GOVERNMENT BUDGET PREDICTIONS WITH MIXED FREQUENCY ANALYSIS

Nazire Ozkan

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Economics.

Chapel Hill
2014

Approved by:

Eric Ghysels

Michael Aguilar

Saraswata Chaudhuri

Jonathan B. Hill

Neville Francis

ABSTRACT

Nazire Ozkan: Government Budget Predictions with Mixed Frequency Analysis (Under the direction of Eric Ghysels)

Based on the growing literature of Mixed Data Sampling (MIDAS) analysis, this dissertation proposes forecasting procedures for the U.S. federal and state government budgets and output growth. Mixed frequency analysis elucidates the information content of data sampled at different frequencies and, hence, enables more accurate forecasts than the conventional approach that aggregates all time series into the lowest common frequency.

This dissertation consists of three essays, each of which is examined in a separate chapter. The first chapter proposes a real-time forecasting procedure involving a combination of MIDAS-type regression models constructed with predictors of different sampling frequencies to predict the annual U.S. federal government current expenditures and receipts. Evidence shows that forecast combinations of MIDAS regression models provide forecast gains over the traditional models, suggesting the use of mixed frequency data consisting of fiscal series and macroeconomic indicators in forecasting the annual federal budget. It is also shown that, although not statistically significant, MIDAS regressions with quarterly leads that are employed to have real-time forecast updates of the current year federal expenditures and receipts are found to have improved forecast performance compared to MIDAS regressions without leads.

Using a sample of 48 mainland U.S. states, the second chapter considers the problem of forecasting state and local governments' expenditures and revenues. It first proposes a forecasting procedure that involves a simple mixed frequency data regression approach, namely combinations of Augmented Distributed Lag–Mixed Data Sampling (ADL–MIDAS) regression models. With this approach, for almost all states, it is found that the use of high frequency state-specific and national variables combined with a low frequency budget series provides forecast performance gains over the traditional models where all data are of the same low/annual sampling frequency. This chapter then proposes a procedure with a multiple equation regression

model, specifically a Mixed Frequency–Bayesian Vector Autoregressive (MF–BVAR) model. The predictive ability of the proposed model is assessed against the forecast performance of a traditional, low frequency Bayesian Vector Autoregressive (BVAR) model. Although the forecast performance varies at the state level, the overall empirical forecast performance of the MF–BVAR is better than that of the traditional BVAR model. Finally, predictive abilities of the two proposed forecasting procedures are empirically examined and the results suggest that one cannot be chosen over the other. While the ADL–MIDAS model provides better forecasts for expenditure series across states, forecasts for revenue series are more accurately obtained via the MF–BVAR model.

The third chapter proposes a method for producing current-quarter forecasts of the U.S. real Gross Domestic Income (GDI) growth with a range of available within the quarter monthly/weekly/daily observations of macroeconomic and financial indicators, such as employment, industrial production, and stock prices. The real-time forecasting procedure involves a combination of MIDAS-type regression models constructed with predictors of different sampling frequencies. Evidence shows that forecast combinations of MIDAS regression models with monthly leads that are employed to have real-time forecast updates of the current-quarter GDI growth provide forecast gains over the traditional models, suggesting the use of readily available within-quarter data in forecasting current-quarter output growth.

ACKNOWLEDGEMENTS

First and foremost, I would like to sincerely thank my dissertation advisor, Eric Ghysels, for his great direction and patience throughout my Ph.D. research. I would also like to express gratitude to my dissertation committee members for their invaluable guidance. The completion of this dissertation would not have been possible without their support.

I would like to thank my family– my parents, and my sister and brother for their love and support throughout my life. I would also like to thank my niece, Elif Ada, for never failing to lift my spirits.

TABLE OF CONTENTS

LIST OF TABLES . . . . **viii**

LIST OF FIGURES . . . . **ix**

1 Real-Time Forecasting of the U.S. Federal Government Budget **1**

1.1 Introduction 1

1.2 Methods 3

1.2.1 ADL–MIDAS Regression Model 4

1.2.2 Forecast Combinations Method 6

1.2.3 A Test of Predictive Accuracy 8

1.3 Data 9

1.4 Empirical Results 10

1.4.1 MIDAS Regressions vs. AR and ADL Regressions 11

1.4.2 MIDAS Regressions with Leads 11

1.5 Conclusion 13

2 Real-Time Forecasting of the U.S. State and Local Governments' Budgets **23**

2.1 Introduction 23

2.2 Models 27

2.2.1 Single Equation Regression Model 28

2.2.2 Multiple Equation Regression Model 31

2.3 Data 36

2.3.1 Single–Equation Regressions 36

2.3.2 Vector Autoregressions 37
2.4 Empirical Results 37
2.4.1 Single–Equation Regression Results 37
2.4.2 Multiple–Equation Regression Results 39
2.5 Conclusion 43
3 Real-Time Forecasting of the Current-Quarter U.S. GDI Growth 72
3.1 Introduction 72
3.2 Methods 74
3.2.1 ADL–MIDAS with Leads 75
3.2.2 Forecast Combinations Method 77
3.3 Data 78
3.4 Empirical Results 79
3.4.1 MIDAS Regressions versus AR and ADL Regressions 80
3.4.2 MIDAS Regressions with Leads 80
3.5 Conclusion 81
BIBLIOGRAPHY 92

LIST OF TABLES

1.1 Data Sets 14

1.2 Comparison of RMSFEs: AR and ADL Models vs. MIDAS Regressions. 15

1.3 Time Series Test for Predictive Ability. 16

1.4 Comparison of Forecast Performances: ADL-MIDAS with and without Leads . . 17

1.5 Testing Equal Forecasting Accuracy between MIDAS with and without Leads . 18

2.1 Data Sets 44

2.2 Forecasts of State Government Expenditures and Revenues Using ADL-MIDAS, ADL and AR models. (Relative RMSFEs). 45

2.3 Forecasts of State Government Expenditures and Revenues Using MF-BVAR and BVAR models. (Relative RMSFEs). 47

2.4 Forecasts of State Government Expenditures and Revenues Using MF–BVAR and ADL–MIDAS models. (Relative RMSFEs). 49

2.5 Panel DM Statistics for MF–BVAR Forecasts of State Government Expenditures and Revenues Relative to the BVAR model. 51

3.1 Data Set 82

3.2 Comparison of RMSFEs: MIDAS Regressions vs. AR and ADL Models. 83

3.3 Time Series Test for Predictive Ability. 84

LIST OF FIGURES

1.1 Boxplots for Forecast Performance Comparisons between MIDAS and AR regression models . . . 19

1.2 Forecast Combination Weights over the Forecasting Period for Receipts . . . 20

1.3 Forecast Combination Weights over the Forecasting Period for Expenditures . . . 21

1.4 Predictions from Forecast Combinations of MIDAS regressions with and without Leads . . . 22

2.1 Forecast Combination Weights on Expenditures . . . 52

2.1 Forecast Combination Weights on Expenditures – continued figure . . . 53

2.2 Forecast Combination Weights on Revenues . . . 54

2.2 Forecast Combination Weights on Revenues – continued figure . . . 55

2.3 Impulse Responses of Revenues to Monetary Shocks . . . 56

2.3 Impulse Responses of Revenues to Monetary Shocks – continued figure . . . 57

2.4 Impulse Responses of Expenditures to Monetary Shocks . . . 58

2.4 Impulse Responses of Expenditures to Monetary Shocks – continued figure . . . 59

2.5 Impulse Responses of Revenues to Fiscal Shocks . . . 60

2.5 Impulse Responses of Revenues to Fiscal Shocks – continued figure . . . 61

2.6 Impulse Responses of Expenditures to Fiscal Shocks . . . 62

2.6 Impulse Responses of Expenditures to Fiscal Shocks – continued figure . . . 63

2.7 Impulse Responses of Revenues to Monetary Shocks: MF-BVAR versus BVAR . 64

2.7 Impulse Responses of Revenues to Monetary Shocks: MF-BVAR versus BVAR – continued figure . . . 65

2.8 Impulse Responses of Expenditures to Monetary Shocks:MF-BVAR versus BVAR 66

2.8 Impulse Responses of Expenditures to Monetary Shocks: MF-BVAR versus BVAR – continued figure . . . 67

2.9 Impulse Responses of Revenues to Fiscal Shocks: MF-BVAR versus BVAR . . . 68

2.9 Impulse Responses of Revenues to Fiscal Shocks: MF-BVAR versus BVAR – continued figure 69

2.10 Impulse Responses of Expenditures to Fiscal Shocks: MF-BVAR versus BVAR . . 70

2.10 Impulse Responses of Expenditures to Fiscal Shocks: MF-BVAR versus BVAR – continued figure 71

3.1 Quarterly Real GDI and GDP Growth Over the Period of 1986Q1-2013Q3. . . . 85

3.2 Boxplots for Forecast Performance Comparison between MIDAS, ADL and AR regressions 86

3.3 Predictions from Forecast Combinations of MIDAS regressions with Leads, ADL and AR models 87

3.4 Forecast Combination Weights 88

CHAPTER 1

REAL-TIME FORECASTING OF THE U.S. FEDERAL GOVERNMENT BUDGET

1.1 Introduction

The United States has been experiencing profound budgetary challenges and the uncertainty of government budget deficit forecasts has become an important public issue. As stated by the Congressional Budget Office (CBO), a nonpartisan government agency whose main goal is to provide accurate forecasts of the federal budget, the deficit of $1.089 trillion for fiscal year 2012 was the fourth year in a row with a deficit of more than $1 trillion.[1] The federal deficit has decreased sharply from previous shortfalls and the CBO estimates that under current laws it will total $514 billion in fiscal year 2014, which will be $166 billion smaller than the figure posted in 2013.[2]

Predictions for federal expenditures and receipts would differ from the actual outcomes, even if federal laws remained unchanged, due to unanticipated changes in economic conditions and factors that affect federal spending and revenues. That is, fiscal policy is surrounded by uncertainties, both in legislative and economic terms. Forward-looking decision makers should react before policy changes actually occur since fiscal policy changes entail time lags. First, there will be time lags between when the economy is dipping into a recession and when the U.S. government figures out what is happening and necessary actions. Second, there will be legislative lags between when legislation is first proposed and enacted since it takes time to develop a fiscal package on which the majority will agree and pass it through Congress. Finally, there will also be implementation lags between when additional government spending

[1] Congressional Budget Office (2012); An Update to the Budget and Economic Outlook: Fiscal years 2012 to 2022.

[2] Congressional Budget Office (2014); The Budget and Economic Outlook: Fiscal years 2014 to 2024.

is passed, contracts are extended and the spending actually occurs. Since it is not possible in this context to deal with legislative uncertainties, this paper proposes a forecasting procedure that deals with economic uncertainties by employing econometric models incorporating fiscal and macroeconomic indicators sampled at mixed frequencies with annual budget data to predict current federal expenditures and receipts.

Federal government current expenditures and receipts are considered measures of the federal fiscal position. To forecast and monitor annual fiscal position, it is essential to assess the implications of intra-annual fiscal data. Several papers indicated the importance of information contained by intra-annual fiscal data for forecasting and monitoring annual budgetary outcomes. Onorante et al. (2010), for example, use a mixed-frequency state-space model to integrate readily available monthly/quarterly fiscal data with annual government series. Paredes et al. (2009) and Pedregal and Perez (2010) show the usefulness of intra-annual fiscal data for real-time fiscal policy surveillance by estimating models with annual and quarterly national accounts fiscal data. Asimakopoulos, Paredes, and Warmedinger (2013) use quarterly fiscal data to forecast a disaggregated set of fiscal series at annual frequency. Our paper utilizes a set of mixed-frequency macroeconomic indicators, in addition to a set of quarterly fiscal data, to forecast the annual federal budget. The results confirm that quarterly fiscal data together with higher frequency macroeconomic indicators include significant information and they should be taken into account when predicting the annual budget.

In this paper, the proposed forecasting procedure involves a combination of Augmented Distributed Lag-Mixed Data Sampling (ADL-MIDAS) regression models with which the annual U.S. federal government current expenditures and receipts are predicted using a set of mixed frequency variables.[3] Our analysis uses quarterly fiscal variables consisting of subcomponents of current expenditures and receipts, and higher frequency macroeconomic and financial indicators. The objective is to obtain annual forecasts of the federal government current expenditures and receipts via MIDAS-type regressions and compare them with forecasts from more traditional models, namely autoregressive (AR) and augmented distributed lag (ADL) regression

[3]As shown in Bai, Ghysels and Wright (2013), one can view the proposed MIDAS regression approach as a computationally simple way of approximating the state space approach suggested by Onorante et al. (2010), among others.

models. MIDAS-regression methodology relies on a more parsimonious approach to regression analysis with data of different frequencies. Low-frequency time-series data is combined with higher-frequency data without imposing a priori aggregation schemes, with minimal model restrictions. Using this approach, we find that the use of mixed frequency fiscal and macroeconomic data to obtain annual forecasts of federal expenditures and receipts provides significant forecast performance gains over traditional models in which all data are of the same low sampling frequency. Furthermore, to take the advantage of readily available higher frequency data in providing real-time forecast updates of the current year federal expenditures and receipts, MIDAS regression with leads is employed by incorporating real-time information using weekly-monthly-quarterly variables. Recent evidence also shows that ADL-MIDAS regressions with quarterly leads provide forecast improvements over ADL-MIDAS regressions without leads suggesting the use of mixed frequency macroeconomic indicators and quarterly fiscal data as they become available throughout the year to improve the end-year forecast.

The remainder of the paper proceeds as follows. In Section 1.2, the econometric methods employed in this paper are presented within subsections devoted to the descriptions of the ADL-MIDAS regression model, a method to combine forecasts and a test of predictive accuracy. Section 1.3 introduces the data and Section 1.4 presents empirical results with the evidence in favor of forecast combinations of ADL-MIDAS regression models when compared to AR and ADL regression models, then compares performances of one-year-ahead forecasts from ADL-MIDAS models versus ADL-MIDAS with quarterly leads. The conclusions are presented in Section 1.5.

1.2 Methods

In addition to the quarterly subcomponents of federal expenditures and receipts, a set of macroeconomic and financial indicators, with different sampling frequencies, that can be representative of the economy are picked and each is used separately as a predictor for both U.S. federal government current expenditures and receipts. In order to deal with data sampled at different frequencies, Mixed Data Sampling (MIDAS) type regression models, specifically Augmented Distributed LagMixed Data Sampling (ADL-MIDAS) regression models are employed. They include lags of the low frequency dependent variable, here annual expenditures or receipts, and lags of the higher frequency predictor. The analysis is conducted in real-time; at each point in time, the models are estimated using only the data for time periods up to that point in time. In-sample estimations are performed with the vintages of data restricted to those available at that time while forecasts are obtained with the latest values from the most recent vintage.[4] After constructing the longest possible samples with available data, individual forecasts with each single indicator are obtained, resulting in multiple forecasts of expenditures and receipts. Stock and Watson (2001) find that combined forecasts generally outperform forecast performance of the best individual model by employing numerous types of models and variables. Also, as indicated by Timmermann (2006), it is not reasonable to think that the same individual model dominates all the others at all points in time since forecasting models are considered as local approximations. Therefore, to obtain more accurate forecasts by using evidence from all individual models, the forecast combinations method is considered rather than using the best single model.[5]

Subsection 1.2.1 describes the ADL-MIDAS regression models and subsection 1.2.2 shortly introduces the forecast combination method employed in this paper. The test of predictive accuracy is discussed in subsection 1.2.3.

1.2.1 Augmented Distributed Lag - Mixed Data Sampling (ADL-MIDAS) Regression Model

In terms of the improvement of low frequency macroeconomic predictions using high frequency data, the advantages of MIDAS-type regressions have been documented by many recent papers. MIDAS regression models are suggested by Andreou, Ghysels, and Kourtellos (2013), Clements and Galvão (2008), and Ghysels, Santa-Clara, and Valkanov (2006), among others.[6]

---

[4]See Clements and Galvão (2013) for a detailed explanation of real-time vintage data analysis compared to end-of-sample vintage data analysis with autoregressive models.

[5]Timmermann (2006) suggests forecast combinations across different models to have more robust forecasts against misspecification biases and measurement errors in the data underlying the individual forecasts.

[6]See Ghysels, Sinko, and Valkanov (2007) for various descriptions MIDAS regressions. The initial work on MIDAS focused on volatility predictions. See Ghysels, Santa-Clara, and Valkanov (2005), among others.

ADL–MIDAS regression models are employed in this paper to forecast annual federal current expenditures and receipts and their forecast performances are compared with forecasting abilities of augmented distributed lag and autoregressive regression models constructed with variables of the same low frequency.

Let the annual variable of interest to be predicted for h-step ahead horizon be Y^A_{t+h}, say federal government current receipts, and the higher frequency predictor series, for instance, quarterly personal current taxes, be X^Q_t. Then, the $ADL - MIDAS(p^A_Y, q^Q_X)$ regression model can simply be defined as

$$Y^A_{t+h} = c + \sum_{j=0}^{p^A_Y - 1} \alpha_{j+1} Y^A_{t-j} + \beta \sum_{j=0}^{q^Q_X - 1} \sum_{i=0}^{N_Q - 1} \omega_{i+j*N_Q}(\theta^Q) X^Q_{N_{Q-i,t-j}} + u_{t+h} \tag{1.2.1}$$

where N_Q denotes the quarterly lags per year and the weighting scheme, ω^Q, involves a low dimensional vector of unknown parameters.[7] Following Ghysels, Sinko, and Valkanov (2007), for the weighting polynomial, one particular specification based on beta function with two parameters is utilized, which is normalized to add up to one to allow for the identification of the slope coefficient β.[8] Beta function is known to be flexible. It can take many shapes, including flat weights, gradually decreasing and hump-shaped patterns. Normalized beta probability density function with unrestricted (u) and restricted (r) cases and with non-zero (nz) and zero (z) last lag specifications can be written as

[7] MIDAS regressions involve NLS and feasible GLS estimation procedures. Errors are not necessarily i.i.d., error process is a linear process with absolute summable Wold decomposition moving average representations. Identification of coefficients rests on the assumption that the regressor is exogenous up to second moments. With MIDAS regressions, since regressors are sampled at higher frequencies, consistency (absence of discretization bias) can be of concern; Ghysels, Santa-Clara, Valkanov (2006) show that discretization bias is eventually eliminated. That is, aggregation bias disappears as the dependent variable is sampled at a fixed frequency while the regressor is sampled more frequently. They note that MIDAS regressions appear like skip-sampled distributed lag models causing autocorrelated residuals (not preventing OLS/NLS to be consistent). And, to ensure correct specification of MIDAS polynomials, the assumption of $E(u_{t+1}|X_\tau; \tau \leq t) = 0$ is required. See, Ghysels, Santa-Clara, Valkanov (2006) for further information.

[8] For alternative polynomial weight specifications and details, please refer to Ghysels et al. (2007).

$$\omega_i^{u,nz} = \omega_i(\theta_1, \theta_2, \theta_3) = \frac{x_i^{\theta_1 - 1}(1 - x_i)^{\theta_2 - 1}}{\sum_{i=1}^{N} x_i^{\theta_1 - 1}(1 - x_i)^{\theta_2 - 1}} + \theta_3$$

$$\omega_i^{u,z} = \omega_i(\theta_1, \theta_2, 0), \quad \omega_i^{r,nz} = \omega_i(1, \theta_2, \theta_3), \quad \omega_i^{r,z} = \omega_i(1, \theta_2, 0) \ , \quad \text{where } x_i = \tfrac{i-1}{N_D - 1}.$$

Autoregressive, $AR(p_Y)$, and augmented distributed lag, $ADL(p_Y, p_X)$, regression models, employed as competing forecasting models, can be represented, respectively, with the following equations:

$$Y_{t+h} = c + \sum_{j=0}^{p_Y - 1} \alpha_{j+1} Y_{t-j} + u_{t+h} \tag{1.2.2}$$

$$Y_{t+h} = c + \sum_{j=0}^{p_Y - 1} \alpha_{j+1} Y_{t-j} + \sum_{j=0}^{p_X - 1} \beta_{j+1} X_{t-j} + u_{t+h} \tag{1.2.3}$$

All variables in these competing models are of annual frequency. The higher frequency data series in ADL regressions are aggregated to construct their corresponding annual series. Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are used to select the optimal number of lags for all regressions.

1.2.2 Forecast Combinations Method

As pointed out by Timmermann (2006), forecast combinations have been viewed as a simple and effective way to obtain more accurate forecasts by using evidence from all models considered rather than using the best single model. Forecast combinations have been used in many areas such as output growth (Stock and Watson (2004), Clements and Galvão (2008), Andreou, Ghysels, and Kourtellos (2013)), inflation (Stock and Watson (2008)) and exchange rates (Wright (2008)). Specifically, Andreou, Ghysels, and Kourtellos (2013) combine a large cross-section of daily financial data instead of using a single series to produce real-time MIDAS regression forecasts of output growth. They combine the MIDAS-regression predictions based on a single series to produce improved predictions that exploit the entire cross-section of financial series. Timmermann (2006) also lists key reasons for using forecast combinations which provide hedging against model uncertainty. Since forecasters may have different information sets, different predictors and modeling structures, forecast combinations can work well under structural breaks while individual forecasts may be very differently affected by them. Combining forecasts across different models results in robust forecasts against misspecification biases and measurement errors in the data set underlying the individual forecasts. Also, with a small set of parameters to obtain a linear projection of high frequency data onto the low frequency dependent variable, the MIDAS setup allows one to compare the forecasting performances of different regressors and choose the predictors with high predicting abilities.

Given N individual forecasting models, forecast combinations are time-varying weighted averages of the individual forecasts,

$$\hat{f}_{N,t+h|t} = \sum_{j=1}^{N} \hat{\omega}_{j,t} \hat{y}_{j,t+h|t}$$

where the weights $\hat{\omega}_{j,t}$ on the j^{th} forecast in period t depends on the historical performance of the individual forecasts.[9]

In this paper, Squared Discounted Mean Square Forecast Errors (henceforth dMSFE) forecast combination method is utilized (see Stock and Watson (2004) and (2008)). Each individual predictor is given a weight according to its historical performance and the weight is inversely proportional to the predictor‘s dMSFE. The discount factor attaches greater weight to the recent predictive ability of the individual predictor. The weights are given as

$$\hat{\omega}_{j,t} = \frac{(\lambda_{j,t}^{-1})^{\kappa}}{\sum_{i=1}^{N} (\lambda_{i,t}^{-1})^{\kappa}}, \qquad \lambda_{j,t} = \sum_{m=T_0}^{t-h} \delta^{t-h-m} (y_{m+h}^{h} - \hat{y}_{j,m+h|m}^{h})^2$$

[9] In this study, $\hat{y}_j$ is the individual prediction from ADL-MIDAS regression estimated one at a time with each individual high-frequency, $j = 1, , N$. N is the number of individual models. Estimating ADL-MIDAS regressions one at a time -as is typical in forecast combination settings- involves efficiency losses compared to systems based on Kalman Filter. See Bai, Ghysels, Wright (2013) for further information.

where $\delta = 0.9$ and $\kappa = 2$ for dMSFE.[10]

The forecast combinations method, in this paper, is employed in three steps in order to combine forecasts from ADL regressions and combine ADL-MIDAS predictions. These three steps, for instance, to obtain annual forecasts of budget data with MIDAS regressions, can be summarized as follows:

- Forecasts are computed for annual federal government expenditures and receipts with single predictors of higher frequency, i.e., forecasts of expenditures and receipts are obtained by estimating ADL-MIDAS regression models with single predictors.[11]

- For both variables, expenditures and receipts, best predictors are picked according to their out-of-sample performance measured by their root mean squared forecast errors (RMSFEs).

- Forecasts obtained from individual ADL-MIDAS regressions are combined according to the rule defined above.

1.2.3 A Test of Predictive Accuracy

The rolling forecasting scheme is employed in this paper. Let T be the total sample size, h be the forecast horizon, R denotes the size of the estimation window and $P = T - R - h + 1$ is the out-of-sample size. Consider a sequence of h-step ahead and time-t rolling window out-of sample forecast $\hat{f}_t(\hat{\beta}_t)$, which corresponds to in sample fitted values $\hat{y}_j(\hat{\beta}_t)$ with $j = h + 1, , T$ for $t = R + 1, R + 2, ..., T - h$.

Let the out-of-sample errors be $e_{t+h|t} = y_{t+h} - \hat{y}_{t+h|t}$, and the quadratic loss function be $L(y_{t+h} - \hat{y}_{t+h|t}) = e^2_{t+h|t}$. For the GW (Giacomini and White) test (Giacomini and White (2006)), the losses depend on the estimated in-sample parameters, and the expectation is taken to be conditional/unconditional on some information set G_t. Testing the null hypothesis of

$$H_0 : E[L(y_{t+h} - f_{t+h|t}(\hat{\beta}_t)) - L(y_{t+h} - g_{t+h|t}(\hat{\theta}_t))|G_t] = 0$$

[10] Weights with other discount factors of $\delta = 1$ and 0.95, and $\kappa = 1$ are also calculated, but $\delta = 0.9$ with $\kappa = 2$ is found to be the best giving the highest forecast gains.

[11] Forecasts are obtained by employing the rolling windows forecast method explained in section 1.4.

GW test statistic is a Wald-type statistic of the following form

$$GW^{\eta_t}_{R,P} = P\left(P^{-1}\sum_{t=R}^{T-h}\eta_t\triangle L_{t+h}\right)'\hat{\Omega}_P^{-1}\left(P^{-1}\sum_{t=R}^{T-h}\eta_t\triangle L_{t+h}\right) = P\bar{Z}'_{R,P}\hat{\Omega}_P^{-1}\bar{Z}_{R,P}$$

where $\triangle L_{t+h}$ is the difference of loss functions at $t+h$ and η_t is referred to as the vector of test functions.[12] $\hat{\Omega}_P$ is a consistent estimator of the asymptotic variance of $\bar{Z}_{R,P}$.[13] Under the null of equal conditional forecast performances, the GW test statistic follows a $\chi^2_{dim(\eta_t)}$ distribution. In this paper, the GW test is employed to test equal conditional predictive ability of forecast combinations of ADL-MIDAS regression models and their competitors.

1.3 Data

The data set includes the U.S. federal government current expenditures and receipts at both annual and quarterly frequencies and their corresponding quarterly subcomponents, together with macroeconomic and financial indicators sampled at higher mixed frequencies. The data set is of real-time vintages. At each point in time, the vintages of data used for the estimation are restricted to those that would have been available at that time. The sample covers the period of 1956-2012, chosen as wide as possible depending on the availability of data having long enough observations to use as in-sample and out-of-sample periods. The annual budget data is combined with higher frequency predictors, such as weekly initial jobless claims, monthly total non-farm employment, and quarterly real GDP growth, to construct ADL-MIDAS regressions. All the data series used in this study are listed in Table 1.1 and are seasonally adjusted, in real quantities, and transformed to induce stationarity, if necessary.

[12]Following Giacomini and White (2006), $\eta_t = (1 \quad \triangle L_t)'$ is used as the test function in this paper, having potential explanatory power for the future difference in forecast performances. It is also indicated by Elliot and Timmermann (2008) that past forecast errors have often been found to have predictive power over future errors. Thus, here the test is constructed conditional on past loss functions.

[13]In this study, for h-step ahead forecasts, $h > 1$, $\hat{\Omega}_P$ is a Newey-West HAC estimator. See Newey and West (1987)

1.4 Empirical Results

Using a rolling forecasting method, pseudo out-of-sample forecasts are obtained in order to evaluate predictive ability of the models for various forecasting horizons, $h = 1, 2, 3$ and 4 years. The total sample size is $T + h$ years, the fixed rolling window size is R and, for each window, forecasts of 4-years-out are used to calculate the root mean squared forecast error for the corresponding window. Hence, the period used to evaluate annual forecasts is $P = T+h-R-4$. The initial estimation period for the data set is 1956-1992 while the forecasting period is 1993-2012. The forecast accuracy of each model is assessed using the root mean squared forecast error, RMSFE, which is obtained as follows:

$$RMSFE_t = \sqrt{\frac{1}{t - T_0 + 1} \sum_{\tau=T_0}^{t} \left(y^h_{\tau+h} - \hat{y}^h_{\tau+h|\tau} \right)^2}$$

where $t = T_1, , T_2$. T_0 is the point at which the first individual pseudo out-of-sample forecast is computed. For the longest sample, $T_0 = 1993$, $T_1 = 1993 + h$ and, $T_2 = 2012 - h$.

Figure 1.1 provides a concise preview of the forecasting gains from one-step-ahead annual expenditures and receipts by displaying two boxplots, one for the forecast combinations of ADL-MIDAS regression models and the other for the competitor AR regression model.[14] These boxplots present predictive abilities of the two competing models, which are measured in terms of RMSFEs. Each point in boxplots is attached to each out-of-sample rolling window. Since smaller RMSFEs reflect better forecast performance, Figure 1.1 indicates that the forecast combinations of ADL-MIDAS models outperform the traditional AR models for both expenditures and receipts, that is, forecast combinations of ADL-MIDAS regression models, each of which is constructed by using a single higher frequency predictor, provide forecast gains for both federal government expenditures and receipts over their autoregressive regression model counterparts.

[14] A boxplot is a way of summarizing a set of data measured on an interval scale and is a type of graph which is used to show the shape of the distribution, its central value, and variability consisting of the most extreme values in the data set (maximum and minimum values), the lower and upper quartiles, and the median.

1.4.1 Forecast Comparison: MIDAS Regressions vs. AR and ADL Regressions

Table 1.2 presents RMSFEs of the models for the annual federal current expenditures and receipts. Each panel in the table shows RMSFEs of AR models, forecast combinations of ADL regression models, and forecast combinations of ADL-MIDAS models for the corresponding variable of interest. The evidence shows that forecast combinations of MIDAS regression models having smaller RMSFEs, for both budget series, provide strong forecast gains over both their AR and ADL counterparts. These results hold for all forecast horizons. In order to evaluate this evidence on forecasting performance the GW test (Giacomini and White (2006)) is employed to test the null hypotheses of equal forecasting accuracy between two different models. Table 1.3 presents these GW test statistics for h-step ahead forecasts, testing for equal forecasting accuracy between forecast combinations of ADL-MIDAS regression models vis-à-vis AR and ADL regression models. It is found that for both expenditures and receipts, MIDAS regression models yield significant forecast gains over their AR and ADL model counterparts. Individual contributions of each predictor to the superior forecast performance of combinations of ADL-MIDAS can be seen in Figures 1.2 and 1.3.

1.4.2 MIDAS Regressions with Leads

The gains of real-time forecast updating are of particular interest to policy makers. Therefore, our analysis in this part is designed to elucidate the value of weekly-monthly-quarterly information in providing real-time forecast updates of the current-year federal expenditures and receipts.[15]

The use of readily available higher frequency data, such as weekly initial claims, monthly industrial production index and quarterly real GDP growth, allows us to obtain weekly, monthly or quarterly updates of the annual forecasts of federal expenditures and receipts. That is to

[15] Additional information contained in extra data expands forecasters' information set, thus, the timing changes information available to forecasters and could affect forecast accuracy. For example, Artis and Marcellino (2001) examine relative accuracy of the budget deficit forecasts from three major international agencies and conclude that there is no single best agency, but the timing of forecasts, hence available information sets might explain the differences in forecast performances between those agencies.

say, it becomes possible to predict current year well ahead of the official figure releases. MIDAS with leads can be employed to update the current year forecasts together with future horizon predictions. In this paper, MIDAS regression with leads is utilized by incorporating real-time information using the weekly-monthly-quarterly data series. For instance, suppose we are two quarters into year $t+1$, implying that we have two quarters of weekly data (e.g., initial claims, federal funds rate, 10-year treasury bond rate), monthly data (e.g., industrial production index, oil price), and quarterly data (e.g., personal current taxes, current transfer payments). Then, if we stand on the last day of the second quarter of the year and aim to have a forecast for the current year, we could use two-quarter leads of the higher frequency data. That is, the notion of leads here implies that the information between year t and year $t+1$ is used. More precisely, the forecaster's information set is extended by using readily available data at the end of the second quarter of a given year to make a forecast. Consider the $ADL - MIDAS(p_Y^A$, $q_X^H, J_X^H)$ regression equation, allowing for J_X^H leads for the high frequency predictor, expressed in multiples of quarters, $J_X^H = 1, 2$ and 3.

$$Y_{t+1}^A = \mu + \sum_{k=0}^{p_Y^A-1} \alpha_k Y_{t-k}^A + \gamma \left[\sum_{i=0}^{J_X^H-1} \omega_i(\theta_X^H) X_{J_X^H-i,t+1}^H + \sum_{j=0}^{q_X^H-1} \sum_{i=0}^{N_H-1} \omega_{i+j*N_M}(\theta_X^H) X_{N_H-i,t-j}^H \right] + u_{t+1}$$

where N_H denotes the high frequency (weekly/monthly/quarterly) lags per year and the weighting scheme, $\omega(\theta^H)$, which involves a low dimensional vector of unknown parameters as discussed in Section 1.2.1.

Using the same data sets, the rolling window forecasting method is utilized with an initial estimation in-sample period of 1956-1992 and an out-of-sample period of 1993-2012 to obtain projections using two- and three- quarter leads of high frequency data to improve the current year's forecast. Real-time forecast updates are obtained for annual expenditures and receipts with single predictors of high frequency by estimating ADL-MIDAS regression models with two- and three- quarter leads of each predictor, separately. Then, the dMSFE forecast combination method is employed as discussed in Section 1.2.2.

Figure 3.3 presents one-year-ahead forecasts from MIDAS regressions without leads for expenditures and receipts and real-time forecast updates from MIDAS regressions with three-quarter-leads together with the actual values over the evaluation period of 1993-2012. It can be concluded that, on average, real-time forecasts from the combinations of MIDAS regressions with quarterly leads follow the actual data a little bit more closely compared to one-year-ahead forecasts from the combinations of MIDAS regressions without leads for both expenditures and receipts. Moreover, forecast performances over the 1993-2012 period are measured by their RMSFEs and presented in Table 1.4. Having smaller RMSFEs, combinations of real-time forecasts from MIDAS regressions with leads provide forecast gains over the combinations of forecasts from MIDAS regressions without leads.

In order to evaluate the evidence on forecasting performance presented above, the hypothesis of equal forecasting accuracy between forecast combinations of ADL-MIDAS regression models with and without quarterly-leads, the GW test is employed. Table 1.5 presents GW test statistics for one-year-ahead forecasts; although there are forecast gains, as measured by RMSFEs, that favor the model with leads, it is found that for both expenditures and receipts, the difference in forecasting performance is not statistically significant.

1.5 Conclusion

This chapter employs ADL-MIDAS regression models to obtain annual forecasts for U.S. federal government current expenditures and receipts. The forecasts from ADL-MIDAS regression models constructed with single predictors of higher sampling frequencies are combined. Evidence shows that these forecast combinations of ADL-MIDAS regression models outperform both autoregressive regression models and forecast combinations of ADL regression models for both expenditures and receipts. It is also shown that, although not statistically significant, ADL-MIDAS regressions with quarterly leads are found to provide forecast improvements over the ADL-MIDAS regressions without leads, suggesting the use of intra-annual data as they become available throughout the year to improve current-year forecasts. To sum up, the empirical results in this chapter support the use of mixed frequency data consisting of quarterly fiscal series and macroeconomic indicators of higher sampling frequency in forecasting annual federal budget data.

Table 1.1: Data Sets

Title	*Frequency*
Federal Government Current Receipts	Annual
Federal Government Current Receipts	Quarterly
Current tax receipts	Quarterly
Personal current taxes	Quarterly
Taxes on corporate income	Quarterly
Taxes on production and imports	Quarterly
Contributions for government social insurance	Quarterly
Income receipts on assets	Quarterly
Current transfer receipts	Quarterly
Current surplus of government enterprises	Quarterly
Federal Government Current Expenditures	Annual
Federal Government Current Expenditures	Quarterly
Consumption expenditures	Quarterly
National defense	Quarterly
Nondefense	Quarterly
Current transfer payments	Quarterly
Government social benefits	Quarterly
To persons	Quarterly
Other transfer payments	Quarterly
Grants-in-aid to state and local governments	Quarterly
Interest payments	Quarterly
Subsidies	Quarterly
Macroeconomic Data Series	
Real GDP growth	Quarterly
Consumer price index for all urban consumers: all items	Monthly
Industrial production index	Monthly
Spot oil price: West Texas intermediate	Monthly
All employees: Total nonfarm	Monthly
Initial claims	Weekly
3-Month treasury bill: Secondary market rate	Weekly
10-Year treasury constant maturity rate	Weekly
Term spread (10yTB minus 3mTB)	Weekly
Effective federal funds rate	Weekly
S&P 500 Stock price index	Weekly

Notes: This table lists the data series used in this study, two budget data series together with their corresponding subcomponents, and macroeconomic data series. All are seasonally adjusted, real quantities, and transformed to ensure stationarity, if necessary. Data are obtained from FRED (Federal Reserve Economic Data), (http://research.stlouisfed.org/fred2/), and from ALFRED (ArchivaL Federal Reserve Economic Data) for the real-time vintages (http://alfred.stlouisfed.org/). Sample period includes the annual time period of 1956-2012; chosen as long as possible depending to the availability of predictors. Some data in the series were not available as early as the annual sample period, thus, their sample period starts whenever they became available.

Table 1.2: Comparison of RMSFEs: AR and ADL Models vs. MIDAS Regressions.

	RMSFE			
	Forecast horizon			
Expenditures	h=1	h=2	h=3	h=4
AR	2.601	2.683	2.676	2.733
ADL	2.386	2.486	2.594	2.640
MIDAS	1.648	1.928	2.246	2.311
Receipts	h=1	h=2	h=3	h=4
AR	5.907	5.854	5.958	5.993
ADL	5.747	5.824	5.758	5.783
MIDAS	4.244	4.907	5.283	5.267

Notes: This table presents Root Mean Squared Forecast Errors (RMSFEs) of autoregressive (AR) models, forecast combinations of augmented distributed lag (ADL) models, and forecast combinations of MIDAS regressions for the annual U.S. federal government current receipts and expenditures for $h = 1-, 2-, 3-$ and 4-step ahead forecasts. The estimation period is 1956-1992 while the forecasting period is 1993-2012.

Table 1.3: Time Series Test for Predictive Ability.

	AR versus MIDAS			
	Forecast horizon			
	h=1	h=2	h=3	h=4
Expenditures				
GW test statistic	2.702	3.420	3.337	1.944
p-value	0.007	0.001	0.001	0.052
Receipts				
GW test statistic	2.032	2.187	3.301	0.952
p-value	0.042	0.029	0.001	0.341
	ADL versus MIDAS			
	Forecast horizon			
	h=1	h=2	h=3	h=4
Expenditures				
GW test statistic	1.987	2.021	2.006	2.359
p-value	0.047	0.043	0.045	0.018
Receipts				
GW test statistic	1.854	1.669	2.962	1.012
p-value	0.064	0.095	0.003	0.312

Notes: This table presents Giacomini-White (GW) statistics for h-step ahead forecasts and their corresponding p-values to test for equal forecasting accuracy between forecast combinations of MIDAS regressions when compared to AR models and forecast combinations of ADL models for the annual U.S. federal government current expenditures and receipts. The estimation period is 1956-1992 while the forecasting period is 1993-2012.

Table 1.4: Comparison of Forecast Performances: ADL-MIDAS with and without Leads

	RMSFEs		
	MIDAS_J=0	**MIDAS_J=2**	**MIDAS_J=3**
Expenditures	1.648	1.617	1.609
Receipts	4.244	4.219	4.204

Notes: This table presents gains of real-time forecast updating, i.e., MIDAS regressions with quarterly leads (J=2 quarters and J=3 quarters) over the one-year ahead forecasts of the ADL-MIDAS regression models (J=0) for the U.S. federal government current expenditures and receipts. Forecasting performance is measured by the root mean squared forecast errors, RMSFEs. Out-of-sample period is 1993-2012.

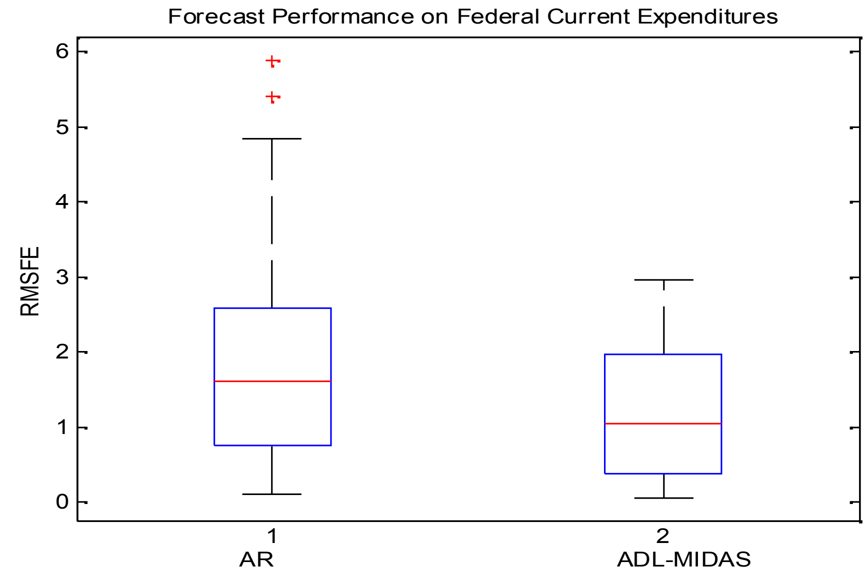
Table 1.5: Testing Equal Forecasting Accuracy between MIDAS with and without Leads

	MIDAS with Leads (J=2) vs. without Leads (J=0)	**MIDAS with Leads (J=3) vs. without Leads (J=0)**
Expenditures		
GW test statistic	1.0856	1.256
p-value	0.278	0.260
Receipts		
GW test statistic	0.984	1.0037
p-value	0.325	0.316

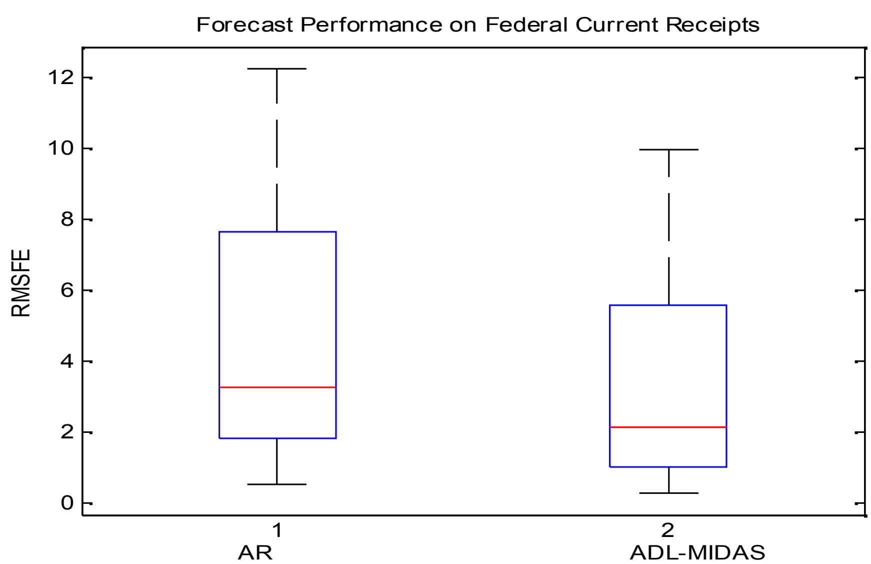
Notes: This table presents Giacomini-White (GW) statistics and their corresponding p-values for testing for equal forecasting accuracy between forecast combinations of ADL-MIDAS regressions with quarterly leads (J=2 and 3 quarters) and without leads (J=0) for the annual U.S. federal government current expenditures and receipts. Out-of-sample period is 1993-2012.

Figure 1.1: Boxplots for Forecast Performance Comparisons
between MIDAS and AR regression models

One-year-ahead forecast performance on the U.S federal government current expenditures



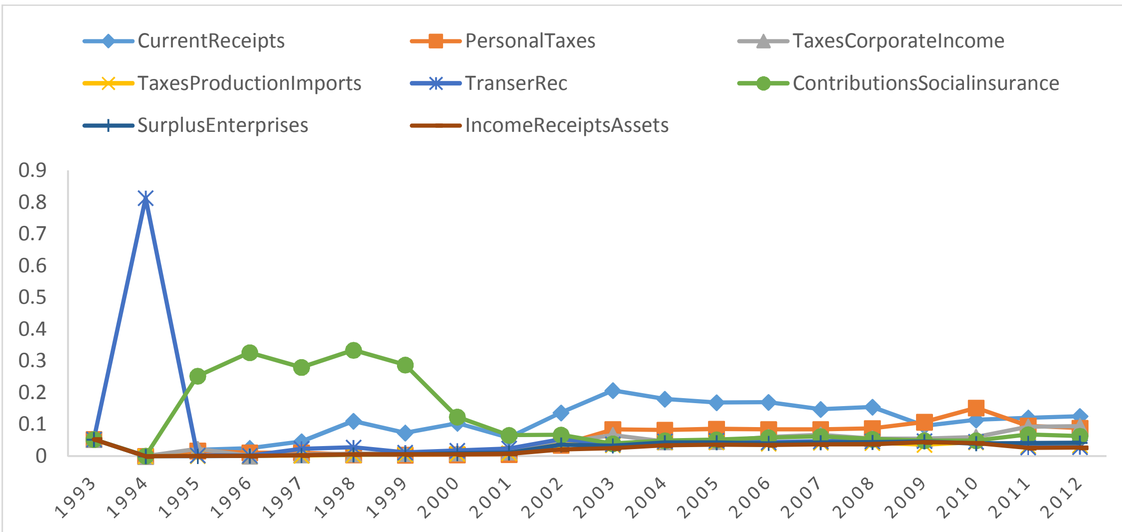
One-year-ahead forecast performance on the U.S federal government current receipts



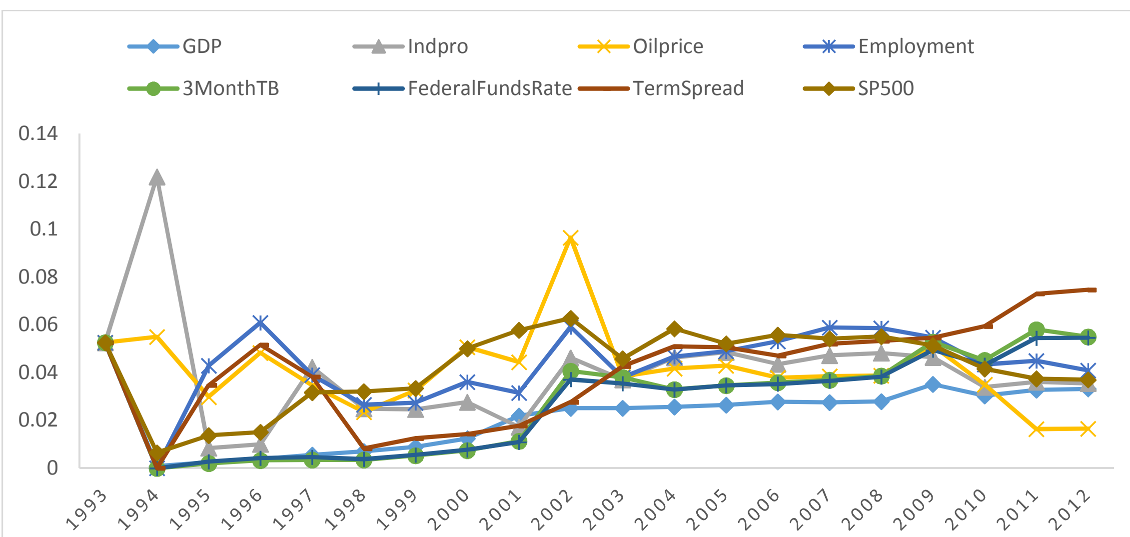
Note: These boxplots present predictive abilities of the two competing models, forecast combination of ADL-MIDAS regressions vs. AR regressions by displaying their forecast performances measured in terms of Root Mean Squared Forecast Errors (RMSFEs). Each point in the boxplots is attached to each out-of-sample rolling window.

Figure 1.2: Forecast Combination Weights over the Forecasting Period for Receipts

Forecast Combination Weights of Fiscal Variables



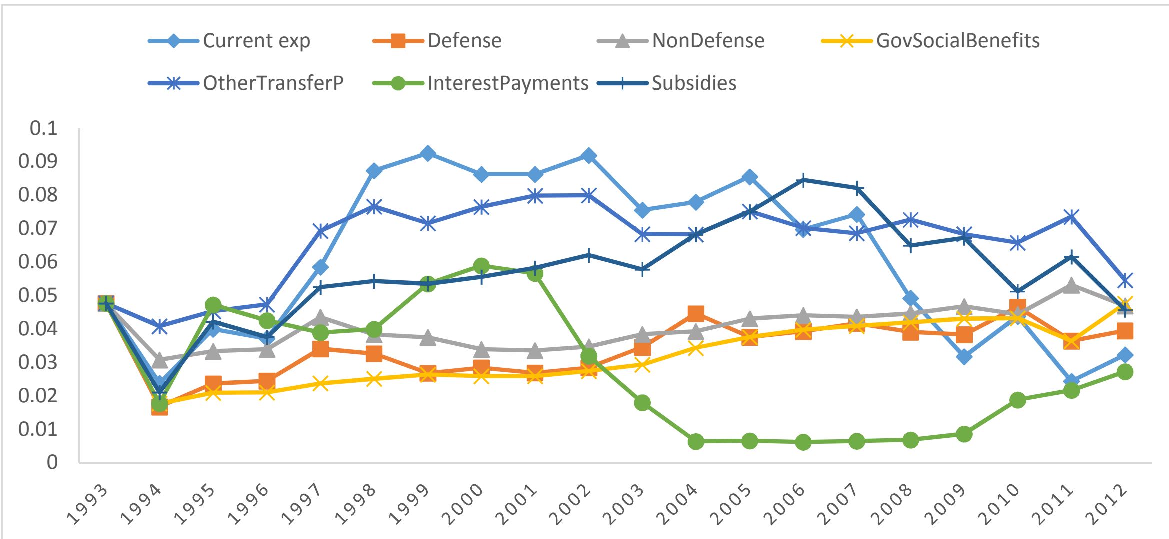
Forecast Combination Weights of Macroeconomic Variables



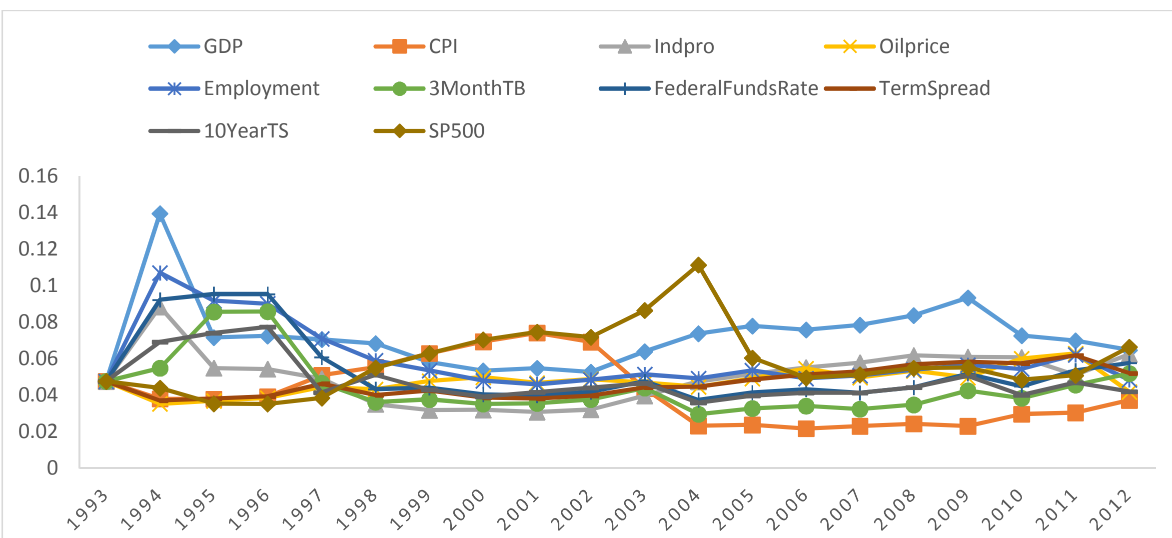
Note: These figures display forecast combination weights of the predictors used in ADL-MIDAS regressions for the U.S. federal government current receipts over the forecasting period of 1993-2012. Note that for a year, all weights add up to one. Weights for the intra-annual fiscal series and for the macroeconomic predictors are displayed with two separate figures to make them easily identifiable and comparable.

Figure 1.3: Forecast Combination Weights over the Forecasting Period for Expenditures

Forecast Combination Weights of Fiscal Variables



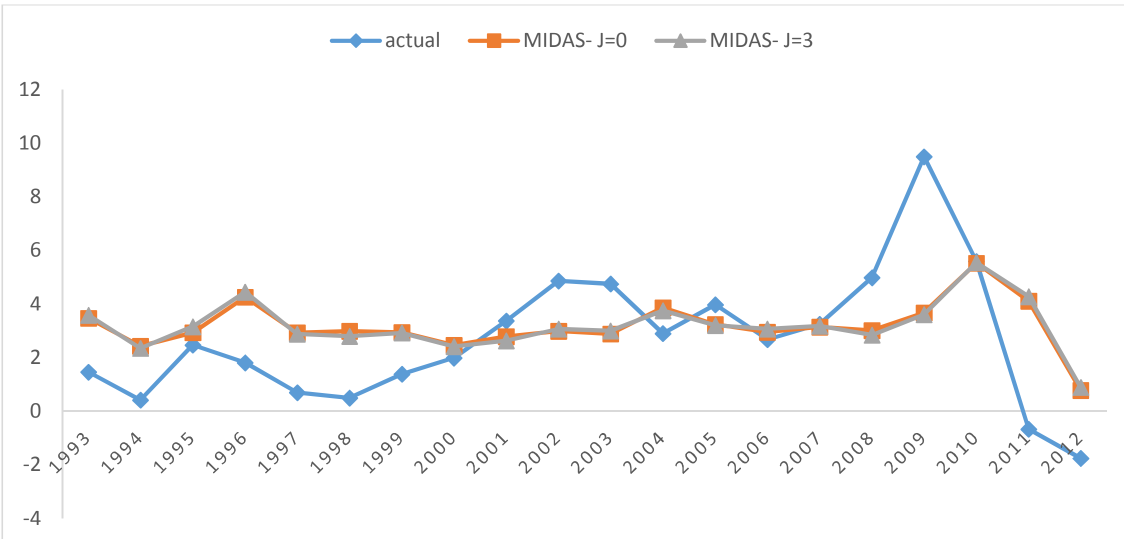
Forecast Combination Weights of Macroeconomic Variables



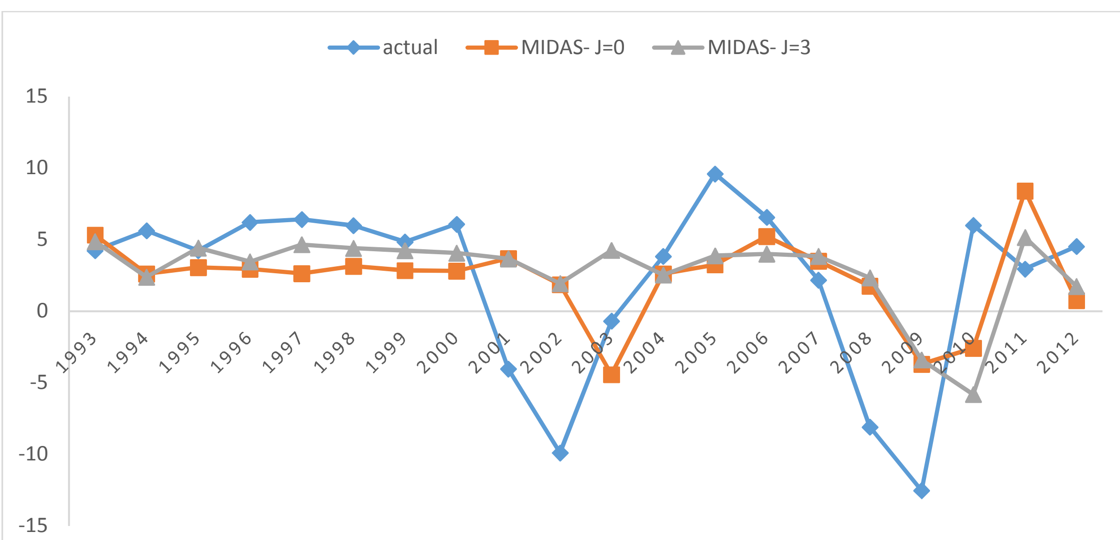
Note: These figures display forecast combination weights of the predictors used in ADL-MIDAS regressions for the U.S. federal government current expenditures over the forecasting period of 1993-2012. Note that for a year, all weights add up to one. Weights for the intra-annual fiscal series and for the macroeconomic predictors are displayed with two separate figures to make them easily identifiable and comparable.

Figure 1.4: Predictions from Forecast Combinations
of MIDAS regressions with and without Leads

Comparison for Federal Current Expenditures



Comparison for Federal Current Receipts



Note: These figures present one-year-ahead predictions from ADL-MIDAS regressions for the U.S. federal government current expenditures and receipts (MIDAS-J=0) and their corresponding real-time forecast updates with quarterly leads (MIDAS-J=3 quarters) over the out-of-sample period of 1993-2012 together with the actual observations. Note that all values are real growth rates.

CHAPTER 2

REAL-TIME FORECASTING OF THE U.S. STATE AND LOCAL GOVERNMENTS' BUDGETS

2.1 Introduction

Fiscal sustainability has become a national challenge at all levels of government with the latest recession. State and local governments, like the federal government, have been experiencing the fiscal stress that is closely tied to the national business cycle, adding to the nation's overall fiscal challenges. Furthermore, the U.S. Government Accountability Office, an independent nonpartisan agency, predicts that under current policies, state and local government budgets, at the aggregate level, will experience an even greater gap between projected revenues and expenditures in the coming years.[1]

A majority of the states have balanced budget requirements, which stipulate that the states raise enough money to cover the costs of estimated expenditures. However, it can be an overwhelming issue to operate a balanced budget in times of fiscal stress. Hence, during recessions deficits might occur since planned revenues are not likely to be generated as predicted while the demand for services could exceed planned expenditures. On the other hand, unlike the federal government, states cannot run prolonged budget deficits due to their balanced budget requirement limitations. Therefore, it is crucial to have reliable budget projections. By forecasting revenues and expenditures accurately for periods beyond prevailing fiscal periods, better informed decision makers can deal with impending budgetary issues sooner and conduct budgeting accordingly before fiscal problems get worse. In fact, the fiscal shocks caused by the

[1] U.S. Government Accountability Office (2013): State and Local Governments Fiscal Outlook, April 2013 Update.

recession combined with a slow recovery have led state governments to search for more reliable budgeting tools that can promise fiscal sustainability. This includes developing long-term economic and revenue trend analyses to avoid service disruptions and widespread budget cuts.[2]

Agencies involved in fiscal forecasting aim to assess whether state finances are developing in accordance with official budgetary plans and to provide a timely warning when they are moving away from those plans. Fiscal policy agencies, in general, implement forecasting procedures based on judgment and/or econometric models including simple regressions, time series methods, and structural macroeconometric models. Many papers in this literature deal with fiscal projections on revenues, particularly tax revenues.[3] While the literature on revenue forecasting is quite well developed, it is also common to implement integrated approaches for both revenues and expenditures by employing forecasting models for expenditures as well (e.g., expenses on unemployment insurance funds).[4] Grizzle and Klay (1994) compare forecasting techniques including judgmental methods, econometric modeling and a combination of both. They find evidence in favor of combining judgment and econometric models as that combination produces more accurate revenue forecasts. Fullerton (1989) also develops a composite predictor of sales tax revenues for the state of Idaho and finds that the proposed combined forecasts technique provides forecast accuracy gains over the individual forecasts obtained separately by using the two approaches.

There are also studies that implement macroeconomic models to ensure that forecasts are internally consistent by taking the interactions between economic cycles and fiscal variables into account. Since the business cycles of the economy significantly affect tax revenues and expenditure developments, macroeconomic models for fiscal forecasting procedures are justified by many studies in the literature.[5] The Congressional Budget Office (CBO), a nonpartisan government agency whose main goal is to provide accurate forecasts of the federal budget,

---

[2] National Association of State Budget Officers, Summer 2013 Report. State Budgeting and Lessons Learned From The Economic Downturn.

[3] See, Lawrence et al. (1998), Fullerton (1989), among others.

[4] See, for example, Pike and Savage (1998), Sentance et al. (1998), and Giles and Hall (1998).

[5] See, for example, Pike and Savage (1998), Holloway (1989), Sentance et all. (1998),and Dalsgaard and De Serres (2001), among others.

employs a macroeconomic model with fiscal parts and combines them with detailed fiscal models and judgment. The model developed by the CBO to study long-run budget issues contains equations that account for the feedback between the budget and the economy by tracing the way in which output depends on capital and labor and hence on the budget and population.[6]

Fiscal decision makers are usually interested in detailed projections for individual revenue and expenditure items in the preparation of budgetary forecasts. In this paper, however, the fiscal side of the proposed procedures is not detailed for a comprehensive analysis of fiscal policies. In that regard, the proposed econometric forecasting methods that are shown to provide forecast accuracy gains relative to traditional econometric models can be used by fiscal decision makers to complement their fiscal and judgment models. In this respect, the current paper contributes to the fiscal policy literature by proposing novel econometric forecasting procedures for the U.S. state budget series that can serve as effective econometric decision-making tools to state and local governments when developing their long- term budget outlooks by allowing them to periodically update budget forecasts and increase transparency in budget projections.

This chapter proposes forecasting procedures to deal with the problem of forecasting state and local governments' expenditures and revenues on a state-by-state basis. Since the finances of state and local governments are closely tied to prevailing national economic conditions, this paper intends to utilize readily available national macroeconomic series as indicators for state and local government budget predictions. The state-specific budget series are observed at annual frequency while the national series are sampled at higher mixed frequencies. Mixed frequency data regressions are employed for forecasting purposes to make use of mixed frequency observations and retain the information content of the higher frequency series. Several papers have indicated the importance of information contained in intra-annual data for forecasting and monitoring annual budgetary outcomes. Onorante et al. (2010), for instance, use a mixed frequency state-space model to integrate readily available monthly/quarterly fiscal data with the annual budget series. Paredes et al. (2009) and Pedregal and Perez (2010) show the usefulness of incorporating intra-annual fiscal data for real-time fiscal policy surveillance by estimating

[6]See, CBO Memorandum (1997), "An Economic Model for Long-Run Budget Simulations".

models with annual and quarterly national accounts fiscal data. Ghysels and Ozkan (2013) employ a simple mixed frequency data regression approach, which is shown to provide forecast gains over traditional models. This suggests that there are potential gains from incorporating mixed frequency data consisting of quarterly fiscal series and macroeconomic indicators of higher sampling frequency in forecasting annual federal budgets.

There has been considerable interest in the development of econometric forecasting models based on mixed frequency data. The mixed data sampling (MIDAS) approach proposed by Ghysels, Santa-Clara and Valkanov (2006) and Ghysels, Sinko and Valkanov (2007) has proven to provide advantages for different forecasting purposes. This approach was first used for financial applications and volatility predictions (Ghysels et al. (2006)), and quickly gained popularity among macroeconomists for improving the real-time forecasting of key economic variables.[7] As a multivariate extension of the MIDAS approach, Ghysels, Foroni, and Marcellino (2013) introduce Mixed-Frequency Vector Autoregressive (MF–VAR) models, which provide an alternative to commonly used parameter-driven state-space models containing latent processes that rely on filtering to extract unobserved states. On the other hand, the MF–VAR model is an observation-driven model formulated in terms of observable data and, thus, impulse response functions are constructed in terms of observables rather than shocks to latent processes. In Ghysels et al. (2013), a Bayesian approach is also covered in addition to the classical one as the parameter proliferation could be a hurdle for both MF–VAR and traditional VAR models.

Mixed frequency data regression models in this paper are employed by means of both single and multiple equation approaches. Therefore, consisting of two portions associated with these approaches, this paper first examines a single-equation regression forecasting procedure which involves a simple mixed frequency data regression model, namely combinations of Augmented Distributed Lag–Mixed Data Sampling (ADL–MIDAS) regression models. The second portion dealing with multiple-equation regression models employs the Mixed Frequency–Bayesian Vector Autoregressive (MF–BVAR) model to forecast each state's government budget series.

In order to assess the empirical forecast performance of the proposed models, traditional models that are common in the forecasting literature are employed as benchmarks. For the

[7] See Clements and Galvao (2008, 2009), Marcellino and Schumacher (2010), and Kuzin et al. (2011), among others.

ADL–MIDAS regressions, the benchmarks are the random walk (RW), autoregressive (AR), and augmented distributed lag (ADL) models. For the MF–BVAR model, low-frequency Bayesian VAR (BVAR) and RW models are used as benchmarks. Although it varies at the state level, the overall empirical forecast performances of the proposed mixed frequency data regressions are found to be better than their traditional low-frequency competitors.

As an additional exercise, the two proposed approaches are also compared in terms of their empirical predictive abilities. Both the MF–BVAR, as the system approach, and the ADL–MIDAS, as the single equation approach, can tackle data series sampled at mixed frequencies. The MIDAS regression is employed mainly in the context of economic forecasting, the MF–BVAR model, on the other hand, can be used for both forecasting and structural analysis. The results suggest that the relative performance of these two approaches varies depending on the predictors and forecast horizons and there is no clearly dominant approach.

The remainder of the chapter proceeds as follows. In Section 2.2, the models employed are described with subsections 2.2.1 and 2.2.2 devoted to the single and multi-equation approaches by outlining the empirical methodologies that are used to obtain forecasts. Section 2.3 contains a description of the data, and empirical results are gathered in Section 2.4. Section 2.5 presents conclusions.

2.2 Models

High-frequency data have become more readily available, but still many series of interest remain at lower sampling frequencies. At the state level, government finance data are available only at annual frequency with considerable release lags.[8] The introduction of Mixed Data Sampling (MIDAS) time series methods has allowed researchers to efficiently exploit information available at different frequencies. The MIDAS regression method introduced by Ghysels et al. (2006, 2007) extracts high-frequency information in a lower frequency regression by employing time-varying tightly-parameterized polynomial weighting schemes that maximize the retained information from each of the high frequency observations rather than a fixed weight for all observations, as in the case of temporal aggregation. Recent literature has shown the advantages

[8] The U.S. Census Bureau conducts annual surveys of state and local government finances. The latest available one is from 2011. Census Bureau, "2011 Annual Surveys of State and Local Government Finances".

of MIDAS-type regressions in terms of the improvement of low-frequency macroeconomic predictions using high-frequency data.[9] As a multivariate extension of MIDAS regressions, Ghysels et al. (2013) introduce a parsimonious mixed sampling frequency VAR model that does not involve latent shocks and also allows for the analysis of the impact of high-frequency data onto low-frequency ones. The following subsections describe the two approaches utilized to obtain point forecasts of the annual state and local government budget series. The first approach uses single equation regressions while the second employs multiple equation regressions.

2.2.1 Single Equation Regression Model: Augmented Distributed Lag–Mixed Data Sampling (ADL-MIDAS) Regression

In the first empirical portion of this chapter, the proposed forecasting procedure involves combinations of Augmented Distributed Lag–Mixed Data Sampling (ADL-MIDAS) regression models in which the 48 mainland U.S. state and local governments' expenditures and revenues are predicted using a set of mixed frequency variables.[10] This set includes quarterly state-specific variables and monthly/quarterly national variables.

The objective of this section is to use MIDAS-type regressions to obtain annual forecasts of the 48 state and local government expenditures and revenues, and compare them with the forecasts from traditional models, namely the random walk (RW), autoregressive (AR), and augmented distributed lag (ADL) regression models. ADL–MIDAS regression models are constructed to include lags of the low frequency dependent variable (e.g., annual state and local government expenditures or revenues) and lags of the high frequency predictor (e.g., quarterly real gdp growth). Individual forecasts with each single indicator are calculated, thus, multiple forecasts of expenditures and revenues are obtained. In order to obtain more accurate forecasts, the forecast combinations method is considered rather than using the best single model (this

[9] MIDAS-type regressions were suggested in recent papers by Ghysels, Santa-Clara, and Valkanov (2006), Ghysels, Sinko, and Valkanov (2007), and Andreou, Ghysels, and Kourtellos (2013). The initial work on MIDAS focused on volatility predictions. See Ghysels, Santa-Clara, and Valkanov (2006), among others.

[10] As shown in Bai, Ghysels and Wright (2013), one can view the proposed MIDAS regression approach as a computationally simple way of approximating the state-space approach suggested by Onorante et al. (2010), among others.

method incorporates evidence from all individual models).[11]

Let the annual variable of interest to be predicted for an h-step ahead horizon be $Y^A_{(s,t+h)}$, for example, real per capita revenues for state s and let the high frequency predictor series be X^Q_t, for example, quarterly per capita real GDP growth. Then, the $ADL - MIDAS(p^A_Y, q^Q_X)$ regression model can be defined as

$$Y^A_{s,t+h} = c + \sum_{j=0}^{p^A_Y - 1} \alpha_{j+1} Y^A_{s,t-j} + \beta \sum_{j=0}^{q^Q_X - 1} \sum_{i=0}^{N_Q - 1} \omega_{i+j*N_Q}(\theta^Q) X^Q_{N_Q - i, t-j} + u_{t+h} \tag{2.2.1}$$

where N_Q denotes the quarterly lags per year and the weighting scheme, $\omega_t(\theta^Q)$, involves a low dimensional vector of unknown parameters. Following the weighting scheme used by Ghysels et al. (2007), all available polynomial specifications are utilized.[12] Then, the best specification for each ADL-MIDAS model is chosen according to its forecast performances. The following subsections describe the methods employed in this portion of the paper.

1. *Benchmark Models:* Random walk and autoregressive $AR(p_Y)$ models, standard benchmarks in most forecasting exercises, and the augmented distributed lag, or $ADL(p_Y, p_X)$, regression model are employed as benchmark forecasting models. These models can be represented, respectively, with the following equations:

$$Y_{s,t+h} = Y_{s,t-1} + u_{t+h}, \tag{2.2.2}$$

$$Y_{s,t+h} = c + \sum_{j=0}^{p_Y - 1} \alpha_{j+1} Y_{s,t-j} + u_{t+h}, \tag{2.2.3}$$

[11] See Timmermann (2006) for a comprehensive survey on forecast combination methodology.

[12] For polynomial weight specifications and details, please refer to Ghysels et al. (2007). Additional information regarding the empirical results obtained by using different weighting polynomials is available upon request from the authors.

$$Y_{s,t+h} = c + \sum_{j=0}^{p_Y-1} \alpha_{j+1} Y_{s,t-j} + \sum_{j=0}^{p_X-1} \beta_{j+1} X_{t-j} + u_{t+h}. \tag{2.2.4}$$

All variables in these competing models are at the annual frequency. High-frequency data series in ADL regressions are aggregated to construct annual series.[13]

2. *Estimation:* MIDAS regressions involve NLS and feasible GLS estimation procedures. Errors, denoted as u_{t+h} s, are not necessarily i.i.d.. The error process is a linear process with absolute summable Wold decomposition moving average representations. Identification of the coefficients rests on the assumption that the regressor is exogenous up to the second moments.[14] Furthermore, to ensure correct specification of MIDAS polynomials, the assumption of $E(u_{t+h}|X_\tau, \tau \le t) = 0$ is required.[15]

   To simulate real-time forecasting, a pseudo out-of-sample experiment is conducted. Using a rolling window method, out-of-sample forecasts are obtained in order to evaluate the predictive ability of the models for h-year ahead horizons. The total sample size is $T + h$ years and the fixed rolling window size is R. For each window, 4-years-out forecasts are obtained in order to calculate the root mean squared forecast error for the corresponding window; hence, the period used to evaluate each annual forecast is $P = T + h - R - 4$.[16] The forecast accuracy of each model is assessed by using the root mean squared forecast error, or RMSFE. [17] RMSFEs for competing models are presented relative to the RMSFEs from the RW model.

---

[13] Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are used to select the optimal number of lags for all regressions.

[14] With MIDAS regressions, since regressors are sampled at higher frequencies, consistency (absence of discretization bias) can be of concern; Ghysels, Santa-Clara, Valkanov (2006) show that discretization bias is eventually eliminated. That is, aggregation bias disappears as the dependent variable is sampled at a fixed frequency while the regressor is sampled more frequently. They note that MIDAS regressions appear like skip-sampled distributed lag models causing autocorrelated residuals (not preventing OLS/NLS to be consistent).

[15] See, Ghysels, Sinko, Valkanov (2007) for further information.

[16] The initial estimation window for the single regression models is 1957-1998 while the pseudo out-of-sample covers the period from 1999-2011.

[17] $RMSFE_t = \sqrt{1/(t - T_0 + 1) \sum_{\tau=T_0}^{t} (y^h_{\tau+h} - \widehat{y}^h_{\tau+h|\tau})^2}$ where $t = T_1, , T_2$ and T_0 is the point at which the first individual pseudo out-of-sample forecast is computed.

3. *Forecast Combinations Method:* To obtain annual combined forecasts of budget data from ADL and ADL–MIDAS regressions, individual forecasts for each annual state and local government expenditures and revenues with single predictors of higher frequency are obtained. Then best predictors are chosen based on their out-of-sample forecast performance assessed by their RMSFEs and combined according to the forecast combination method described in Section 1.2.2.

2.2.2 Multiple Equation Regression Model: Mixed Frequency-Bayesian Vector Autoregression (MF-BVAR)

The Vector Autoregression (VAR) proposed by Sims (1980) is a workhorse model for forecasting as well as characterizing dynamic relations among macroeconomic variables. VAR models are rich in parameters and may suffer from parameter proliferation; hence, Bayesian methods with VAR models have become popular as the use of prior information offers a way of shrinking parameters.

This paper employs Bayesian estimation of Mixed Frequency Vector Autoregressions, or MF-BVARs, by means of a method similar to the one developed in Chiu, Eraker, Foerster, Kim, and Seoane (2011), which utilizes an algorithm for Markov Chain Monte Carlo (MCMC) to sample from the posterior distributions of the VAR parameters. In that paper, the posterior is conditioned on data observed at mixed frequencies, assuming that the high frequency realizations of the low frequency data are not observed and, thus, treated as missing values (i.e., a traditional missing data problem is assessed without temporal aggregation). Hence, their Bayesian mixed frequency algorithm relies on a Gibbs sampler that produces alternate draws from missing data and unknown parameters. As opposed to their mixed frequency approach, which draws from missing data, this chapter utilizes a method following Ghysels, Foroni, and Marcellino (2013) by treating the mixed frequency data as skip-sampled processes. The model is then formulated exclusively in terms of observable data. Therefore, the mixed frequency VAR approach introduced by Ghysels (2012) and Ghysels et al. (2013) does not rely on latent process representations, which is the case for the common state-space models for mixed frequency data regressions.[18] Furthermore, it allows the analysis of the impact of high frequency data on low frequency with impulse response functions that are formulized in terms of observable data.

MF-BVAR models are estimated for each state considered. When setting the prior distributions, the priors introduced in Ghysels et al. (2013), which are very close to the standard procedure developed in Litterman (1986) with modifications proposed by Kadiyala and Karlsson (1997), are utilized. After constructing the stacked vectors of mixed frequency data, a standard MCMC algorithm is used as in Chiu et al. (2011) to draw from the posterior distributions. Consistent with the notation of Ghysels et al. (2013), the model to obtain point forecasts for each state is constructed as follows.

Let $x_L(\tau_L)$ denote the low frequency multivariate process for state s (i.e., the state-specific vector of annual state and local government expenditures (EXP) and revenues (REV)) and $x_H(\tau_L, k_H)$ be the high frequency multivariate process, vector of quarterly national indicators, including the federal funds rate (FFR), federal government deficit (DEF), inflation (INF), industrial production ($INDPRO$), and gross domestic product (GDP). The number of variables in the stack is $m = 4 * K_H + K_L$, where K_H is the number of quarterly national series and K_L is the number of annual state-specific budget series. The mixed frequency $VAR(1)$ model for annual state government budget series and quarterly national macroeconomic variables can be formulized as follows:

$$\begin{bmatrix} x_H(\tau_L, 1) \\ x_H(\tau_L, 2) \\ x_H(\tau_L, 3) \\ x_H(\tau_L, 4) \\ x_L(\tau_L) \end{bmatrix} = A_0 + A_1 \begin{bmatrix} x_H(\tau_L - 1, 1) \\ x_H(\tau_L - 1, 2) \\ x_H(\tau_L - 1, 3) \\ x_H(\tau_L - 1, 4) \\ x_L(\tau_L - 1) \end{bmatrix} + \epsilon(\tau_L) \tag{2.2.5}$$

[18] See, for example, Zadrozny (1990), Mariano and Murasawa (2003), Mittnik and Zadrozny (2005), Aruoba, Diebold, and Scotti (2009), Kuzin, Marcellino, and Schumacher (2011), and Marcellino and Schumacher (2010), among others.

with

$$x_H(\tau_L, j) = \begin{bmatrix} DEF(\tau_L, j) \\ INDPRO(\tau_L, j) \\ GDP(\tau_L, j) \\ INF(\tau_L, j) \\ FFR(\tau_L, j) \end{bmatrix}, \ \ j = 1, 2, 3, 4 \ \ and \quad x_L(\tau_L) = \begin{bmatrix} REV(\tau_L) \\ EXP(\tau_L) \end{bmatrix}$$

where the last two equations in the system can be read as ADL-MIDAS regressions for expenditures and revenues; hence, the VAR model includes MIDAS regressions. To obtain real-time forecasts of the annual budget data, the two low frequency data series are assumed to be released at the same time, appearing at the end of year τ_L since the order of appearance in the vector determines the timing of the τ_L releases. More than one series is released at the high/quarterly frequency, so the order is subject to the same considerations as in traditional VAR models. The real-time forecast updating works as follows: a) first, the MF-BVAR model is estimated; b) second, the Choleski factorization of the errors are computed; c) third, the $4 * K_H$ lower triangular truncations of the original factorization are taken to add the information innovation of the high frequency series to the equations and reweight the old information. Finally, the last two equations in the system become ADL-MIDAS regression models with leads giving the updates of the expenditure and revenue forecasts.

The following subsections describe the methods employed in this portion of the chapter.

1. *Estimation:* The mixed frequency VAR system is assumed to be covariance stationary. The stacked vector of variables in equation (2.2.5) is of dimension $K_L + 4 * K_H$ and has a finite order covariance stationary VAR representation. The variables in the stacked vector are transformed to induce stationarity. The quarterly data are in four-quarter change and the budget data series are in annual change. The error term,or ϵ_{τ_L}, has the variance-covariance matrix represented by $\Sigma = E[\epsilon(\tau_L)\epsilon(\tau_L)^{'}]$. Σ is an unknown positive definite matrix rather than a fixed diagonal matrix. Under these assumptions, the model is estimated by Bayesian techniques. The formulation of the priors are explained in the

following section.

In order to deal with the mixed frequency data to be incorporated into the VAR analysis, as suggested by Ghysels et al. (2013), the step function approach to MIDAS is utilized as in Ghysels, Sinko, and Valkanov (2007) which is equivalent to the U-MIDAS, or unrestricted MIDAS, approach by Foroni, Marcellino, and Schumacher (2013).[19] With the U-MIDAS approach, the last two rows of the matrix A_1 are unrestricted and the priors are specified for the step functions where the steps are multiples of four (the number of quarters in a year).[20] The models are estimated by means of the recursive window method. Forecasts are also calculated recursively.

2. *The Priors:* The prior utilized here is introduced in Ghysels et al. (2013), which follows Doan, Litterman (1984), and Sims (1984), Litterman (1986), Sims ans Zha (1998), among others. The slope coefficient matrix is of dimension $m = 4 * K_H + K_L$, where K_H is the number of quarterly national series and K_L is the number of state specific variables, namely state expenditures and revenues, at the annual frequency. It should be noted that the sample size allows for only one lag in the VAR specification. The system in equation (2.2.5) can be written as

$$\begin{bmatrix} x_H(\tau_L, 1) \\ x_H(\tau_L, 2) \\ x_H(\tau_L, 3) \\ x_H(\tau_L, 4) \\ x_L(\tau_L) \end{bmatrix} = A_0 + \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{bmatrix} \begin{bmatrix} x_H(\tau_L - 1, 1) \\ x_H(\tau_L - 1, 2) \\ x_H(\tau_L - 1, 3) \\ x_H(\tau_L - 1, 4) \\ x_L(\tau_L - 1) \end{bmatrix} + \epsilon(\tau_L) \qquad (2.2.6)$$

[19]U-MIDAS is a completely unrestricted specification where each weight is estimated separately. Foroni, Marcellino, and Schumacher (2011) show that it works for small numbers of high frequency series per low frequency observation. For the data set utilized here, there are four quarterly observations per year.

[20]Large dimensional VAR models are obtained with the use of the U-MIDAS approach, which results from the stacks of the same high frequency series. The shrinkage method could be considered, but the priors already offer a great deal of shrinkage. For further information, see Ghysels et al. (2013)

with the priors set as

$$
\begin{array}{lll}
E[a_{k,l}]{=}0_{K_H^2} & V[a_{k,l}]{=}\ \frac{\lambda^2}{[4-l+k]^2} 1_{K_H^2} & k = 1, ..., 4,\ l = 1, ..., 3 \\
E[a_{k,4}]{=}diag(\rho^k)_{K_H^2} & V[a_{k,4}]{=}\ \frac{\lambda^2}{k^2} 1_{K_H^2} & k = 1, ..., 4 \\
E[a_{k,5}]{=}0_{K_H \times K_L} & V[a_{k,5}]{=}\ \upsilon_{HL}\ \frac{\lambda^2}{k^2}\ 1_{K_H \times K_L}\ S_{HL} & k = 1, ..., 4 \\
E[a_{5,l}]{=}0_{K_L \times K_H} & V[a_{5,l}]{=}\ \upsilon_{LH}\ \lambda^2\ 1_{K_L \times K_H}\ S_{LH} & l = 1, ...4 \\
E[a_{5,5}]{=}\ 0_{K_L^2} & V[a_{5,5}]{=}\ \lambda^2\ 1_{K_L^2} &
\end{array}
$$

where the notation $V[\,]$ stands for a matrix of variances, 0 and 1 are, respectively, zeros and ones matrices, $diag(x)$ is a diagonal matrix with elements x with the dimension as subscript, and S_{HL}=$[\sigma^2_{i,H}/\sigma^2_{j,L};\ i = 1, ..., K_H, j = 1, ..., K_L]$ and S_{LH}=$[\sigma^2_{j,L}/\sigma^2_{i,H};\ i = 1,$ $..., K_H, j = 1, ..., K_L]$ capture the differences in scaling between high and low frequency data. The hyperparameter λ governs the overall tightness of the prior distributions around the $AR(1)$ specification for the high frequency process with ρ as the autoregressive parameter, which is common among all series. The hyperparameters υ_{HL} and $\upsilon_{LH} \in (0, 1)$ govern the extent to which the low/high frequency data affect high/low frequency data, respectively. Note that within the high frequency series prior distribution is uniform by treating the dependence within the vector of high frequency data as uniform. The priors imply that the high/low frequency data do not have any impact on the low/high frequency data. The priors for the parameters pertaining to the covariance matrix of the errors are set to the Normal-Inverse Wishart as in Kadiyala and Karlsson (1997) and the posteriors are employed as in that paper. The MCMC procedure, therefore, relies on the posterior distributions.

3. *Benchmark Models:* As a benchmark model, the traditional Bayesian VAR (BVAR) model is employed. The estimation procedures described above, including the assumptions on priors, are utilized for this model as well except the one for the mixed frequency data (the data for this model are all at annual frequency). The forecast performances of both

models are assessed via RMSFEs relative to the RMSFEs from the random walk (RW) model. The RW model is based on priors with the hyperparameters $\lambda = 0$ and $\rho = 1$.

4. *Impulse Response Functions:* In order to examine the advantages of the use of the MF–BVAR model against the traditional BVAR model by comparing their dynamics, annual responses of budget variables to the shocks to national series are analyzed with impulse response functions. Note that the parameters governing the variance-covariance matrix $E[\epsilon(\tau_L)\epsilon(\tau_L)^{'}]$ and, hence, its Choleski factorization are tied to the autoregressive parameter ρ governing the VAR dynamics. Since the MF-BVAR model is constructed with observable data, impulse response functions are obtained in terms of observable high- and low-frequency data rather than shocks to latent processes. The impulse responses for the BVAR models are also constructed with observable data, but all are at an annual frequency. Since the focus of this paper is on annual state budget series, the impacts of the shocks to national series on these variables are analyzed. With the traditional BVAR model, one can obtain annual responses to the annual shocks only. On the other hand, the MF–BVAR model allows for the analysis of annual responses to the quarterly shocks. Both models operate at the annual frequency, thus, the impulse responses are represented in terms of annual time ticks. The Figures 2.3, 2.4, 2.5, and 2.6 depict the impacts of shocks to $x_H(\tau_L, j)$ (i.e., shocks to quarter j national series) for $j = 1, .., 4$ quarters on future annual $x_L(\tau_L + k)$ for $k = 1, ..., 15$ years as determined by the MF–BVAR.

2.3 Data

2.3.1 Single–Equation Regressions

The data set employed for the single equation regression analysis includes 48 U.S. state and local government expenditures and revenues at an annual frequency and their corresponding quarterly personal income series. Each state-specific set is incorporated with national data involving higher mixed frequency macroeconomic and financial indicators. The data set is of real-time vintages; at each point in time, the vintages of data used for each estimation are restricted to those that would have been available at that time. The sample includes data from 1958-2011, which was the widest time span chosen based on the available data. The data series used in this study are listed in Table 2.1. These data are seasonally adjusted, denoted in real per capita quantities and transformed to induce stationarity, if necessary. In-sample estimation covers the period from 1958-1998 while the pseudo-out-of-sample time span is 1999-2011.

2.3.2 Vector Autoregressions

In the VAR setting, all data series except the annual state specific government budget data are at a quarterly frequency. The quarterly data are in four-quarter change and the budget data series are in annual change, all expressed as percentage points. For every quarter, the quarterly data are observed while the annual data are observed only during the last quarter of each year. The in-sample estimation covers the period from 1958-2007 while the pseudo-out-of-sample time period is 2008-2011.

2.4 Empirical Results

2.4.1 Single–Equation Regression Results

Table 2.2 presents RMSFEs for one- and two-year-ahead forecasts from the three single equation regression models for the annual expenditures and revenues for each state relative to the RMSFEs from the RW model over the forecast evaluation period from 1999-2011. From a general perspective considering the forecast performance of the three models across the 48 states, the forecast combinations of ADL–MIDAS regression models have smaller RMSFEs, for both budget series at all horizons and, thus, provide forecast gains over their AR and ADL counterparts. Since relative RMSFEs that are smaller than one indicate better predictions over the RW model, this result also holds for the RW model. It is also noteworthy that the forecast performances of the three models vis-à-vis the RW model for the revenue series are better than those for the expenditures. This is an expected result since the expenditures are somewhat under control with the budgeting process as the decision makers aim to keep them stable and below estimated revenues over fiscal years. The relative RMSFEs for the revenue series for all 48 state governments are below one for the ADL–MIDAS forecasts whereas for the expenditures, the RW model outperforms the ADL–MIDAS models for four states, which are North Carolina, North Dakota, Ohio and Oregon.

In order to attribute the individual contributions of the predictors that are included in the combinations to the superior performance of the ADL–MIDAS models, one can look at the time-varying forecast combination weights since they are assigned relative to their historical forecast performance. Figures 2.1 and 2.2 summarize these forecast combinations weights over two particular time periods. Maps on the first column visualize these weights for the period of 2002-2007 and those on the second column are from 2008-2011. Each map shows the forecast combination weights for the variable of interest for each state averaged over the corresponding time period. The first time period, 2002-2007, is chosen in order to analyze weights when the national economy is not in a recession. The second period from 2008-2011, covering the latest recessionary period of 2008-2009, is taken to examine the differences in the recessionary environment. As the color over a state gets darker, the forecast combination weight of the predictor depicted on the map gets larger for the budget series forecast of the respective state on a scale from 0 to 30%. For instance, the map on the third panel of Figure 2.1 presents weights for the ADL–MIDAS forecasts obtained by using GDP as the predictor for the state government expenditures. For example, the figure on the left indicates that the average share of the GDP from 2002-2007 for the expenditure forecast of Wyoming is above 25% while it declines to 10% during the recessionary period of 2008-2011 as depicted by the corresponding figure on the right. Since these combination weights for a forecast sum up to one, there has to be about a 15% increase on the weights of other predictors in the combination. In fact, the first two panels on the same figure show a 5% increase for the average weight of the federal funds rate (FFR) from 10% to 15% and a 10% increase for the federal government budget deficit from 5% to 15%. It should be noted that the shares of the federal funds rate and the federal government budget deficit as predictors of the state government expenditure forecasts increase for almost all states in the second period, which covers the latest recession. These two series can be considered, respectively, as instruments for monetary and fiscal policies. Hence, it could be informative to analyze possible responses of the state budget series to the shocks to these instruments in the VAR setting. Associated results are summarized in the following section.

The empirical findings in this part support the use of mixed frequency data sampling methods in order to utilize additional information contained in national macroeconomic indicators of higher sampling frequency in forecasting annual state and local governments' budget series.

2.4.2 Multiple–Equation Regression Results

In this section, the point forecasts obtained from the MF–BVAR and BVAR models are empirically compared. The forecasts are computed recursively using the posterior mean of the parameters. As in the single equation regression model portion of this paper, a real-time experiment is conducted over the pseudo out-of sample period, then out-of-sample forecast accuracy is measured in terms of RMSFEs. The results are reported relative to the benchmark RW model. Table 2.3 presents these relative RMSFEs for the forecast horizons of $h = 1, 2$ years.

For the majority of states, the MF–BVAR forecasts of the budget series outperform their BVAR counterparts as assessed by smaller relative RMSFEs. This result indicates that adding extra information with the use of higher frequency data helps improve the forecasts for the two low frequency state and local government budget series for almost all states and across all horizons. It should, however, be noted that some of the states' expenditure forecasts do not improve over the simple random walk predictions, especially beyond the first year; however, the MF–BVAR forecasts for all revenue series are better than RW forecasts. This lower performance for the expenditure series relative to the RW model can be attributed to the states' budget process. Since states are required to run balanced budgets, attempts are made to keep expenditures below revenue estimations, which makes expenditures somewhat stable across fiscal years.

In this exercise, there are only four one-year-ahead forecasts for each of the budget series per state; hence, they are not sufficient for statistical testing. Nevertheless, a panel version of the Diebold Mariano, or DM, test (Diebold and Mariano (1995)), developed by Pesaran, Schuermann, and Smith (2008), can be utilized in order to statistically test MF–BVAR forecasts against the forecasts from the benchmark BVAR model for a given budget series across all states. Peseran et al. indicate that it is possible to carry out the panel DM test by pooling forecast errors for the same variable across different individuals, provided that it is appropriately adapted to take account of the panel nature of the pooled series. Table 2.5 presents these test statistics for state government expenditures and revenues, which are obtained by pooling forecast errors across all states. Although not statistically significant, negative test statistics imply that the MF–BVAR models for expenditures and revenues provide forecast gains over the traditional BVAR models, indicating the benefit of employing quarterly national series instead of their annual counterparts when forecasting state governments' budgets.

In addition to the natural ordering in terms of release time, for the structural analysis, the order that is subject to similar considerations as in traditional VAR models is also taken into account. Because, the national series are observed at the same high/quarterly frequency. Monetary and fiscal policy shocks are simultaneously identified by means of a recursive/Choleski identification scheme where the high frequency variables are categorized as slow and fast moving following Christiano, Eichenbaum, and Evans (1999) and Bernanke, Boivin, and Eliasz (2005). The assumption is that the slow-moving variables, such as federal government budget deficit, industrial production, GDP and inflation, do not contemporaneously respond to a shock to the fast moving variable, e.g., monetary policy instrument such as federal funds rate. Low frequency/annual state specific government budget series are assumed to respond to the changes in the high frequency/quarterly national variables after the information over all quarters is arrived; hence, these variables are ordered last. On the other hand, quarterly national series do not respond to the changes in annual state-specific variables sooner than a year. Similarly, the most common identifying assumption in the fiscal policy literature is the ordering restriction; hence, to identify federal deficit shocks at the high/quarterly frequency, federal deficit is ordered first among high frequency national series (following Blanchard and Perotti (2002) and Fatás and Mihov (2001), who assume that federal government spending does not contemporaneously react to macroeconomic variables). Owyang and Zubairy (2013) also simultaneously identify fiscal and monetary policy shocks, federal spending and federal funds rate as the instruments, respectively, via ordering restrictions. They order the federal spending first and the federal funds rate last having state-specific variables in between, all at the same quarterly frequency.

The interest here is on the individual low frequency/annual responses of state and local government expenditures and revenues to the high frequency/quarterly shocks to the federal deficit and federal funds rate. For comparison, the responses to these shocks' low frequency/annual counterparts that are obtained from the traditional BVAR model are also presented in Figures 2.7, 2.8, 2.9 and 2.10. It is important to note that, except for relatively small differences in magnitude, the shapes of the responses of both variables to either high or low frequency shocks are similar. The responses to the 1-, 2-, 3- and 4-quarter federal funds rate shocks could be considered as the average of these four shocks to obtain the aggregate annual shock, while the sum of the four quarterly federal deficit shocks could be taken into account to obtain the annual version, depending on the stock or flow nature of the variable of interest. Figures 2.3, 2.4, 2.5, and 2.6 depict the annual point responses of state and local government revenues and expenditures to the quarterly monetary and fiscal shocks. Although the magnitude and timing of the responses vary across states, the typical response of revenues to the fiscal shock is positive while that of expenditures is oscillating and, on average, four times smaller in magnitude compared to the revenue responses. This can be explained by the fact that, in general, states plan their expenditure budget based on their revenue estimates. They can make budget revisions to reduce expenditures whenever it appears that actual revenues for a fiscal period fall below the estimates (i.e., expenditures are targeted to stay below the realized revenues via interventions whenever possible). Hence, expenditure responses to the shocks should be more volatile than revenue responses, but smaller in magnitude. Similarly, the common response of revenues to the monetary shock is positive for the first year, negative for the following two/three years, and then dies out afterwards. The expenditure responses are negative for the first three years.

When it comes to the assessment of the differences in responses to quarterly shocks, it is important to note that Figures 2.4 and 2.4 show that the responses to the first two quarters' shocks are greater in magnitude than those to the shocks that hit after the third quarter of the corresponding year. This could be attributed to the timing of the fiscal year for the state and local governments, which ends with the second calendar quarter for the majority of states.[21] For example, if the national economy is hit by a monetary shock in the first quarter of 2010, running the third quarter of the 2010 fiscal year, these states are not likely to control expenses that increased due to the shock from the previously planned budget. However, when considering the budget for the next fiscal year, which starts just a quarter later, expenditures can be reduced accordingly. On the contrary, if the shock hits in the fourth quarter of 2010, running the second quarter of the fiscal year, states tend to have revisions on the previously planned budget to

[21] The fiscal year ends on June 30^{th} for all states except Alabama, Michigan (both end on September 30^{th}), Texas (ends on August 31^{st}), and New York (ends on March 31^{st}).

take expenditures under control.[22][23] With the exception of the fourth quarter that results in a smaller magnitude of responses, high frequency fiscal shocks are similar in received responses. This implies that it does not matter in which quarter a fiscal shock hits the national economy for the individual annual responses of state and local government revenues. It can be concluded that high frequency shocks are better than those of low frequency in capturing the dynamics of the responses of the annual state government budget series. Because the responses to low frequency/annual monetary and fiscal shocks are greater in magnitude and exhibit no variation in dynamics. These results may be misleading since the reaction could be smaller and/or even veer in a different direction, depending on the quarter during which the shock is observed.

It is also worth looking into the relative forecast performance of the two proposed forecasting procedures, namely the single equation regression model, ADL–MIDAS, and the multiple equation regression model, MF–BVAR. To carry out this comparison, ADL–MIDAS forecasts that are included in the forecast combinations are chosen so as to obtain their equivalent counterparts, which are located at the last two equations of the MF–BVAR model (equation 2.2.5). Moreover, since the out-of-sample forecasting period for the ADL–MIDAS regressions is longer, it is set to match with the MF-BVARs. Table 2.4 presents the relative RMSFEs of these two models across all states. It can be concluded that there is no single winner across states; the forecast combinations of the ADL–MIDAS regressions provide relative forecast gains over the MF–BVAR model for the expenditure series while, for the revenue series, the overall forecast performance of the MF–BVAR model is better than that of the ADL–MIDAS regressions. A similar result is also obtained in Kuzin, Marcellino, and Schumacher (2011), which compares forecasting performance of MIDAS regressions and MF–VARs estimated by maximum likelihood approach. The authors similarly find that these models' predictive ability varies depending on the predictors and forecasting horizon.

---

[22]For states starting their fiscal year on July 1^{st}, the budget process starts around the third calender quarter of the previous year with strategic planning of state agencies. The budget is then proposed by the governor in December, and, during the second quarter, it is passed by the legislature and signed by the governor by the end of that quarter. The budget becomes ready to for enactment on July 1^{st}.

[23]U.S. Census Bureau, State and Local Government Finances Summary Report 2010: State and Local governments' revenue consists of about 25% federal support, 50% tax receipts, 17% charges, and 8% other while about 30% of state and local government expenditures go to education, 15% to public welfare, 12% to insurance trust expenditures, 7% to utility expenditures, 5% to highways, 5% to hospitals and remaining funds to other expenditures.

2.5 Conclusion

This chapter considers the problem of forecasting state and local governments' expenditures and revenues by utilizing econometric forecasting methods. Two mixed-frequency data regression approaches are proposed for forecasting purposes, namely single-equation and system of equations approaches. Using a sample of the 48 mainland U.S. states, a forecasting procedure is proposed that involves a simple mixed frequency data regression approach, combinations of Augmented Distributed Lag–Mixed Data Sampling (ADL–MIDAS) regression models, and examines its relative forecasting performance against traditional benchmark models. The empirical results indicate that for almost all 48 states, the use of higher frequency state-specific and national variables combined with low frequency budget series provides forecast performance gains over the traditional models when all data are of the same low/annual sampling frequency. The second empirical portion of the paper proposes a forecasting procedure with a multiple equation regression model, namely a Mixed Frequency–Bayesian VAR (MF–BVAR) model, and assesses its predictive ability against a traditional low frequency Bayesian VAR (BVAR) model. It is found that the overall empirical forecast performance of the MF–BVAR is better than its counterpart, the BVAR model, as measured by root mean squared forecast errors relative to the random walk (RW) model. Finally, empirical predictive abilities of the two proposed forecasting procedures are compared and the results suggest that one cannot be chosen over the other. The ADL–MIDAS model provides better forecasts for the expenditure series across states whereas forecasts for the revenue series are more accurately obtained with the help of the MF–BVAR model.

Table 2.1: Data Sets

Title	*Frequency*
State-Specific Data Series	
State and Local Government Total Revenues	Annual
State and Local Government Total Expenditures	Annual
State personal income	Quarterly
National Data Series	
Real GDP	Quarterly
Federal Government Budget Deficit	Quarterly
Effective federal funds rate	Monthly / Quarterly
Consumer price index for all urban consumers: All items	Monthly / Quarterly
Industrial production index	Monthly / Quarterly
Spot oil price: West Texas intermediate	Monthly / Quarterly
3-Month treasury bill: Secondary market rate	Monthly / Quarterly
10-Year Treasury Bond Rate	Monthly / Quarterly

Notes: This table lists the data series used in this paper. There are 48 sets of state and local governments' budget data series together with state-specific and national macroeconomic series, all are seasonally adjusted, real per capita quantities, and transformed to ensure stationarity if necessary. The data set is of real-time vintages, i.e., at each point in time, the vintages of data used for estimation are restricted to those that would have been available at that time. In the VAR setting, all data series except the annual state specific government budget data are at quarterly frequency. The quarterly data are in four-quarter change and the budget data series are in annual change, all expressed as percentage points. The state government budget series are obtained from the U.S. Census Bureau. The national macroeconomic series are from FRED (Federal Reserve Economic Data) (http://research.stlouisfed.org/fred2/), ALFRED (ArchivaL Federal Reserve Economic Data) (http://alfred.stlouisfed.org/). Sample period includes the annual time period from 1958-2011, which was the longest period possible given the available predictors.

Table 2.2: Forecasts of State Government Expenditures and Revenues Using ADL-MIDAS, ADL and AR models. (Relative RMSFEs).

States	*Expenditures*						*Revenues*					
	ADL-MDS		ADL		AR		ADL-MDS		ADL		AR	
	h=1	h=2	h=1	h=2	h=1	h=2	h=1	h=2	h=1	h=2	h=1	h=2
Alabama	0.87	0.93	0.94	0.93	0.81	0.80	0.60	0.61	0.61	0.61	0.61	0.62
Arizona	0.72	0.70	0.71	0.71	0.73	0.84	0.79	1.16	0.92	0.98	0.93	0.91
Arkansas	0.60	0.64	0.65	0.65	0.63	0.70	0.70	0.78	0.72	0.76	0.74	0.77
California	0.88	0.93	0.88	0.97	0.98	1.11	0.71	0.74	0.72	0.80	0.73	0.78
Colorado	0.71	0.75	0.73	0.75	0.78	1.03	0.57	0.59	0.59	0.71	0.61	0.67
Connecticut	0.67	0.82	0.75	0.81	1.72	2.13	0.62	0.63	0.65	0.65	0.66	0.66
Delaware	0.61	0.64	0.62	0.69	0.70	0.89	0.72	1.24	0.92	0.86	0.75	0.80
Florida	0.71	0.82	0.77	0.77	0.77	0.80	0.72	0.78	0.74	0.73	0.74	0.76
Georgia	0.86	0.89	0.87	0.82	1.70	1.51	0.69	0.66	0.72	0.74	0.74	0.75
Idaho	0.71	0.76	0.76	0.78	1.14	1.87	0.64	0.71	0.66	0.69	0.67	0.69
Illinois	0.82	0.90	0.80	0.84	0.91	0.93	0.67	0.70	0.68	0.75	0.69	0.75
Indiana	0.79	0.92	0.85	0.86	0.90	0.87	0.84	0.98	0.87	0.93	0.86	0.86
Iowa	0.70	0.80	0.80	0.86	0.81	1.01	0.59	0.68	0.60	0.65	0.63	0.64
Kansas	0.80	0.81	0.78	0.83	0.82	0.92	0.59	0.79	0.63	0.74	0.61	0.65
Kentucky	0.77	0.81	0.80	0.83	0.88	1.29	0.78	0.92	0.80	0.83	0.79	0.80
Louisiana	0.84	0.86	0.84	0.92	0.89	1.07	0.71	0.69	1.00	0.76	0.73	0.74
Maine	0.72	0.83	0.78	0.88	1.55	1.15	0.76	0.83	0.78	0.81	0.80	0.83
Maryland	0.89	0.97	0.96	1.05	0.95	1.02	0.64	0.69	0.67	0.73	0.68	0.70
Massachusetts	0.72	0.82	0.82	0.83	0.77	0.78	0.65	0.71	0.69	0.74	0.72	0.69
Michigan	0.87	0.88	0.89	0.95	1.05	1.18	0.59	0.64	0.61	0.77	0.62	0.63
Minnesota	0.93	1.04	0.99	0.96	1.06	1.05	0.72	0.78	0.74	0.76	0.74	0.76
Mississippi	0.75	0.74	0.72	0.72	0.86	0.83	0.59	0.68	0.60	0.59	0.59	0.60
Missouri	0.76	0.82	0.80	0.80	1.67	1.36	0.63	0.65	0.65	0.67	0.67	0.69
Montana	0.59	0.69	0.61	0.65	0.90	1.34	0.59	0.65	0.60	0.65	0.61	0.65
Nebraska	0.80	0.78	0.77	0.79	0.85	0.84	0.66	0.66	0.53	0.71	0.69	0.71
Nevada	0.54	0.57	0.55	0.53	0.55	0.50	0.75	0.85	0.86	0.80	0.78	0.78
New Hampshire	0.59	0.68	0.43	0.77	1.97	1.47	0.70	0.79	0.74	0.71	0.72	0.70

Continued on next page

Table2.2 – continued from previous page

States	*Expenditures*						*Revenues*					
	ADL-MDS		ADL		AR		ADL-MDS		ADL		AR	
	h=1	h=2	h=1	h=2	h=1	h=2	h=1	h=2	h=1	h=2	h=1	h=2
New Jersey	0.70	0.75	0.68	0.67	0.74	1.75	0.88	1.17	0.92	0.86	0.91	0.90
New Mexico	0.53	0.67	0.67	0.77	0.65	0.87	0.70	0.77	0.72	0.74	0.72	0.73
New York	0.74	0.77	0.74	0.84	1.96	2.56	0.68	0.74	0.71	0.75	0.76	0.79
North Carolina	1.29	1.36	1.38	1.36	1.52	1.47	0.60	0.75	0.65	0.76	0.67	0.73
North Dakota	1.26	1.29	1.33	1.35	1.27	1.97	0.58	1.63	0.55	0.65	0.63	0.64
Ohio	1.03	1.05	0.99	0.99	1.09	1.02	0.75	0.86	0.78	0.89	0.84	0.90
Oklahoma	0.67	0.70	0.70	0.71	0.73	0.86	0.55	0.67	0.62	0.65	0.59	0.59
Oregon	1.09	1.21	1.14	1.17	1.39	1.42	0.57	0.67	0.60	0.71	0.61	0.65
Pennsylvania	0.63	0.80	0.67	0.84	1.16	1.18	0.60	0.67	0.63	0.73	0.67	0.72
Rhode Island	0.60	0.65	0.59	0.64	0.68	0.75	0.59	0.65	0.62	0.69	0.66	0.68
South Carolina	0.86	0.99	1.00	0.99	1.00	0.86	0.70	0.83	0.71	0.75	0.73	0.73
South Dakota	0.56	0.73	0.66	0.71	0.79	1.91	0.77	0.87	0.82	0.79	0.78	0.79
Tennessee	0.65	0.67	0.64	0.63	0.64	0.79	0.70	0.78	0.77	0.82	0.78	0.81
Texas	0.75	0.90	0.87	0.97	1.67	1.09	0.69	0.69	0.71	0.75	0.73	0.75
Utah	0.65	0.76	0.72	0.70	0.70	1.01	0.50	0.61	0.51	0.64	0.55	0.60
Vermont	0.52	0.58	0.54	0.68	0.59	0.89	0.65	0.71	0.67	0.73	0.68	0.71
Virginia	0.70	0.83	0.82	0.83	0.83	0.93	0.66	0.68	0.67	0.73	0.70	0.73
Washington	0.76	0.82	0.76	0.77	0.81	1.41	0.65	0.82	0.67	0.76	0.68	0.71
West Virginia	0.98	1.02	0.96	0.96	1.04	1.04	0.82	1.04	0.86	0.86	0.85	0.83
Wisconsin	0.92	0.97	0.95	0.99	1.31	1.27	0.57	0.61	0.59	0.66	0.61	0.65
Wyoming	0.90	1.06	0.88	1.42	1.04	1.75	0.79	0.95	0.82	0.96	0.81	0.81
q25	0.66	0.73	0.69	0.72	0.77	0.87	0.60	0.67	0.62	0.70	0.65	0.66
Median	0.74	0.82	0.78	0.83	0.89	1.03	0.66	0.73	0.69	0.74	0.71	0.73
q75	0.86	0.93	0.88	0.95	1.15	1.39	0.72	0.83	0.77	0.79	0.76	0.79

Notes: Numbers are root mean squared forecast errors, or RMSFEs, relative to RW model, for the corresponding forecast horizon. $h = 1$ represents the one-year-ahead and $h = 2$ denotes the two-year-ahead forecast horizon. *q25* and *q75* are the first and third quartiles, respectively. Out-of-sample period is 1999–2011 while the in-sample estimation covers the 1958–1998 period. Rolling window forecasts are obtained.

Table 2.3: Forecasts of State Government Expenditures and Revenues Using MF-BVAR and BVAR models. (Relative RMSFEs).

States	*Expenditures*				*Revenues*			
	MF-BVAR		BVAR		MF-BVAR		BVAR	
	h=1	h=2	h=1	h=2	h=1	h=2	h=1	h=2
Alabama	0.69	0.68	0.63	0.48	0.64	0.34	0.74	0.45
Arizona	0.77	0.81	1.07	1.57	0.82	1.20	0.88	1.31
Arkansas	0.85	1.14	0.89	1.37	0.77	0.63	0.80	0.64
California	0.39	0.75	0.37	0.78	0.77	0.59	0.82	0.70
Colorado	0.67	0.63	0.61	0.65	0.58	0.74	0.78	0.92
Connecticut	0.87	0.39	1.52	1.08	0.68	0.60	0.76	0.65
Delaware	0.91	0.82	1.14	1.78	0.71	0.54	0.79	0.59
Florida	1.19	2.14	1.72	2.16	0.73	0.59	0.80	0.64
Georgia	0.83	0.79	1.00	1.07	0.71	0.60	0.82	0.62
Idaho	0.95	0.65	0.99	0.59	0.68	0.43	0.74	0.50
Illinois	1.10	1.09	1.08	1.08	0.68	0.66	0.81	0.82
Indiana	1.12	1.43	1.35	1.51	0.72	0.46	0.79	0.56
Iowa	0.81	0.56	0.89	0.61	0.75	0.63	0.80	0.67
Kansas	0.84	0.61	0.55	0.27	0.63	0.48	0.77	0.64
Kentucky	1.25	1.32	1.70	1.33	0.75	0.45	0.81	0.55
Louisiana	0.89	0.54	0.87	0.65	0.61	0.46	0.74	0.57
Maine	0.62	0.48	1.42	1.04	0.67	0.50	0.76	0.58
Maryland	1.06	1.06	1.40	1.21	0.70	0.53	0.77	0.63
Massachusetts	0.87	0.85	0.92	0.81	0.88	0.86	0.89	0.87
Michigan	0.94	1.26	0.57	0.76	0.64	0.36	0.72	0.45
Minnesota	0.92	0.63	0.76	0.48	0.72	0.54	0.79	0.65
Mississippi	0.87	0.64	1.13	1.02	0.71	0.41	0.77	0.45
Missouri	0.92	0.81	1.31	1.04	0.71	0.59	0.81	0.69
Montana	0.77	1.82	0.50	1.11	0.65	0.62	0.74	0.69
Nebraska	1.11	1.09	1.03	0.94	0.60	0.55	0.83	0.79
Nevada	1.08	1.51	0.88	1.08	0.77	0.56	0.82	0.58
New Hampshire	0.99	1.17	0.80	0.66	0.60	0.53	0.74	0.70

Continued on next page

Table2.3 – continued from previous page

States	*Expenditures*				*Revenues*			
	MF-BVAR		BVAR		MF-BVAR		BVAR	
	h=1	h=2	h=1	h=2	h=1	h=2	h=1	h=2
New Jersey	0.77	0.97	0.70	0.76	0.93	0.38	0.85	0.35
New Mexico	0.89	0.99	1.30	1.25	0.72	0.56	0.76	0.59
New York	0.64	0.55	0.81	0.83	0.68	0.69	0.79	0.84
North Carolina	1.33	1.21	1.15	1.25	0.64	0.80	0.79	0.94
North Dakota	2.05	1.61	1.75	1.92	0.71	0.81	0.84	0.81
Ohio	0.33	0.27	1.49	1.02	0.71	0.73	0.83	0.89
Oklahoma	0.70	0.64	0.85	0.84	0.71	0.50	0.76	0.54
Oregon	0.91	0.63	1.13	0.73	0.74	0.58	0.79	0.68
Pennsylvania	0.48	0.42	0.76	0.71	0.64	0.68	0.79	0.83
Rhode Island	0.90	1.73	1.87	1.98	0.68	0.60	0.75	0.67
South Carolina	1.07	0.87	1.07	1.33	0.71	0.58	0.79	0.69
South Dakota	0.94	0.90	1.47	1.24	0.77	0.49	0.78	0.51
Tennessee	0.71	1.41	0.52	0.65	0.71	0.52	0.80	0.62
Texas	0.80	0.64	0.50	0.24	0.53	0.58	0.75	0.75
Utah	0.93	0.33	0.92	0.40	0.60	0.67	0.75	0.80
Vermont	0.86	1.18	1.20	1.46	0.55	0.50	0.74	0.69
Virginia	1.31	1.37	1.39	1.45	0.68	0.53	0.78	0.68
Washington	0.62	0.31	0.71	0.37	0.66	0.51	0.79	0.65
West Virginia	0.79	0.40	0.74	0.32	0.73	0.43	0.75	0.49
Wisconsin	0.83	0.74	1.66	1.50	0.66	0.73	0.74	0.87
Wyoming	0.87	0.60	1.06	0.69	0.50	0.85	0.74	1.19
q25	0.77	0.62	0.76	0.66	0.64	0.50	0.75	0.58
Median	0.87	0.81	1.02	1.02	0.70	0.57	0.79	0.66
q75	0.97	1.18	1.33	1.29	0.72	0.65	0.80	0.79

Notes: Numbers are root mean squared forecast errors, RMSFEs, relative to RW model, for the corresponding forecast horizon. $h = 1$ represents the one-year-ahead and $h = 2$ denotes the two-year-ahead forecast horizon. *q25* and *q75* are the first and third quartiles, respectively. Out-of-sample period is $2008 - 2011$ while the in-sample estimation covers the $1958 - 2007$ period. Forecasts are obtained recursively.

Table 2.4: Forecasts of State Government Expenditures and Revenues Using MF–BVAR and ADL–MIDAS models. (Relative RMSFEs).

States	*Expenditures*				*Revenues*			
	MF–BVAR		ADL–MDS		MF–BVAR		ADL–MDS	
	h=1	h=2	h=1	h=2	h=1	h=2	h=1	h=2
Alabama	0.69	0.68	0.63	1.02	0.64	0.34	0.64	0.60
Arizona	0.77	0.81	0.80	0.80	0.82	1.20	1.17	1.12
Arkansas	0.85	1.14	0.93	0.90	0.77	0.63	1.30	1.17
California	0.39	0.75	1.24	0.54	0.77	0.59	1.36	1.27
Colorado	0.67	0.63	0.59	1.16	0.58	0.74	0.92	0.73
Connecticut	0.87	0.39	0.66	0.94	0.68	0.60	1.03	0.78
Delaware	0.91	0.82	1.85	0.63	0.71	0.54	1.28	1.03
Florida	1.19	2.14	0.53	0.30	0.73	0.59	1.20	1.10
Georgia	0.83	0.79	0.90	0.85	0.71	0.60	0.89	0.67
Idaho	0.95	0.65	0.73	0.80	0.68	0.43	0.98	0.95
Illinois	1.10	1.09	0.42	1.10	0.68	0.66	1.29	1.20
Indiana	1.12	1.43	0.77	1.81	0.72	0.46	2.13	2.03
Iowa	0.81	0.56	0.80	1.31	0.75	0.63	0.95	0.91
Kansas	0.84	0.61	1.22	1.80	0.63	0.48	0.90	0.82
Kentucky	1.25	1.32	1.19	1.34	0.75	0.45	0.82	0.79
Louisiana	0.89	0.54	0.88	1.06	0.61	0.46	1.06	0.68
Maine	0.62	0.48	0.40	1.12	0.67	0.50	1.21	1.08
Maryland	1.06	1.06	0.89	1.95	0.70	0.53	0.95	0.80
Massachusetts	0.87	0.85	0.60	0.70	0.88	0.86	0.95	0.90
Michigan	0.94	1.26	0.90	1.64	0.64	0.36	0.61	0.59
Minnesota	0.92	0.63	0.92	1.72	0.72	0.54	1.05	0.96
Mississippi	0.87	0.64	0.39	0.48	0.71	0.41	0.71	0.64
Missouri	0.92	0.81	0.46	1.47	0.71	0.59	1.18	1.07
Montana	0.77	1.82	0.65	0.95	0.65	0.62	1.28	1.19
Nebraska	1.11	1.09	0.88	0.78	0.60	0.55	1.30	1.10
Nevada	1.08	1.51	0.13	0.57	0.77	0.56	1.13	1.04
New Hampshire	0.99	1.17	0.56	3.44	0.60	0.53	0.95	0.99

Continued on next page

Table2.4 – continued from previous page

States	*Expenditures*				*Revenues*			
	MF–BVAR		ADL–MDS		MF–BVAR		ADL–MDS	
	h=1	h=2	h=1	h=2	h=1	h=2	h=1	h=2
New Jersey	0.77	0.97	0.53	0.50	0.93	0.38	0.93	0.72
New Mexico	0.89	0.99	0.53	1.02	0.72	0.56	1.16	1.05
New York	0.64	0.55	0.57	0.94	0.68	0.69	1.33	1.30
North Carolina	1.33	1.21	0.95	2.67	0.64	0.80	0.99	0.90
North Dakota	2.05	1.61	1.79	7.10	0.71	0.81	0.46	0.66
Ohio	0.33	0.27	0.82	0.63	0.71	0.73	1.29	1.25
Oklahoma	0.70	0.64	0.62	0.90	0.71	0.50	0.69	0.68
Oregon	0.91	0.63	0.93	1.84	0.74	0.58	1.32	1.11
Pennsylvania	0.48	0.42	0.36	0.80	0.64	0.68	1.30	1.11
Rhode Island	0.90	1.73	1.32	1.14	0.68	0.60	1.25	1.08
South Carolina	1.07	0.87	0.49	2.19	0.71	0.58	1.06	0.83
South Dakota	0.94	0.90	0.70	2.88	0.77	0.49	0.93	0.87
Tennessee	0.71	1.41	0.61	0.91	0.71	0.52	1.10	0.85
Texas	0.80	0.64	0.67	1.28	0.53	0.58	1.18	1.12
Utah	0.93	0.33	0.88	1.65	0.60	0.67	0.93	0.98
Vermont	0.86	1.18	0.28	0.81	0.55	0.50	0.99	0.98
Virginia	1.31	1.37	0.31	0.71	0.68	0.53	1.07	0.94
Washington	0.62	0.31	0.47	1.60	0.66	0.51	1.17	1.09
West Virginia	0.79	0.40	0.35	0.52	0.73	0.43	0.84	0.81
Wisconsin	0.83	0.74	0.43	0.59	0.66	0.73	1.31	1.03
Wyoming	0.87	0.60	1.99	1.90	0.50	0.85	0.79	0.62
q25	0.77	0.62	0.51	0.79	0.64	0.50	0.93	0.79
Median	0.87	0.81	0.66	1.04	0.70	0.57	1.06	0.97
q75	0.97	1.18	0.91	1.82	0.72	0.65	1.26	1.10

Notes: Numbers are root mean squared forecast errors, RMSFEs, relative to RW model, for the corresponding forecast horizon. $h = 1$ represents the one-year-ahead and $h = 2$ denotes the two-year-ahead forecast horizon. *q25* and *q75* are the first and third quartiles, respectively. Out-of-sample forecast period is $2008 - 2011$ while the in-sample estimation covers the $1958 - 2007$ period and forecasts are obtained recursively. Note that the ADL–MIDAS model utilized here to compare with the MF–BVAR model is different from those employed in section 2.2.1 in that the out-of-sample period is shorter and the number of models combined are reduced to obtain comparable forecasts as the number of variables in MF–BVAR models is limited.

Table 2.5: Panel DM Statistics for MF–BVAR Forecasts of State Government Expenditures and Revenues Relative to the BVAR model.

Benchmark Model	*Expenditures*	*Revenues*
BVAR	-0.1315	-0.2423

Notes: $z_{sgt} = [e^A_{sgt}(1)]^2 - [e^B_{sgt}(1)]^2$ is the loss differential of forecasting the budget series g, either expenditures or revenues, in state s, using method A=MF-BVAR relative to method B=BVAR, where $e_{sgt}(1)$ is the one-year-ahead forecast error. Given g, z_{sgt}=α_{sg}+ϵ_{sgt}, H_0: α_s=0 vs. H_1: α_s<0 for some s. Under the null and assuming $\epsilon_{sgt} \sim iid(0, \sigma^2_{sg})$, the test statistic is $\overline{DM}$=$\bar{z}/\sqrt{V(\bar{z})} \sim N(0, 1)$.

Figure 2.1: Forecast Combination Weights on Expenditures

Forecast Combination Weights for State Government Expenditures: As the color over a state gets darker, the forecast combination weight of the predictor depicted on the map gets larger for the forecast of the budget series of the respective state on a scale from 0 to 30%. FF stands for the federal funds rate, DEF for the federal deficit and GDP for the gross domestic product.

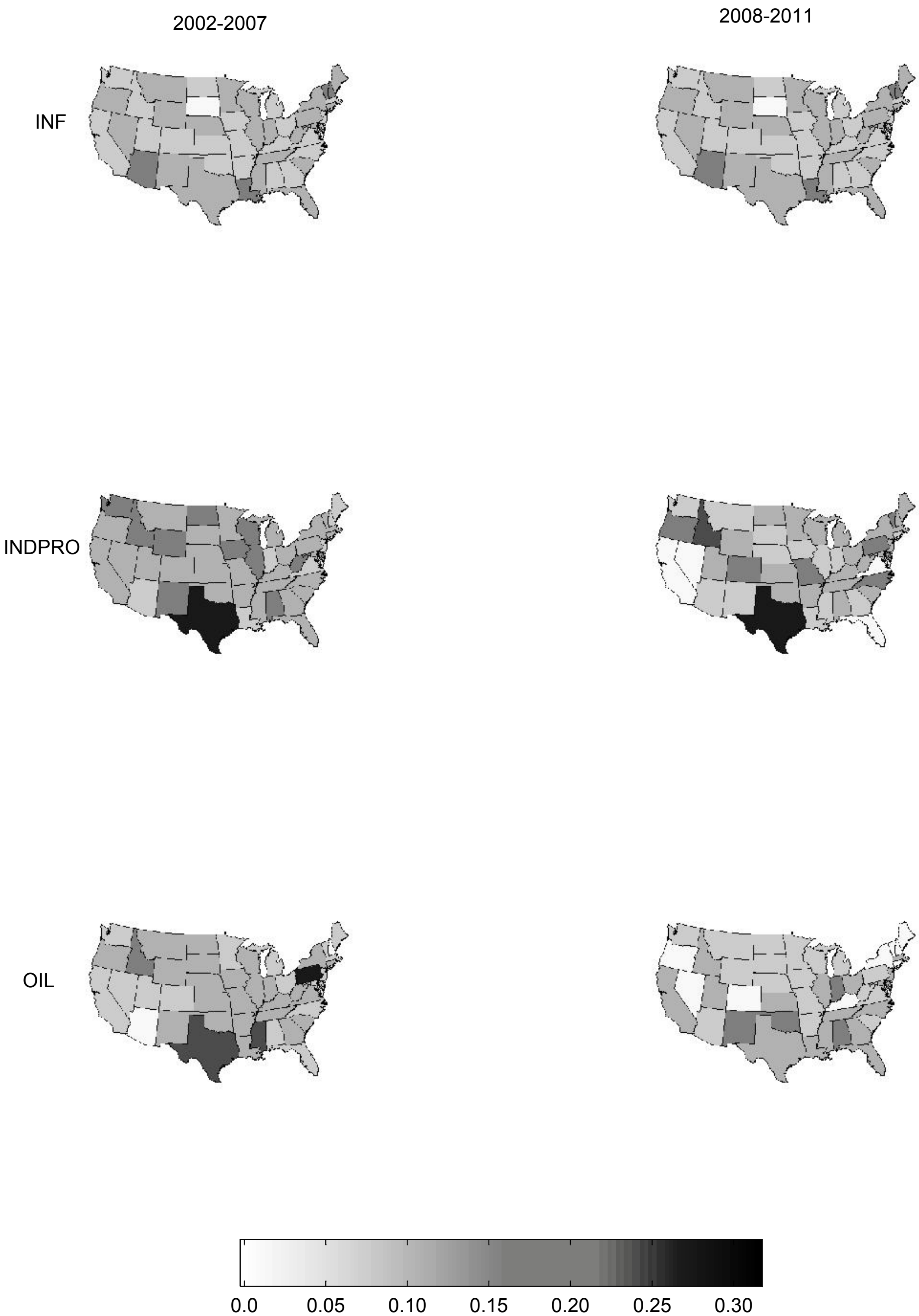


Figure 2.1: Forecast Combination Weights on Expenditures – continued figure

Forecast Combination Weights for State Government Expenditures: As the color over a state gets darker, the forecast combination weight of the predictor depicted on the map gets larger for the forecast of the budget series of the respective state on a scale from 0 to 30%. INF stands for the inflation, INDPRO for the industrial production, and OIL for the oil price.

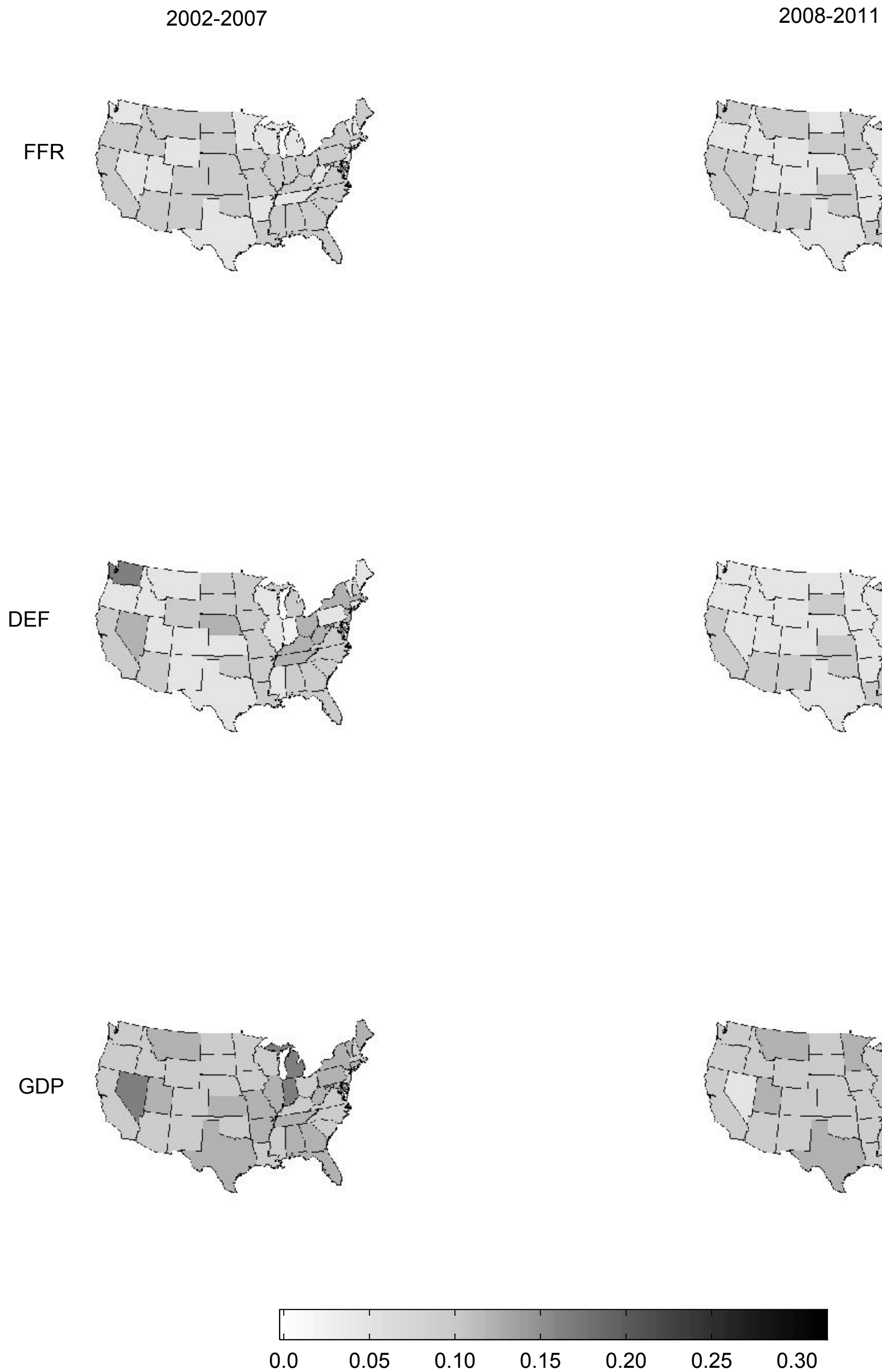


Figure 2.2: Forecast Combination Weights on Revenues

Forecast Combination Weights for State Government Revenues: As the color over a state gets darker, the forecast combination weight of the predictor depicted on the map gets larger for the forecast of the budget series of the respective state on a scale from 0 to 30%. FF stands for the federal funds rate, DEF for the federal deficit and GDP for the gross domestic product.

Figure 2.2: Forecast Combination Weights on Revenues – continued figure

Forecast Combination Weights for State Government Revenues: As the color over a state gets darker, the forecast combination weight of the predictor depicted on the map gets larger for the forecast of the budget series of the respective state on a scale from 0 to 30%. INF stands for the inflation, INDPRO for the industrial production, and OIL for the oil price.

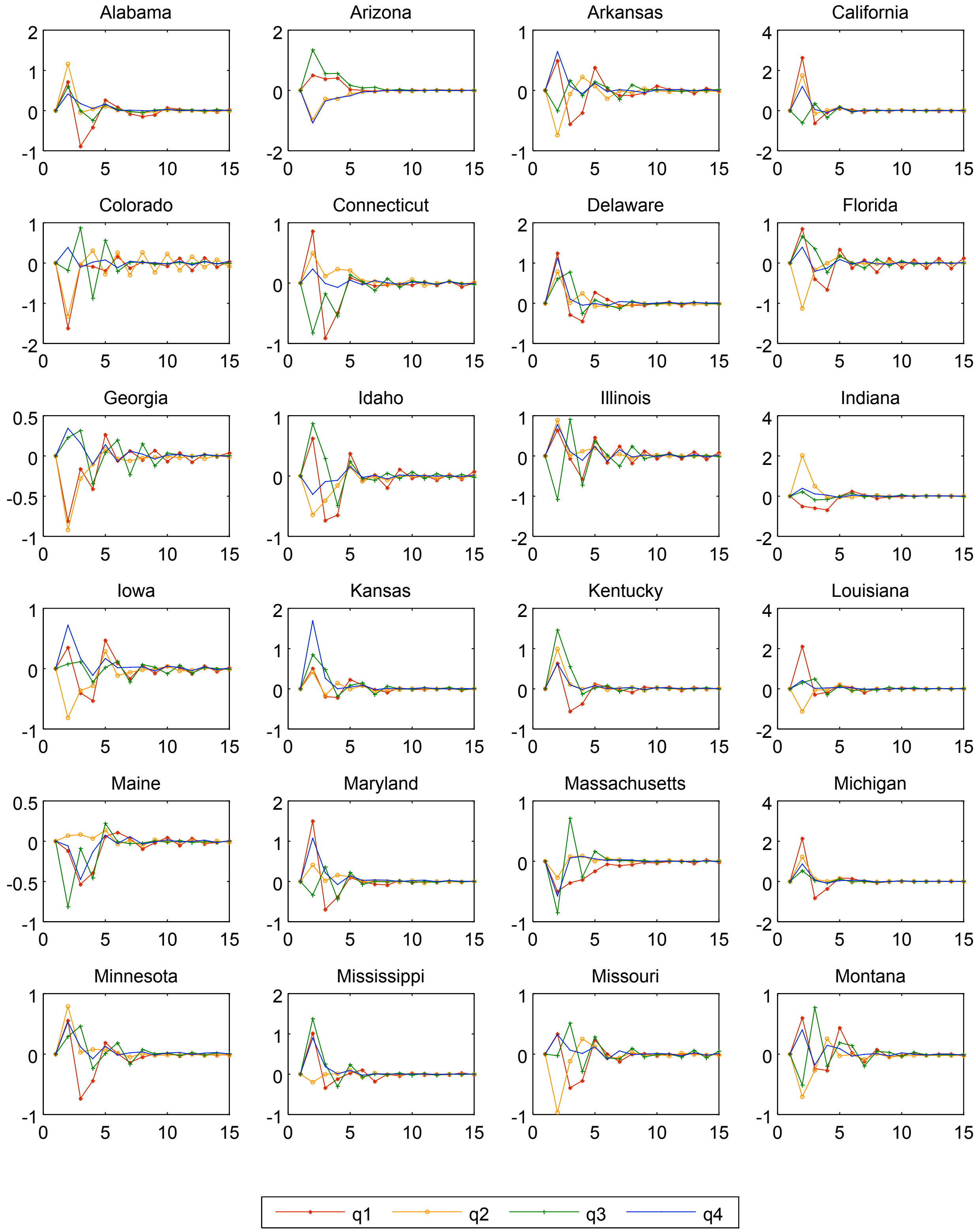


Figure 2.3: Impulse Responses of Revenues to Monetary Shocks

Impulse Responses of State Government Revenues to the Shocks to Federal Funds Rate: 1, 2, 3 and 4 quarter shocks from the MF-BVAR model.

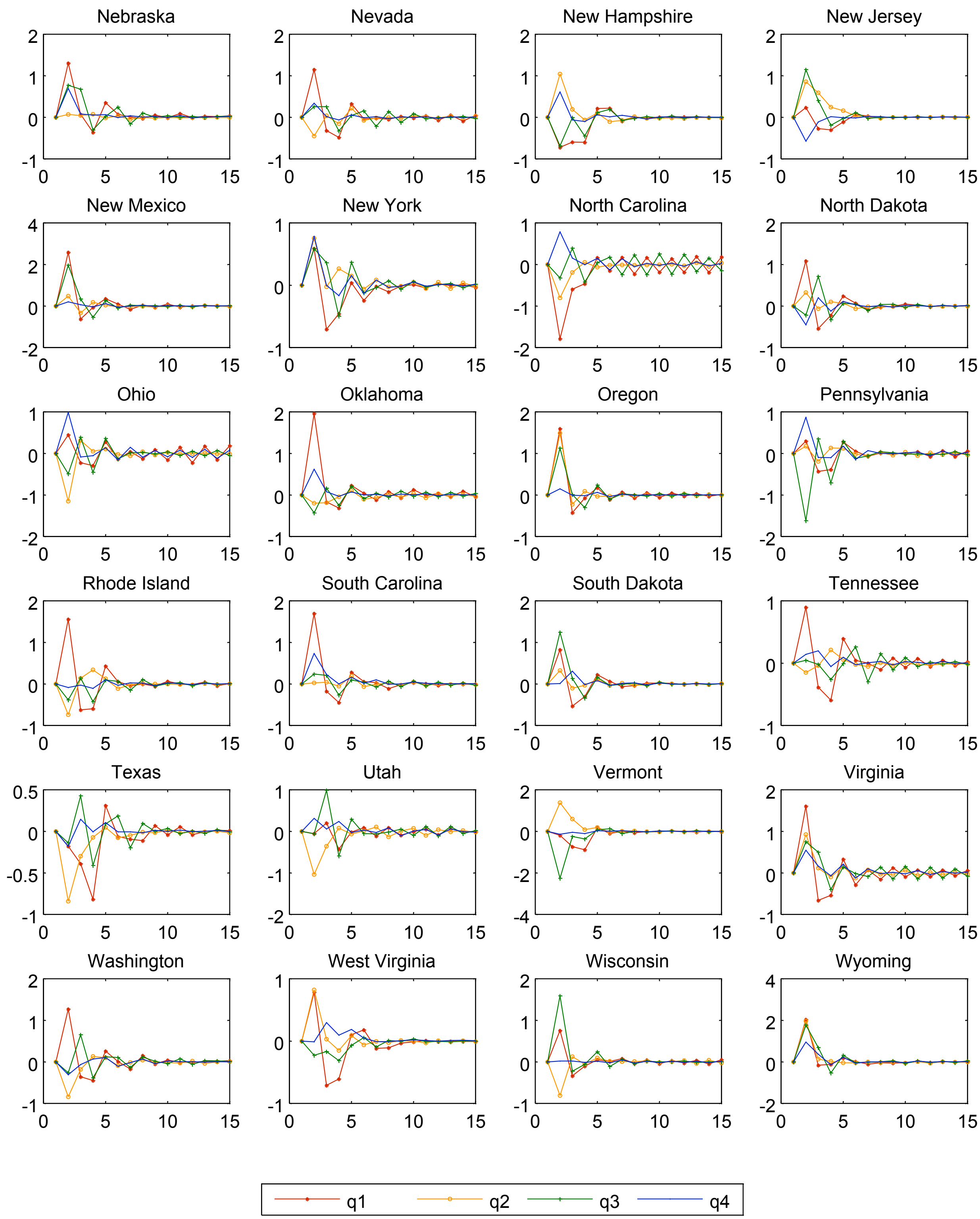


Figure 2.3: Impulse Responses of Revenues to Monetary Shocks – continued figure

Impulse Responses of State Government Revenues to the Shocks to Federal Funds Rate: 1, 2, 3 and 4 quarter shocks from the MF-BVAR model.

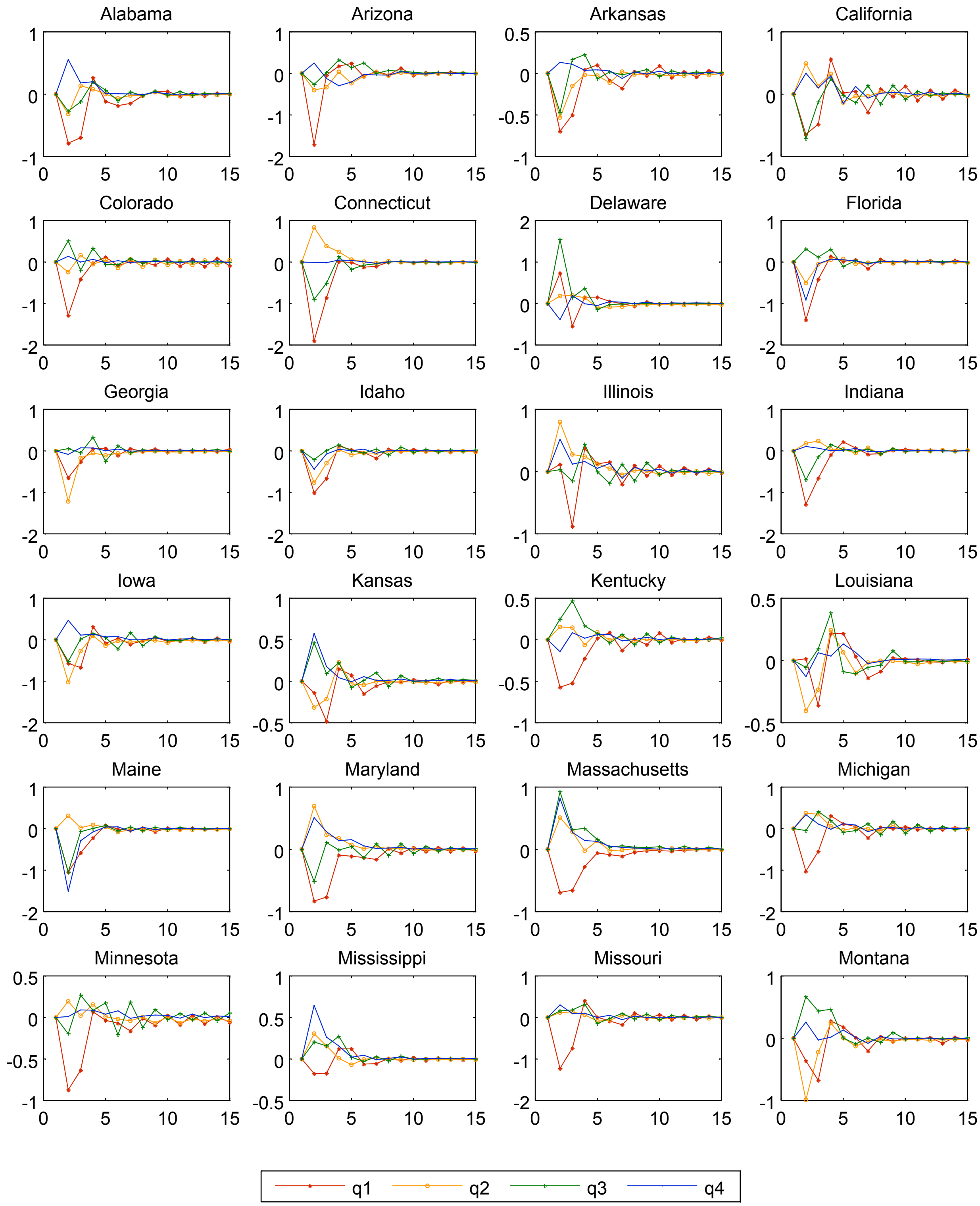


Figure 2.4: Impulse Responses of Expenditures to Monetary Shocks

Impulse Responses of State Government Expenditures to the Shocks to Federal Funds Rate: 1, 2, 3 and 4 quarter shocks from the MF-BVAR model.

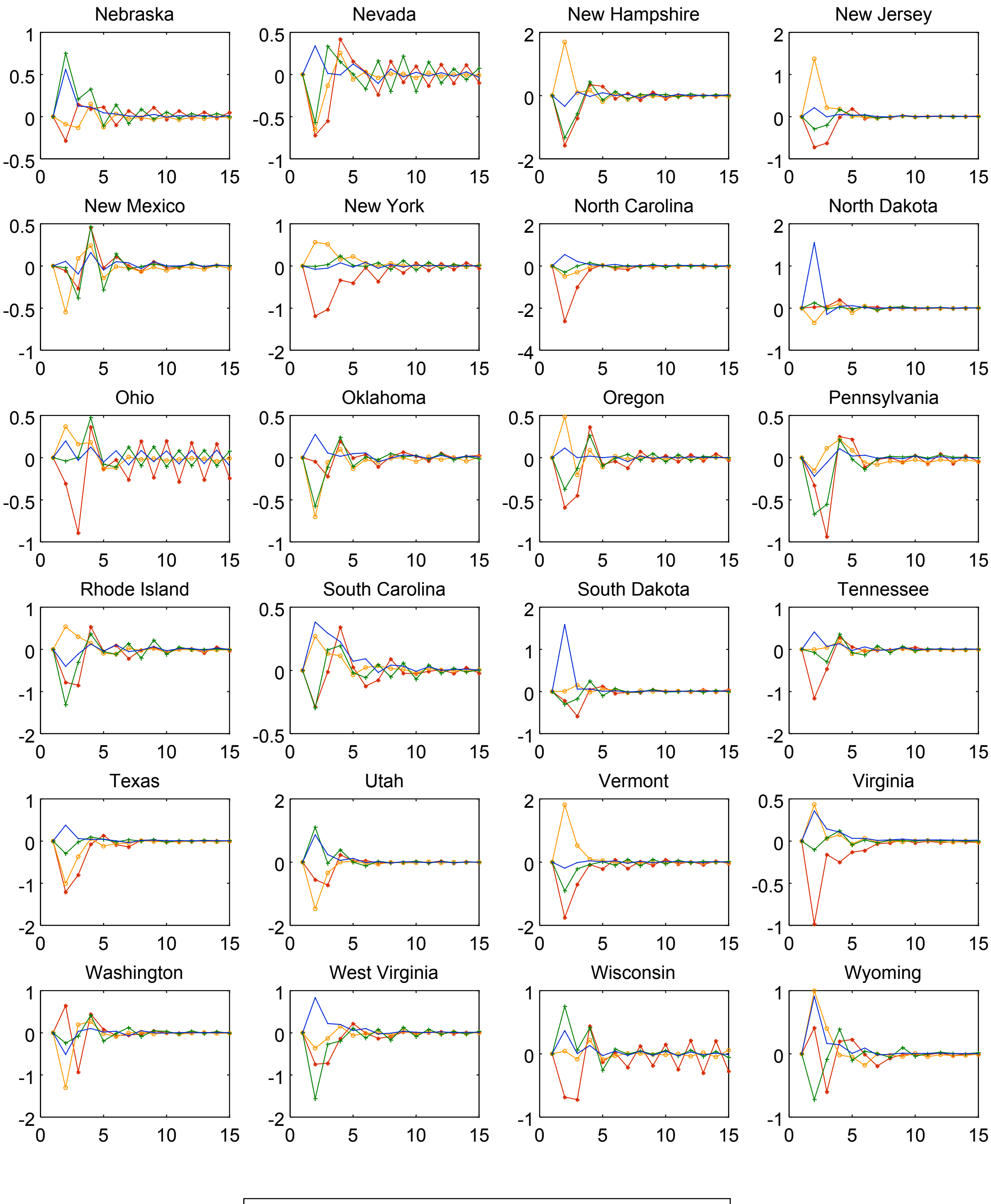


Figure 2.4: Impulse Responses of Expenditures to Monetary Shocks – continued figure

Impulse Responses of State Government Expenditures to the Shocks to Federal Funds Rate: 1, 2, 3 and 4 quarter shocks from the MF-BVAR model.

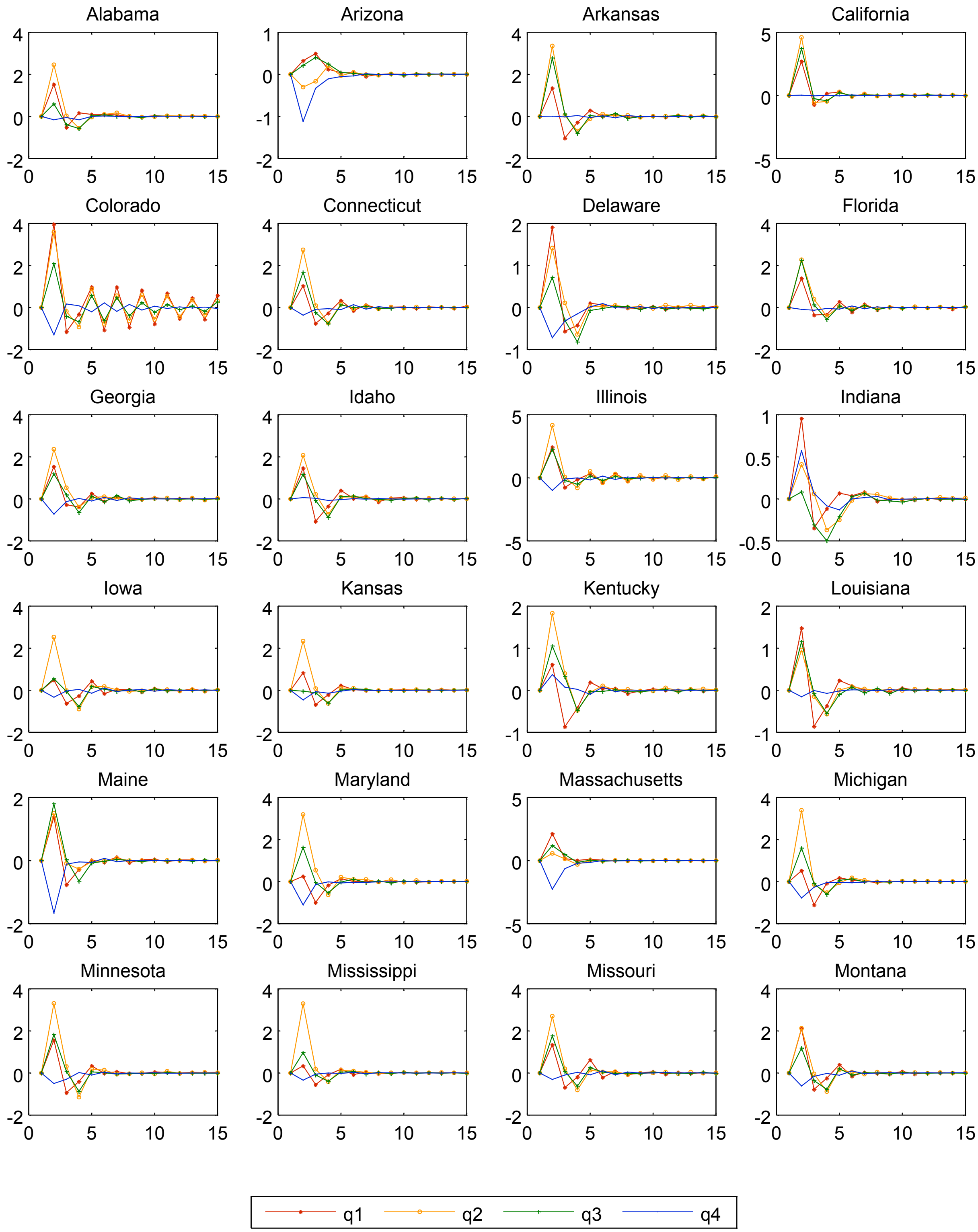


Figure 2.5: Impulse Responses of Revenues to Fiscal Shocks

Impulse Responses of State Government Revenues to the Shocks to Federal Deficit: 1,2,3,and 4 quarter shocks from the MF-BVAR model.

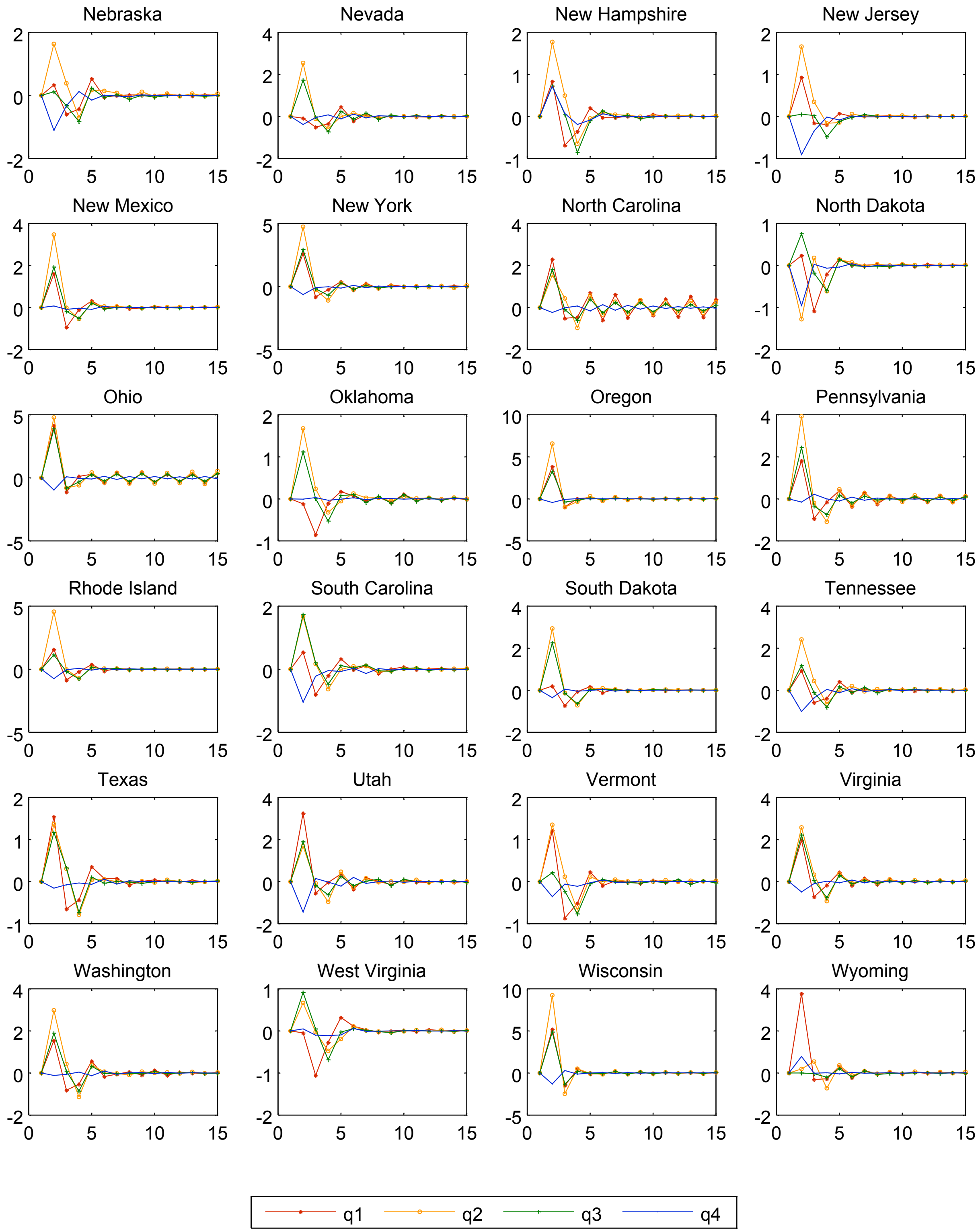


Figure 2.5: Impulse Responses of Revenues to Fiscal Shocks – continued figure

Impulse Responses of State Government Revenues to the Shocks to Federal Deficit: 1, 2, 3 and 4 quarter shocks from the MF-BVAR model.

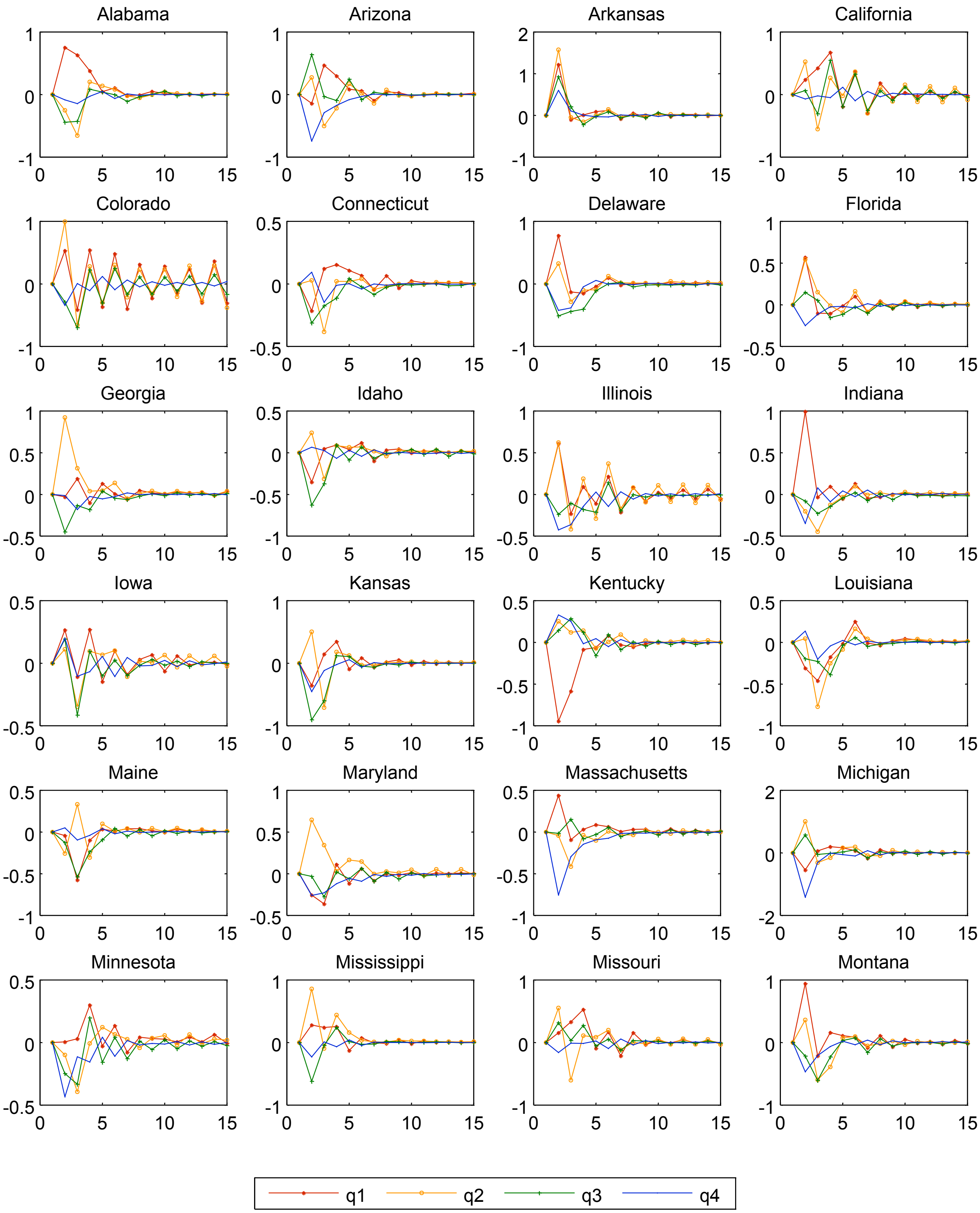


Figure 2.6: Impulse Responses of Expenditures to Fiscal Shocks

Impulse Responses of State Government Expenditures to the Shocks to Federal Deficit: 1,2,3,and 4 quarter shocks from the MF-BVAR model.

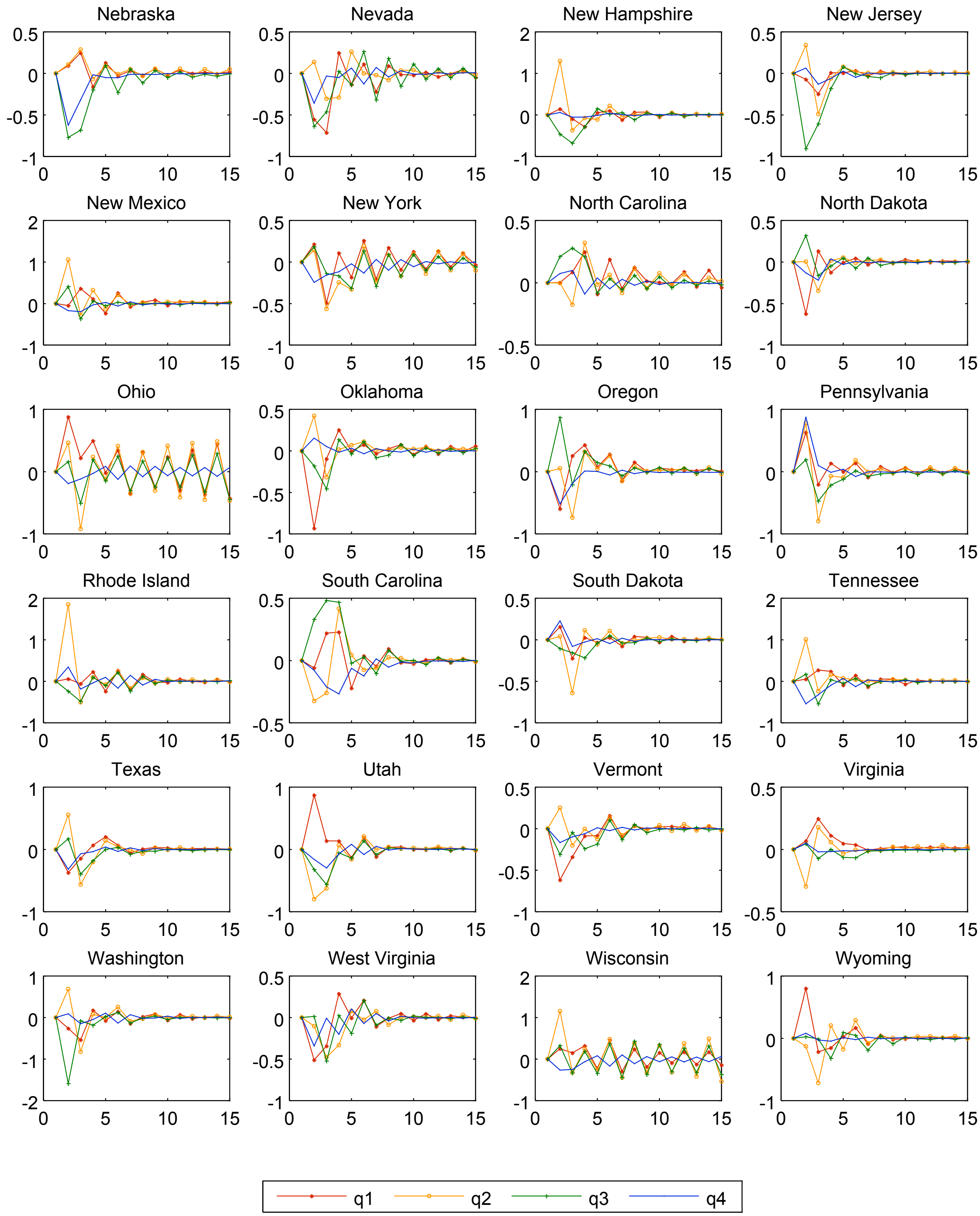


Figure 2.6: Impulse Responses of Expenditures to Fiscal Shocks – continued figure

Impulse Responses of State Government Expenditures to the Shocks to Federal Deficit: 1,2,3,and 4 quarter shocks from the MF-BVAR model.

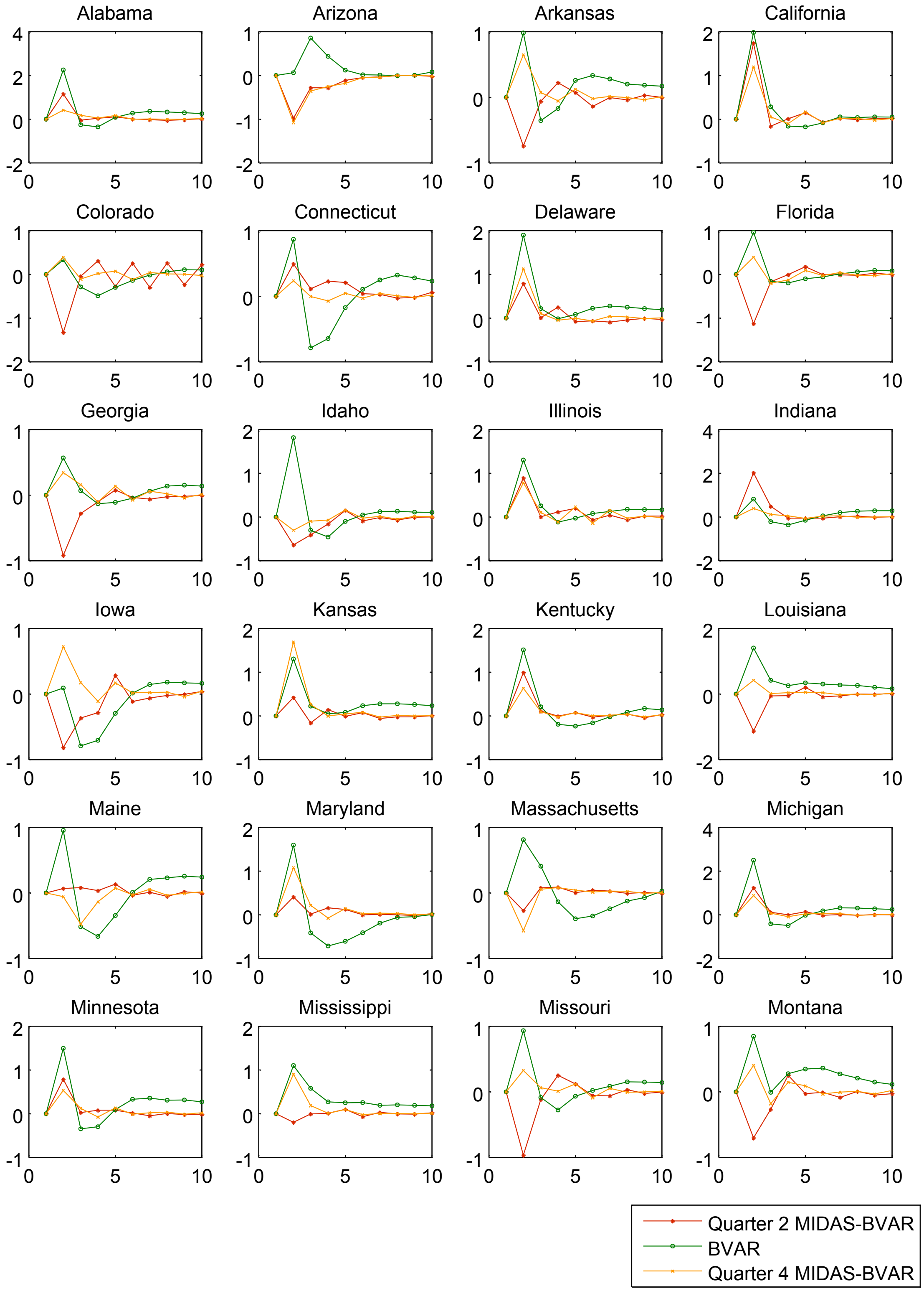


Figure 2.7: Impulse Responses of Revenues to Monetary Shocks: MF-BVAR versus BVAR

Impulse Responses of State Government Revenues to the Shocks to Federal Funds Rate: Second and Fourth Quarter Shocks from the MF-BVAR model versus aggregate annual shocks from the annual BVAR model.

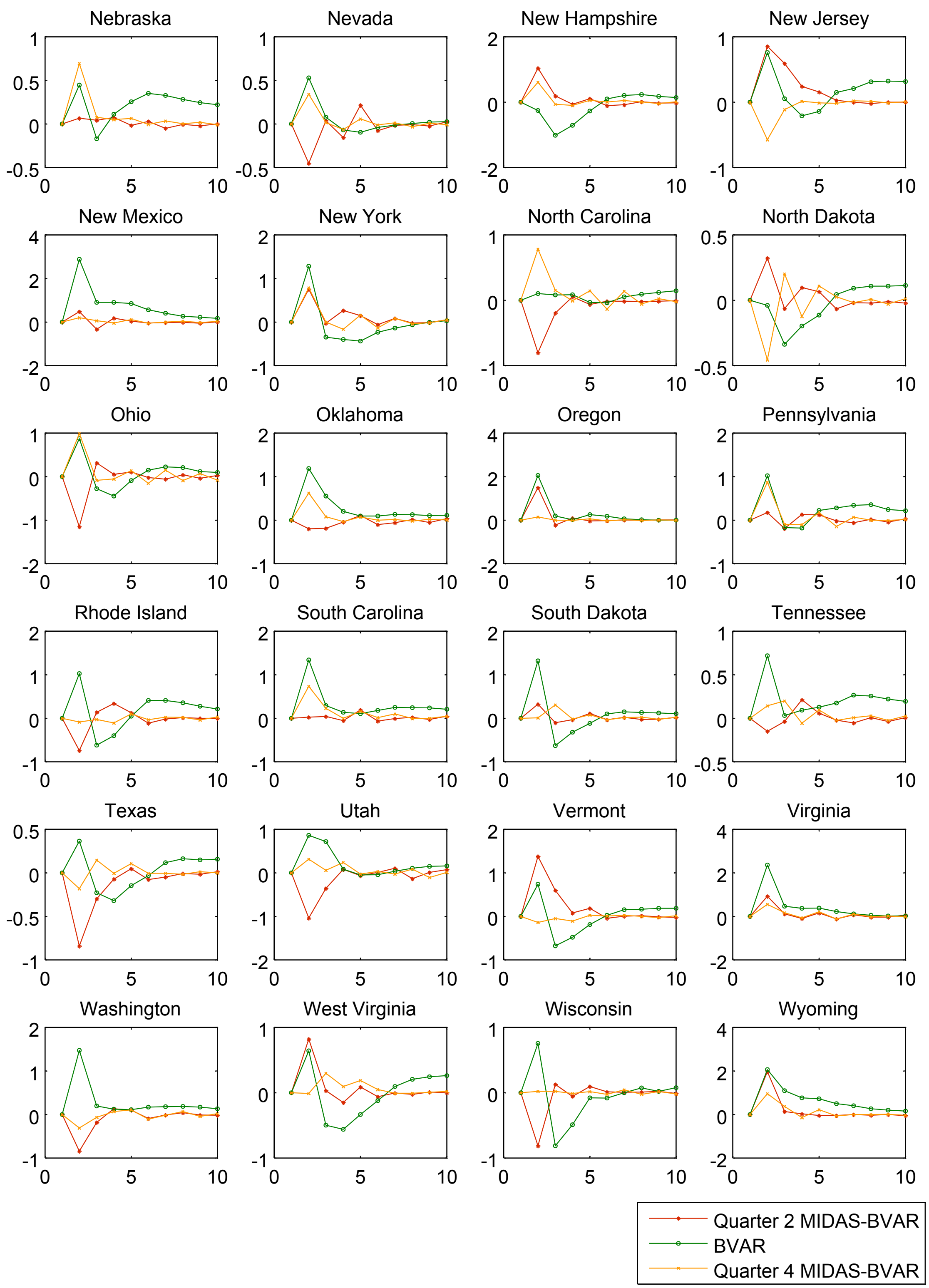


Figure 2.7: Impulse Responses of Revenues to Monetary Shocks:
MF-BVAR versus BVAR – continued figure

Impulse Responses of State Government Revenues to the Shocks to Federal Funds Rate: Second and Fourth Quarter Shocks from the MF-BVAR model versus aggregate annual shocks from the annual BVAR model.

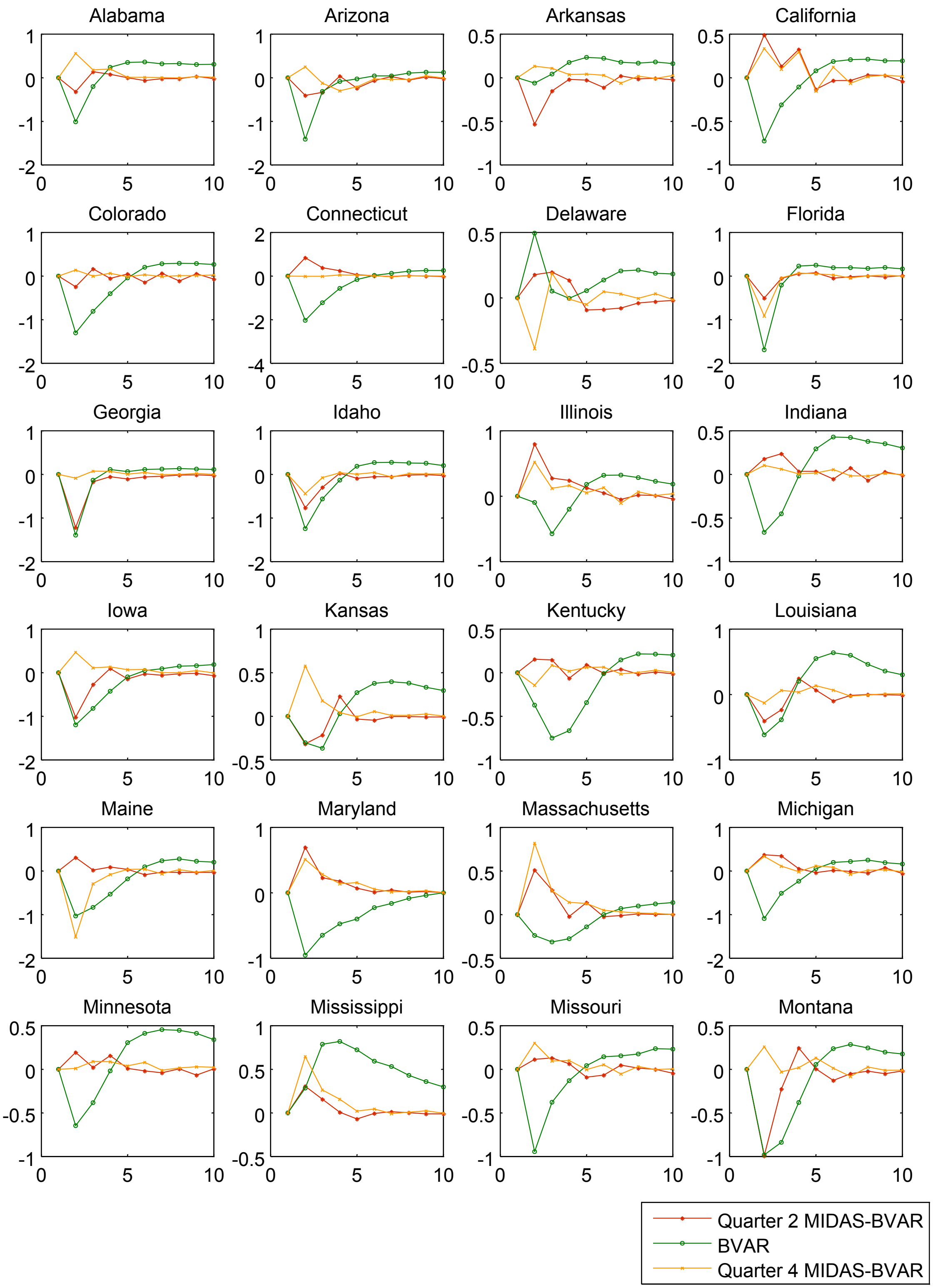


Figure 2.8: Impulse Responses of Expenditures to Monetary Shocks:MF-BVAR versus BVAR

Impulse Responses of State Government Expenditures to the Shocks to Federal Funds Rate: Second and Fourth Quarter Shocks from the MF-BVAR model versus aggregate annual shocks from the annual BVAR model.

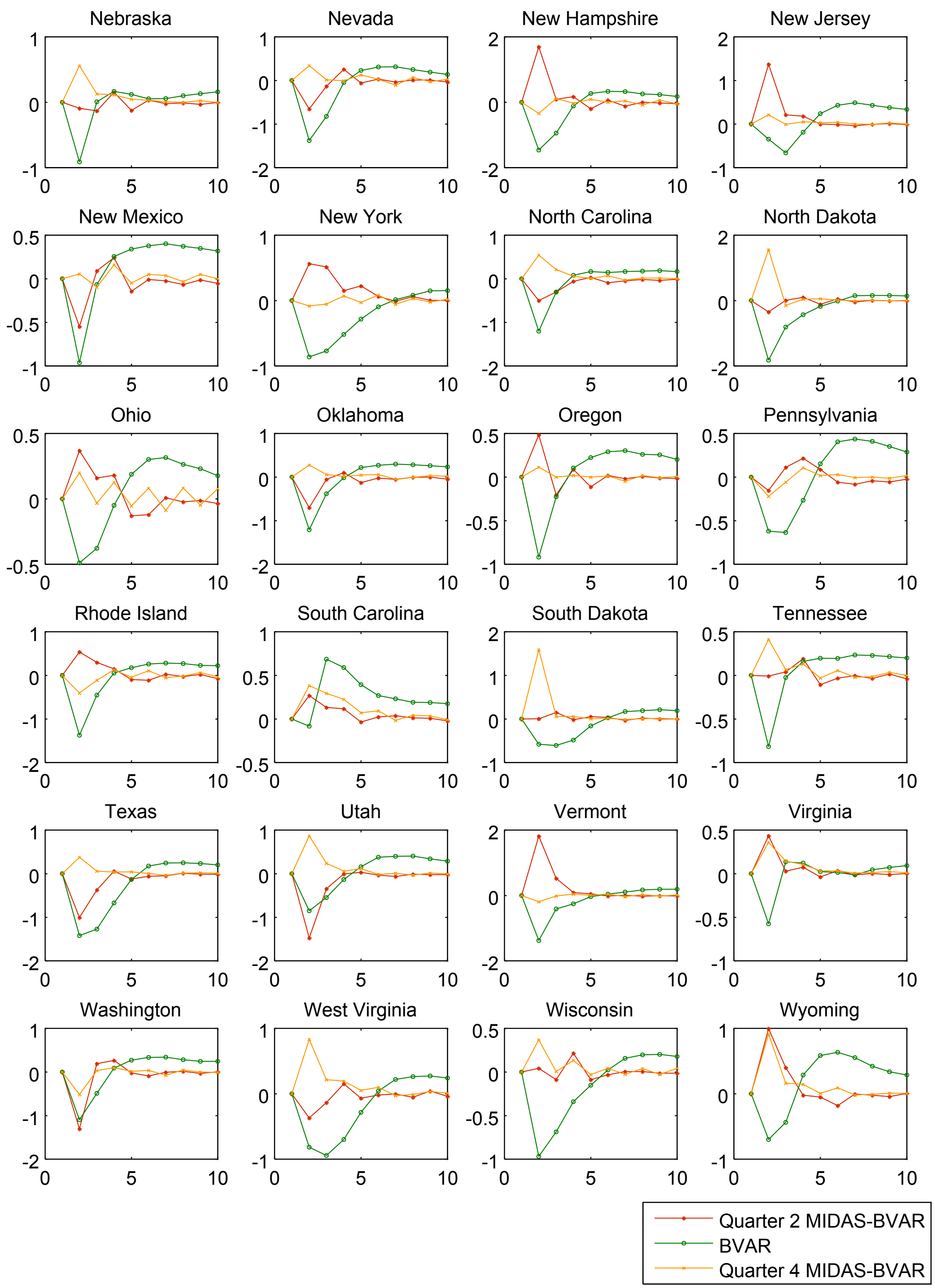


Figure 2.8: Impulse Responses of Expenditures to Monetary Shocks: MF-BVAR versus BVAR – continued figure

Impulse Responses of State Government Expenditures to the Shocks to Federal Funds Rate: Second and Fourth Quarter Shocks from the MF-BVAR model versus aggregate annual shocks from the annual BVAR model.

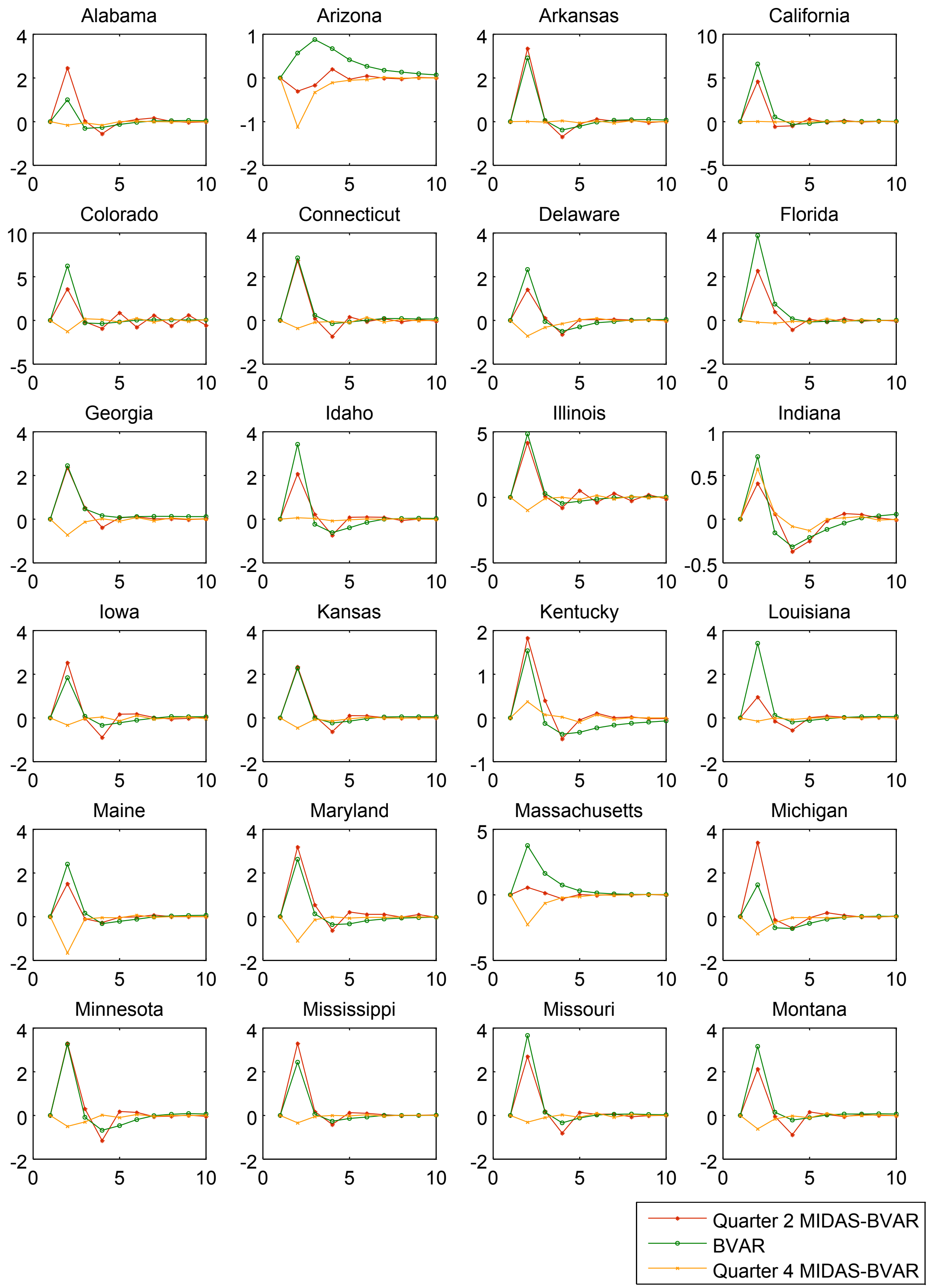


Figure 2.9: Impulse Responses of Revenues to Fiscal Shocks: MF-BVAR versus BVAR

Impulse Responses of State Government Revenues to the Shocks to Federal Government Deficit: Second and Fourth Quarter Shocks from the MF-BVAR model versus aggregate annual shocks from the annual BVAR model.

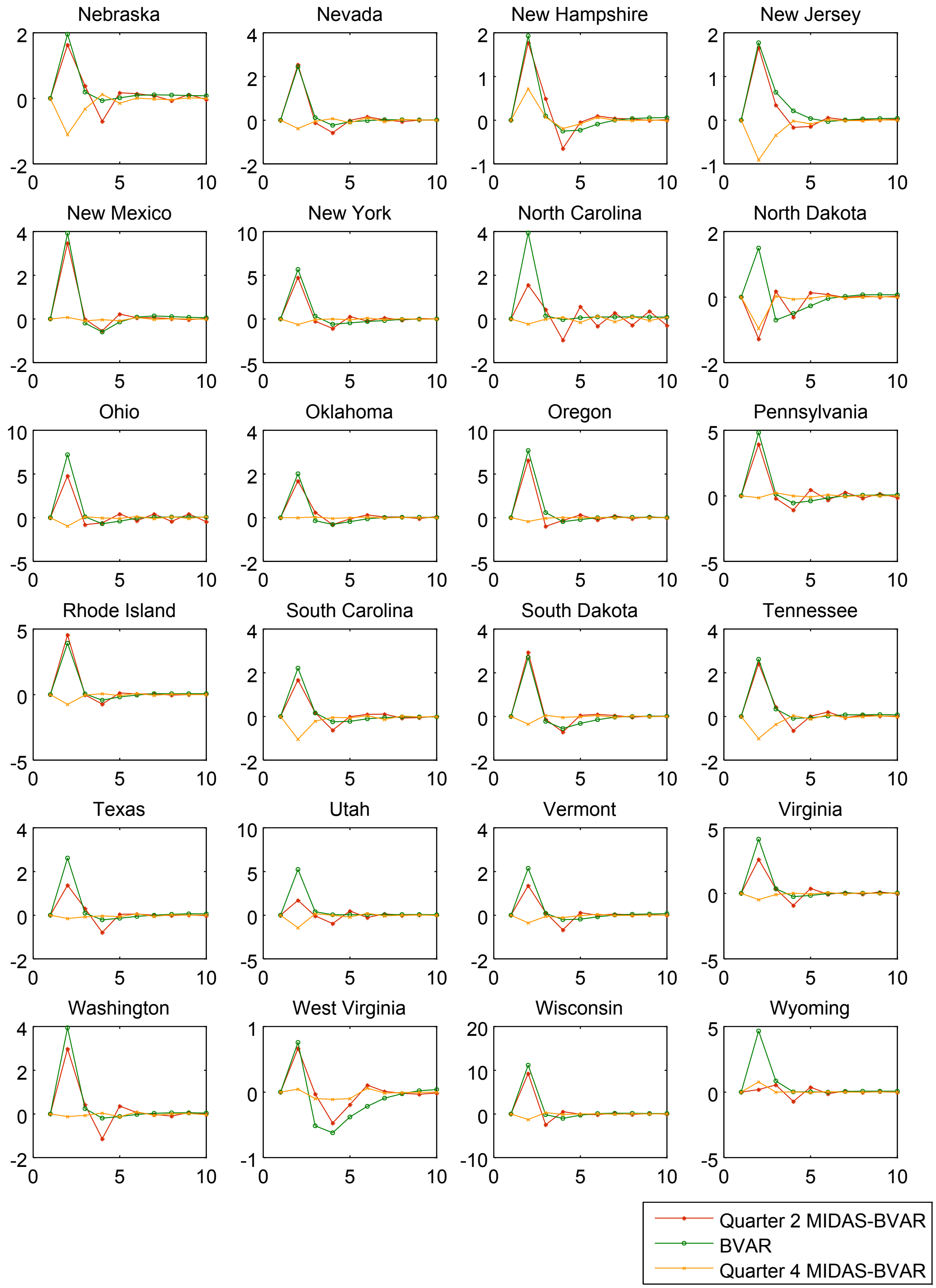


Figure 2.9: Impulse Responses of Revenues to Fiscal Shocks:
MF-BVAR versus BVAR – continued figure

Impulse Responses of State Government Revenues to the Shocks to Federal Government Deficit: Second and Fourth Quarter Shocks from the MF-BVAR model versus aggregate annual shocks from the annual BVAR model.

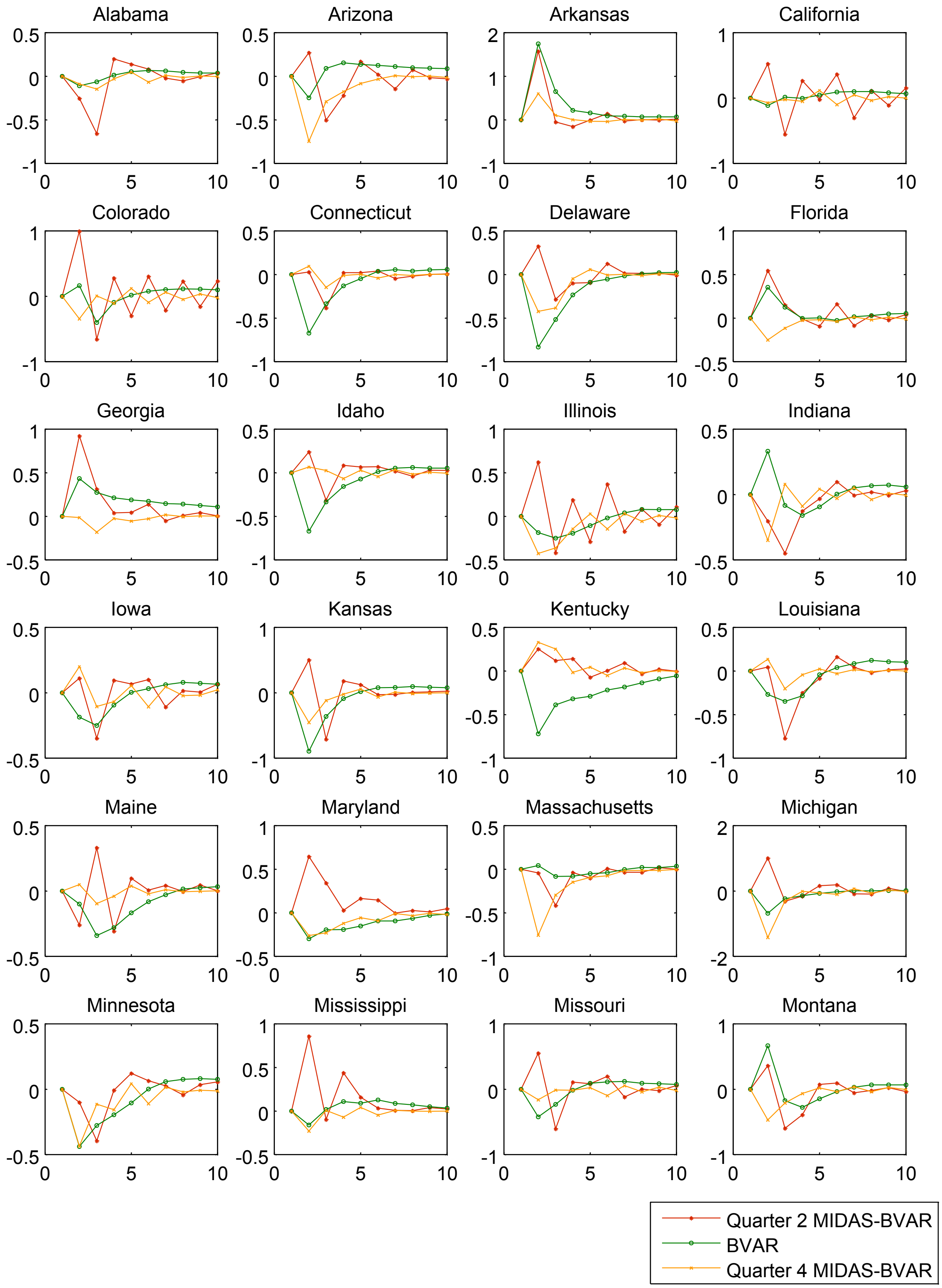


Figure 2.10: Impulse Responses of Expenditures to Fiscal Shocks: MF-BVAR versus BVAR

Impulse Responses of State Government Expenditures to the Shocks to Federal Government Deficit: Second and Fourth Quarter Shocks from the MF-BVAR model versus aggregate annual shocks from the annual BVAR model.

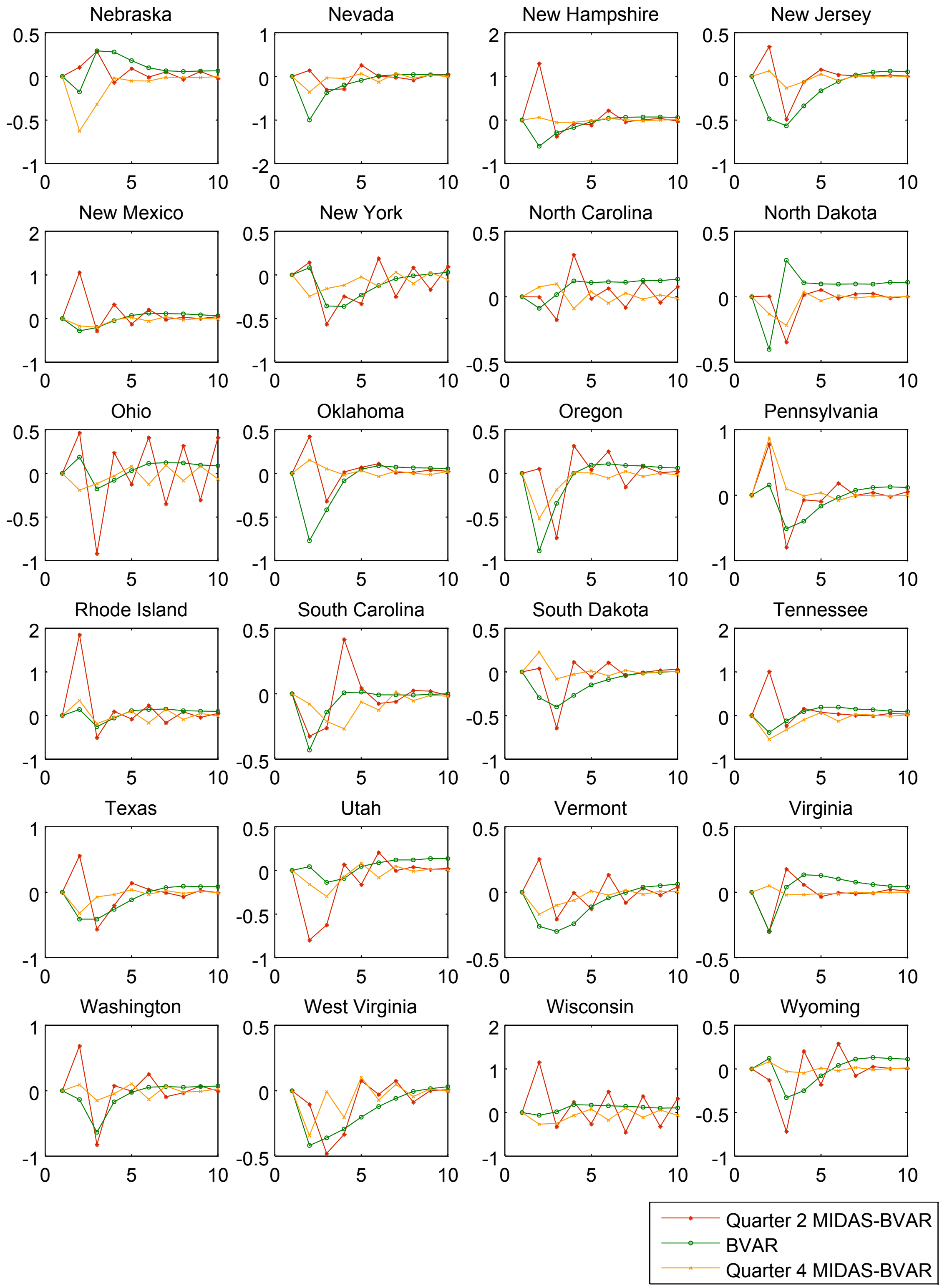


Figure 2.10: Impulse Responses of Expenditures to Fiscal Shocks:
MF-BVAR versus BVAR – continued figure

Impulse Responses of State Government Expenditures to the Shocks to Federal Government Deficit: Second and Fourth Quarter Shocks from the MF-BVAR model versus aggregate annual shocks from the annual BVAR model.

CHAPTER 3

REAL-TIME FORECASTING OF THE CURRENT-QUARTER U.S. GDI GROWTH

3.1 Introduction

Output growth is undoubtedly the most fundamental concept in macroeconomics. The Bureau of Economic Analysis (BEA) produces two conceptually identical measures of output, which are a commonly used expenditure-side estimate, Gross Domestic Product (GDP) and a less commonly used income side estimate, Gross Domestic Income (GDI). Despite being identical in principle, these two measures of output differ in practice because each is constructed from different source data and has been exhibiting different cyclical properties over the past twenty five years, with GDI showing a more pronounced cycle than GDP.

Many recent papers have examined GDI and its properties together with GDP. Both Fixler and Nalewaik (2009), and Nalewaik (2010) point out that GDI deserves serious attention since it has properties that may be superior to those of GDP from a variety of perpectives.[1] Landefeld (2010) provides accompanying comments on Nalewaik (2010) and indicates that the relative merits of GDP and GDI rely mainly on the quality of the underlying source data which, in turn, depend on the vintage that estimates are examined.

In consideration of the divergence between the two measures of output, Fixler et al. (2011) provide a detailed review on both GDI and GDP as well as their properties, and conclude that both output estimates are accurate and that GDI provides additional and valuable information about the true state ofconomic activity, which is never observed. Taking these merits of GDI into account, Aruoba et al.

[1] Nalewaik (2010) provides a broad range of results in favor of GDI arguing that it is better in reflecting the business cycle fluctuations in true output growth. He shows that the divergences between GDP and GDI estimates became highly cyclical around mid-1980s; GDI rose faster than GDP through the 1990s expansions and exhibited a relatively short boom period during 2004-2006; in contrast, GDI growth fell below GDP growth in the 2001 recession and in the latest cyclical contraction of 2007-2009, all lining up with the NBERs peak- and trough- dates. Furthermore, he points out that GDI growth is relatively more correlated with cyclically sensitive business cycle indicators such as change in unemployment rate and employment growth. See Figure 3.1 providing time series plots of quarterly real GDI and GDP growth rates together with their descriptive statistics for the 1986Q1-2013Q3 period.

(2012) and (2013) propose new measures that pool the information in the two versions, featuring GDI prominently. The authors examine the superiority of their new measure of output in reflecting business cycle fluctuations and suggest that it should be used as the benchmark output estimate. In this regard, the Federal Reserve Bank of Philadelphia has started to produce real-time data on this new measure, naming it GDPplus.

Despite its prominent feature, GDI is not timely available as GDP due to insufficiency of the source data on the income side. GDP is usually released one month after the end of the quarter, whereas GDI is not available until two months after the end of the quarter and, in the case of the fourth quarter, three months after the end of the quarter. Producing the current-quarter forecast of output growth is crucial as it provides useful information on recent news on the economy and is used as an input to central banks medium-term forecasting models. The literature on forecasting the contemporaneous value of real GDP growth is pretty rich.[2] Producing predictions of real GDP growth has always been the fundamental exercise for any forecasting method developed. On the other hand, recent literature does not provide methods for producing current-quarter forecasts of GDI growth as it has been a less widely used measure of output. Therefore, this chapter aims to propose a method to provide current-quarter forecasts of GDI growth.

The proposed forecasting procedure involves a combination of Augmented Distributed Lag-Mixed Data Sampling (ADL-MIDAS) regression models.[3] A potentially important property of the MIDAS approach is that it can be used to incorporate readily available within-the-quarter information to update the current-quarter forecast, as described in Clements and Galvão (2008) and Kuzin, Marcellino, and Schumacher (2013), who have introduced MIDAS regressions with *leads*. The authors have shown that the use of current-quarter monthly information leads to significant improvements in forecasting real GDP growth. Similar to their methodology, MIDAS models with *leads* are employed in this paper to incorporate readily available within-the-quarter, real-time information to forecast current-quarter real GDI growth. That is, current-quarter GDI growth is predicted with the use of monthly/weekly/daily observations of macroeconomic and financial indicators, such as employment, industrial production, and stock prices available within that quarter. The initial estimation period is 1986:Q1-2003:Q4 while

[2]See, for example, Andreou, Ghysels and Kourtellos (2013) take advantage of information content of daily financial data to update current-quarter forecasts of real GDP growth. Nunes (2005) and Giannone, Reichlin, and Small (2008) formalize the process of updating the forecast on GDP growth and inflation as new releases of data become available. Carriero, Clark, and Marcellino (2012) develop a method for producing current-quarter forecasts of GDP growth with a large range of within-the-quarter monthly economic indicators, among others.

[3]As shown in Bai, Ghysels and Wright (2013), one can view the proposed MIDAS regression approach as a computationally simple way of approximating the state-space approach suggested by Onorante et al. (2010), among others.

the forecasting period is 2004:Q1-2013:Q3. The predictive ability of MIDAS regression models are compared with the corresponding traditional models that simply take an equally weighted average of monthly/weekly/daily indicators. The assessment of the forecast accuracy is done with the root mean squared forecast errors (RMSFE). Evidence shows that ADL-MIDAS regressions with leads provide significant forecast improvements over the traditional models, suggesting the use of within-quarter mixed frequency macroeconomic and financial indicators as they become available throughout the quarter to improve the current-quarter forecast of real GDI growth. These findings are based on a recursive out-of-sample forecasting exercise that uses real-time data.

The remainder of this chapter proceeds as follows. In Section 3.2, the econometric methods employed in this paper are presented within subsections devoted to the descriptions of the ADL-MIDAS regression models, a method to combine forecasts and a test of predictive accuracy. Section 3.3 introduces the data and Section 3.4 presents empirical results with the evidence in favor of the forecast combinations of ADL-MIDAS regression models with monthly leads against AR and ADL regression models. Section 3.5 presents conclusions.

3.2 Methods

In order to deal with data sampled at different frequencies, Mixed Data Sampling -MIDAS- type regression models, specifically Augmented Distributed LagMixed Data Sampling (ADL-MIDAS) regression models, are employed. Following Koenig et al. (2003), who suggest that the explanatory variables should be measured at each date within a sample, the analysis in this paper is in real-time. At each point in time, the models are estimated using only the data available for time periods up to that point, in-sample estimations are performed with the vintages of data restricted to those available at that time, and the forecasts are obtained with the latest values from the most recent vintage, which is 2013Q3.[4] For example, to produce the 2004Q1 GDI growth forecast, all the explanatory variables should be measured as they appeared in 2004Q1. Clements and Galvão (2008) use a similar method with MIDAS regressions with leads to incorporate real-time information on monthly indicators to produce current-quarter forecasts of GDP growth. In this paper, for each economic indicator, current-quarter forecasts of real GDI growth are obtained using a recursive out-of-sample forecasting exercise, resulting in multiple predictions for each out-of-sample observation. It is well known that pooling forecasts produces a robust tool in

[4]See Koenig et al. (2003) and Clements and Galvão (2013) for a detailed explanation of real-time vintage data analysis compared to end-of-sample vintage data analysis.

the presence of misspecification and parameter instability.[5] Hence, to obtain more accurate forecasts by using evidence from all individual models, the forecast combinations method is considered rather than using the best single model.

Subsection 3.2.1 describes the ADL-MIDAS regression models with *leads* and subsection 1.2.2 briefly introduces the forecast combination method employed in this paper. Then, the test of predictive accuracy is discussed in subsection 1.2.3.

3.2.1 ADL–MIDAS with Leads

In terms of the improvement of low-frequency macroeconomic predictions using high-frequency data, the advantages of MIDAS-type regressions have been discussed by many recent papers. MIDAS regression models are suggested by Andreou, Ghysels, and Kourtellos (2013), Clements and Galvão (2008) and (2009), and Ghysels, Santa-Clara, and Valkanov (2006), among others.[6] ADL-MIDAS regression models with *leads*, similar to the MIDAS-with-*leads* models introduced by Clements and Galvão (2008) and Kuzin, Marcellino, and Schumacher (2013), are employed in this paper to produce current-quarter forecasts of real GDI growth. Then their forecast performances are compared with predictive abilities of traditional models, namely augmented distributed lag (ADL) and autoregressive (AR) regression models.

Within each quarter, the contemporaneous value of GDI growth is not available, but it can be predicted using higher frequency variables, which are more timely available. In order to produce the current-quarter forecast of real GDI growth, ADL-MIDAS models with leads are employed by incorporating within-the-quarter information contained in monthly/weekly/daily economic and financial indicators. For instance, suppose we are two months into quarter t+1, implying that we have two-months worth of daily data, e.g., stock price index or monthly data, such as the unemployment rate. Then, if we stand on the last day of the second month of the quarter and aim to produce a forecast for the current quarter, we could use two-month leads of these higher frequency data series. That is, the notion of leads here implies that the information between quarter t and quarter t+1 is utilized. More precisely, the forecaster's information set is extended by using readily available real-time data at the end of the second month of a quarter. Consider the $ADL-MIDAS(p_Y^Q, q_X^H, J_X^H)$ regression equation allowing for J_X^H leads for the higher frequency indicator, expressed in multiples of months, $J_X^H = 1, 2$.

[5] See, for example, Timmermann (2006), who suggests forecast combinations across different models to have more robust forecasts against misspecification biases and measurement errors in the data underlying the individual forecasts. Stock and Watson (2001, 2004) also find that combined forecasts generally outperform forecast performances of the best individual model by employing numerous types of models and variables.

[6] See Ghysels, Sinko, and Valkanov (2007) for various descriptions of MIDAS regressions. The initial work on MIDAS focused on volatility predictions (see Ghysels, Santa-Clara, and Valkanov (2005), among others).

$$Y^Q_{t+1} = \mu + \sum_{k=0}^{p^Q_Y - 1} \alpha_k Y^Q_{t-k} + \gamma \left[\sum_{i=0}^{J^H_X - 1} \omega_i(\theta^H_X) X^H_{J^H_X - i, t+1} + \sum_{j=0}^{q^H_X - 1} \sum_{i=0}^{N_H - 1} \omega_{i+j*N_M}(\theta^H_X) X^H_{N_H - i, t-j} \right] + u_{t+1}$$

where N_H denotes the higher frequency, daily/weekly/monthly lags per quarter and the weighting scheme, $\omega(\theta^H)$, involves a low dimensional vector of unknown parameters.[7,8] One of the specifications for the weighting scheme introduced by Ghysels, Sinko, and Valkanov (2007), for example, is based on a beta function with two parameters, which is normalized to add up to one to allow for the identification of the slope coefficient γ. Beta function is known to be flexible; it can take many shapes, including flat weights, gradually decreasing and hump-shaped patterns. Normalized beta probability density function with unrestricted (u) and restricted (r) cases and with non-zero (nz) and zero (z) last lag specifications can be written as

$$\omega_i^{u,nz} = \omega_i(\theta_1, \theta_2, \theta_3) = \frac{x_i^{\theta_1 - 1}(1 - x_i)^{\theta_2 - 1}}{\sum_{i=1}^{N} x_i^{\theta_1 - 1}(1 - x_i)^{\theta_2 - 1}} + \theta_3$$

$$\omega_i^{u,z} = \omega_i(\theta_1, \theta_2, 0), \quad \omega_i^{r,nz} = \omega_i(1, \theta_2, \theta_3), \quad \omega_i^{r,z} = \omega_i(1, \theta_2, 0) \ , \quad \text{where } x_i = \tfrac{i-1}{N_D - 1}.$$

Autoregressive, $AR(p_Y)$, and augmented distributed lag, $ADL(p_Y, p_X)$, regression models, employed

---

[7]MIDAS regressions involve NLS and feasible GLS estimation procedures. Errors are not necessarily i.i.d., error process is a linear process with absolute summable Wold decomposition moving average representations. Identification of coefficients rests on the assumption that the regressor is exogenous up to second moments. With MIDAS regressions, since regressors are sampled at higher frequencies, consistency (absence of discretization bias) can be of concern. Ghysels, Santa-Clara, Valkanov (2006) show that discretization bias is eventually eliminated. That is, aggregation bias disappears since the dependent variable is sampled at a fixed frequency while the regressor is sampled more frequently. The authors note that MIDAS regressions appear like skip-sampled distributed lag models, causing autocorrelated residuals (not preventing OLS/NLS to be consistent). To ensure correct specification of MIDAS polynomials, the assumption of $E(u_{t+1}|X_\tau; \tau \leq t) = 0$ is required. See Ghysels, Santa-Clara, Valkanov (2006) for further information.

[8]For the weighting polynomial, all specifications introduced by Following Ghysels, Sinko, and Valkanov (2007) are employed, then the one providing the best historical forecast performance is chosen.

as competing forecasting models, can be represented, respectively, with the following equations:

$$Y_{t+h} = c + \sum_{j=0}^{p_Y-1} \alpha_{j+1} Y_{t-j} + u_{t+h}, \tag{3.2.1}$$

$$Y_{t+h} = c + \sum_{j=0}^{p_Y-1} \alpha_{j+1} Y_{t-j} + \sum_{j=0}^{p_X-1} \beta_{j+1} X_{t-j} + u_{t+h}. \tag{3.2.2}$$

All variables in these competing models are sampled at the quarterly frequency, i.e., equally-weighted averages of higher frequency data are used to construct quarterly series. Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are used to select the optimal number of lags for all regressions.

3.2.2 Forecast Combinations Method

As pointed out by Timmermann (2006), forecast combinations have been viewed as a simple and effective way to obtain more accurate forecasts by using evidence from all models considered rather than using the best single model. Stock and Watson (2004), Clements and Galvão (2008), and Andreou, Ghysels, and Kourtellos (2013), among others, have used forecast combination techniques in the context of real GDP growth. Specifically, Andreou, Ghysels, and Kourtellos (2013) combine a large cross-section of daily financial indicators to produce real-time MIDAS regression forecasts of GDP growth.

Kuzin, Marcellino, and Schumacher (2013) also find that there is considerable uncertainty with respect to the appropriate specification of the econometric tools required to deal with large data sets of macroeconomic variables; and in this regard, they suggest pooling many specifications within and across the MIDAS models as it is superior to selecting a single model.

Given N individual forecasting models, forecast combinations are time-varying weighted averages of the individual forecasts and can be represented as

$$\hat{f}_{N,t+h|t} = \sum_{j=1}^{N} \hat{\omega}_{j,t} \hat{y}_{j,t+h|t}$$

where the weights $\hat{\omega}_{j,t}$ on the j^{th} forecast in period t depends on the historical performance of the individual forecasts.[9] Each individual predictor is given a weight according to its historical performance

[9] In this study, $\hat{y}_j$ is the individual prediction from ADL-MIDAS regression estimated one at a time with each

and the weight is inversely proportional to the predictor's dMSFE. The discount factor attaches greater weight to the recent predictive ability of the individual predictor. The weights are given as follows:

$$\hat{\omega}_{j,t} = \frac{(\lambda_{j,t}^{-1})^{\kappa}}{\sum_{i=1}^{N}(\lambda_{i,t}^{-1})^{\kappa}}, \qquad \lambda_{j,t} = \sum_{m=T_0}^{t-h} \delta^{t-h-m}(y_{m+h}^{h} - \hat{y}_{j,m+h|m}^{h})^2$$

where $\delta = 0.9$, 0.95 or 1 and $\kappa = 1$ or 2 for dMSFE, depending on the forecast gains that could be achieved.

3.3 Data

GDI and GDP are measures of the quarter-over-quarter rate of growth of real output in annualized percentage points.[10] Figure 3.1 presents time series plot of quarterly growth rates of real GDI and GDP over the period of 1986Q1-2014Q3 and provides descriptive statistics. The simple correlation between the two series is 0.697. Median GDI growth is a bit higher than that of GDP, and GDI growth is noticeably more persistent than GDP. Since GDP is relatively more timely available and, in light of the findings of Nalewaik (2010) on the correlation between GDI and GDP estimates, GDP is employed as an indicator to predict GDI.

In applying the proposed method to forecast current-quarter GDI growth, various combinations of quarterly/monthly/weekly/daily indicators are considered. These particular indicators are chosen to be broadly representative of major economic and financial indicators, with some consideration regarding their timeliness. The indicators employed are as follows: a) stock price changes as measured by S&P 500 index; b) slope of the yield curve as the differential between 10- and 2-year Treasury bonds; c) spread between high-yield corporate bonds and Treasury bonds; d) changes in unemployment rate; e) employment growth; and f) the manufacturing index of ISM (Institute for Supply Management), all of which are found by Nalewaik (2010), who carefully selects the underlying source data to construct GDI to be highly correlated with GDI growth. In addition to these indicators, industrial production index,

---

indicator, $j = 1, , N$. N is the number of individual models. Estimating ADL-MIDAS regressions one at a time -as is typical in forecast combination settings- involves efficiency losses compared to systems based on Kalman Filter. See Bai, Ghysels, Wright (2013) for further information.

[10] Note that all annualized growth rates are calculated using the formula for continuous compounding. The formula for annualizing quarterly data is straightforward, $g_q = \left[\left(\frac{X_q}{X_{q-1}}\right)^4 - 1\right] * 100$.

housing starts, 10-year Treasury bond yield, and the 3-month Treasury bill rate are also included as they are also commonly used indicators utilized in forecasting GDP growth.[11]

The data set consists of real-time vintages. At each point in time, the vintages of data used for estimation are restricted to those that would have been available at that time. The real-time data consist of quarterly vintages of GDI growth and monthly vintages of the indicators. Following Koenig et al. (2003) and Clements and Galvão (2008), the aim is to forecast the latest vintage data on GDI growth, which is 2013Q3 in this paper. All the data series used in this study are listed in Table 3.1 and are seasonally adjusted, in real quantities, and transformed to induce stationarity, if necessary. The estimation period is chosen to be 1986Q1-2003Q4 and the forecasting period is 2004Q1-2013Q3. The cutoff point between in- and out-of-samples is determined considering the availability of data because the BEA provides real-time data for the revised estimates of GDI starting from 2002 and, thus, 2004Q1 is picked as the first out-of-sample observation to allow enough time for the latest estimates to pass through all the annual revisions.

3.4 Empirical Results

Using a rolling window method, models are estimated then, with a recursive out-of-sample forecasting exercise, current-quarter forecasts are obtained in order to evaluate predictive ability of the models. The total sample size is T years, the fixed rolling window size is R and, for each window, forecasts of 8-quarters-out are used to calculate the root mean squared forecast error for the corresponding window; hence, the period used to evaluate annual forecasts is P=T-R-8. The initial estimation period for the data set is 1986Q1-2003Q4 while the forecasting period is 2004Q1-2013Q3. For each of these quarters, forecasts are produced at the zero horizon with monthly steps, i.e., h=1/3, 2/3. The forecast accuracy of each model is assessed using the root mean squared forecast error, RMSFE.[12]

Figure 3.2 provides a concise preview of the predictive gains of real-time forecast updating of the current-quarter real GDI growth by displaying three boxplots, one for the forecast combinations of ADL-MIDAS regression models with leads and the other two for the traditional ADL and AR regression models.[13] These boxplots present predictive abilities of the competing models, which are measured in

[11]See, for example, Clements and Galvão (2009), who employ these indicators to produce current-quarter forecasts of GDP growth.

[12]RMSFEs are obtained as follows: $RMSFE_t = \sqrt{\frac{1}{t-T_0+1}\sum_{\tau=T_0}^{t}\left(y^h_{\tau+h} - \hat{y}^h_{\tau+h|\tau}\right)^2}$ where $t = T_1, , T_2$. T_0 is the point at which the first individual pseudo out-of-sample forecast is computed. $T_0 = 2004Q1$, $T_1 = 2004Q1+8$ and, $T_2 = 2013Q3$.

[13]A boxplot is a way of summarizing a set of data measured on an interval scale, a type of graph, which is used

terms of RMSFEs. Each point on the boxplots is attached to each out-of-sample rolling window. Since smaller RMSFEs reflect better forecast performance, Figure 3.2 indicates that the forecast combinations of ADL-MIDAS models with leads outperforms ADL and AR models, that is, updating the current-quarter forecast by incorporating within-the-quarter information via ADL-MIDAS with leads models provide forecast gains over their traditional counterparts.

3.4.1 MIDAS Regressions versus AR and ADL Regressions

Table 3.2 presents RMSFEs from the models predicting quarterly real GDI growth. It reports RMSFEs from the AR model, the forecast combinations of ADL regression models and the forecast combinations of ADL-MIDAS models. The evidence shows that forecast combinations of MIDAS regression models having smaller RMSFEs provide strong forecast gains over both of their AR and ADL counterparts. In order to evaluate this evidence on forecasting performance, Giacomini and White (GW) test (Giacomini and White (2006)), which is described in Section 1.2.3 is employed to test the null hypotheses of equal forecasting accuracy between two different models. Table 3.3 presents these GW test statistics, testing for equal forecasting accuracy between forecast combinations of ADL-MIDAS regression models vis-à-vis AR and ADL regression models. It is found that MIDAS regression models yield significant forecast gains over their AR and ADL model counterparts. Individual contributions of each predictor to the superior forecast performance of combinations of ADL-MIDAS can be seen in Figure 3.4.

3.4.2 MIDAS Regressions with Leads

Figure 3.3 presents one-quarter-ahead forecasts from AR and ADL regressions and real-time forecast updates for current-quarter GDI growth from MIDAS regressions with two-month-leads, together with the actual values over the evaluation period from 2004Q1-2013Q3. It can be concluded that, on average, real-time forecasts from the combinations of MIDAS regressions with monthly leads follow the actual data more closely compared to one-quarter-ahead forecasts from the AR and ADL regressions.

---

to show the shape of the distribution, its central value, and variability consisting of the most extreme values in the data set (maximum and minimum values), the lower and upper quartiles, and the median.

3.5 Conclusion

This chapter employs ADL-MIDAS regression models with monthly leads to produce real-time forecast updates of the current-quarter real GDI growth. Evidence shows that the forecast combinations of ADL-MIDAS regression models outperform both AR and forecast combinations of ADL regression models. The proposed forecasting method provides significant predictive performance over these traditonal models, suggesting the use of within-quarter information on daily/weekly/monthly indicators as they become available throughout the quarter to improve the current-quarter real GDI forecasts, and MIDAS is an effective way of incorporating within-quarter data.

Table 3.1: Data Set

Title	Frequency
Real GDI Growth	Quarterly
Real GDP Growth	Quarterly
Consumer Price Index for All Urban Consumers: All Items	Monthly
Industrial Production Index	Monthly
All Employees: Total Nonfarm	Monthly
Housing Starts	Monthly
Unemployment Rate	Monthly
ISM Manufacturing Index	Monthly
Initial Claims	Weekly
S&P 500 Stock Price Index	Daily
3-Month Treasury Bill: Secondary market rate	Daily
10-Year Treasury Constant Maturity Rate	Daily
Yield Curve (10yearTB minus 2yearTB)	Daily
High-Low Spread (High Yield minus 10yearTB)	Daily

Notes: This table lists the data series used in this study, all are seasonally adjusted, real quantities, and transformed to ensure stationarity, if necessary. Output growth rates are annualized quarterly rates. They are obtained from Bureau of Economic Analysis, http://www.bea.gov, and from FRED (Federal Reserve Economic Data) (http://research.stlouisfed.org/fred2/), and from ALFRED (ArchivaL Federal Reserve Economic Data) (http://alfred.stlouisfed.org/) and Real-Time Data Research Center of Federal Reserve Bank of Philadelphia, (http://www.phil.frb.org/research-and-data/real-time-center/real-time-data/data-files/) for the real-time vintages. Sample period covers the quarterly time period of 1986Q1-2013Q3, the longest possible time span selected based the available predictors.

Table 3.2: Comparison of RMSFEs: MIDAS Regressions vs. AR and ADL Models.

	Models				
	MIDAS_J=2	MIDAS_J=1	MIDAS_J=0	ADL	AR
RMSFE	1.6209	1.7516	1.7999	1.9727	2.1321

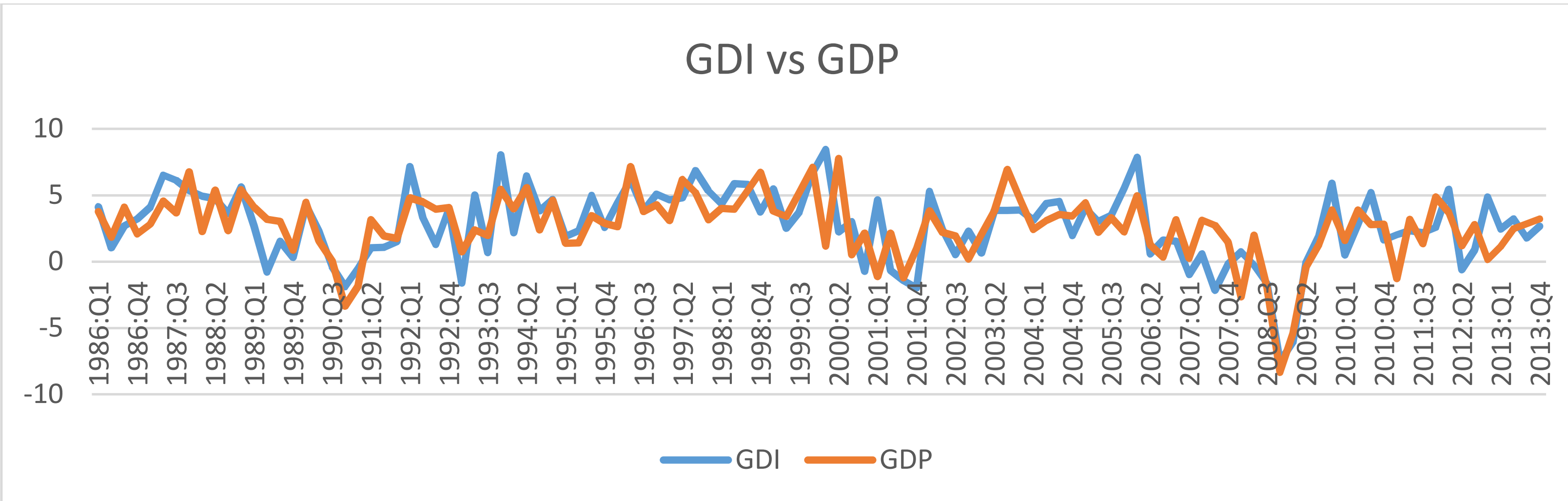
Notes: This table presents Root Mean Squared Forecast Errors (RMSFEs) of autoregressive (AR) models, forecast combinations of augmented distributed lag (ADL) models, and forecast combinations of MIDAS regressions for the quarterly U.S. GDI growth over the out-of-sample period of 2004Q1-2013Q3. MIDAS_J=j represents the ADL-MIDAS regression with j-monthly leads. Forecast horizon is one quarter for ADL, AR and MIDAS_J=0 models.

Table 3.3: Time Series Test for Predictive Ability.

	AR versus MIDAS		
	Forecast horizon		
	J=0	J=1	J=2
GW test statistic	2.702	3.420	3.337
p-value	0.007	0.001	0.001
	ADL versus MIDAS		
	Forecast horizon		
	J=0	J=1	J=2
GW test statistic	1.987	2.021	2.006
p-value	0.047	0.043	0.045

Notes: This table presents Giacomini-White (GW) statistics for current-quarter forecasts and their corresponding p-values to test for equal forecasting accuracy between forecast combinations of MIDAS regressions with J=j monthly leads against AR models and against forecast combinations of ADL models for the quarterly U.S. GDI growth. The estimation period is 1986Q1-2003Q4 while the forecasting period is 2004Q1-2013Q3.

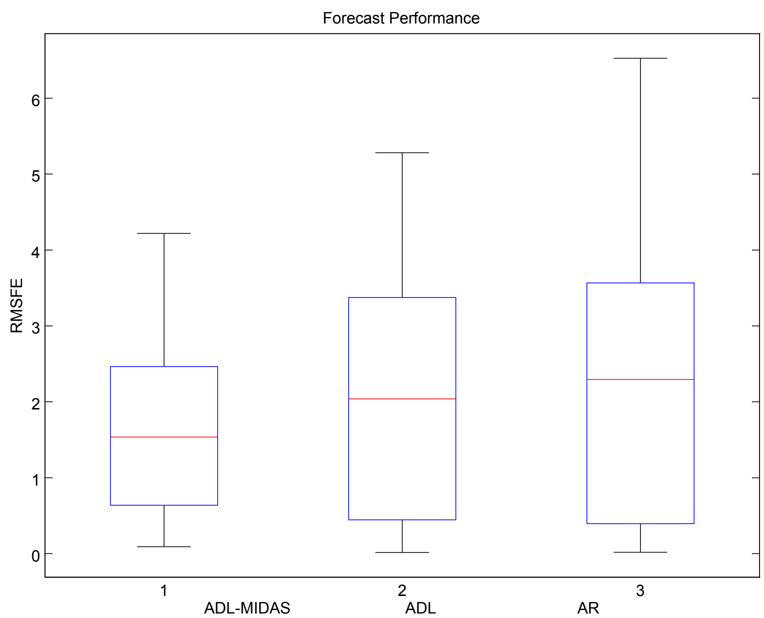
Figure 3.1: Quarterly Real GDI and GDP Growth Over the Period of 1986Q1-2013Q3.



This figure presents time series plots for quarter-over-quarter growth rates of real GDI and GDP in annualized percentage points over the period of 1986Q1-2014Q3. The correlation between the two series is 0.697. The table below provides descriptive statistics. $\hat{\rho}_j$ is a sample autocorrelation at a displacement of j-quarters, Q_{12} is the Ljung-Box serial correlation test statistic calculated using $\hat{\rho}_1$ and $\hat{\rho}_2$, ..., $\hat{\rho}_{12}$.

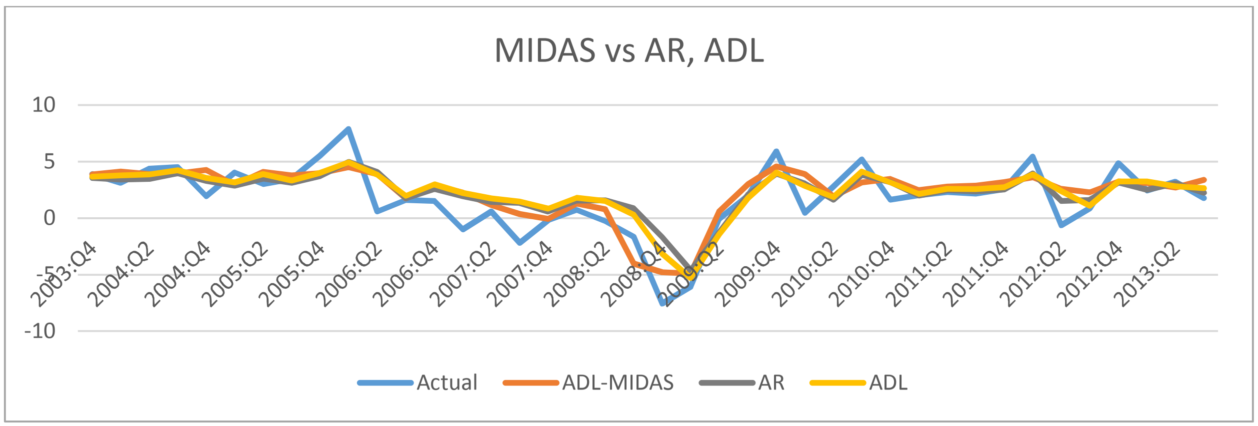
	Mean	*Median*	*Stdev.*	*Skew.*	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\rho}_3$	$\hat{\rho}_4$	$\hat{\rho}_8$	Q_{12}
GDI	2.74	2.74	2.73	-0.70	0.43	0.38	0.22	0.23	0.07	78.64
GDP	2.66	2.85	2.45	-1.20	0.41	0.38	0.16	0.19	0.02	56.71

Figure 3.2: Boxplots for Forecast Performance Comparison
between MIDAS, ADL and AR regressions



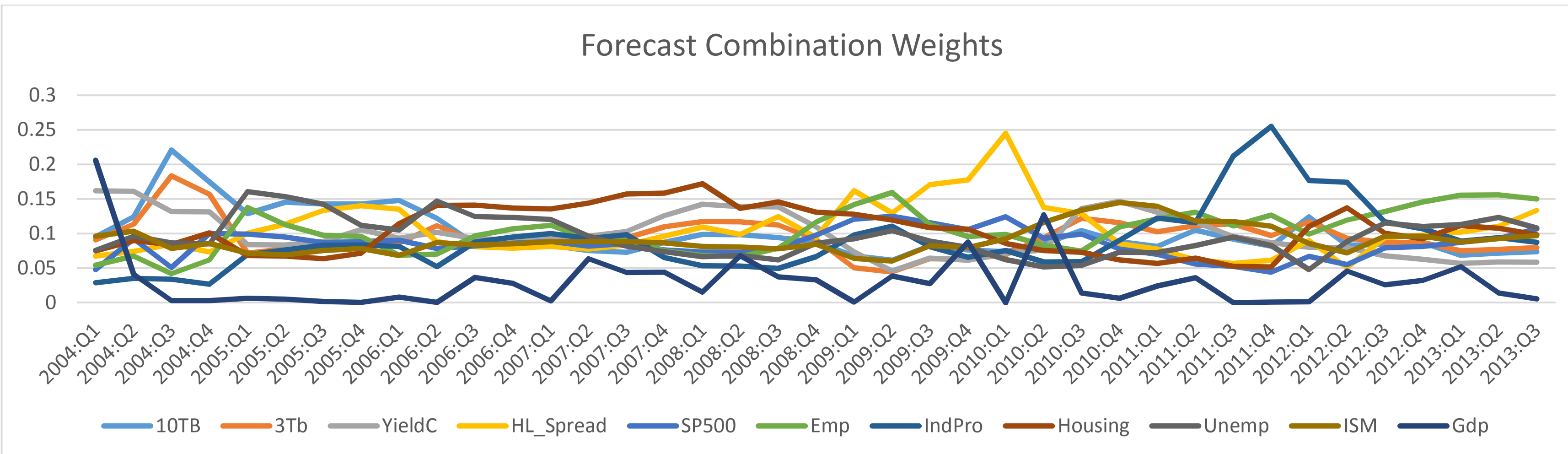
These boxplots present predictive abilities of the competing models and forecast combinations of ADL-MIDAS regressions vs. ADL and AR regressions by displaying their forecast performances measured in terms of Root Mean Squared Forecast Errors (RMSFEs). Each point in the boxplots is attached to each out-of-sample rolling window.

Figure 3.3: Predictions from Forecast Combinations of MIDAS regressions with Leads, ADL and AR models



This figure presents one-quarter-ahead predictions from ADL and AR models and current quarter forecasts from the ADL-MIDAS regressions with monthly leads (MIDAS-J=2 months) for the U.S. GDI growth over the out-of-sample period of 2004Q1-2013Q3 together with the actual observations. Note that all values are annualized quarterly real growth rates.

Figure 3.4: Forecast Combination Weights



This figure displays forecast combination weights of the predictors used in ADL-MIDAS regressions with two-month leads for the quarterly U.S. GDI growth over the forecasting period of 2004Q1-2013Q3. Note that for a quarter, all weights add up to one.

BIBLIOGRAPHY

Andreou, E., E. Ghysels, and A. Kourtellos (2013): "Should Macroeconomic Forecasters use Daily Financial Data and How?," *Journal of Business and Economic Statistics*, 31, 240–251.

Artis, M., and M. Marcellino (2001): "Fiscal Forecasting: The Trach Record of the IMF, OECD and EC," *Econometrics Journal*, 4, 20–36.

Aruoba, S., F. Diebold, J. Nalewaik, F. Schorfheide, and D. Song (2012): "Improving GDP Measurement: A Forecast Combination Perspective," Recent Advances and Future Directions in Causality, Prediction, and Specification Analysis: Essays in Honour of Halbert L. White Jr.

——— (2013): "Improving GDP Measurement: A Measurement-Error Perspective," NBER Working Paper, 18954.

Aruoba, S., F. Diebold, and C. Scotti (2009): "Real-Time Measurement of Business Conditions," *Journal of Business and Economics Statistics*, 27, 417–427.

Asimakopoulos, S., J. Paredes, and T. Warmedinger (2013): "Forecasting Fiscal Time Series Using Mixed Frequency Data," ECB Working Paper Series, 1550.

Bai, J., E. Ghysels, and J. Wright (2013): "State Space Models and MIDAS Regressions," *Econometric Reviews*, 32, 779–813.

Bernanke, B., J. Boivin, and P. Eliasz (2005): "Measuring Monetary Policy: A Factor Augmented Autoregressive (FAVAR) Approach," *Quarterly Journal of Economics*, 120, 387–422.

Blanchard, O., and R. Perotti (2002): "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output," *Quarterly Journal of Economics*, 117, 1329–68.

Carriero, A., T. Clark, and M. Marcellino (2012): "Real-Time Nowcasting with a Bayesian Mixed Frequency Model with Stochastic Volatility," Federal Reserve Bank of Cleveland Working Paper, 1227.

CBO (1997): "An Economic Model for Long-Run Budget Simulations," Memorandum, Congressional Budget Office.

——— (2012): "An Update to the Budget and Economic Outlook: Fiscal years 2012 to 2022," Discussion paper, Congressional Budget Office.

——— (2014): "The Budget and Economic Outlook: Fiscal Years 2014 to 2024," Discussion paper, Congressional Budget Office.

Chiu, C., B. Eraker, A. Foerster, T. Kim, and H. Seoane (2011): "Estimating VARs Sampled at Mixed or Irregular Spaced Frequencies: A Bayesian Approach," Federal Reserve Bank of Kansas City, RWP 11-11.

Christiano, L., M. Eichenbaum, and C. Evans (1999): "Monetary Policy shocks: What Have We learned and to What End?," *Handbook of Macroeconomics*, 1, 65–148.

Clements, M., and A. Galvão (2008): "Macroeconomic Forecasting with Mixed-Frequency Data: Forecasting Output Growth in the United States," *Journal of Business and Economic Statistics*, 26, 546–554.

Clements, M., and A. Galvao (2009): "Forecasting US Output Growth Using Leading Indicators: An Appraisal Using MIDAS Models," *Journal of Applied Econometrics*, 24, 1187–1206.

Clements, M., and A. Galvão (2013): "Real-Time Forecasting of Inflation and Output Growth with Autoregressive Models in the Presence of Data Revisions," *Journal of Applied Econometrics*, 28, 458–477.

Dalsgaard, T., and A. De Serres (2001): "Estimating Prudent Budgetary Margins for EU Countries: A Simulated SVAR Model Approach," in *The Stability and Growth Pact. The Architecture of Fiscal Policy in EMU*. Palgrave Publisher, U.K.

Diebold, F., and R. Mariano (1995): "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, 13, 253–265.

Doan, T., R. Litterman, and C. Sims (1984): "Forecasting and Conditional Projection Using Realistic Prior Distributions," *Econometric Reviews*, 3, 1–100.

Elliott, G., and A. Timmermann (2008): "Economic Forecasting," *Journal of Economic Literature*, 46, 3–56.

Fatás, A., and I. Mihov (2001): "The Effects of Fiscal Policy on Consumption and Employment: Theory and Evidence," Cepr discussion paper, 2760.

Fixler, D., R. Greenaway-Mcgrevy, and B. Grimm (2011): "Revisions to GDP, GDI and Their Major Components," *Survey of Current Business*, 91, 9–31.

Fixler, D., and J. Nalewaik (2009): "News, Noise, and Estimates of the "True" Unobserved State of the Economy," BEA Working Paper, 0068.

Foroni, C., M. Marcellino, and C. Schumacher (2013): "U-MIDAS: MIDAS Regressions with Unrestricted Lag Polynomials," *Journal of the Royal Statistical Society*, Series A (fortcoming).

Fullerton, T. (1989): "A Composite Approach to Forecasting State Government Revenues: Case Study of the Idaho Sales Tax," *International Journal of Forecasting*, 5, 373–380.

GAO (2013): "State and Local Governments Fiscal Outlook, April 2013 Update," Discussion paper, U.S. Government Accountability Office.

Ghysels, E. (2012): "Macroeconomics and the Reality of Mixed Frequency Data," Working paper, University of North Carolina at Chapel Hill.

Ghysels, E., C. Foroni, and M. Marcellino (2013): "Mixed Frequency Vector Autoregressive Models," *Advances in Econometrics* forthcoming.

Ghysels, E., and N. Ozkan (2013): "Real-Time Forecasting of the U.S. Federal Government Budget: A Simple Mixed Frequency Data Regression Approach," Working paper, UNC Chapel Hill.

Ghysels, E., P. Santa-Clara, and R. Valkanov (2005): "There is a Risk-Return Trade-off After All," *Journal of Financial Economics*, 76, 509–548.

——— (2006): "Predicting volatility: Getting the Most out of Return Data Sampled at Different Frequencies," *Journal of Econometrics*, 131, 59–95.

Ghysels, E., A. Sinko, and R. Valkanov (2007): "MIDAS Regressions: Further Results and New Directions," *Econometric Reviews*, 26, 53–90.

Giacomini, R., and H. White (2006): "Tests of Conditional Predictive Ability," *Econometrica*, 74, 1545–1578.

Giannone, D., L. Reichlin, and D. Small (2008): "Nowcasting: The Real-Time Informational Content of Macroeconomic Data," *Journal of Monetary Economics*, 55, 665–676.

Giles, C., and J. Hall (1998): "Forecasting the PSBR Outside Government: The IFS Perspective," *Fiscal Studies*, 19, 83–100.

Grizzle, A., and W. Klay (1994): "Forecasting State Sales Tax Revenues: Comparing the Accuracy of Different Metods," *State and Local Government Review*, 26, 142–152.

Holloway, T. (1989): "Measuring the Cyclical Sensitivity of Federal Receipts and Expenditures: Simplified Estimation Procedures," *International Journal of Forecasting*, 5, 347–360.

Kadiyala, K., and S. Karlsson (1997): "Numerical Methods for Estimation and Inference in Bayesian VAR-models," *Journal of Applied Econometrics*, 12, 99–132.

Koenig, E., S. Dolmas, and J. Piger (2003): "The Use and Abuse of Real-Time Data in Economic Forecasting," *Review of Economics and Statistics*, 85, 618–628.

Kuzin, V., M. Marcellino, and C. Schumacher (2011): "MIDAS versus Mixed-Frequency VAR: Nowcasting GDP in the Euro Area," *International Journal of Forecasting*, 27, 529–542.

——— (2013): "Pooling versus Model Selection for Nowcasting GDP with Many Predictors: Empirical Evidence for Six Industrialized Countries," *Journal of Applied Econometrics*, 28, 392–411.

Landefeld, J. (2010): "Comments to J.J. Nalewaik, "The Income- and Expenditure-Side Estimates of U.S. Output Growth"," Discussion paper, Brookings Papers on Economic Activity.

Lawrence, K., A. Anandarjan, and G. Kleinman (1998): "Forecasting State Tax Revenues: A New Approach," *Advances in Business and Management Forecasting*, 2, 157–170.

Litterman, R. (1986): "Forecasting with Bayesian Vector Autoregressions: Five Years of Experience," *Journal of Business and Economic Statistics*, 4, 25–38.

Marcellino, M., and C. Schumacher (2010): "Factor MIDAS for Nowcasting and Forecasting with Ragged-Edge Data: A Model Comparison for German GDP," *Oxford Bulletin of Economics and Statistics*, 72, 518–550.

Mariano, R., and Y. Murasawa (2003): "A New Coincident Index of Business Cycles Based on Monthly and Quarterly Series," *Journal of Applied Econometrics*, 18, 427–443.

Mittnik, S., and A. Zadrozny (2005): "Forecasting Quarterly German GDP at Monthly Intervals Using Monthly IFO Business Conditions Data," Ifo Survey Data in Business Cycle and Monetary Policy Analysis.

Nalewaik, J. (2010): "The Income- and Expenditure-Side Estimates of U.S. Output Growth," *Brookings Papers on Economic Activity*, 1, 71–127.

Newey, W., and K. West (1987): "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.

Nunes, L. (2005): "Nowcasting Quarterly GDP Growth in a Monthly Coincident Indicator Model," *Journal of Forecasting*, 24, 575–592.

Onorante, L., D. Pedregal, J. Perez, and S. Signorini (2010): "The Usefulness of Intra-Annual Government Cash Budgetary Data for Fiscal Forecasting in The Euro Area," *Journal of Policy Modeling*, 32, 98–119.

Owyang, M., and S. Zubairy (2013): "Who Benefits from Increased Government Spending? A State-Level Analysis," *Regional Science and Urban Economics*, 43.

Paredes, J., D. Pedregal, and J. Perez (2009): "A Quarterly Fiscal Database for The Euro Area Based on Intra-Annual Fiscal Information," ECB Working Paper Series, 1132.

Pedregal, D., and J. Perez (2010): "Should Quarterly Government Finance Statistics Be Used for Fiscal Surveillance in Europe?," *International Journal of Forecasting*, 26, 794–807.

Peseran, M., T. Schuerann, and L. Smith (2008): "Forecasting Economic and Financial Variables with Global VARs," Federal Reserve Bank of New York Staff Reports, 317.

Pike, T., and D. Savage (1998): "Forecasting the Public Finances in the Treasury," *Fiscal Studies*, 19, 49–62.

Sentance, A., S. Hall, and J. O'Sullivan (1998): "Modelling and Forecasting UK Public Finances," *International Journal of Forecasting*, 19, 63–81.

Sims, C. A. (1980): "Macroeconomics and Reality," *Econometrica*, 48, 1–48.

Sims, C. A., and T. Zha (1998): "Bayesian Methods for Dynamic Multivariate Models," *International Economic Review*, 39, 949–968.

Stock, J., and M. Watson (2001): "A Comparison of Linear and Nonlinear Univariate Models for Forecasting Macroeconomic Time Series," pp. 1–44. Festschrift in Honour of Clive Granger , Cambridge University Press, Cambridge.

——— (2004): "Combination Forecasts of Output Growth in a Seven-Country Data Set," *Journal of Forecasting*, 23, 405–430.

——— (2008): "Phillips Curve Inflation Forecasts," NBER Working paper, 14322.

Timmermann, A. (2006): "Forecast Combinations," *Handbook of Economic Forecasting*, 1, 135–196.

Wright, J. (2008): "Bayesian Model Averaging and Exchange Rate Forecasting," *Journal of Econometrics*, 146, 329–341.

Zadrozny, P. (1990): "Forecasting US GNP at Monthly Intervals with an Estimated Bivariate Time Series Model," *Federal Reserve Bank of Atlanta Economic Review*, 75, 2–15.