nicholas Holschuh

February 13, 2010
Comprehensive Exercise in Economics
Advisor: Pavel Kapinos

Abstract

The business cycle has been a subject of great economic interest over the past century. Decision making in both the public and private sector is influenced by the phase of the business cycle, and as a result, our ability to understand and model real economic activity is incredibly important. This study presents a group of linear models that attempt to explain the evolution of real economic activity, in an effort to determine how the inclusion of leading indicators affects out-of-sample predictive power. I focus on 10 potential leading indicators: interest rate spread, producer price index, hours worked, corporate profits, M1, M2, the Federal Funds Rate, the S&P 500, and the Dow Jones industrial average. Using a rolling vector autoregressive structure and two different forecasting methods, all possible combinations of these leading indicators were analyzed. I found that including any of the viable leading indicator candidates in the model improves performance, however interest rate spread, the producer price index, and M1 yield the best results. With every additional variable beyond two included in the regression, the loss in degrees of freedom results in worse forecasts despite better in-sample fit.

1.   Introduction

Understanding the business cycle has been a subject of great interest to both the public and private sector over the last 100 years. Much of the debate during the 20th century focused on merely defining the business cycle and determining how best to detrend it. Before the quarterly gross domestic product series was established, there was no obvious measure of real economic activity in the United States. As a result, economists began developing composite indices: collections of multiple economic variables that attempt to characterize the state of the economy as a whole. Currently, the National Bureau of Economic Research (NBER) uses information from observable economic variables to classify the United States business cycle, determining each month if the economy is in a contractionary period or expansionary period. The use of composite indices is still common in the field, although real gross domestic product is now the benchmark series for economic activity.

When modeling business cycles, there are two general ways of determining how a model performs. The first is done by comparing the ex post predictions of the model to business cycle observations. This consists of relating characteristics of the fitted series (e.g. cycle period, cycle curvature, amplitude differences between cycles, and turning point locations) to those of the observed reference series. This type of analysis is more common in the literature, but does not necessarily provide any information on a model's ability to provide accurate future predictions. In order to determine the ex ante predictive power of a model, it requires omitting some data when estimating the regression. By estimating a model using a restricted data set, it is possible to compare model generated "future" forecasts to the observed values that were omitted during the model estimation process. This type of ex ante study has been relatively infrequent, although it is becoming more common in the wake of the recent financial crisis.

Recent research has focused primarily on the co-cyclical nature of individual economic time series, as well as the change in economic behavior between expansions and contractions. In attempting to model business cycles, these two relationships become very important. Researchers must use the correct mathematical model to reflect the complex dynamics that drive the business cycles, while also including variables that contain enough information for accurate ex ante prediction. The use of leading indicators (variables whose cyclic behavior precedes that of the business cycle) is therefore an important aspect of model specification. As macroeconomists continue to analyze new parametric models, non-parametric models, and leading variables, they can more accurately predict business cycle turning points in real time.

Using a similar methodology to that used by Diebold and Rudebusch (1989), I test a wide variety of leading indicator combinations in an effort to forecast business cycle dynamics. While most of the literature focuses on individual leading indicators or mathematical specifications, very few take a systematic look at multiple leading indicators within a common mathematical framework. There is no literature that analyzes the relative performance of different combinations of known leading indicators in predicting business cycles. The focus of this paper is to determine, in a linear framework, which variable combinations yield the most accurate ex ante forecasts.

The following five sections of this paper examine more carefully the problems of developing reliable forecasts for future movements in the business cycle. Section 2 provides a comprehensive look at the existing literature as it relates to business cycles and leading indicator models, while Sections 3 through 5 focus on the nature of this specific study. Section 3 details the data collection, preliminary data manipulation, as well as the empirical model used for

forecasting. Section 4 enumerates the results of the regression analysis, Section 5 indicates possible avenues for further research, and Section 6 concludes.

2. *Review of Existing Literature*

The literature about business cycle modeling centers largely on questions of model specification. There are two major issues being addressed: which mathematical framework most accurately reflects business cycle properties, and what variables are most useful for future prediction. The objective of this paper is to perform a comprehensive analysis of variable selection. That being said, I largely ignore the specifics of model selection in this literature review. A more comprehensive overview of the regime switching, Markov chain, and transition models prominent in the study of business cycle turning points can be found in Rudebusch (1996).

The business cycle is a difficult thing to quantify, and defining a reference series for real economic activity can be highly subjective. Previous to the development of the quarterly gross domestic product series, economists analyzing business cycle characteristics were forced to construct their own series to describe real activity. As a result, there has been debate amongst academics about what data best describe true business cycle fluctuations. Layton (1997) believe that a composite coincident index is best for analyzing Australian business cycles, while Hamilton (1989) uses gross national product for the United States. Yamada et al. (2010) and Forni et al. (2001) believe that using band pass filtering techniques on composite indices provide the best measures of business cycle activity. Harding and Pagan (2002), however, show that growth rate in output provides a much better tool for analysis than trend deviations like those created using data filters. This study uses growth in real gross domestic product as the reference series for business cycle fluctuations.

Many macroeconomic time series exhibit cyclical behavior. When these cycles are out of phase, past turning points in one series may be able to predict future turning points in others. Those data whose turning points occur in advance of the business cycle are called leading indicators. To macroeconomists, these cycles are of great interest, given that they have the potential to provide warning for future macroeconomic fluctuations.

The act of analyzing leading indicator performance has been done a number of ways in the literature. Both parametric and non-parametric studies have been conducted, and leading indicators have been scored using event based prediction (turning points) and specific value prediction. This study takes a regression based approach to predicting business cycle turning points; much like the work done by Wecker (1979) and Kling (1987) in that I translate forecasted values into turning point predictions. This is different than much of the parametric analysis that has been done in the field, starting with the works by Auerbach (1982) and Neftiçi (1982), which look at value prediction. I chose to analyze model performance based on predicted turning points instead of values because the regressions estimated in this study are linear[1]. Linear models have the potential to define accurate turning points despite failing to capture intra-cycle dynamics that non-linear models are designed to reflect.

Diebold and Rudebusch (1989) performed a study evaluating individual and composite leading indicators using a non-parametric Bayesian sequential probability recursion. Their methods, however, are not entirely ex ante in predicting business cycle turning points. Probability densities used in the turning point estimation were calculated over the entire period, meaning they used information a forecaster would not have in real time to make their predictions. Despite this flaw in their prediction methods, the evaluation method for model performance in this study was logically sound. Using a quadratic probability score, they were

---

[1] The reason for excluding non-linear models is discussed in detail in section 5.

able to quantify leading indicator model performance using the observed turning points during the period of interest. I adapted their probability score method to analyze rolling regression forecasts for use in this study.

Variable selection for this paper was motivated primarily by previous business cycle forecasting studies and theory. The remainder of this section of the paper briefly examines the leading indicator selections of other economists, as well as their findings. Boehm and Summers (1999) included a wide variety of potential leading indicators in their analysis, making their paper a good baseline for variable selection. While their study focuses on in-sample behavior, they also perform a qualitative analysis of forecasting using a composite leading index. In the construction of their leading index, they chose to include data on hours worked, changes in producer prices, stock prices, changes in profitability, and price/cost ratios, believing these all to be useful leading indicators for business cycle dynamics. They found that using this composite leading index provided accurate results when forecasting future movements in real activity, prompting me to include many of the same series they analyzed.

Ghent and Owyang (2010) looked at regional and national housing cycles to determine whether or not they behaved as leading indicators to the business cycle. They found no statistical relationship between city house prices and local employment levels. Ghent and Owyang actually found that national housing permits exhibit stronger leading indicator qualities for city level employment than a city's own permits, indicating that any added forecasting strength from housing permit data is merely due to collinearity between housing permits and other economic variables. As a result, I chose not to include housing data in my analysis.

One of the most widely studied leading indicators is the yield curve. The interest rate term spread, defined by the slope of the yield curve, contains information on both the real interest rate

and expected future inflation, two variables central to predicting movements in the macroeconomy. Anderson, Athanasopoulos, and Vahid (2007) found that interest rate spread helps significantly when trying to replicate business cycle length, amplitude, and curvature for the G-7 countries. Fritsche and Kuzin (2005) found that term spread, along with the real effective exchange rate, can provide valuable insight into future movements of the German business cycle both in and out-of-sample.

De Bondt and Hahn (2010) set out to define a new composite leading indicator for the Euro Area. Using a set of three statistical criteria, they settled on nine different time series for the index: 10 year nominal bond yield, nominal stock prices, US unemployment, M1, German IFO (business expectations), building permits, economic sentiment indicator, consumer confidence indicator, and the manufacturing new orders-stocks ratio. They found that using these leading indicators they were able to produce reliable predictions, with the best forecasts being 4 to 8 months in advance. Given the optimal prediction period found in their study, I structured my analysis to focus on results during that same range of forecasts.

The financial crisis of 2008 has prompted quite a bit of research on business cycle forecasting, forcing many macroeconomists to ask whether or not leading indicators could have provided any advanced warning of the recently experienced economic volatility. Schrimpf and Wang (2010) re-examined the predictive power of the yield curve, and found that in recent years yield curve based forecast accuracy has been falling. Frankel and Saravelos (2010) found that the real exchange rate and central bank reserves proved the most useful in predicting the current crisis. Bunda and Zorzi (2010) found that price competitiveness and the public debt-to-GDP ratio both provided valuable information in predicting the tensions found in today's financial markets.

Drechsel and Scheufele (2010) performed a comprehensive analysis of leading indicators for the German business cycle using individual time series, pooled time series, and composite series. They found that most models including only a single leading indicator performed poorly both in and out of the current recession. They do find, however, that including financial indicators like interest rate term spread does result in better forecasts.

In their work, Fichtner, Ruffer, and Schnatz (2009) performed a temporal analysis of leading indicator models. They believed that, as a result of an increased level of globalization, the power of country specific leading indicators would fall over time. Their results show that leading indicator models were in-fact more effective at predicting changes in output in the past. By including international data in predictive models they were able to significantly improve current forecast accuracy.

While the literature has taken a wide variety of approaches to leading indicator analysis, there are a few areas where current research is clearly lacking. Systematic comparisons of leading indicators have been performed frequently in a non-parametric setting, but rarely in a regression framework. Many studies have compared individual time series and composite leading indices, but rarely have studies considered using many individual time series in the forecasting process. By taking a parametric approach to leading indicator analysis, this study manages to fill in some of the gaps currently present in the literature. Not only do I study the relative performance of single leading indicator regressions and multivariate ones, the structure of this study allows me to determine if the different combinations of leading indicators can predict better than the sum of their parts.

3. Data and Methodology[2]

In this paper, I study ten different economic variables in an effort to quantify their ability to predict changes in real GDP. The data is at a quarterly frequency, with a time period spanning from 1964 (Q2) to 2010 (Q2). This period was chosen based on data availability for the leading indicator candidates, which range from micro-level data to US monetary aggregates. For stock market activity, I look at the monthly closing price for the Dow Jones industrial average (*DOW*), and the S&P 500 (*SP500*). For commercial activity, I selected an index for the aggregate number of hours worked per month (*HOURS*), the average monthly volume of commercial and industrial loans at all commercial banks (*LOANS*), and corporate profits (*PROFITS*). Finally, for macro-level and financial data I use the producer price index (*PPI*), the federal funds rate (*FEDFUNDS*), interest rate spread between 10 year and 1 year treasury notes (*SPREAD*), M1 (*M1*), and M2 (*M2*). The stock market data were collected from Lexis Nexus, while all other micro and macroeconomic data were collected from the Federal Reserve of St. Louis.

3.1 Preliminary Analysis and Transformation

In order to use these data for regression estimation, it is important to first determine if they satisfy the conditions required for time series analysis. These ten series were tested for non-stationarity using an Augmented Dickey Fuller (ADF) test. This is done because the presence of a unit root has significant implications for time series regression analysis. For example, when a series is non-stationary, the variance of the series will increase indefinitely as it evolves over time. Also, parameter estimates will be biased for regressions which contain integrated time series.

---

[2] Variable names will be italicized throughout the rest of the paper. Levels of those variables are capitalized, log levels are lowercase, and the first difference is denoted with a Δ. A more thorough discussion of data and data collection can be found in Appendix A.

The ADF test is conducted by first estimating the regression

$$\Delta y_t = \alpha + \beta t + \delta y_{t-1} + \sum_{=1}^{l} \gamma_i * \Delta y_{t-i} + \epsilon_t \qquad (1)$$

Using the regression results, it is possible to determine whether or not a particular time series (y) evolves according to a stochastic process with a drift (α) and a trend (β). Lag length was selected for each test using the Hannon-Quinn Information Criterion. After estimating the regression defined by equation (1), I test the null hypothesis that $\delta=0$ with an alternative hypothesis of $\delta<0$. Rejection of the null hypothesis indicates that no unit root is present, and the series is stationary. The intuition behind the structure of the ADF test lies in the definition of stationarity; high levels in the previous period (represented by a positive value for the lagged term) should be, on average, correlated with a negative change into the next period. By finding a value for δ significantly less than zero, the series is shown to be stationary.

Because of the potential for error introduced by the presence of a unit-root, performing the Augmented Dickey Fuller test is a necessary step before doing any further regression analysis. For each variable that fails to reject the null hypothesis, I take the difference of that series and then re-estimate the ADF regression. This process can be repeated until each of the time series are stationary, however for these data I did not have to go beyond the first difference to make all variables covariance stationary.

Table 1 displays the results of the ADF test on the variables in levels. Note that *WORK* and *SPREAD* were the only two series that were able to reject the null hypothesis of a unit root at the 5% level. To eliminate the unit root of the other time series, either the log first difference (in the case of exponentially growing series) or first difference (for linearly growing series) were taken and the ADF test was performed again. I decided to take the first difference instead of using levels of the *WORK* series based on a visual inspection of the data. The results of the ADF

test for the differenced series are presented in Table 2. All series now reject the null of a unit root

at the 5% level, and can be used for further regression analysis.

A preliminary search for leading indicator behavior is carried out using a Granger

causality test. The use of "causality" in the name of this test is somewhat of a misnomer; what

the Granger causality test is good for is indicating temporal relationships between two time

series. By definition, the Granger causality test determines whether or not the lags of one

variable improve prediction of the current values of a second variable. While the test seems well

suited for the type of forecasting analysis central to this study, there is no indication that

evidence of strong Granger causality necessarily leads to better turning point prediction. In

addition to studying leading indicator combinations, I also hope to determine whether or not it is

accurate to assume Granger causality test results are correlated with forecast reliability.

For each combination of series y ($\Delta rgdp$) and z (one of the leading indicator candidates),

the Granger causality test is performed by estimating the following two regressions:

$$y_t = a_0 + \sum_{i=1}^{lmax} a_i y_{t-i} + \sum_{j=1}^{lmax} b_j z_{t-j} + e_{yt} \qquad (2)$$

and

$$z_t = c_0 + \sum_{i=1}^{lmax} c_i y_{t-i} + \sum_{j=1}^{lmax} d_j z_{t-j} + e_{zt} \qquad (3)$$

I defined the lag order (lmax) using the Hannon-Quinn criterion. The purpose of estimating these

regressions is to test the null hypotheses $b_1 = b_2 = \cdots = 0$ and $c_1 = c_2 = \cdots = 0$. If the b

coefficients are found to be significant, it indicates that z partially explains the evolution of y, a

quality referred to as "Granger causing" y. Significance in the c coefficients indicates that y

Granger causes z. If a variable is shown to Granger cause $\Delta rgdp$, this likely means it behaves as

a leading indicator to the business cycle. If both the potential leading indicator and $\Delta rgdp$ appear

to Granger cause one another, they are likely coincident time series. Only variables that exhibit one of these two qualities will be included in the final forecasting model; those series which do not Granger cause *Δrgdp* will be excluded from further analysis.

The results of the Granger causality test over the entire study period are presented in Table 3. There are several aspects of the results to take note of. The first is that Δ*profits* are shown to neither Granger cause Δ*rgdp* nor are they Granger caused by Δ*rgdp*. This indicates that there is neither a leading nor lagged relationship between them, and Δ*profits* will not increase predictive power in the final regressions. I also found that Δ*WORK* does not Granger cause Δ*rgdp,* so it will also be excluded from further regression analysis. *Δdow*, *Δsp500*, *Δm1*, *Δm2*, and *Δppi* are all shown to Granger cause Δ*rgdp*, while none of them are Granger caused by Δ*rgdp*. This is the ideal result, indicating that these five series should be strong leading indicators for the business cycle. *SPREAD* is shown to Granger cause the reference series, but is also Granger caused by it at the 5% confidence level. Although this could mean that *SPREAD* is more coincident than leading, it will still be included in the final regressions, in order to determine whether or not Granger causality test results reliably predict forecast accuracy for leading indicators.

Following the procedure used by Wells (1999), I also performed a series of rolling Granger causality tests. The goal of these tests is to see whether or not the potential leading indicators exhibit different periods of strength in forecasting, or if the temporal relationship between Δ*rgdp* and the leading indicators are constant over time. By looking at the eventual forecast results in conjunction with these rolling Granger causality tests, it might be possible to determine why certain variables perform better when used together in forecasting future business cycle turning points. I expect that using variables with different periods of relative strength will result in improved forecast performance.

The rolling Granger causality tests were performed two ways; once with a fixed start date, and once with a rolling start date. Both begin with the same sample period (Q2 1964 – Q3 1970), and continue to a final ending period of Q2 2010. The results are plotted in Figure 1. The first important trend shown in these plots is indicated by the fixed start date rolling Granger causality test. For most series, the leading indicators do not appear to Granger cause $\Delta rgdp$ until the mid 1970's. I believe this is due to the small data set used for those Granger causality tests, however as a result I would expect that forecasts before 1980 will tend to be less accurate.

Note how the periods of strongest Granger causality as shown by the rolling window differ for each series. The *FEDFUNDS* series shows strongest causality during the 40 quarters ending between 1980 and 1990, and then again in the mid 1990s, while indicators like *Δppi* and *SPREAD* show very weak causality in the ten years following 2000. Despite showing a Granger causality test p-value under 5% when including the entire data set, many of these series do not seem to exhibit leading indicator qualities over the majority of the study period.

3.2 The Empirical Model and Estimation

Time-series analysis requires the use of autoregressive (AR) models to fully capture the evolution of the data. The basic AR model takes the form

$$y_t = \alpha + \sum_{l=1}^{lmax} \beta_i y_{t-l} + u_t \qquad (4)$$

This equation describes a linear relationship between y and its own lags, where *lmax* is the number of lags selected. To extend the above autoregressive model into a multivariate context, I used vector autoregressive (VAR) leading indicator models. These take the form

$$x_t^i = \alpha + \sum_{l=1}^{lmax} \gamma_l X_{t-l} + u_t \qquad (5)$$

where i is the number of endogenous variables (x) analyzed in the regression, $\mathbf{X}_{t-l}$ is a vector containing lags of those variables, and $\boldsymbol{\gamma}_{t-l}$ represents a vector of parameter estimates associated with each of those lagged variables. By including more lags, the in-sample fit of these models improves, however the standard error associated with the parameter estimates increases. This is something that must be kept in mind when attempting to forecast future data, because in-sample fit and out-of-sample predictive power are not necessarily correlated. The high standard errors associated with estimates involving many regressors will result in very little consistency for predicted future values.

In order to determine which combination of leading indicators can best predict future turning points in RGDP, I must first define what qualifies as a turning point. Using one of the criteria laid out by Bry and Boschan (1971) for the NBER, we define RGDP peaks as

$$\Delta y_t > 0 \, , \Delta y_{t+1} < 0 \, , \Delta y_{t+2} < 0$$

and troughs as

$$\Delta y_t < 0 \, , \Delta y_{t+1} > 0 \, , \Delta y_{t+2} > 0.$$

Having set this definition for business cycle turning points, the next step is to establish a method for turning point prediction.

Because I am concerned with predictive power for these models, I only evaluate the out-of-sample forecasting results for each regression. To do this, I developed a rolling regression structure that allows me to simulate real time forecasts for all periods between 1970 (Q4) and 2010 (Q2). The process starts by estimating a VAR using a restricted dataset, containing only data from the first 25 quarters of available data (1964 Q2 – 1970 Q3). Although this starting period is short, it was chosen to ensure that all documented turning points between 1964 and 2010 are included in the forecasted section of the total study period.

The values for $\Delta rgdp$ and the leading indicator variables were forecasted one period ahead by two methods. The first method of forecasting started by estimating equation (5). From that result I collected a vector of the parameter estimates, as well as the standard error for the regression residuals ($\hat{\sigma}$). Using those values, I reinserted the leading indicator values, observed $\Delta rgdp,$ and the parameter estimates back into the equation

$$y_t = \alpha^e + \sum_{l=1}^{lmax} \gamma^e{}_l \mathbf{X}_{t-l} + u_t \tag{6}$$

where t is the forecast period (the first one being 26), and $\mathbf{X}_{t-l}$ is the vector of observed values leading up to the end of the restricted regression period. After taking the sum of the product of parameters and variables, I added to that a stochastic error term defined by $u_t \rightarrow N(0, \hat{\sigma}^2)$. This forecasting technique will subsequently be referred to as Method 1.

The second method used a different process to incorporate a probabilistic component into the forecasting. After estimating the regression for the restricted data set, I stored both the parameter estimates and the standard errors associated with each parameter estimate ($\hat{\sigma}_\iota$). Then, I forecasted time period t using the equation

$$y_t = \alpha + \sum_{l=1}^{lmax} \gamma_l \mathbf{X}_{t-l} \tag{7}$$

Instead of using the defined parameter estimates for α and $\boldsymbol{\gamma_l}$, those values are drawn from the distributions $\alpha \rightarrow N(\alpha^e, \hat{\sigma}_\alpha{}^2)$ and $\boldsymbol{\gamma_l} \rightarrow N(\boldsymbol{\gamma_l}^e, \hat{\sigma}_l{}^2)$ where the $\hat{\sigma}$ values are the standard errors associated with each estimated parameter. A fundamental assumption of this study is that the regression residuals are normally distributed, so drawing the parameter estimates from a normal distribution as opposed to a uniform or f-distribution is done to maintain consistency. This forecasting technique will be referred to later in the paper as Method 2.

Using the forecast results for $\Delta rgdp$ and the leading indicators at time period 26, the next time period (27) was then forecasted using equations (6) and (7). This process was repeated until 5 periods had been forecasted and any observed turning points in the $\Delta rgdp$ forecast were recorded. For this restricted, 25 period data set, the 5 period projections are performed 1000 times to get a distribution of forecast values[3]. Figure 2 provides a graphical representation of this process. If t is the last period included in the regression, turning points can only be found for the 1000 forecasts of (t+1), (t+2), and (t+3), because by definition there must be data for two periods following to define a turning point.

When 1000 5-period forecasts have been made for the 25 period VAR, the restricted data set is expanded by one to include the 26th time period, and the process starts over. The regressions described in equations (6) and (7) are estimated using 26 data points, and periods 27-31 are forecasted. The ending period for the restricted data set continues to roll forward until all available data is included in the initial regression.

This means that for every combination of possible leading indicators and $\Delta rgdp$, 151 regressions are estimated. From each regression, it is possible to predict as many as 1000 turning points for each of the three time periods following the restricted dataset. Because turning points for each time period starting with the 28th are forecasted in three separate steps of the rolling regression (once as t+1, t+2, and t+3), as many as 3000 turning points can be predicted for each quarter. This analysis is performed for all 256 combinations of leading indicators, resulting in a just under two billion forecasted values in total[4].

---

[3] Accuracy of forecast results would improve with a higher number of replications, however using this method, the forecasting process already took upwards of 30 days computing time.

[4] The rolling regressions and associated forecasts were computed using the statistical programming language R. The code used for analysis can be found in Appendix E.

Previous studies using a rolling regression framework have taken different approaches to lag selection for each individual stage of the VAR. For consistency across forecasts, I chose to use a constant lag structure for all steps of the rolling regression and all variable combinations. Plotted in Figure 3 is a histogram of selected lag orders for a random draw of the regressions in the forecasting process, determined using the Hannon-Quinn criterion and normalized by the total number of regressions analyzed. This plot indicates that, in general, the models perform best when few lags are included. I chose to use one lag in each regression based on this random sampling.

In order to compare the predictive power of each of the leading indicator combinations, I need to be able to determine the accuracy of model forecasts. To do this, I calculate quadratic probability scores (QPS) for each model specification. This score compares the probability of predicting a turning point at each time period (defined by the number of turning points forecasted divided by 3000) to the probability of a turning point actually occurring. The probability scores can be directly compared between models to determine which variable combinations provide the most accurate forecasts. The equation that defines the QPS is

$$QPS = \frac{2}{T}\sum_{t=1}^{T}(P_t - D_t)^2 \qquad\qquad (0 < QPS < 2) \qquad (8)$$

where P_t is the probability at every point in time of predicting a business cycle turning point, and D_t is a vector containing index values for whether or not a turning point should have been predicted (1=yes, 0=no). Accurate predictions result in lower probability scores.

4. *Results and Discussion*

This study is multifaceted in nature, so it is important to keep in mind the study objectives when discussing the results in detail. The primary goal of this paper is to determine which

individual leading indicators, and which combinations of those leading indicators, yield the best forecasts for future economic turning points. There are other conclusions, however, that can be drawn from this analysis. This paper has the potential to provide insight into the value of Granger causality tests in model selection, as well as methods for economic forecasting in general.

The results from the two forecasting methods yielded very similar results. The QPS for the bivariate models are shown in Table 4[5]. *SPREAD* performs the best of all the variables when using only one leading indicator. *Δppi* and *Δm1* also performed well across both forecasting methods. Curiously, *Δm2* and *Δloans* performed quite well using Method 2, but under Method 1 they were two of the worst predictors. *FEDFUNDS* and *Δdow* performed the worst of all leading indicator candidates.

Given the theoretical understanding for why the yield curve would act as a leading indicator, it is no surprise that it performed the best of all candidates. The fact that both money supply variables performed well while the federal funds rate did not seems strange, given the causal link between movements in M1 and the interest rate. It is also surprising that there is a difference in performance between the S&P 500 and the Dow Jones industrial average, given that these markets are expected to move roughly together.

By definition, the Granger causality test determines whether or not a particular time series is useful when trying to predict future values of a second time series. Based on the p-values from the Granger causality tests performed, I would have expected *Δdow, Δsp500, Δppi, Δm1,* and *Δm2* to provide the best results. This was not the case; the best forecasts were generated using a variable that did not even show Granger causality at the 1% level. This indicates that,

---

[5] The full set of QPS for all tested models can be found in Appendix D.

while the Granger causality test may provide a good preliminary test for forecasting power, stronger Granger causality does not necessarily correlate with greater accuracy in prediction.

On average, Method 2 provides significantly more accurate forecasts than Method 1. This was true across all quantities and combinations of variables. While 42 variable combinations performed worse than the naïve model when forecasting using Method 1 (16.4% of all variable combinations), only one performed worse when using Method 2. This model included only *Δdow* and *Δsp500* as leading indicators.

There was only one multivariate model across both forecasting methods which provided better predictions than the bivariate VARs. This model included *SPREAD, Δdow,* and *Δm2* as leading indicators. Looking at the rolling Granger causality results plotted in Figure 1, there appear to be some relationships that might explain the increase in predictive power when they are all included in a single regression. *Δm2* shows strong Granger causality up until around 1990, offsetting the spikes of no Granger causality found in the other two series. *Δm2* shows high p-values after the year 2000. This period of no Granger causality by *Δm2* is compensated for by the high degree of Granger causality found from 2001-2006 in the *Δdow* series. While this analysis is qualitative and anecdotal, it may give credence to the idea that combining series with different periods of relative strength as defined by a rolling Granger causality test might improve model performance. This is something that should be pursued in more depth during future study.

Table 5 provides a breakdown of model performance by number of included variables across both forecasting methods. Based on the information in this table, it is clear that the inclusion of variables beyond one leading indicator only serves to reduce overall model performance. I believe that this is a result of the reduction in degrees of freedom for the regression as the number of parameter estimates increases. The errors associated with each

estimator go up, resulting in a higher variance in the predictions, which negatively offset any improvement in results which might come from any additional information those leading indicators contain.

There are a number of caveats that should be taken into consideration when looking at the results of this study. There are only 13 turning points experienced during the study period, which will make it difficult to discern the relative performances of the different models. Even a model that fails to predict any turning points will produce a QPS of 0.1625. By expanding the time frame to include a greater number of data points, this analysis would more clearly indicate differences in model performance. This is difficult to do, however, because it raises the question of what reference series should be used prior to the construction of the quarterly RGDP series.

A second thing to keep in mind is that this study only performed 1000 replications of each forecasting process. I completed the rolling regression again for several variable combinations in order to get a sense for the possible variability in results. After recalculating the QPS for these models, I found that the score changed for some models by as much as .01. While these changes are not dramatic, re-forecasting for all variable combinations may result in slightly different results. That being said, I believe the trends found in this study are robust: interest rate spread outperforms the other leading indicators, the more variables included in the regression the higher the QPS, and the relationship between Granger causality and forecast performance may not be as strong as theory would indicate.

5. *Directions for Future Research*

The obvious way to extend this research is to expand it into a non-linear framework. On the outset of this study I had hoped to include both linear and non-linear models, however I encountered a number of logical barriers to when attempting to do it. Throughout the rolling

regression, I found that not all periods tested for it exhibited non-linear behavior. Without being able to justify the inclusion of a nonlinearity in all steps of the rolling regression, I chose to only perform linear analysis. By selecting a different sample period, or including more time periods in the initial restricted data set of the rolling regression, this problem may be avoided.

It would also be valuable to further constrain the quadratic probability scores for each possible model. This could be done one of two ways: by increasing the number of repetitions in the forecasting process, or by repeating the entire forecasting process multiple times to get a sense of the distribution of QPS for each model. This will result in a better comparison between model specifications, as well as help to determine with a higher level of certainty which forecasting method provides the better results.

Continued analysis of Granger causality test results would help substantially in the variable selection process. If it can be found that using rolling or full period Granger causality tests can indicate forecast accuracy for leading indicators, it will reduce the need for systematic leading indicator studies like the one presented in this paper. While I provide some preliminary insight into the meaning of Granger causality test results, there is still quite a bit that is unknown.

One other thing this study did not address is whether or not a combined index of these variables might provide better predictions than using the individual time series in the regression. A composite index would solve any problems associated with a reduction in degrees of freedom caused by including more variables, while still providing any information that might be contained in the individual time series. It would be interesting to produce different composite indices using these 8 leading indicators and determine whether or not they perform better than the individual time series.

6. Concluding Remarks

The focus of this study was to determine which combinations of variables provide the most information about future turning points in real economic activity. I found that when interest rate spread, producer price index, or money supply is included in a regression, RGDP forecasts are significantly better than those produced using the naïve framework. I also found that for each additional variable included in the model beyond one, prediction performance declines.

Not only did this study address the issue of potential leading indicators, it also provides interesting insights into Granger causality and forecasting methods. I found that the temporal relationships tested using a Granger causality framework do not necessarily indicate anything about the accuracy of out-of-sample forecasts those variables might generate. While there might be potential in using rolling Granger causality tests to determine time periods of weak forecasts, more work must be done to accurately determine exactly what information Granger causality test results provide. I also found that of the two forecasting methods used in this study, Method 2 produces better forecast results.

With the advancement of computing technology and econometric techniques, there is a growing burden on macroeconomists to provide information to businesses and policy makers about expected future movements in the real activity. Given the recent volatility in the US economy, producing reliable forecasts becomes an increasingly difficult task. I believe that this study, along with the existing literature described earlier in the paper, have contributed a great deal to our understanding and help to develop methods for more reliable business cycle forecasting.

7. *References Cited*

Anderson, H. M., G. Athanasopoulos, and F. Vahid. 2007. "Nonlinear autoregressive leading indicator models of output in G-7 countries." *J Appl Econom*, 22:1, pp. 63-87.

Auerbach, Alan J. 1982. "The Index of Leading Indicators: "Measurement without Theory," Thirty-Five Years Later." *The Review of Economics and Statistics*, 64:4, pp. 589-95.

Boehm, E. A. and P. M. Summers. 1999. "Analysing and Forecasting Business Cycles with the Aid of Economic Indicators." *International Journal of Management Reviews*, 1:3, pp. 245-77.

Bry, Gerhard and Charlotte Boschan. 1971. *Cyclical Analysis of Time Series: Selected Procedures and Computer Programs*: National Bureau of Economic Research, Inc.

Bunda, Irina and Michele Ca' Zorzi. 2010. "Signals from Housing and Lending Booms." *Emerging markets review*, 11:1, pp. 1-20.

de Bondt, Gabe and C. K. Hahn. 2010. "Predicting recessions and recoveries in real time: The euro area-wide leading indicator (ALI)." *European Central Bank Working Paper Series*, No. 1246, pp. 50.

Diebold, F. X. and G. D. Rudebusch. 1989. "Scoring the Leading Indicators." *J Bus*, 62:3, pp. 369-91.

Diebold, F. X. and G. D. Rudebusch. 1996. "Measuring business cycles: A modern perspective." *Rev Econ Stat*, 78:1, pp. 67-77.

Drechsel, Katja and Rolf Scheufele. 2010. "Should We Trust in Leading Indicators? Evidence from the Recent Recession." Halle Institute for Economic Research.

Fichtner, Ferdinand, Rasmus Ruffer, and Bernd Schnatz. 2009. "Leading indicators in a globalised world." *European Central Bank Working Paper Series*, No. 1125, pp. 26.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin. 2001. "Coincident and leading indicators for the EURO area." *Econ J*, 111:471, pp. C62-C85.

Frankel, Jeffrey A. and George Saravelos. 2010. "Are Leading Indicators of Financial Crises Useful for Assessing Country Vulnerability? Evidence from the 2008-09 Global Crisis." *National Bureau of Economic Research Working Paper Series*, No. 16047.

Fritsche, U. and V. Kuzin. 2005. "Prediction of business cycle turning points in Germany." *Jahrb Natl Stat*, 225:1, pp. 22-43.

Ghent, Andra and Michael Owyang. 2010. "Is Housing the Business Cycle? Evidence from US Cities." *Journal of urban economics*, 67:3, pp. 336-51.

Hamilton, James D. 1989. "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle." *Econometrica*, 57:2, pp. 357-84.

Harding, D. and A. Pagan. 2002. "Dissecting the cycle: a methodological investigation." *J Monetary Econ*, 49:2, pp. 365-81.

Kling, John L. 1987. "Predicting the Turning Points of Business and Economic Time Series." *The Journal of Business*, 60:2, pp. 201-38.

Layton, A. P. 1997. "Do leading indicators really predict Australian business cycle turning points?" *Econ Rec*, 73:222, pp. 258-69.

Neftiçi, Salih N. 1982. "Optimal prediction of cyclical downturns." *Journal of Economic Dynamics and Control*, 4, pp. 225-41.

Schrimpf, Andreas and Qingwei Wang. 2010. "A Reappraisal of the Leading Indicator Properties of the Yield Curve under Structural Instability." *Int J Forecasting*, 26:4, pp. 836-57.

Wecker, William E. 1979. "Predicting the Turning Points of a Time Series." *The Journal of Business*, 52:1, pp. 35-50.

Wells, J. M. 1999. "Seasonality, leading indicators, and alternative business cycle theories." *Appl Econ*, 31:5, pp. 531-38.

Yamada, H., S. Nagata, and Y. Honda. 2010. "A comparison of two alternative composite leading indicators for detecting Japanese business cycle turning points." *Appl Econ Lett*, 17:9, pp. 875-79.

Appendix A: Data Description [1]

RGDP – Seasonally adjusted real gross domestic product, in billions of chained 2005 dollars. Can be found at *[http://research.stlouisfed.org/fred2/series/GDPC96]*.

DOW –The Dow Jones Industrial Average, an index of the stock prices for 30 large US public companies. Converted from daily to quarterly using the final closing price of each quarter. Data can be found at *[http://www.lnstatistical.com/Main.jsp;jsessionid=449E9AE6E6637B4D5E 1680C23106AEF8#datasets3&]*.

SP500 – The S&P 500, an index of the prices of 500 large-cap common stocks actively traded in the United States. Converted from daily to quarterly using the final closing price of each quarter. The data can be found at *[http://www.lnstatistical.com/Main.jsp;jsessionid=449E9AE6E6637B 4D5E1680C23106AEF8#datasets3&]*.

PPI – The producer price index for finished goods. Seasonally adjusted, and converted from monthly to quarterly frequency. The series can be found at *[http://research.stlouisfed.org/fred 2/series/PPIFGS]*.

WORK – An index of the aggregate weekly hours worked in private industry. Values are seasonally adjusted, and the data has been converted from monthly to. The index is based on a survey conducted by the Bureau of Labor Statistics, and can be found at *[http://research. stlouisfed.org/fred2/series/AWHI]*.

SPREAD – The difference between the yield on 1-year constant maturity treasury bills (found at *[http://research.stlouisfed.org/fred2/series/WGS1YR]*) and 10-year constant maturity rate bonds

---

[1] Descriptions collected from Lexis Nexus and the Federal Reserve of Bank of St. Louis Websites.

(found at *[http://research.stlouisfed.org/fred2/series/WGS10YR]*). Both were weekly series, converted to quarterly using the final rate of each quarter.

PROFITS – Corporate profits after tax, in billions of dollars. Collected by the Bureau of Economic Analysis. Can be found at [http://research.stlouisfed.org/fred2/series/CP.

FEDFUNDS – Effective federal funds rate, averaged to form the quarterly series. Can be found at *[http://research.stlouisfed.org/fred2/series/FEDFUNDS]*.

M1 – M1 Money Stock in billions of dollars. Converted from monthly to quarterly. M1 consists of: (1) currency outside the U.S. Treasury, Federal Reserve Banks, and the vaults of depository institutions; (2) traveler's checks of nonbank issuers; (3) demand deposits; and (4) other checkable deposits. The series can be found at *[http://research.stlouisfed.org/fred2/series/M1]*.

M2 – M2 money stock in billions of dollars. M2 includes a broader set of financial assets held principally by households. M2 consists of M1 plus: (1) savings deposits (which include money market deposit accounts, or MMDAs); (2) small-denomination time deposits (time deposits in amounts of less than $100,000); and (3) balances in retail money market mutual funds (MMMFs). Can be found at *[http://research.stlouisfed.org/fred2/series/M2]*.

LOANS – Commercial and industrial loans at all commercial banks, converted from monthly to quarterly. Can be found at *[http://research.stlouisfed.org/fred2/series/BUSLOANS]*.

Appendix B: Tables

Table 1. – Augmented Dickey Fuller test results for levels of all leading indicators

	RGDP	DOW	SP500	PPI	SPREAD	PROFITS	WORK	FEDFUNDS	M1	M2	LOANS
P-Value	0.63	0.53	0.21	0.56	0.00***	0.68	0.27	.02**	0.38	0.38	0.12

(Asterisks represent rejection of the null hypothesis [unit root present] at the 1 [***], 5 [**], and 10 [*] percent levels)

Table 2. – Augmented Dickey Fuller test results for first differences and log first differences (when appropriate).

	Δrgdp	Δdow	Δsp500	Δppi	Δprofits	ΔWORK	Δm1	Δm2	Δloans
P-Value	.00***	.00***	.00***	.00***	.00***	.00***	0.02**	0.04**	.00***

(Asterisks represent rejection of the null hypothesis [unit root present] at the 1 [***], 5 [**], and 10 [*] percent levels)

Table 3. – Granger causality results over the period 1964 (Q2) to 2010 (Q2).

	Δdow	Δsp500	Δppi	SPREAD	Δprofits	ΔWORK	FEDFUNDS	Δm1	Δm2	Δloans
Granger Cause Δrgdp	0.00***	0.00***	0.02**	0.01**	0.16	0.26	0.00***	0.03**	0.02**	0.09*
Are Granger Caused by Δrgdp	0.26	0.61	0.14	0.00***	0.44	0.00***	0.00***	0.13	0.24	0.00***

(Asterisks represent rejection of the null hypothesis [unit root present] at the 1 [***], 5 [**], and 10 [*] percent levels)

Table 4. – Quadratic probability scores for the bivariate leading indicator models across both forecasting methods.

	Δdow	*Δsp500*	*Δppi*	*SPREAD*	*FEDFUNDS*	*Δm1*	*Δm2*	*Δloans*
Method 1	0.144	0.138	0.135	0.131	0.151	0.137	0.145	0.140
Method 2	0.153	0.139	0.137	0.127	0.145	0.138	0.129	0.122

Table 5. – Breakdown of model performance by number of included leading indicators.

Number of Leading Indicators								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
# of models which perform worse than the naïve model	0	1	4	12	14	7	5	0
(% of total in each category)	*(0.00)*	*(1.79)*	*(3.57)*	*(8.57)*	*(12.50)*	*(12.50)*	*(31.25)*	*(0.00)*
Average QPS	0.1382	0.1438	0.1516	0.1567	0.1598	0.1615	0.1713	0.1622

Appendix C: Graphs
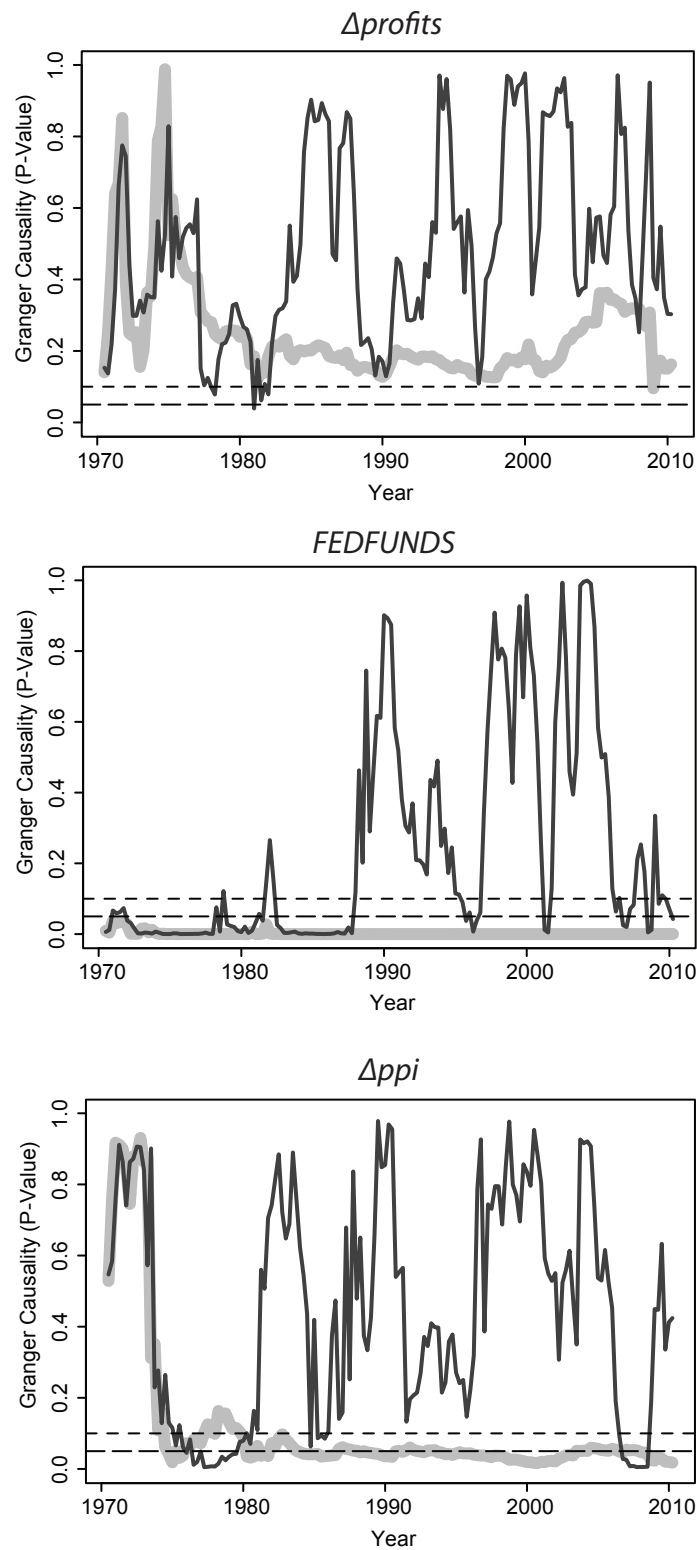
Figure 1. - (continues on to next page)

Δm1

Δm2

SPREAD

Figure 1. - (continues on to next page)

Δdow

Δsp500

ΔWORK

Figure 1. - (continues on to next page)

Δprofits
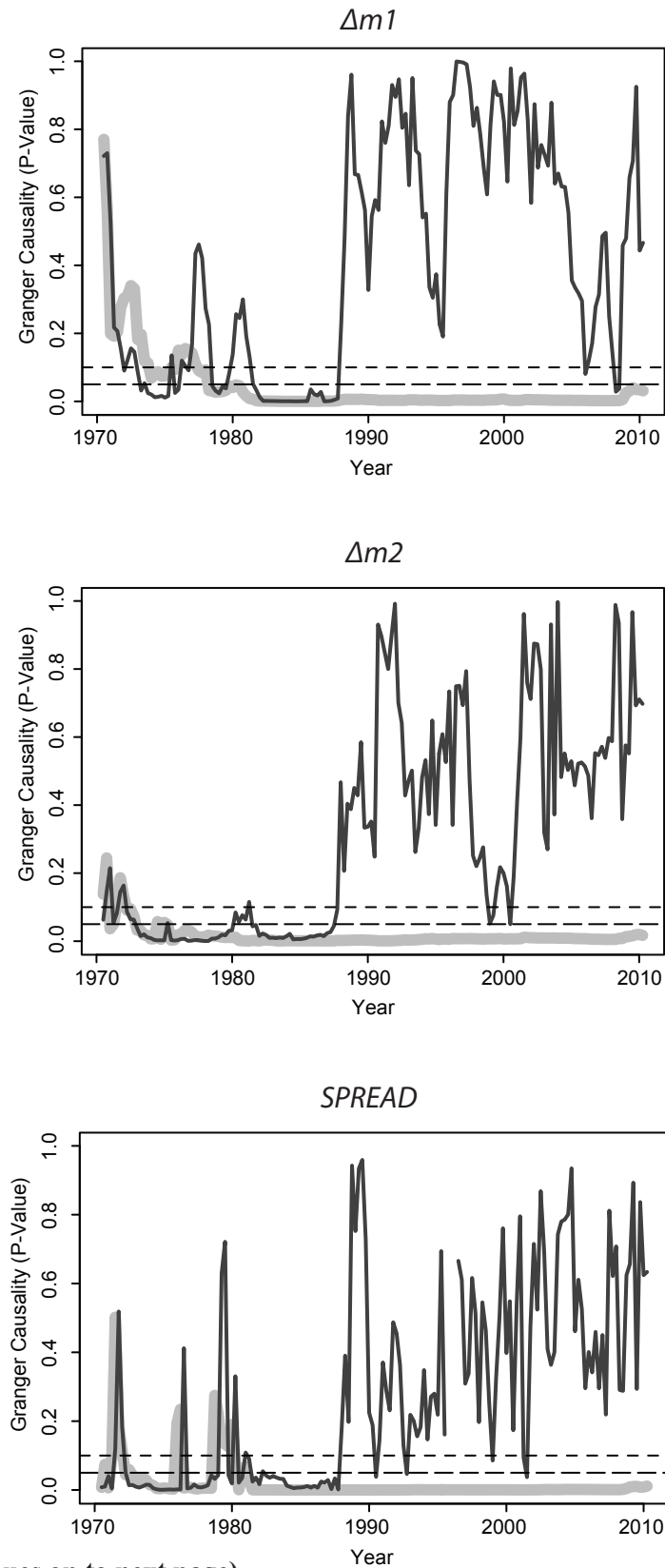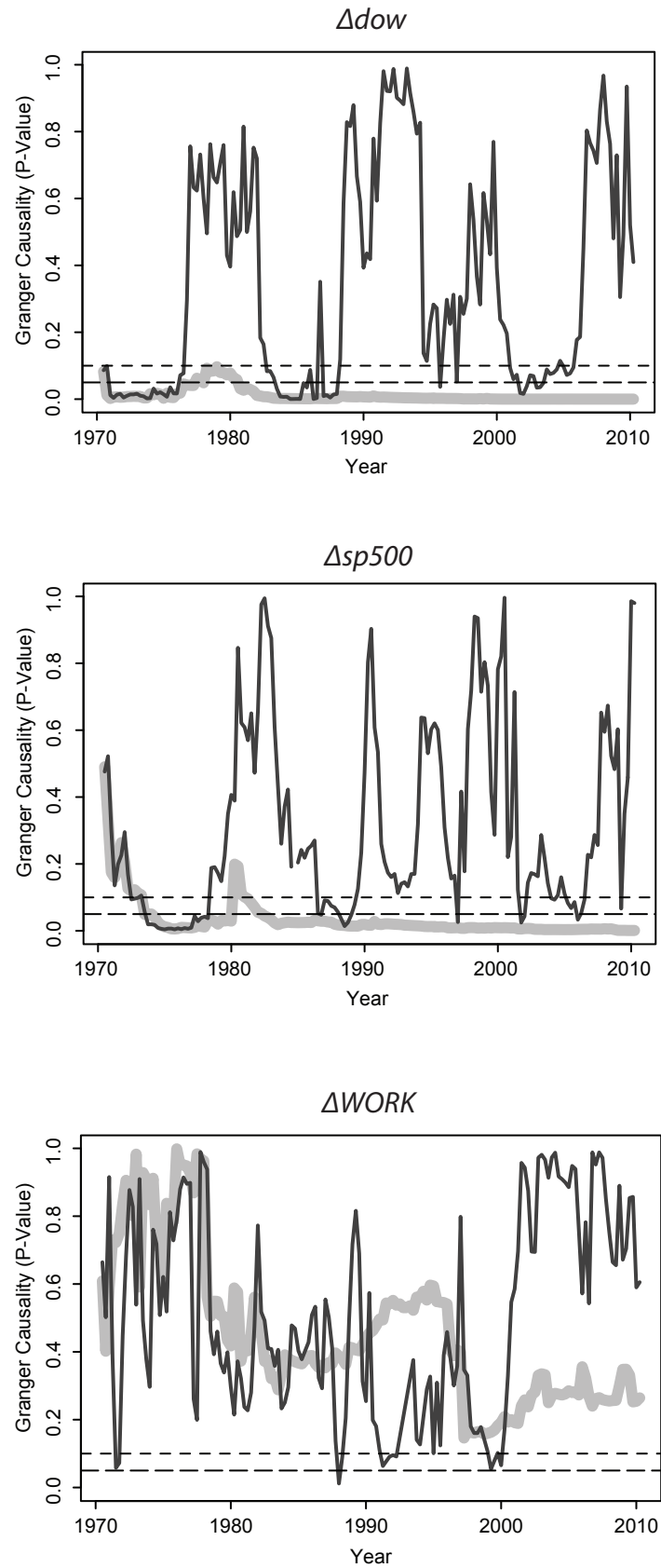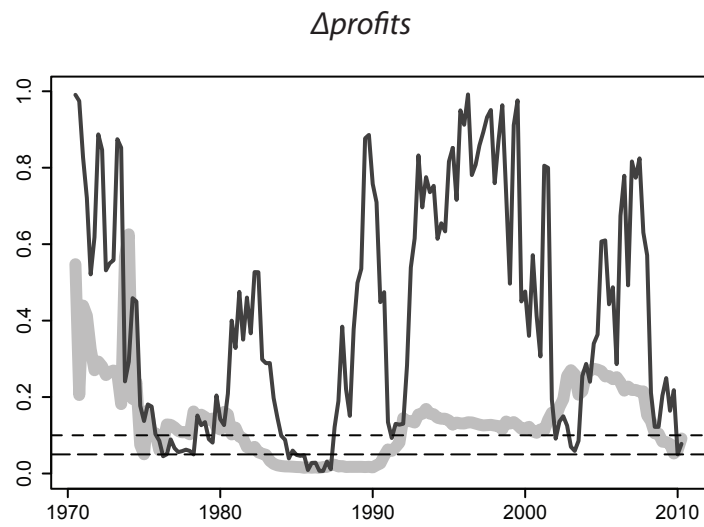
Figure 1. - Results of the rolling Granger causality tests for all potential leading indicators. Black line shows the moving window results, while the bold grey line shows the fixed start date results.
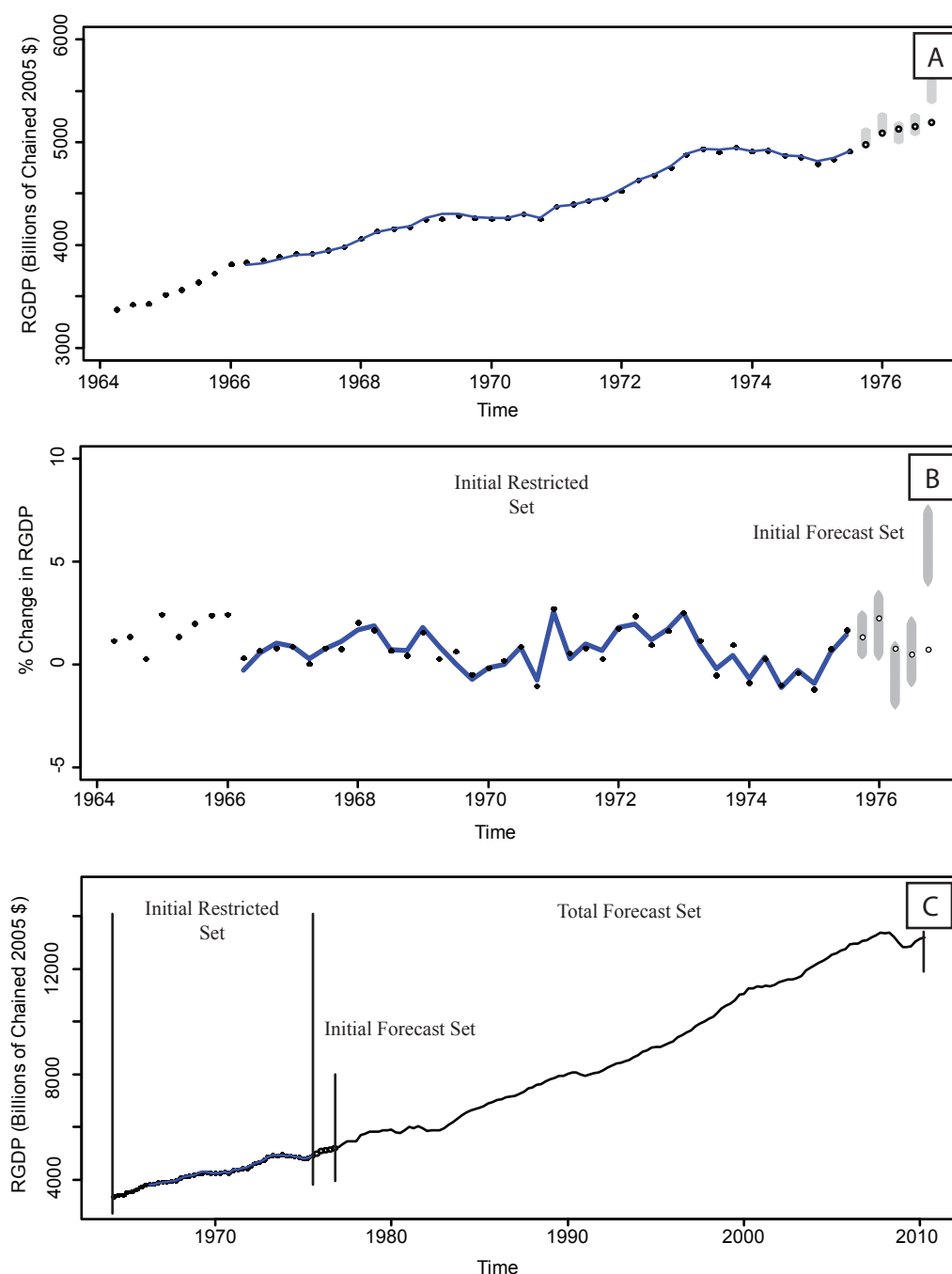
Figure 2. -

These plots help illustrate the forecasting process. The black dots represent observed data points for RGDP (Graph A) and the % change in RGDP (Graph B). Using the first 25 data points, a regression is estimated. The restricted set used for this one stage of the regression is shown on Graph C. The in-sample fit (or ex post prediction) is plotted in blue. Using the regression paramater estimates, data is forecasted five periods into the future (ex ante prediction). The distribution of the 1000 forecasts are plotted in grey. These distributions are compared to the observed data for those time periods, which are plotted in white. Once that series of forecasts is completed, the subsample can be expanded by one. This process repeats until the entire data set plotted in Graph C is included in the model.
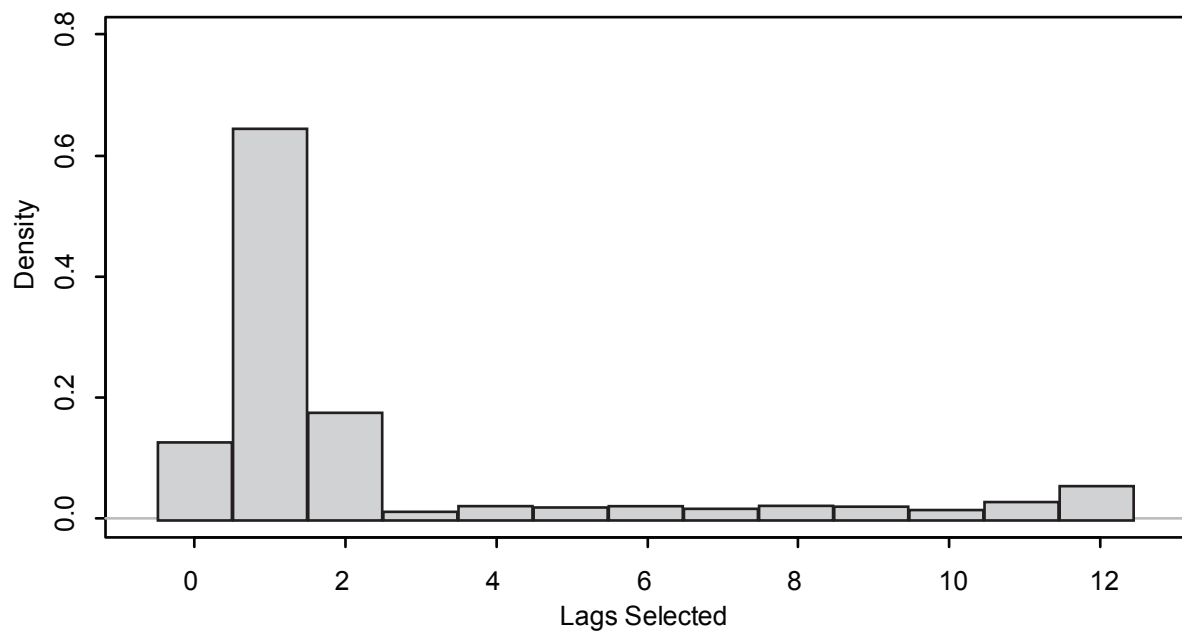
Figure 3. - Plot indicating the density of lags selected for a sample regression. Based on this plot, I decided to uniformly use 1 lag for all regressions

Appendix D: Quadratic Probability Scores

Leading Indicator Shorthand:
A - sp500
B - PPI
C - Interest Rate Spread
D - Dow Jones
E - Federal Funds Rate
F - M1
G - M2
H - Business Loans

2 Variable VARs - Included Leading Indicators:

	A	B	C	D	E	F	G	H
Method 1	0.1383	0.1348	0.1313	0.1438	0.1513	0.1373	0.1447	0.1399
Method 2	0.1387	0.1366	0.1268	0.1534	0.1454	0.1383	0.1288	0.1221

3 Variable VARs - Included Leading Indicators:

	A,B	A,C	A,D	A,E	A,F	A,G	A,H	B,C
Method 1	0.1405	0.1505	0.1364	0.1487	0.1546	0.1532	0.1490	0.1387
Method 2	0.1356	0.1256	0.1596	0.1439	0.1393	0.1306	0.1302	0.1359

	B,D	B,E	B,F	B,G	B,H	C,D	C,E	C,F
Method 1	0.1390	0.1543	0.1432	0.1455	0.1622	0.1381	0.1482	0.1321
Method 2	0.1430	0.1442	0.1404	0.1356	0.1308	0.1446	0.1388	0.1325

	C,G	C,H	D,E	D,F	D,G	D,H	E,F	E,G
Method 1	0.1412	0.1462	0.1658	0.1419	0.1554	0.1480	0.1654	0.1603
Method 2	0.1261	0.1248	0.1548	0.1506	0.1430	0.1435	0.1474	0.1427

	E,H	F,G	F,H	G,H
Method 1	0.1664	0.1623	0.1425	0.1546
Method 2	0.1428	0.1286	0.1305	0.1237

4 Variable VARs - Included Leading Indicators:

	A,B,C	A,B,D	A,B,E	A,B,F	A,B,G	A,B,H	A,C,D	A,C,E
Method 1	0.1596	0.1655	0.1703	0.1697	0.1656	0.1666	0.1904	0.1715
Method 2	0.1329	0.1432	0.1424	0.1385	0.1330	0.1332	0.1400	0.1386

	A,C,F	A,C,G	A,C,H	A,D,E	A,D,F	A,D,G	A,D,H	A,E,F
Method 1	0.1727	0.1780	0.1557	0.1556	0.1403	0.1739	0.1535	0.1987
Method 2	0.1305	0.1243	0.1277	0.1523	0.1505	0.1432	0.1514	0.1464

	A,E,G	A,E,H	A,F,G	A,F,H	A,G,H	B,C,D	B,C,E	B,C,F
Method 1	0.1954	0.1733	0.1613	0.1661	0.1765	0.1635	0.1681	0.1517
Method 2	0.1416	0.1431	0.1296	0.1313	0.1265	0.1421	0.1396	0.1382

	B,C,G	B,C,H	B,D,E	B,D,F	B,D,G	B,D,H	B,E,F	B,E,G
Method 1	0.1591	0.1780	0.1626	0.1427	0.1546	0.1666	0.1651	0.1735
Method 2	0.1354	0.1346	0.1512	0.1445	0.1386	0.1375	0.1453	0.1434

	B,E,H	B,F,G	B,F,H	B,G,H	C,D,E	C,D,F	C,D,G	C,D,H
Method 1	0.1438	0.1574	0.1835	0.1546	0.1429	0.1689	0.1569	0.1525
Method 2	0.1429	0.1371	0.1375	0.1338	0.1464	0.1439	0.1345	0.1410

	C,E,F	C,E,G	C,E,H	C,F,G	C,F,H	C,G,H	D,E,F	D,E,G
Method 1	0.1727	0.1349	0.1341	0.1436	0.1671	0.1934	0.1758	0.1673
Method 2	0.1407	0.1401	0.1393	0.1305	0.1294	0.1255	0.1513	0.1471

	D,E,H	D,F,G	D,F,H	D,G,H	E,F,G	E,F,H	E,G,H	F,G,H
Method 1	0.1292	0.1794	0.1708	0.2056	0.1470	0.1379	0.1479	0.1663
Method 2	0.1532	0.1387	0.1407	0.1352	0.1460	0.1469	0.1423	0.1268

5 Variable VARs - Included Leading Indicators:							
	A,B,C,D	*A,B,C,E*	*A,B,C,F*	*A,B,C,G*	*A,B,C,H*	*A,B,D,E*	*A,B,D,F*	*A,B,D,G*
Method 1	0.2030	0.1933	0.1807	0.2000	0.2011	0.1561	0.1739	0.1805
Method 2	0.1380	0.1377	0.1354	0.1334	0.1354	0.1496	0.1416	0.1370

	A,B,D,H	*A,B,E,F*	*A,B,E,G*	*A,B,E,H*	*A,B,F,G*	*A,B,F,H*	*A,B,G,H*	*A,C,D,E*
Method 1	0.1829	0.1694	0.1965	0.1943	0.1600	0.1871	0.1982	0.1612
Method 2	0.1408	0.1430	0.1417	0.1407	0.1345	0.1348	0.1318	0.1458

	A,C,D,F	*A,C,D,G*	*A,C,D,H*	*A,C,E,F*	*A,C,E,G*	*A,C,E,H*	*A,C,F,G*	*A,C,F,H*
Method 1	0.198	0.201	0.186	0.188	0.175	0.156	0.181	0.203
Method 2	0.139	0.133	0.140	0.138	0.139	0.138	0.129	0.130

	A,C,G,H	*A,D,E,F*	*A,D,E,G*	*A,D,E,H*	*A,D,F,G*	*A,D,F,H*	*A,D,G,H*	*A,E,F,G*
Method 1	0.1778	0.1777	0.1729	0.1979	0.2162	0.1765	0.1798	0.1581
Method 2	0.1258	0.1505	0.1474	0.1517	0.1396	0.1412	0.1380	0.1455

	A,E,F,H	*A,E,G,H*	*A,F,G,H*	*B,C,D,E*	*B,C,D,F*	*B,C,D,G*	*B,C,D,H*	*B,C,E,F*
Method 1	0.1755	0.1526	0.1640	0.1696	0.1434	0.1553	0.1815	0.1509
Method 2	0.1444	0.1412	0.1281	0.1444	0.1412	0.1372	0.1395	0.1396

	B,C,E,G	*B,C,E,H*	*B,C,F,G*	*B,C,F,H*	*B,C,G,H*	*B,D,E,F*	*B,D,E,G*	*B,D,E,H*
Method 1	0.1510	0.1801	0.1669	0.1573	0.1885	0.1653	0.1893	0.1549
Method 2	0.1406	0.1389	0.1362	0.1366	0.1346	0.1485	0.1465	0.1483

	B,D,F,G	*B,D,F,H*	*B,D,G,H*	*B,E,F,G*	*B,E,F,H*	*B,E,G,H*	*B,F,G,H*	*C,D,E,F*
Method 1	0.1534	0.1610	0.1709	0.1323	0.1616	0.1790	0.1696	0.1649
Method 2	0.1395	0.1400	0.1389	0.1464	0.1456	0.1425	0.1346	0.1437

	C,D,E,G	*C,D,E,H*	*C,D,F,G*	*C,D,F,H*	*C,D,G,H*	*C,E,F,G*	*C,E,F,H*	*C,E,G,H*
Method 1	0.1655	0.1580	0.1789	0.1742	0.1820	0.1389	0.1824	0.1742
Method 2	0.1422	0.1461	0.1368	0.1381	0.1343	0.1410	0.1399	0.1394

	C,F,G,H	*D,E,F,G*	*D,E,F,H*	*D,E,G,H*	*D,F,G,H*	*E,F,G,H*		
Method 1	0.1634	0.1607	0.1477	0.1734	0.1547	0.1616		
Method 2	0.1290	0.1500	0.1506	0.1470	0.1333	0.1467		

6 Variable VARs - Included Leading Indicators:

	A,B,C,D,E	A,B,C,D,F	A,B,C,D,G	A,B,C,D,H	A,B,C,E,F	A,B,C,E,G	A,B,C,E,H
Method 1	0.1758	0.1646	0.2024	0.1740	0.1758	0.1721	0.1986
Method 2	0.1435	0.1383	0.1364	0.1387	0.1379	0.1396	0.1381

	A,B,C,F,G	A,B,C,F,H	A,B,C,G,H	A,B,D,E,F	A,B,D,E,G	A,B,D,E,H	A,B,D,F,G
Method 1	0.1676	0.1992	0.1724	0.1895	0.1722	0.1768	0.1696
Method 2	0.1348	0.1350	0.1329	0.1484	0.1464	0.1493	0.1381

	A,B,D,F,H	A,B,D,G,H	A,B,E,F,G	A,B,E,F,H	A,B,E,G,H	A,B,F,G,H	A,C,D,E,F
Method 1	0.2151	0.1933	0.1517	0.1829	0.1849	0.1933	0.1723
Method 2	0.1395	0.1372	0.1450	0.1429	0.1406	0.1338	0.1418

	A,C,D,E,G	A,C,D,E,H	A,C,D,F,G	A,C,D,F,H	A,C,D,G,H	A,C,E,F,G	A,C,E,F,H
Method 1	0.1900	0.1907	0.2018	0.1692	0.1917	0.1608	0.1642
Method 2	0.1427	0.1475	0.1349	0.1379	0.1336	0.1383	0.1374

	A,C,E,G,H	A,C,F,G,H	A,D,E,F,G	A,D,E,F,H	A,D,E,G,H	A,D,F,G,H	A,E,F,G,H
Method 1	0.1773	0.1782	0.1672	0.1621	0.2014	0.1606	0.1673
Method 2	0.1396	0.1301	0.1494	0.1492	0.1473	0.1362	0.1446

	B,C,D,E,F	B,C,D,E,G	B,C,D,E,H	B,C,D,F,G	B,C,D,F,H	B,C,D,G,H	B,C,E,F,G
Method 1	0.1646	0.1819	0.1528	0.1908	0.2044	0.1923	0.1696
Method 2	0.1428	0.1425	0.1434	0.1382	0.1400	0.1381	0.1410

	B,C,E,F,H	B,C,D,E,F	B,C,D,E,G	B,C,D,E,H	B,C,D,F,G	B,C,D,F,H	B,C,D,G,H
Method 1	0.2084	0.1646	0.1819	0.1528	0.1908	0.2044	0.1923
Method 2	0.1388	0.1428	0.1425	0.1434	0.1382	0.1400	0.1381

	B,C,E,F,G	B,C,E,F,H	B,C,E,G,H	B,C,F,G,H	B,D,E,F,G	B,D,E,F,H	B,D,E,G,H
Method 1	0.1696	0.2084	0.2003	0.2023	0.1741	0.1802	0.1393
Method 2	0.1410	0.1388	0.1395	0.1352	0.1479	0.1492	0.1468

	B,D,F,G,H	B,E,F,G,H	C,D,E,F,G	C,D,E,F,H	C,D,E,G,H	C,D,F,G,H	C,E,F,G,H
Method 1	0.1456	0.1565	0.1967	0.2013	0.2174	0.1646	0.1374
Method 2	0.1375	0.1451	0.1439	0.1433	0.1429	0.1354	0.1400

	D,E,F,G,H
Method 1	0.1396
Method 2	0.1494

7 Variable VARs - Included Leading Indicators:

	A,B,C,D,E,F	A,B,C,D,E,G	A,B,C,D,E,H	A,B,C,D,F,G	A,B,C,D,F,H	A,B,C,D,G,H
Model 1	0.1576	0.1864	0.2024	0.1837	0.2109	0.1982
Model 2	0.1416	0.1431	0.1446	0.1378	0.1384	0.1365

	A,B,C,E,F,G	A,B,C,E,F,H	A,B,C,E,G,H	A,B,C,F,G,H	A,B,D,E,F,G	A,B,D,E,F,H
Model 1	0.1838	0.1669	0.2028	0.1720	0.1898	0.1774
Model 2	0.1395	0.1382	0.1387	0.1345	0.1483	0.1484

	A,B,D,E,G,H	A,B,D,F,G,H	A,B,E,F,G,H	A,C,D,E,F,G	A,C,D,E,F,H	A,C,D,E,G,H
Model 1	0.1985	0.2089	0.1900	0.1633	0.1516	0.2031
Model 2	0.1467	0.1382	0.1436	0.1425	0.1432	0.1424

	A,C,D,F,G,H	A,C,E,F,G,H	A,D,E,F,G,H	B,C,D,E,F,G	B,C,D,E,F,H	B,C,D,E,G,H
Model 1	0.1752	0.1721	0.1555	0.1901	0.1614	0.1738
Model 2	0.1359	0.1403	0.1481	0.1425	0.1417	0.1435

	B,C,D,F,G,H	B,C,E,F,G,H	B,D,E,F,G,H	C,D,E,F,G,H
Model 1	0.1622	0.1674	0.1871	0.1859
Model 2	0.1374	0.1401	0.1482	0.1442

8 Variable VARs - Included Leading Indicators:

	A,B,C,D,E,F,G	A,B,C,D,E,F,H	A,B,C,D,E,G,H	A,B,C,D,F,G,H	A,B,C,E,F,G,H
Model 1	0.1896	0.2122	0.2116	0.1903	0.1934
Model 2	0.1428	0.1419	0.1426	0.1385	0.1403

	A,B,D,E,F,G,H	A,C,D,E,F,G,H	B,C,D,E,F,G,H
Model 1	0.2229	0.2004	0.1790
Model 2	0.1486	0.1441	0.1433

Appendix E: Regression Estimation Code

```
####################################################################
#                                                                  #
#              Nick Holschuh - Integrative Exercise                #
#           Business  Cycle  Turning  Point  Codebook              #
#      Rolling Regressions and Probability Score Calculation       #
#                            1/20/2010                             #
#                                                                  #
####################################################################

Colnames=list(list(),list("QLDrgdp","QDLsp500","QDLppi","Qspread","QDLdow","Qfedfunds","QDLm1","QDLm2","QDLbusinessloans"))

Timeseries=matrix(nrow=length(QDLrgdp),ncol=length(Colnames[[2]]),dimnames=Colnames)
Timeseries[,1]=QDLrgdp
Timeseries[,2]=QDLsp500
Timeseries[,3]=QDLppi
Timeseries[,4]=Qspread
Timeseries[,5]=QDLdow
Timeseries[,6]=Qfedfunds
Timeseries[,7]=QDLm1
Timeseries[,8]=QDLm2
Timeseries[,9]=QDLbusinessloans

Timeseries=ts(Timeseries,start=c(1964,2),frequency=4)


##############################################################
# We define here all possible combinations of the 9 variables #
##############################################################

combin=c(1)
combin[2]=length(combinations2[1,])
combin[3]=length(combinations3[1,])
combin[4]=length(combinations4[1,])
combin[5]=length(combinations5[1,])
combin[6]=length(combinations6[1,])
combin[7]=length(combinations7[1,])
combin[8]=length(combinations8[1,])
combin[9]=length(combinations9[1,])

combinlist=list()
combinlist[[1]]=1
combinlist[[2]]=combinations2
combinlist[[3]]=combinations3
combinlist[[4]]=combinations4
combinlist[[5]]=combinations5
combinlist[[6]]=combinations6
combinlist[[7]]=combinations7
combinlist[[8]]=combinations8
combinlist[[9]]=combinations9
```

```
options(warn=-1)

###############################################################
# Setting up the initial conditions for the rolling regression #
###############################################################

seedvalue=25            #Number of time periods in first regression
forecastlength=5        #The number of time periods forecasted
loops=1000              #The number of repitions for the forecasting process

turningpointarray=array(data=0,dim=c(9,length(QDLrgdp)+3,max(combin)))
#The array in which the number of forecasted turning points are stored


for(i in 2:9)                   #i represents the number of variables used in each regression
        {
        for(j in 1:combin[i])  #j represents the number of combinations of i-1 variables (RGDP is always included)
                {
                for(k in 1:(length(QDLrgdp)-seedvalue))     #k represents the number of steps in the rolling regression
                        {
                        tempnames=list("QDLrgdp")      #Selecting the variable names for each regression
                        if(i>1)
                                {
                                for(l in 1:(i-1))
                                        {tempnames[[1]][1+l]=Colnames[[2]][(combinlist[[i]][l,j]+1)]}
                                }
                        nameslist=list(list(),tempnames[[1]])
                        tempmatrix=matrix(ncol=i,nrow=seedvalue+k,dimnames=nameslist)#creating a matrix containing the data from QDLrgdp
                        for(m in 1:(seedvalue+k))
                                {tempmatrix[m,1]=Timeseries[m,1]}
                        maxlag=1
                        tempts=ts(tempmatrix,start=c(1964,2),frequency=4)
                        tempvar=VAR(tempmatrix,ic="HQ",lag.max=maxlag,type="const") #Here we run the VAR on the relevant variables
                        lags=tempvar$p                                             #We extract the number of lags selected by the VAR
                        tempmodellist=list()                                       #List storing individual ARs from the VAR
                        tempnameslist=list()                 #The variable names and lagged variable names used in the regression
                        for(m in 1:i)
                                {
                                tempmodellist[[m]]=tempvar$varresult[[m]]
                                }
                        for(m in 1:length(tempmodellist[[1]]$coefficients))
                                {
                                tempnameslist[m]=names(tempmodellist[[1]]$coefficients)[m]
                                }
                        forecastnameslist=list(list(),tempnameslist)
                        forecastmatrix=matrix(ncol=length(tempnameslist),nrow=(forecastlength+1),dimnames=forecastnameslist)
                                                #Matrix for forecast calculations
                        for(m in 1:length(tempnameslist))
                                {
```

```
                        forecastmatrix[1,m]=tempmodellist[[1]]$model[length(tempmodellist[[1]]$model[,1]),m+1]
                                                  #Extracting the last data point used in the regression
                }
        for(o in 1:loops)
                {
                matrixsum=matrix(ncol=i,nrow=length(tempmodellist[[1]]$coefficients))
                #A matrix used to store the product of the data points and the coefficients,
                for(p in 1:forecastlength)    #The sum of which is the forecasted value
                        {
                        for(q in 1:length(tempmodellist[[1]]$coefficients))
                                           #Forecasting each variable subject to their individual AR
                            {
                            for(r in 1:i)
                                    {
                                    matrixsum[q,r]=tempmodellist[[r]]$coefficients[q]*forecastmatrix[p,q]
                                    }
                            }
                        for(r in 1:i)
                                {

forecastmatrix[p+1,r]=sum(na.exclude(matrixsum[,r]))+rnorm(1,mean=0,sd=sd(tempmodellist[[r]]$residuals))
                                                  #Generating the random error
                                }
                        if(lags>1)              #logic to populate the lagged values into the forecast matrix
                                {
                                for(q in 1:i)
#For Each variable used (q), you muust populate all lagged cells of that variable (i*r+q), which in total is r+1
                                        {
                                        for(r in 1:(lags-1))
                                                {
                                                forecastmatrix[1+p,(i*r+q)]=forecastmatrix[p,(i*(r-1)+q)]
                                                }
                                        }
                                }

forecastmatrix[p+1,length(tempmodellist[[1]]$coefficients)]=forecastmatrix[p,length(tempmodellist[[1]]$coefficients)]
                        }
                        for(s in 1:3)#This looks 3 time periods in the future, predicting if any will be a turning point
                            {
                            if(forecastmatrix[s,1]<0)
                                    {if(forecastmatrix[s+1,1]>0)
                                            {if(forecastmatrix[s+2,1]>0)
                                                    {turningpointarray[i,seedvalue+k+s-
                                                    1,j]=turningpointarray[i,seedvalue+k+s-1,j]+1}
                                            }
                                    }
                            if(forecastmatrix[s,1]>0)
                                    {if(forecastmatrix[s+1,1]<0)
                                            {if(forecastmatrix[s+2,1]<0)
                                                    {turningpointarray[i,seedvalue+k+s-
                                                    1,j]=turningpointarray[i,seedvalue+k+s-1,j]+1}
```

```
                                                  }
                                              }
                                          }
                                      }
                                  }
print(i)
print(j)
                          }
              }


################################################################
#       Setting up the probability score calculations          #
################################################################


turningpoints1=matrix(ncol=combin[1],nrow=length(QDLrgdp)+3)
turningpoints2=matrix(ncol=combin[2],nrow=length(QDLrgdp)+3)
turningpoints3=matrix(ncol=combin[3],nrow=length(QDLrgdp)+3)
turningpoints4=matrix(ncol=combin[4],nrow=length(QDLrgdp)+3)
turningpoints5=matrix(ncol=combin[5],nrow=length(QDLrgdp)+3)
turningpoints6=matrix(ncol=combin[6],nrow=length(QDLrgdp)+3)
turningpoints7=matrix(ncol=combin[7],nrow=length(QDLrgdp)+3)
turningpoints8=matrix(ncol=combin[8],nrow=length(QDLrgdp)+3)
turningpoints9=matrix(ncol=combin[9],nrow=length(QDLrgdp)+2)



for(i in 1:9)
        {
        for(j in 1:combin[i])
                {
                for(k in 1:length(QDLrgdp))
                        {
                        if(k<27)
                                {
                                turningpoints1[k,j]=0
                                }
                        else{
                        if(i==1)
                                {
                                turningpoints1[k,j]=turningpointarray[i,k,j]/300
                                }
                        if(i==2)
                                {
                                turningpoints2[k,j]=turningpointarray[i,k,j]/300
                                }
                        if(i==3)
```

```
                                  {
                                  turningpoints3[k,j]=turningpointarray[i,k,j]/300
                                  }
                          if(i==4)
                                  {
                                  turningpoints4[k,j]=turningpointarray[i,k,j]/300
                                  }
                          if(i==5)
                                  {
                                  turningpoints5[k,j]=turningpointarray[i,k,j]/300
                                  }
                          if(i==6)
                                  {
                                  turningpoints6[k,j]=turningpointarray[i,k,j]/300
                                  }
                          if(i==7)
                                  {
                                  turningpoints7[k,j]=turningpointarray[i,k,j]/300
                                  }
                          if(i==8)
                                  {
                                  turningpoints8[k,j]=turningpointarray[i,k,j]/300
                                  }
                          if(i==9)
                                  {
                                  turningpoints9[k,j]=turningpointarray[i,k,j]/300
                                  }}
                          }
                  }
          }


probtsqps=probts[27:length(probts)]
QPSdata=turningpointarray[,27:length(QDLrgdp),]/300
QPSdata2=array(data=0,dim=c(9,length(QDLrgdp)-seedvalue,max(combin)))


for(i in 1:9)
          {
          for(j in 1:combin[i])
                  {
                  for(k in 1:(length(QDLrgdp)-27))
                          {
                          QPSdata2[i,k,j]=(QPSdata[i,k,j]-probtsqps[k])^2
                          }
                  }
          }

QPSscore=matrix(0,ncol=9,nrow=max(combin))
```

```
for(i in 1:9)
        {
        for(j in 1:combin[i])
                {
                QPSscore[j,i]=(2/length(probtsqps))*sum(QPSdata[i,,j])
                }
        }

save.image("postrun_workspace_1000loops.RData")
```