CONSIDERING POTENTIAL BENEFITS:
THE FREQUENCY BAND DECOMPOSITION
IN U.S. RECESSION PROBABILITY FORECASTING

John Summers
B.A., California State University, Sacramento, 2006

THESIS

Submitted in partial satisfaction of
the requirements for the degree of

MASTER OF ARTS

in

ECONOMICS

at

CALIFORNIA STATE UNIVERSITY, SACRAMENTO

FALL
2010
CONSIDERING POTENTIAL BENEFITS:
THE FREQUENCY BAND DECOMPOSITION
IN U.S. RECESSION PROBABILITY FORECASTING

A Thesis

by

John Summers

Approved by:

___________________________________, Committee Chair
Stephen Perez, Ph.D.

___________________________________, Second Reader
Mark Siegler, Ph.D.

___________________________
Date
Student: John Summers

I certify that this student has met the requirements for format contained in the University format manual, and that this thesis is suitable for shelving in the Library and credit is to be awarded for the thesis.

____________________________, Graduate Coordinator  _____________________
Jonathan D. Kaplan, Ph.D.      Date

Department of Economics
Abstract of
CONSIDERING POTENTIAL BENEFITS:
THE FREQUENCY BAND DECOMPOSITION
IN U.S. RECESSION PROBABILITY FORECASTING
by
John Summers

Abstract: Against the backdrop of The Great Moderation, The Great Recession is a remarkable departure. Economic losses have been severe. The U.S. business cycle is decidedly not dead and the justifications for continued research into U.S. recession forecasting are myriad. In this study, I test whether targeting the business-cycle frequency band of explanatory data can, ceteris paribus, increase U.S. recession probability forecast model performance. This study is the first to combine frequency decomposed time series with probit models using the General-to-Specific search methodology. Using data from the NBER and The Conference Board, I compared test criteria across model estimation results derived from General-to-Specific specification selection methodologies, and concluded that, on average, those derived from the decomposed explanatory datasets outperform their full-spectrum dataset rivals. Additionally, I’ve concluded that useful explanatory information is left behind by the filtering processes utilized here and that further research into differing decomposition schemes is therefore warranted.

________________________
, Committee Chair
Stephen Perez, Ph.D.

________________________
Date
ACKNOWLEDGEMENTS

I wish to thank Professors Stephen Perez and Mark Siegler for their particular help in the completion of this thesis. I would also like to thank all my professors, past and present, for their invaluable inspiration and guidance. Finally, I thank my wife Betsy, whose love and support made this thesis and the better part of my academic career possible.
# TABLE OF CONTENTS

| Acknowledgments                                                                 | v   |
| List of Tables                                                                  | viii|
| List of Figures                                                                 | ix  |

# Chapter

## 1. INTRODUCTION

1.1 A Particular Way of Looking at Economic Time Series                     2
1.2 The Impetus for This Study                                               4
1.3 My Research Question and a Secondary Follow-up                           4
1.4 Relevance                                                                6
1.5 Empirical Methodology                                                     11
1.6 Principal Findings                                                        12
1.7 Organization                                                             13

## 2. LITERATURE

2.1 A Brief Survey of the U.S. Recession Probability Forecasting Literature  14
2.2 The Empirical Models Used Within the Surveyed Literature               14
2.3 The Explanatory Data Used Within the Surveyed Literature          15
2.4 Best-Case Specification Selection                                      16
2.5 The Usual Uses for Frequency Decomposed Economic Time-Series Data      20

## 3. DATA

3.1 Data Sources and Sample Periods                                          22
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1</td>
<td>Specification Scoring – Primary Research Question – Full Data Sample</td>
<td>34</td>
</tr>
<tr>
<td>Table 2</td>
<td>Specification Scoring – Primary Research Question – 1959M01:2002M07</td>
<td>35</td>
</tr>
<tr>
<td>Table 3</td>
<td>Specification Scoring – Primary Research Question – 1959M01:1991M11</td>
<td>36</td>
</tr>
<tr>
<td>Table 4</td>
<td>Specification Scoring – Primary Research Question – All Sample Periods</td>
<td>37</td>
</tr>
<tr>
<td>Table 5</td>
<td>Specification Scoring – Secondary Question – Full Data Sample</td>
<td>39</td>
</tr>
<tr>
<td>Table 6</td>
<td>Specification Scoring – Secondary Question – 1959M01:2002M07</td>
<td>39</td>
</tr>
<tr>
<td>Table 8</td>
<td>Specification Scoring – Secondary Question – All Sample Periods</td>
<td>40</td>
</tr>
<tr>
<td>Table 9</td>
<td>Robustness Testing – Primary Research Question – All Sample Periods</td>
<td>43</td>
</tr>
<tr>
<td>Table 10</td>
<td>Robustness Testing – Secondary Question – All Sample Periods</td>
<td>44</td>
</tr>
<tr>
<td>Table 11</td>
<td>The J-Test – Results</td>
<td>47</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>Conference Board Leading Economic Indicators</td>
<td>3</td>
</tr>
<tr>
<td>Figure 2</td>
<td>U.S. Annualized Quarterly Growth Rates</td>
<td>7</td>
</tr>
<tr>
<td>Figure 3</td>
<td>U.S. Recession Period Output and Employment</td>
<td>9</td>
</tr>
</tbody>
</table>
Chapter 1
INTRODUCTION

In this study, I examine whether targeting the business-cycle frequency band of explanatory data can, ceteris paribus, increase U.S. recession probability forecast model performance. My intuition is founded upon the simple notion that the portion of an explanatory data series which cycles within the same frequency band as the phenomena being modeled should be the portion of that data which best describes the phenomena, all else held constant.

The business-cycle frequency band range used in this study is 18 to 128 monthly observations. These values are the peak-to-peak and trough-to-trough minimum and maximum monthly observation range values for U.S. recession periods post-World War II, as defined by the National Bureau of Economic Research (NBER).

I found no previous research in which this particular component of data, or frequency-decomposed time series in general, are used as explanatory data within this class of regression analyses.

To be clear, the relative differences tested in this study are entirely contained within particular transformations of the explanatory data. The intention of this study is not to develop a better forecasting model per se. Rather, the intention is to determine if the use of frequency-decomposed explanatory data can, ceteris paribus, add explanatory value. Though it is beyond the scope of this study, next steps should include research into leveraging what is learned here into generally better performing forecast models.
1.1 – A Particular Way of Looking at Economic Time Series

For the purposes of this study it is helpful to think of a given economic time series as the sum of three distinct components: a short-run (SR) component, a medium-run or business-cycle component (BC), and a long-run (LR) component.1 Within this context, the full spectrum (FS) of an individual data series may be defined as follows: \( FS = SR + BC + LR \).

1.1.1 – Frequency Decomposition of Economic Time-Series and the Band-Pass Filter

Unless otherwise indicated, frequency decomposition is the process of breaking apart the full spectrum of explanatory data into its BC and (SR + LR) components. The theoretical justification for this process comes from the spectral representation theorem. The practical application of this decomposition process is made possible by using a band-pass data filter.2

---

1 See Iacobucci (2003) for discussion of this particular decomposition scheme.
Figure 1 shows the post-frequency decomposition components of the full spectrum of The Conference Board’s Leading Economic Indicator Composite Leading Index. The gray vertical bars indicate NBER-defined U.S. recession periods. The heavy solid line is the full-spectrum index series and is measured on the right-hand scale. The coarse dashed line is the isolated business-cycle component of the index and is measured on the left-hand scale. Also included and measured on the right-hand scale is the fine dashed line which represents the decomposition process residual series. Within the
context explicated in Section 1.1, this residual series is the sum of the SR and LR components of the FS. It is my intuition that the isolated business-cycle component of the data should, ceteris paribus, better describe impending NBER-defined U.S. recession conditions than does the full-spectrum series.

1.2 – The Impetus for This Study

This research stems from the convergence of two analytical methods: co-movement analysis of the business-cycle components of leading indicator series relative to the same component of a proxy for real gross domestic product, and impending recession probability forecasting using discrete choice regression models. The convergence of these two methods is manifested here by the use of transformed data from the former method within the empirical modeling structure of the latter method. This technique is unique within the literature. The empirical model and some of the processes used in this study follow those used by Silvia, Bullard, and Lai (2008).

1.2.1 – Silvia, Bullard and Lai (2008)

The study done by Silvia, et al. (2008) featured a probit model used to forecast the probability of impending NBER-defined U.S. recession conditions as a function of The Conference Board’s U.S. Composite Leading Index (shown in Figure 1), and other variables. The authors’ preferred model was selected from a data pool of more than five hundred variables using a Stepwise specification selection methodology.

1.3 – My Research Question and a Secondary Follow-up

The business-cycle component of a given economic time series is, in theory, the portion of that series which cycles within the same frequency band as the greater macro
economy. Therefore, using the business-cycle components of the explanatory data in place of the full-spectrum series within a given unrestricted recession probability forecasting model should, ceteris paribus, result in superior derived best-case specification model performance statistics. To test this thesis I established a series of questions.

My primary research question is the following. For a given baseline set of data, does the use of that data’s isolated business-cycle components on average result in a better recession probability forecasting model than if the full spectrum of the same data is used? Referring back to Figure 1, my primary question asks if the data represented by the coarse dashed line does a better job of anticipating the gray vertical bars than does the heavy solid line. My intuition is that it should.

My secondary question flows from my primary question. I do not refer to it as a research question per se because it does not directly follow from my intuition or my thesis. The question is as follows. Is there any value contained within the decomposition process residual series? Recall from Figure 1 that the residual series is, in theory, the sum of the SR and LR components of the data. My original intuition held that excluding these components of the data should increase descriptive power. This would seem to infer that there is negligible explanatory value contained within the residual series. The purpose of this secondary question is to appropriately qualify that portion of my intuition.

3 In this study what constitutes a better model was determined by comparing reported regression estimation output statistics and results derived from tests against the generated fitted series.
4 To test this condition the same statistics and tests against the generated fitted series were compared for the appropriate expressions of the explanatory data.
1.4 – Relevance

This research is relevant now because, given the severity of The Great Recession, the relative economic calm enjoyed in the U.S. during The Great Moderation may be over. This research is relevant going forward because the U.S. economy is constantly evolving and the factors which determine and/or best describe cyclical behaviors need to be continually reevaluated. The processes developed within this study can be extended into systems which address these ongoing needs.

1.4.1 – The Great Moderation

From the mid 1980s through 2007 the U.S. economy enjoyed diminished levels of economic volatility. Some believed better monetary and fiscal policy had at last tamed the business cycle. Stock and Watson (2002) credit the majority of these realized effects to various forms of “good luck.”

---

5 The name The “Great Recession” is established within the literature. See Mian and Sufi (2010).
6 See Mitchell and Burns (1938), Stock and Watson (2003), and Campos et al. (2005) for discussion of the changing nature of economies and the need to continually reevaluate preferred indicators.
Figure 2 illustrates the Great Moderation by showing the decline in the volatility of several macroeconomic time series beginning in 1984. Following themes explicated within Stock and Watson (2002), the shaded area highlights the period 1984Q1 to 2007Q4. This is the period from the beginning of The Great Moderation to the beginning of The Great Recession. The variables shown are 1) national output, 2) production in manufacturing, 3) a measure of aggregate prices, and 4) non-farm jobs. They represent a
broad cross-section of the greater U.S. macro economy. The reported series are rolling 10-year standard deviations of annualized quarterly growth rates which are used as an indicator of economic volatility. Normalized values are reported in order to provide a clearer picture of how volatility levels across this broad snapshot of the U.S. economy declined similarly within the highlighted time period. The figure illustrates that from the beginning of The Great Moderation to the onset of The Great Recession the U.S. economy as a whole became, and stayed, remarkably less volatile. This decline in volatility creates the context for what happened next.

1.4.2 – The Great Recession

The Great Recession, 2007Q4 through 2009Q2, in both its relative magnitude of contraction and in its duration, was the most severe U.S. recession post-World War II. In addition, when compared with the contractions which occurred during The Great Moderation, The Great Recession was particularly severe.
Figure 3 illustrates the relative severity of The Great Recession when compared with previous U.S. recession periods. The reported values are the peak-to-trough rates of change in real gross domestic product and non-farm payrolls associated with U.S. recessions since World War II. The numbers in parentheses indicate the duration in quarters of each contractionary period. The gray vertical shading highlights the
contractions which occurred following the beginning of The Great Moderation and prior to the onset of The Great Recession.

1.4.3 – The Ever-Changing Economy

The U.S. economy is complicated and ever-changing. According to Campos, et al. (2005) it is “a dynamic, nonlinear, simultaneous, high-dimensional, and evolving entity (p.1).” Its social systems and laws are constantly in flux and technological progress is commonly occurring in one or more forms. Therefore, as a quantifiable entity, the U.S. economy is a moving target at best. Not only is it moving, it “behaves in a distinctly non-stationary manner […] and is] subject to sudden and unanticipated shifts” (Campos et al., 2005, p. 1).

Economists have recognized the difficulties inherent in these complexities since as early as 1938. Mitchell and Burns, in the first section of their much-cited “Statistical Indicators of Cyclical Revivals,” state that “one of the clearest teachings of experience is that every business cycle has features that are peculiar to it. […] Whatever judgments are formed ought to be based, not upon the behavior of one or two indexes of business conditions, but upon the behavior of a considerable number of statistical series that represent a wide variety of economic processes” (Mitchell and Burns, 1938, p. 1).

Sixty-five years later this idea still applies. In their study of leading indicator performance in predicting the U.S. recession of 2001, Stock and Watson (2003) reaffirmed that the variables which best describe impeding recessionary conditions do change from recession to recession, and that these indicators ought not to be used in
isolation. Instead, they should be considered collectively (here also citing Mitchell and Burns (1938)).

The recent experience of The Great Recession shows that the business cycle is not dead and that continued business-cycle research is warranted. Because the U.S. economy is complicated and always changing, continued investigation into new classes of U.S. recession forecasting models is warranted.

1.5 – Empirical Methodology

In this section I briefly discuss the choice of a general empirical model, potential explanatory variables, and the specification search methodology leading to the final specification estimation output statistics presented in this paper.

1.5.1 – An Appropriate Empirical Model

Following Silvia, et al. (2008) I chose to use a probit model. The probit model is strongly preferred within the relevant literature surveyed. I chose to follow Sylvia, et al. (2008) for the use of their empirical model, and because their study contained guidance for my other two methodological requirements. No other study surveyed provided this guidance.

1.5.2 – An Appropriate Baseline Data Set

The data are the most important component of this study. The relative differences being tested are entirely contained within particular transformations of the explanatory data.

I chose to use The Conference Board’s Leading Economic Indicator (LEI) data set because it is a diverse collection of indicators designed to signal peaks and troughs in the
U.S. business cycle.\textsuperscript{8} I also chose it because it encompasses the explanatory variables within Silvia’s preferred model as well as other models found within the relevant recession forecasting literature. The LEI data set constitutes the component series of the Composite Leading Index featured in Figure 1 shown within Section 1.1.1.

\textit{1.5.3 – A Method for Determining Best-Case Specifications}

In order to make claims regarding model performance relative only to differing transformations of a given baseline data set, a method of determining a best-case specification for each data transformation was required. Though Silvia et al. (2008) utilized a Stepwise specification selection method, I chose to use the General-to-Specific methodology because of its documented superior abilities. These include the ability to recover true known specifications within simulated experiments and the ability to return better specifications compared to those recovered by other commonly used specification selection methodologies, including the Extreme Bounds and Stepwise approaches.

\textit{1.6 – Principal Findings}

My research confirms that the business-cycle component of the data does outperform the full spectrum of the same data. And, there is still explanatory value contained within the decomposition process residual series. In robustness testing I found that similar results were achieved when applying a high-pass filter to the explanatory data in place of the band-pass filter. Finally, the results of encompassing tests conducted for rival specifications suggest that explanatory power is often increased without loss of information content for the band-pass decomposition scheme tested here.

\textsuperscript{8} See this description of purpose at The Conference Board web site http://www.conference-board.org/data/bcicountry.cfm?cid=1.
1.7 – Organization

The remainder of this paper is organized as follows. Chapter 2 surveys relevant literature, Chapter 3 presents the data used, Chapter 4 describes my methodology in detail, Chapter 5 presents the results, and Chapter 6 concludes.
Chapter 2
LITERATURE

I found no previous research in which frequency decomposed time series are used as explanatory data within discrete choice regression analyses. Therefore, the contributions of this paper to the literature are unique. The contributions to this paper from the literature primarily consist of justifications for my choice of empirical model, my selection of data, and my method for determining best-case specifications.

2.1 – A Brief Survey of the U.S. Recession Probability Forecasting Model Literature

Within the surveyed literature the probit approach is the preferred method for forecasting the probability of impending recessions. Following Silvia et al. (2008) I chose to use a probit model featuring a cumulative forecast horizon.

There are other studies which have used similar models. In Estrella et al. (2003) the authors used cumulative and marginal forecast horizon probit models to forecast the probability of impending U.S. and German recessions. Wright (2006) used a cumulative horizon probit model to predict financial market perceptions. In a seminal work in the topic area, Estrella and Mishkin (1998) used the marginal probit approach in predicting U.S. recessions.

2.2 – The Empirical Models Used Within the Surveyed Literature

As discussed above, the probit model is the preferred model for recession forecasting within the surveyed literature. Within this sampling of the literature probit models featuring both cumulative and marginal forecast horizon approaches were used. It
is therefore important to understand the difference between the cumulative and marginal approaches.

2.2.1 – *Cumulative and Marginal Forecast Horizons*

Within the surveyed probit literature there are two particular methods of defining true observations of the binary dependent variable: the cumulative method and the marginal method. The cumulative method defines observations of the binary dependent variable as true if the condition of interest is true within the next $h$ observations of the data. For example, if $h$ is set equal to three, and the condition of interest is NBER defined U.S. recession conditions, relative to time observation $t$, if any one of the next three observations is a defined recessionary observation, then observation $t$ of the dependent variable is set to true. If this condition is not met, then observation $t$ defaults to false. The marginal method defines observations of the dependent variable as true if and only if the condition of interest is true at time $t$.\(^9\)

Following Silvia et al (2008) I chose the cumulative forecast horizon method. I also chose this method because it allows for leveraging in-sample explanatory power to make what are essentially out-of-sample forecast predictions.

2.3 – *The Explanatory Data Used Within the Surveyed Literature*

Each study within the surveyed recession forecasting literature used some form of bond-yield characteristics within its preferred model’s explanatory data set. The Leading Economic Indicator data set contains a bond-yield spread series. The other explanatory variables contained within the preferred models surveyed include: 1) The Conference

\(^9\) See Mishkin et al. (2003) for a discussion of forecast horizons.
Board’s Composite Leading Indicator, 2) an employment index and a measure of stock prices (Silvia et al., 2008), 3) the Fed Funds Rate (Wright, 2006), 4) corporate bond credit spreads (King et al., 2007), and 5) other yield curve associated characteristics (Estrella et al., 2003). The Conference Board’s LEI set contains variables similar to most of these. Following these studies, and for the reasons already argued, I chose The Conference Board’s LEI data set for my baseline data set.

2.4 – Best-Case Specification Selection

Within the surveyed recession forecasting literature only Sylvia et al. (2008) used an automated specification selection methodology. The authors used a version of the Stepwise method. A limitation of the Stepwise method is that it utilizes fairly simple reduction criteria and can be manually intensive to apply (Silvia et al., 2008). I chose to use the General-to-Specific methodology because of its documented superior performance abilities. In the three sections which follow I will argue my justification for selecting the General-to-Specific methodology.

2.4.1 – What is the General-to-Specific Methodology?

General-to-Specific methodology is a main tenet of the London School of Economics’ approach to econometrics (Campos et al., 2005). Hoover and Perez (1999) describe the application of the General-to-Specific methodology as follows:

“[General-to-Specific] relies on an intuitively appealing idea. A sufficiently complicated model can, in principle, describe the salient features of the economic world. Any more parsimonious model is an improvement on such a complicated model if it conveys all of the same information in a simpler, more compact form. Such a parsimonious model would necessarily be superior to all other models that are restrictions of the completely general model except, perhaps, to a class of models nested within the parsimonious model itself” (p. 168).
General-to-Specific is similar to the Stepwise method in that they both are reductive processes. As described above, these processes are intended to reveal the most parsimonious model possible subject to the established general unrestricted model.

2.4.2 – General-to-Specific v. the Average Economic Regression Method

Silvia et al. (2008) argue that reductive processes are a better approach than conventional modeling methods because they allow the data to speak for itself in assessing economic theory. These processes also minimize the potential for researcher bias and/or outcomes skewed towards pre-determined expectations.

As pointed out by Hoover and Perez (1999), the General-to-Specific methodology begins with a reasonable approximation of all the data necessary and sufficient to describe the economic phenomena of interest. Then, through an iterative process of judicious restrictions placed upon that data, it selects a best-case specification.

The Average Economic Regression (AER) approach compels the econometrician, by a priori theory, to select an initial set of variables and estimate an equation using those variables. Then she must conduct an iterative process of adding and or subtracting candidates to the variable pool as subsequent estimation results and modifications to the theory dictate and/or allow. This process is repeated until a satisfactory model is produced, a model which meets both theoretical and statistical expectations.

According to Gilbert (1986), the main problem with the AER approach is that it uses econometrics to illustrate independently held a priori theories rather than leaving econometrics to discover and test tenable new views upon economic phenomena. In other
words, with the AER approach a priori theory is imposed upon the data. The reductive approach allows the data to inform the theory.

2.4.3 – *General-to-Specific v. Other Specification Selection Methods*

The General-to-Specific methodology produces specifications which perform as well as or better than those derived by other automated specification selection methods, including the Stepwise method and variants of the Extreme Bounds methodology. It has also derived better specifications than those preferred by econometricians in empirical studies which began with pre-defined explanatory data collections. The brief chronology that follows addresses documentation of these results within the literature.

In their 1999 work “Data Mining Reconsidered,” Hoover and Perez tested their General-to-Specific algorithm against the established baseline data from Lovell (1983). Lovell had tested a small sampling of selection methodologies, including a Stepwise method. Hoover and Perez wanted to see if General-to-Specific methodology would work in a situation where it should work. They found that their algorithm recovered the correct or a closely related specification most of the time. Their results were a dramatic improvement upon the results of Lovell.

Later in 1999 Hendry and Krolzig (1999) introduced their General-to-Specific algorithm within a software application called PcGets. It embodied alternative specification evaluation standards relative to the algorithm of Hoover and Perez (1999). In the cases where the Hoover and Perez (1999) algorithm performed at its worst in recovering the known true specification, PcGets returned significantly better results.
Two years later, using an improved version of PcGets, Krolzig and Hendry (2001) were able to improve upon the preferred specifications of human modelers subject to unknown true specifications. In other words, their General-to-Specific application returned arguably better specifications than those settled upon by human researchers in situations where the available data pool is known but the true specification is not.\(^\text{10}\)

Hoover and Perez (2004) tested an updated version of their General-to-Specific algorithm against cross-sectional data and returned a more concise and parsimonious model compared to that selected by Sali-i-Martin in his 1997 study “I Have Just Run Two Million Regressions.” Sali-i-Martin had used variants of the Extreme Bounds selection methodology in his study. The Hoover and Perez (2004) study is important because it demonstrated that General-to-Specific methodology could be used effectively against data set types other than time series.

Also in 2004 Hendry and Krolzig, in their paper “We Ran One Regression,” confirmed and corroborated the results of Hoover and Perez (2004). They pointed out the obvious efficiency benefits of their approach relative to the Sali-i-Martin (1997) study.

In sum, I used the General-to-Specific methodology as my method for determining best-case specifications because of its documented superior performance abilities. No other methodology within the surveyed literature performed as well.

\(^{10}\) The studies improved upon include Davidson, et al. (1978) and Hendry and Ericsson (1991). See Campos et al. (2005).
2.5 – The Usual Uses for Frequency Decomposed Economic Time-Series Data

The usual use within the surveyed literature for decomposed time-series data is time-series co-movement analysis. The listing of research studies which follows documents particular applications of this methodology.

Hodrick and Prescott (1997) used high-pass filtered data to study the nature of co-movements between the higher frequency components of a variety of U.S. macroeconomic time series. They found that the nature of the relationship between the higher frequency components of the data is different from that of the respective business-cycle and long-term trend components.

Baxter (1994) investigated the link between real exchange and real interest rate differentials. She found evidence of a significant co-movement relationship at both long-term trend and business-cycle frequencies. This was significant because prior studies that had focused upon higher frequency transformations of the data had found no apparent relationship.

Christiano and Fitzgerald (1998) used their band-pass filter to evaluate U.S. business-cycle sectoral relationships for two to eight year frequency bands of the data. They found a substantial degree of co-movement between hours worked across a variety of business sectors.

Christiano and Fitzgerald (2003) studied the relationship between money growth and inflation in the U.S. prior to and after 1960. They uncovered a significant difference in the quality of the relationship relative to the division point. In the early period the relationship was strong and positive at all frequencies. In the latter period the relationship
turned negative at frequencies of twenty years or longer, though it remained positive at all other frequencies tested. Using their band-pass filter the authors decomposed the data into frequency ranges of one and a half to eight years, eight to twenty years and twenty to forty years.

Iacobucci (2003) studied U.S. unemployment and inflation. Using a “windowed” filter she successfully confirmed that a Phillips curve relationship does exist within specific frequency bands of the data, even though no such relationship exists within the full-spectrum series. The frequency range used was one to twenty-one years.
Chapter 3

DATA

The data are the most important component of this study. The relative differences being tested are entirely contained within particular transformations of the explanatory data. For all testing the dependent variable set is held constant. The empirical modeling form is also a constant. It is assumed that each General-to-Specific derived best-case specification is just that – the best case possible. Therefore, that aspect of the analytical process is considered to be held constant as well. The only variables within the entire testing environment are the particular transformations of the explanatory data. It is, then, appropriate to speak of these transformations as rivals.

3.1 – Data Sources and Sample Periods

Unless otherwise indicated the data used in this study are monthly and came from either the NBER or The Conference Board. The dependent variable data is from the former, the explanatory data is from the latter. The study addresses three sample periods of explanatory data. The full-sample range is January 1959 through June 2007. The two sub-sample periods are January 1959 through November 1991 and January 1959 through July 2002. The sub-sample date ranges were selected in relation to the two U.S. recessions immediately prior to The Great Recession. For each of the three sample periods tested, twelve observations of the data were reserved to facilitate out-of-sample forecasting.
3.2 – The Dependent Variables

Four different binary dependent variables were tested. Each dependent variable represents a different cumulative forecast horizon value and each was determined in relation to NBER recession period dating. Using the methodology described in Section 2.2.1, the four variables were constructed for values of $h$ equal to 3, 6, 9, and 12.

3.3 – The Baseline Explanatory Data Set

The baseline explanatory data set is The Conference Board’s LEI data set, which consists of ten individual indices. The indices are all for the U.S. and are as follows:

- The Average Workweek of Production Workers in Manufacturing
- Home Building Permits
- Consumer Expectations
- The Ten Year Treasury Yield Less Fed Funds Rate Interest Rate Spread
- The M2 Money Supply
- New Manufacturing Orders for Consumer Goods
- New Manufacturing Orders for Non-Defense Capital Goods
- The Prices of 500 Common Stocks
- A Diffusion Index of Vendor Delivery Performance
- Average Weekly Initial Jobless Claims

3.4 – The Decomposition Scheme

The decomposition scheme applied against the full spectrum (FS) of the explanatory data isolates the business-cycle (BC) component of the data and leaves behind a residual series. This scheme is based upon an assumption that the BC
component is that portion of the FS which cycles within an 18 to 128 monthly observation range. Recall that these values are the NBER defined U.S. recession peak-to-peak and trough-to-trough minimum and maximum monthly observation range values for U.S. recession periods post-World War II. Further recall that under the assumption that the full spectrum of the data is equal to the sum of its short-run (SR), BC, and long-run (LR) components, the residual component from the decomposition process is by definition the sum of the SR and LR components.

3.4.1 – The Band-Pass Filter

The filter used to implement the decomposition scheme described in Section 3.4 was the Christiano-Fitzgerald Full Sample Asymmetric Frequency Filter, a band-pass filter.¹¹

3.5 – The Post-Decomposition Explanatory Data Components

Prior to the decomposition process there is a single data component: the FS of the data. Following the decomposition process there are three distinct data components: the FS, the BC, and the residual (SR + LR).

Following standard econometric practice, all post-decomposition data components were tested for the presence of a unit root process using the Augmented Dickey Fuller test and then differenced as required until the null hypothesis of unit root could be rejected at better than the 5% level.

¹¹ The Christiano-Fitzgerald frequency filter is included functionality within EViews version 6.0.
3.5.1 – The Full-Spectrum Series

The FS components of the data are The Conference Board LEI data set. There are ten series contained within the data set. The individual series were defined in Section 3.3.

3.5.2 – The Business-Cycle Components

The BC components of the data are the isolated business-cycle frequency band components of the FS series. There are ten of these components, one for each of the ten series in the full-spectrum data set.

3.5.3 – The Residual Components

Again, the residual components of the data are by definition the sum of the SR and LR components of the FS series. There are ten of these components, one for each of the ten series in the full-spectrum data set. Note that in this study the residual is never decomposed into its short and long-run sub-components.

3.6 – Expressions of the Explanatory Data

In total, this study establishes three separate explanatory expressions: an FS expression of the data, the band-pass filtered business-cycle (BPBC) expression, and the band-pass filtered business cycle in union with residual components of the data (BPBCR) expression.

3.6.1 – The Full-Spectrum Components of the Data Expression

The FS expression of the data is the baseline against which the BC expression of the data was tested. The full-spectrum expression contains the ten stationarity-corrected

---

12 Recall here that because all other factors are held constant, the actual rival entities relative to my thesis testing are the differing expressions of the data.
data series from The Conference Board’s LEI data set. It is the set of component series described in Section 3.5.1.

3.6.2 – The Business-Cycle Components of the Data Expression

The BPBC expression of the data is the collection of the isolated BC components of the FS data series. This expression of the data contains the ten series described in Section 3.5.2, one for each of the ten FS series. Comparison of the model specification results derived using this expression of the explanatory data to those derived using the FS expression of the data were used to answer the primary research question.

3.6.3 – The Business Cycle in Union with Residual Components of the Data

The BC in union with residual components expression contains twenty data series. In addition to the isolated BC components of the full-spectrum series, the post-decomposition residual series are included. This expression of the explanatory data is the union of the post-decomposition components described in Sections 3.5.2 and 3.5.3. Comparison of the model specification results derived using this expression of the explanatory data to those derived using the BC components expression of the data were used to answer the secondary question.

Taken together, the four different dependent variables regressed against the three different expressions of the baseline explanatory data set resulted in twelve distinct modeling combinations for each of the three data sample periods. The statistics derived from the best of each of these was used to assess rival performance.
Chapter 4

METHODOLOGY

The methodology used in this study follows a simple analytical process in which statistics from best-case specifications derived from rival expressions of the explanatory data are compared in order to determine the best performing expression, ceteris paribus.

4.1 – The Empirical Model

The specific form of all models estimated follows Silvia et al. (2008). It is shown below in specification (1).

\[
(1) \quad y_{t+h|t} = \beta' x_t + \varepsilon_t
\]

In specification (1) \( y_{t+h|t} \) is a variable which at time \( t \) determines if NBER defined recession conditions are true at any time between times \( t \) and \( t+h \) inclusive. In other words, if recession conditions occur anytime within the next \( h \) periods relative to time \( t \) then \( y_{t+h|t} \) is set equal to 1, otherwise, \( y_{t+h|t} \) defaults to 0. \( x_t \) is a vector which includes the explanatory variables at time \( t \), and \( \beta' \) is a vector of estimated coefficients including a constant term. \( \varepsilon_t \) is the error term.

\[
(2) \quad P(y_{t+h|t} = 1) = \phi(\hat{\beta}' x_t)
\]

As shown in specification (2), the estimated probability that \( y_{t+h|t} \) is equal to 1 for any given \( t \) is a function of the estimated coefficients and the explanatory data subject to \( \phi \), the cumulative density function of the standard normal distribution (Silvia et al., 2008).
4.2 – The Performance Criteria

The relative performance of rival expressions was compared using four different test criteria. Three are in-sample criteria and one is an out-of-sample criteria. Together these four tests constitute the source for the statistics by which relative performance is assessed. The four criteria are 1) the Akaike Information Criteria (AIC/n), 2) the McFadden R-square, 3) the percentage of correct predictions made in-sample, and 4) the percentage of correct predictions made out-of-sample. In the four following sections I briefly describe each criterion.

4.2.1 – AIC/n

The Akaike Information Criteria was used for in-sample testing. It is the information criteria reported by the particular General-to-Specific algorithm used and is similar to other information criteria used by other research within the surveyed literature.13 Lower values of the AIC/n indicate better performance.

4.2.2 – McFadden R-square

The McFadden R-square statistic was used for in-sample testing. It is a pseudo r-square statistic commonly reported for probit model estimations. It was used by several of the studies within the surveyed literature.14 Higher values of this statistic indicate better performance.

4.2.3 – Percentage Correct Predictions In-Sample

The percent of correct in-sample predictions was determined by comparing estimated in-sample probabilities against actual outcomes. Estimated values greater than

---

0.5 were interpreted as positive predictions. Estimated values equal to or less than 0.5 were interpreted as negative predictions. This criterion was selected for its ease of interpretation. Higher values of this statistic indicate better performance.

4.2.4 – Percentage Correct Predictions Out-of-Sample

The percent of correct out-of-sample predictions was also determined by comparing the estimated out-of-sample probabilities against actual outcomes. As with the in-sample criterion, estimated values greater than 0.5 were interpreted as positive predictions and values equal to or less than 0.5 were interpreted as negative predictions. Again, this criterion was selected for its ease of interpretation. Higher values of this statistic indicate better performance.

4.3 – Final Empirical Specifications

To answer my specific questions I compared the appropriate subsets of my performance criteria statistics. For each sample period, particular subsets from the twelve distinct sets of four statistics discussed at the end of Chapter 3 relate to each question.

The primary research question in its testable form compared the performance criteria statistics described in Section 4.2 from the derived best-case specifications for the full-spectrum (FS) and band-pass derived business-cycle component (BPBC) expressions of the explanatory data. If the BPBC expression performed better than the FS expression, then the answer to the primary question is “yes,” the business-cycle component of the explanatory data better describes impending U.S. recession conditions than does the full spectrum of the same data. If the inverse is true, then either the answer to the question is “no,” or no conclusion is possible.
The secondary question in its testable form compared the same performance criteria statistics for the derived best-case specifications from the BPBC and the BC cycle in union with residual-component (BPBCR) expressions of the explanatory data. If the BPBCR component expression performed better than the BC expression, then the answer to the secondary question is “yes,” there is additional explanatory value left within the residual component of the data. If the inverse is true, then either the answer to the question is “no,” or no conclusion is possible.

4.4 – The Analytical Process

For each sample period, and for each combination of the four differing dependent variables and the three differing explanatory expressions, the General-to-Specific specification selection algorithm was used to determine an objective best-case model. Each best-case model was then estimated and the four performance criteria statistics described in Section 4.2 were collected. Finally, the appropriate subsets of these statistics were compared in order to answer the primary and secondary questions.

4.5 – The Software

The econometric software used in this study was EViews version 6.0 and OxMetrics’ PcGive version 12. The Autometrics function within PcGive is the General-to-Specific specification selection algorithm used in this study.

4.5.1 – An Autometrics Issue and the Workaround Implemented

The Autometrics function within PcGive repeatedly failed to execute to completion for thirteen of the thirty-six total dependent variable/explanatory expression combinations related to the robustness testing described in results Section 5.3.
Consequently, in these cases the General-to-Specific algorithm failed to estimate probit model specifications. The error descriptions returned by the software provided no useful information. For each of these thirteen probit model estimation failures, a logit model specification selection was attempted. Six of these estimations failed similarly. In these six cases ordinary linear model specification selection was completed successfully. In all thirty-six cases the specification ultimately selected was successfully estimated in EViews as a probit model. Though the models determined by means of this workaround may not be considered best-case probit specifications per se, under the circumstances they stand as the best possible candidates.
In short, the answer to both questions tested is “yes.” Isolating the business-cycle component of the explanatory data did result in improved model performance. And, explanatory information was left in the residual series by the particular decomposition scheme used.

The relevant test statistics from the General-to-Specific determined best-case specification estimations are reported in the tables which follow. Because the General-to-Specific process was used here as a “black box,”\textsuperscript{15} the selected variables for each derived specification are assumed to be the best combination possible given the data provided. Consequently, which variables were selected is irrelevant for purposes of my testing. In total, sixty models were examined encompassing nine hundred coefficient inclusions (with estimates) or exclusions, as was determined by the process for each. In order to present my results in a more succinct form, the full estimation outputs are not reported.

Recall from the introductory section that the intention of this study is not to develop a better forecasting model, and that the particular questions posed by this study are unique within the literature. Therefore, the results presented in this section are not compared with previous research.

Section 5.1 presents the conclusions reached by comparing the specification results derived using the band-pass business-cycle expression of the explanatory data to those derived using the full-spectrum expression. Section 5.2 presents the results of the

\textsuperscript{15} Here the term “black box” refers to a process which can be viewed solely in terms of its input and output without any knowledge of its internal workings.
comparison of specification results derived using the band-pass business-cycle expression to those derived using the band-pass business-cycle in union with residual series expression of the data. Section 5.3 presents the results of robustness testing conducted and Section 5.4 presents an encompassing test for particular groups of rival specifications.

5.1 – The Primary Research Question

The primary research question addresses whether there is value in the consideration of frequency-decomposed time series as explanatory data within regression analyses. My intuition was that the portion of the explanatory data which cycles within the same frequency band as the U.S. macro economy should, ceteris paribus, describe a particular characteristic of that economy (i.e. recessions) better than the full spectrum of the same data.

5.1.1 – Primary Research Question – Results

Recall from Section 4.3.1 that the primary research question in its testable form compares the performance criteria statistics derived from the best-case specifications of the full-spectrum (FS) and the band-pass filter derived business-cycle component (BPBC) expressions of the explanatory data. Also recall from Section 3.2 that the four differing dependent variables are relative to different values of $h$, the cumulative forecast horizon variable component of the empirical model used.

Table 1 reports results for the full data sample period, January 1959 through June 2009. It uses the following mnemonics. The dependent variable name mnemonics follow the form R_WIN_nn, in which R stands for ‘recession,’ WIN stands for ‘within,’ and nn
stands for the respective value for \( h \), either 3, 6, 9, or 12. The mnemonic R_WIN_03 in the first row of Table 1 indicates that the dependent variable was determined relative to observations in which NBER defined recession conditions occurred within the next three observations. The other three dependent variables follow the same format.

The explanatory expression mnemonics FS and BPBC distinguish between the two rival expressions tested by the primary research question. Again, FS stands for the full-spectrum expression of the explanatory data. BPBC stands for the band-pass filter derived business-cycle component expression of the explanatory data.

### Table 1 – Specification Scoring – Primary Research Question – Full Data Sample

<table>
<thead>
<tr>
<th>Dep Var</th>
<th>Expression</th>
<th>AIC /n (in-sample)</th>
<th>McFadden R-sq (in-sample)</th>
<th>% Correct (in-sample)</th>
<th>% Correct (out-of-sample)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R_WIN_03</td>
<td>FS</td>
<td>0.668</td>
<td>0.310</td>
<td>85.50%</td>
<td>8.33%</td>
</tr>
<tr>
<td></td>
<td>BPBC</td>
<td>0.282</td>
<td>0.724</td>
<td>94.44%</td>
<td>33.33%</td>
</tr>
<tr>
<td>R_WIN_06</td>
<td>FS</td>
<td>0.678</td>
<td>0.375</td>
<td>85.67%</td>
<td>16.67%</td>
</tr>
<tr>
<td></td>
<td>BPBC</td>
<td>0.351</td>
<td>0.686</td>
<td>93.43%</td>
<td>33.33%</td>
</tr>
<tr>
<td>R_WIN_09</td>
<td>FS</td>
<td>0.675</td>
<td>0.431</td>
<td>84.49%</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>BPBC</td>
<td>0.457</td>
<td>0.627</td>
<td>90.40%</td>
<td>25.00%</td>
</tr>
<tr>
<td>R_WIN_12</td>
<td>FS</td>
<td>0.760</td>
<td>0.394</td>
<td>82.80%</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>BPBC</td>
<td>0.568</td>
<td>0.560</td>
<td>86.70%</td>
<td>33.33%</td>
</tr>
</tbody>
</table>

The four performance criteria statistics from each relevant best-case specification model estimation are reported. Each row of statistics is reported relative to its respective dependent variable/explanatory expression combination. Boxes are placed around the best statistic from each rival expression comparison. For example, relative to the
R_WIN_03 dependent variable for the AIC/n performance criteria, the BPBC expression of the explanatory data scored better than the FS expression, 0.282 to 0.668.¹⁶

Notice that for all comparisons, across all criteria and for all four dependent variables, the BPBC expression of the explanatory data performed better than the FS expression.

Table 2 reports results for the January 1959 through July 2002 subsample. The same mnemonic conventions apply. In cases where there is a tie between the compared performance statistics a box is placed around both reported values.

<table>
<thead>
<tr>
<th>Dep Var</th>
<th>Expression</th>
<th>AIC /n (in-sample)</th>
<th>McFadden R-sq (in-sample)</th>
<th>% Correct (in-sample)</th>
<th>% Correct (out-of-sample)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R_WIN_03</td>
<td>FS</td>
<td>0.664</td>
<td>0.321</td>
<td>85.29%</td>
<td>75.00%</td>
</tr>
<tr>
<td></td>
<td>BPBC</td>
<td>0.294</td>
<td>0.727</td>
<td>94.52%</td>
<td>75.00%</td>
</tr>
<tr>
<td>R_WIN_06</td>
<td>FS</td>
<td>0.678</td>
<td>0.381</td>
<td>85.49%</td>
<td>75.00%</td>
</tr>
<tr>
<td></td>
<td>BPBC</td>
<td>0.356</td>
<td>0.698</td>
<td>93.54%</td>
<td>75.00%</td>
</tr>
<tr>
<td>R_WIN_09</td>
<td>FS</td>
<td>0.662</td>
<td>0.444</td>
<td>85.49%</td>
<td>75.00%</td>
</tr>
<tr>
<td></td>
<td>BPBC</td>
<td>0.468</td>
<td>0.632</td>
<td>91.78%</td>
<td>75.00%</td>
</tr>
<tr>
<td>R_WIN_12</td>
<td>FS</td>
<td>0.763</td>
<td>0.397</td>
<td>83.53%</td>
<td>75.00%</td>
</tr>
<tr>
<td></td>
<td>BPBC</td>
<td>0.582</td>
<td>0.559</td>
<td>87.48%</td>
<td>75.00%</td>
</tr>
</tbody>
</table>

The results reported in Table 2 are similar to those reported for the full data sample. For all in-sample test criteria, and for all four dependent variables, the BPBC expression of the explanatory data performed better than the FS expression. For the percentage correct predictions out-of-sample criteria both expressions performed equally well.

¹⁶ Recall from Section 4.2.1 that lower values of the AIC/n statistic indicate better performance.
Table 3 – Specification Scoring – Primary Research Question – 1959M01:1991M11

<table>
<thead>
<tr>
<th>Dep Var</th>
<th>Expression</th>
<th>AIC / n (in-sample)</th>
<th>McFadden R-sq (in-sample)</th>
<th>% Correct (in-sample)</th>
<th>% Correct (out-of-sample)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R_WIN_03</td>
<td>FS</td>
<td>0.762</td>
<td>0.298</td>
<td>83.77%</td>
<td>50.00%</td>
</tr>
<tr>
<td></td>
<td>BPBC</td>
<td>0.305</td>
<td>0.757</td>
<td>95.04%</td>
<td>91.67%</td>
</tr>
<tr>
<td>R_WIN_06</td>
<td>FS</td>
<td>0.835</td>
<td>0.301</td>
<td>82.20%</td>
<td>50.00%</td>
</tr>
<tr>
<td></td>
<td>BPBC</td>
<td>0.350</td>
<td>0.741</td>
<td>93.99%</td>
<td>83.33%</td>
</tr>
<tr>
<td>R_WIN_09</td>
<td>FS</td>
<td>0.898</td>
<td>0.299</td>
<td>81.15%</td>
<td>50.00%</td>
</tr>
<tr>
<td></td>
<td>BPBC</td>
<td>0.478</td>
<td>0.642</td>
<td>89.82%</td>
<td>91.67%</td>
</tr>
<tr>
<td>R_WIN_12</td>
<td>FS</td>
<td>1.063</td>
<td>0.200</td>
<td>75.92%</td>
<td>50.00%</td>
</tr>
<tr>
<td></td>
<td>BPBC</td>
<td>0.583</td>
<td>0.572</td>
<td>89.30%</td>
<td>100.00%</td>
</tr>
</tbody>
</table>

The results reported in Table 3 are for the subsample period January 1959 through November 1991. The results here nearly mirror those reported for the full data sample. For all comparisons, across all criteria, and for all four dependent variables, the BPBC expression of the explanatory data performed better than the FS expression.

Table 4 summarizes the results for the primary research question. The reported values represent the number of times the indicated expression outscored its rival across the three sample periods tested. In the cases of a tie the particular comparison is disregarded. The report is broken down by dependent variable with totals across all dependent variables reported in the bottom row.
Table 4 – Specification Scoring – Primary Research Question – All Sample Periods

<table>
<thead>
<tr>
<th>Dep Var</th>
<th>Expression</th>
<th>AIC /n (in-sample)</th>
<th>McFadden R-sq (in-sample)</th>
<th>% Correct (in-sample)</th>
<th>% Correct (out-of-sample)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R_WIN_03</td>
<td>FS</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>BPBC</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>R_WIN_06</td>
<td>FS</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>BPBC</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>R_WIN_09</td>
<td>FS</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>BPBC</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>R_WIN_12</td>
<td>FS</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>BPBC</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>TOTAL</td>
<td>FS</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>BPBC</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>8</td>
</tr>
</tbody>
</table>

For all three data samples, for each expression comparison, across all in-sample criteria, and for all four dependent variables, the BPBC expression of the explanatory data performed better than the FS expression twelve out of twelve times. For the percentage correct predictions out-of-sample criteria, the BPBC expression out-performed the FS expression eight out of twelve times. In no cases did the FS expression out-perform the BPBC expression.

5.1.2 – Primary Research Question – Conclusions Drawn

I conclude by the results presented above that targeting an appropriate frequency band within the explanatory data can, ceteris paribus, add value to a regression modeling process.
5.2 – *The Secondary Question*

The secondary question considers potential losses in consequence of the particular decomposition scheme undertaken. Due diligence motivated an assessment of the information set aside. Does it contain value? Is it worth keeping?

5.2.1 – *Secondary Question – Results*

Recall from Section 4.3.2 that the secondary question in its testable form compares the performance criteria statistics derived of the best-case specifications from the business-cycle component expression and the band-pass business-cycle in union with residual component expressions of the explanatory data. The report format is the same as for the primary research question. The mnemonics used for the relevant expressions of the explanatory data are BPBC for the band-pass filtered business-cycle component expression and BPBCR for the band-pass filtered business-cycle in union with residual component expression.

As with the primary research question, the results are broken down by sample period. Table 5 presents the results for the full sample period, Table 6 presents the results for the January 1959 through July 2002 sample period, and Table 7 presents the results for the January 1959 through November 1991 sample period. Table 8 summarizes the results for the secondary question using the same convention as Table 4 in Section 5.1.1. Discussion of these results follows Table 8.
Table 5 – Specification Scoring – Secondary Question – Full Data Sample

<table>
<thead>
<tr>
<th>Dep Var</th>
<th>Expression</th>
<th>AIC /n (in-sample)</th>
<th>McFadden R-sq (in-sample)</th>
<th>% Correct (in-sample)</th>
<th>% Correct (out-of-sample)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R_WIN_03</td>
<td>BPBC</td>
<td>0.282</td>
<td>0.724</td>
<td>94.44%</td>
<td>33.33%</td>
</tr>
<tr>
<td></td>
<td>BPBCR</td>
<td>0.282</td>
<td>0.725</td>
<td>94.61%</td>
<td>41.67%</td>
</tr>
<tr>
<td>R_WIN_06</td>
<td>BPBC</td>
<td>0.351</td>
<td>0.686</td>
<td>93.43%</td>
<td>33.33%</td>
</tr>
<tr>
<td></td>
<td>BPBCR</td>
<td>0.298</td>
<td>0.752</td>
<td>94.44%</td>
<td>25.00%</td>
</tr>
<tr>
<td>R_WIN_09</td>
<td>BPBC</td>
<td>0.457</td>
<td>0.627</td>
<td>90.40%</td>
<td>25.00%</td>
</tr>
<tr>
<td></td>
<td>BPBCR</td>
<td>0.389</td>
<td>0.693</td>
<td>92.92%</td>
<td>8.33%</td>
</tr>
<tr>
<td>R_WIN_12</td>
<td>BPBC</td>
<td>0.568</td>
<td>0.560</td>
<td>86.70%</td>
<td>33.33%</td>
</tr>
<tr>
<td></td>
<td>BPBCR</td>
<td>0.494</td>
<td>0.624</td>
<td>92.06%</td>
<td>8.33%</td>
</tr>
</tbody>
</table>

Table 6 – Specification Scoring – Secondary Question – 1959M01:2002M07

<table>
<thead>
<tr>
<th>Dep Var</th>
<th>Expression</th>
<th>AIC /n (in-sample)</th>
<th>McFadden R-sq (in-sample)</th>
<th>% Correct (in-sample)</th>
<th>% Correct (out-of-sample)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R_WIN_03</td>
<td>BPBC</td>
<td>0.294</td>
<td>0.727</td>
<td>94.52%</td>
<td>75.00%</td>
</tr>
<tr>
<td></td>
<td>BPBCR</td>
<td>0.261</td>
<td>0.770</td>
<td>94.51%</td>
<td>75.00%</td>
</tr>
<tr>
<td>R_WIN_06</td>
<td>BPBC</td>
<td>0.356</td>
<td>0.698</td>
<td>93.54%</td>
<td>75.00%</td>
</tr>
<tr>
<td></td>
<td>BPBCR</td>
<td>0.275</td>
<td>0.785</td>
<td>94.90%</td>
<td>75.00%</td>
</tr>
<tr>
<td>R_WIN_09</td>
<td>BPBC</td>
<td>0.468</td>
<td>0.632</td>
<td>91.78%</td>
<td>75.00%</td>
</tr>
<tr>
<td></td>
<td>BPBCR</td>
<td>0.376</td>
<td>0.716</td>
<td>93.53%</td>
<td>75.00%</td>
</tr>
<tr>
<td>R_WIN_12</td>
<td>BPBC</td>
<td>0.582</td>
<td>0.559</td>
<td>87.48%</td>
<td>75.00%</td>
</tr>
<tr>
<td></td>
<td>BPBCR</td>
<td>0.474</td>
<td>0.643</td>
<td>93.35%</td>
<td>75.00%</td>
</tr>
</tbody>
</table>
### Table 7 – Specification Scoring – Secondary Question – 1959M01:1991M11

<table>
<thead>
<tr>
<th>Dep Var</th>
<th>Expression</th>
<th>AIC /(n) (in-sample)</th>
<th>McFadden R-sq (in-sample)</th>
<th>% Correct (in-sample)</th>
<th>% Correct (out-of-sample)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R_WIN_03</td>
<td>BPBC</td>
<td>0.305</td>
<td>0.757</td>
<td>95.04%</td>
<td>91.67%</td>
</tr>
<tr>
<td></td>
<td>BPBCR</td>
<td>0.262</td>
<td>0.799</td>
<td>95.55%</td>
<td>83.33%</td>
</tr>
<tr>
<td>R_WIN_06</td>
<td>BPBC</td>
<td>0.350</td>
<td>0.741</td>
<td>93.99%</td>
<td>83.33%</td>
</tr>
<tr>
<td></td>
<td>BPBCR</td>
<td>0.248</td>
<td>0.848</td>
<td>98.17%</td>
<td>100.00%</td>
</tr>
<tr>
<td>R_WIN_09</td>
<td>BPBC</td>
<td>0.478</td>
<td>0.642</td>
<td>89.82%</td>
<td>91.67%</td>
</tr>
<tr>
<td></td>
<td>BPBCR</td>
<td>0.366</td>
<td>0.755</td>
<td>94.76%</td>
<td>91.67%</td>
</tr>
<tr>
<td>R_WIN_12</td>
<td>BPBC</td>
<td>0.583</td>
<td>0.572</td>
<td>89.30%</td>
<td>100.00%</td>
</tr>
<tr>
<td></td>
<td>BPBCR</td>
<td>0.544</td>
<td>0.602</td>
<td>91.12%</td>
<td>75.00%</td>
</tr>
</tbody>
</table>

### Table 8 – Specification Scoring – Secondary Question – All Sample Periods

<table>
<thead>
<tr>
<th>Dep Var</th>
<th>Expression</th>
<th>AIC /(n) (in-sample)</th>
<th>McFadden R-sq (in-sample)</th>
<th>% Correct (in-sample)</th>
<th>% Correct (out-of-sample)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R_WIN_03</td>
<td>BPBC</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>BPBCR</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>R_WIN_06</td>
<td>BPBC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>BPBCR</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>R_WIN_09</td>
<td>BPBC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>BPBCR</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>R_WIN_12</td>
<td>BPBC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>BPBCR</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>TOTAL</td>
<td>BPBC</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>BPBCR</td>
<td>11</td>
<td>12</td>
<td>11</td>
<td>2</td>
</tr>
</tbody>
</table>

As with the primary question, for all three data samples, for each expression comparison, across all in-sample criteria, and for all four dependent variables, one expression of the explanatory data performed better than the other. Here, the BPBCR
expression clearly performed better, on average, than its rival, the BPBC expression. The out-of-sample test, however, tells a different story. Here the BPBC expression tended to perform better.

5.2.2 – Secondary Question – Conclusions Drawn

Though these results are less one-sided than those of the primary question, I still conclude that there is value within the residual series of the band-pass filter decomposition scheme used in this study. I further conclude that this is evidence in support of a need to explore and test other decomposition schemes. Recall that the structure of the residual series had only to do with it being the remainder from the process of isolating the BC component. It is reasonable to suppose that other decompositions of this information could result in even more beneficial outcomes.

Perhaps an argument could be made that inclusion of the residual components into the modeling process is to be avoided because, on average, it detracts from out-of-sample model performance. However, given the in-sample leverage aspect of the cumulative forecast horizon model mentioned in Section 2.2.1, and the strong in-sample performance of the BPBCR expression reported here, I believe that my conclusion holds.

5.3 – A Test for Robustness – The High-Pass Filter

To test the robustness of my findings, I applied a high-pass filter to the full spectrum of the baseline explanatory data set. I then conducted the same analyses that I had conducted against the band-pass filtered data against the high-pass filtered data.
5.3.1 – An Alternative Decomposition Scheme

The high-pass filter, like the band-pass filter, is a business-cycle filter. However, its decomposition scheme is fixed and does not generate a series directly analogous to the business-cycle component of the data already discussed.

The high-pass filter generates two series. These series are referred to in the literature as representing an underlying economic trend and the associated higher frequency cyclical behaviors of the data.\textsuperscript{17} The sum of these two series is equal to the full spectrum.

For the purposes of this robustness test, the underlying economic trend component was cast into the role previously held by the band-pass filter business-cycle component and the higher frequency cyclical series was cast into the role previously held by the band-pass residual series component. This testing has more to do with consideration of whether allowing the modeling process the opportunity to consider the higher frequency portion of the full spectrum separately from the lower frequency portion can add value.

A summary of the results relative to this recasting of the primary research question is presented in Table 9. Table 10 summarizes results relative to the recast secondary question. The reporting convention utilized is the same for both as for Table 4 in Section 5.1.1 and Table 8 in Section 5.2.1 respectively. The mnemonics used for the relevant expressions of the explanatory data are FS for the full-spectrum expression, HPT for the high-pass filtered trend component expression, and HPTR for the high-pass filtered trend in union with the residual series component expression.

\textsuperscript{17} See Christiano and Fitzgerald (2003).
The results shown in Table 9 fairly closely shadow those for the band-pass filter decomposed data. For all three data samples, for each expression comparison, across all in-sample criteria, and for all four dependent variables, the HPT expression of the explanatory data performed better than the FS expression. The out-of-sample test report too, though not quite as strong in favor of the HPT expression here as it was for the BPBC expression in Table 4, does support a claim of better performance on average.
<table>
<thead>
<tr>
<th>Dep Var</th>
<th>Expression</th>
<th>AIC /n (in-sample)</th>
<th>McFadden R-sq (in-sample)</th>
<th>% Correct (in-sample)</th>
<th>% Correct (out-of-sample)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R_WIN_03</td>
<td>HPT</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>HPTR</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>R_WIN_06</td>
<td>HPT</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>HPTR</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>R_WIN_09</td>
<td>HPT</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>HPTR</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>R_WIN_12</td>
<td>HPT</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>HPTR</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>TOTAL</td>
<td>HPT</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>HPTR</td>
<td>11</td>
<td>12</td>
<td>12</td>
<td>0</td>
</tr>
</tbody>
</table>

The results shown in Table 10 also shadow those for the band-pass filter decomposed data reported in Table 8. For all three data samples, for each expression comparison, across all in-sample criteria, and for all four dependent variables the HPTR expression of the explanatory data clearly performed better than the HPT expression. The out-of-sample test results are also similar to those reported in Table 8, but here the HPT expression performed better without exception.

5.3.2 – A Brief Summary of Results and Conclusions Drawn

I conclude that for both recast questions previously posed relative to the band-pass filtered data the same conclusions hold for the high-pass filtered data. And, though it is beyond the scope of this study, I further conclude that these results do stand as further evidence that there is inherent value in the decomposition process itself.
5.4 – *An Encompassing Test of Rival Specifications*

In light of the results reported above, it is appropriate and useful to include here an encompassing test for particular sets of rival specifications. This test provides insight into the relative information content characteristics of the specifications tested. Though the test results already reported strongly suggest that explanatory information quantity is on average greater within the specifications featuring the frequency decomposed components, an encompassing test can speak to the relative information content across the two rivals. Relative to the original testing of my intuition, this test was conducted for rival specifications derived of the FS and BPBC expressions respectively.

5.4.1 – *What it Means When a Specification Encompasses its Rival*

When a specification encompasses its rival, it contains the explanatory information of its rival specification in a more parsimonious form. According to Campos et al. (2005), “an encompassing test evaluates a given model against the information content [contained within] an alternative specification […]. [It] assesses whether [that] model is able to explain why the alternative explanation obtains the results that it does” (p. 57).

5.4.2 – *The J-Test: Its Empirical Form and Process*

For non-nested, non-linear specifications, the J-Test is an appropriate encompassing test (Chen, 2000). The J-Test is implemented by first estimating fitted values for each rival specification. Next, the fitted series from each specification is added to its rival specification as an additional explanatory variable. Both models are then re-estimated and the coefficient upon the fitted series variable in each model is tested for
statistical significance. Given two possible significance test outcomes for each of the two estimated coefficients, there are four possible interpretations to the J-Test result.

5.4.3 – The J-Test: Its Four Possible Outcomes and Their Interpretations

The four possible outcomes of the J-Test are determined based upon tests of the null hypothesis that the information content of a rival specification cannot add explanatory value to the specification under consideration. The first two possible outcomes are inversely related. Across the two augmented specification estimations described in Section 5.4.2, if the coefficient on the first fitted series is determined to be statistically significant and the coefficient on the second is not, then the original specification, the one which generated the fitted series that is statistically significant, is held to encompass its rival specification. The second possible outcome is simply the inverse result: the first coefficient isn’t statistically significant but the second coefficient is. In this case the second specification encompasses the first.

The third possible outcome occurs when the coefficients on both fitted series are determined to be statistically significant. In this case a better re-specification of the non-redundant union of the two rival specifications is likely possible. In other words, both rival specifications are likely misspecifications of the assumed best-case specification.

The fourth possible outcome occurs when the coefficients on both fitted series are determined to be statistically insignificant. In this unique case it is held that insufficient data is available to make a valid test interpretation.
5.4.4 – The J-Test: Results and Conclusions Drawn

The results of the encompassing tests conducted for two particular groups of rival specifications are shown in the table below. The two groups are the best-case specifications derived for each of the four dependent variables as a function of the full-spectrum (FS) versus band-pass business-cycle (BPBC) expressions of the explanatory data, and the same FS specifications versus the band-pass business-cycle in union with residuals expression (BPBCR). The former group was selected because it relates directly to my primary research question. The latter group test builds upon conclusions reached out of my secondary question.

<table>
<thead>
<tr>
<th></th>
<th>FS encompasses BPBC</th>
<th>0</th>
<th>FS encompasses BPBCR</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>BPBC encompasses FS</td>
<td>3</td>
<td></td>
<td>BPBCR encompasses FS</td>
<td>6</td>
</tr>
<tr>
<td>Misspecification</td>
<td>9</td>
<td></td>
<td>Missspecification</td>
<td>6</td>
</tr>
<tr>
<td>Insufficient data</td>
<td>0</td>
<td></td>
<td>Insufficient data</td>
<td>0</td>
</tr>
</tbody>
</table>

The left-hand side results in Table 11 are not surprising. In short, they reveal that for the comparisons across all three data samples, for the four different dependent variables, in most cases each rival specification is a misspecification of the non-redundant union of both rival specifications. This may make some sense given that the General-to-Specific algorithm selected these specifications from mutually exclusive explanatory data pools. Or, another interpretation could be that the net information content difference between the two explanatory data pools (i.e., the SR and LR
components of the data), does play some important role. Still, in three of the twelve comparisons the BC component expression encompasses the FS expression.

The right-hand side results in Table 11 are more interesting. Though it is true that the same condition of mutually exclusive baseline data pools exists, it is also true that, in theory, the BPBCR expression of the explanatory data is the decomposed equivalent of the full-spectrum expression. In other words, there is no information content difference between the two expressions. The only difference is in the decomposition scheme applied. Here, in half the test cases the decomposed data expression derived specification encompasses the full spectrum of the data derived specification.

Finally, note that across both test groups in no case did the full-spectrum expression derived best-case specification encompass its rival expression derived specification. Though its rival didn’t always outperform, the full-spectrum derived specification never outperformed.

The results presented above are both reasonable and encouraging. In particular, I am encouraged by the results relative to the full spectrum versus business cycle in union with residual component specification comparisons. I believe that these particular results may be the strongest support uncovered within this study for the potential benefits to explanatory information recovery from economic time series frequency band decomposition.
Chapter 6

CONCLUSIONS

In this study I presented a test of my intuition that targeting the business-cycle frequency band of explanatory data can, ceteris paribus, increase U.S. recession probability forecast model performance. I have clearly demonstrated that it can. This study is unique within the literature, and as such I believe it to be a potential starting point for new research topics going forward.

6.1 – My Thesis and The Primary Research Question

Because the business-cycle component of an economic time series is in theory the portion of that series which cycles within the same frequency band as the greater macro economy, I believed that using that component of the explanatory data in place of the full-spectrum series should result in superior model performance statistics. I tested this belief with my primary research question.

6.1.1 – A Summary of Results and Conclusions Drawn

The results of the testing related to my primary research question establish that, ceteris paribus, isolating the business-cycle component of the explanatory data on average results in a better recession probability forecasting model. For in-sample testing, in all cases, the business-cycle component expression of the data outperformed the full-spectrum series expression. For out-of-sample testing, in eight of the twelve comparisons the business-cycle component expression outperformed the full-spectrum analysis. In no test cases, in-sample or out-of-sample, did the full-spectrum expression outperform the business-cycle component expression.
I conclude that further research into the consideration of frequency decomposed time series as explanatory data within regression analyses is worthwhile.

6.2 – The Secondary Question

Because the decomposition of the data associated with the primary question involved the elimination of a significant portion of its potential information content, it reasonably followed that consideration for potential information loss should be given. I therefore tested for this condition with my secondary question.

6.2.1 – A Summary of Results and Conclusions Drawn

The results of the testing related to my secondary question support my conclusion that there is explanatory value contained within the particular decomposition process residual series tested here. For in-sample testing, in thirty-four of the thirty-six comparison cases, the business cycle in union with residual series expression of the data outperformed the business-cycle component expression. The inverse occurred in only one in-sample testing case. For out-of-sample testing the business-cycle component expression slightly outperformed the business cycle in union with residual series expression. In five of the twelve comparisons it scored better. The business-cycle in union with residual component expression outperformed its rival in only two comparisons.

By including the residual series for each decomposed leading indicator within the available explanatory data pool, the modeling process had an opportunity to include any or all individual residual series within its selected best-case specification. Also, the opportunity to assign differing marginal effects to the residual series selected became
available to the process. By the results reported here, I believe that this enhanced flexibility has added value to the overall modeling process.

By these results I have concluded that there is a need to explore and test other frequency band decomposition schemes.

6.3 – The Benefit Realized from Explanatory Data Decomposition

The results of this study clearly demonstrate that a potential role for frequency decomposed explanatory data in regression modeling cannot be summarily dismissed. The potential for the development of new models to describe economic reality within theoretical contexts already understood is immense. The potential roles for new processes which allow the data to better inform theories we are yet to discover is fascinating.

6.4 – Extensions of This Research

It is my hope that this study can be a starting point for future research in the areas of time series frequency decomposition, business-cycle event probability forecasting, and the development of generalized systems which are better equipped to deal with the ever-changing nature of economic structures.

The processes and methods used within this study can be extended into systems which will allow for flexible and powerful real time forecasting systems.

In particular, this study has demonstrated a new combination of existing technologies and methods. When used together as demonstrated here, these technologies and methods can be leveraged into a system that should do a better job in foreseeing the next catastrophic business-cycle event.
REFERENCES


