Real-time forecasting of the US federal government budget: A simple mixed frequency data regression approach

CrossMark

Eric Ghysels [a,b,*], Nazire Ozkan [b]

[a] Kenan-Flagler Business School, Chapel Hill, NC, United States
[b] Department of Economics, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

A R T I C L E   I N F O

A B S T R A C T

This paper proposes a real-time forecasting procedure involving a combination of MIDAS-type regression models constructed with predictors of different sampling frequencies for predicting the annual U.S. federal government current expenditures and receipts. The evidence shows that forecast combinations of MIDAS regression models provide forecast gains over traditional models, which suggests the use of mixed frequency data consisting of fiscal series and macroeconomic indicators for forecasting the annual federal budget. It is also shown that, although this was not statistically significant, MIDAS regressions with quarterly leads that are employed to include real-time forecast updates of the current year federal expenditures and receipts are found to have improved forecast performances compared to MIDAS regressions without leads.

© 2015 International Institute of Forecasters. Published by Elsevier B.V. All rights reserved.

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. According to the Congressional Budget Office (CBO), a nonpartisan government agency whose main goal is to provide accurate forecasts of the federal budget, the 2012 fiscal year's deficit of $1.089 trillion was the fourth year in a row with a deficit of more than $1 trillion (CBO, 2012). The federal deficit has since decreased sharply from previous shortfalls, and the CBO estimates that, under current laws, it will total $514 billion in the 2014 fiscal year, which will be $166 billion lower than the figure posted in 2013 (CBO, 2014).

Predictions for federal expenditures and receipts will differ from the actual outcomes even if federal laws remain unchanged, due to unanticipated changes in economic conditions and the factors that affect federal spending and revenues. That is, fiscal policy is surrounded by uncertainties, in both 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 US government figures out what is happening and what actions are necessary. Second, there will be legislative lags between when legislation is first proposed and when it is enacted, since it takes time both to develop a fiscal package on which the majority will agree and to pass it through Congress. Finally, there will also be implementation lags between the time when additional government spending is passed, and when contracts are extended and the spending actually occurs. Since it is not possible to deal with legislative uncertainties in this context, this paper proposes a forecasting procedure that

---

* Corresponding author.
    E-mail addresses: eghysels@gmail.com (E. Ghysels),
nazireozkan@gmail.com (N. Ozkan).

deals with economic uncertainties by employing econometric models that incorporate fiscal and macroeconomic indicators sampled at mixed frequencies with annual budget data for predicting current federal expenditures and receipts.

Federal government current expenditures and receipts are considered as measures of the federal fiscal position. In order to forecast and monitor the annual fiscal position, it is essential to assess the implications of intra-annual fiscal data. Several papers have indicated the importance of the information contained in intra-annual fiscal data for forecasting and monitoring annual budgetary outcomes. Onorante, Pedregal, Perez, and Signorini (2010), for example, use a mixed-frequency state space model to integrate readily available monthly/quarterly fiscal data with annual government series. Pedregal and Perez (2010) and Paredes, Pedregal, and Perez (2009) show the usefulness of intra-annual fiscal data for real-time fiscal policy surveillance by estimating models with fiscal data from annual and quarterly national accounts. Asimakopoulos, Paredes, and Warmedinger (2013) use quarterly fiscal data to forecast a disaggregated set of fiscal series at an annual frequency. Our paper utilizes a set of mixed-frequency macroeconomic indicators, in addition to a set of quarterly fiscal data, for forecasting the annual federal budget. The results confirm that quarterly fiscal data, together with higher frequency macroeconomic indicators, include significant information, and 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 US federal government current expenditures and receipts are predicted using a set of mixed frequency variables.[1] 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 to compare them with the forecasts from more traditional models, namely autoregressive (AR) and augmented distributed lag (ADL) regression models. The MIDAS-regression methodology relies on a more parsimonious approach to regression analysis with data of different frequencies. Low-frequency time series data are combined with higher-frequency data without imposing any a priori aggregation schemes, and with minimal model restrictions. Using this approach, we find that the use of mixed frequency fiscal and macroeconomic data for obtaining annual forecasts of federal expenditures and receipts provides significant forecast performance gains over traditional models which use only data of one low sampling frequency. Furthermore, in order to take advantage of readily-available higher frequency data when providing

real-time forecast updates of the current year federal expenditures and receipts, a MIDAS regression with leads is employed by incorporating real-time information using weekly, monthly and 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 year-end forecast.

The remainder of the paper proceeds as follows. In Section 2, the econometric methods employed in this paper are presented within subsections that are devoted to the description of the ADL-MIDAS regression model, a method for combining forecasts, and a test of predictive accuracy, respectively. Section 3 introduces the data and Section 4 presents empirical results, showing the evidence in favor of forecast combinations of ADL-MIDAS regression models relative to AR and ADL regression models, then compares the performances of one-year-ahead forecasts from ADL-MIDAS models with those from an ADL-MIDAS with quarterly leads. The conclusions are presented in Section 5.

2. Methods

In addition to the quarterly subcomponents of federal expenditures and receipts, we select a set of macroeconomic and financial indicators, with different sampling frequencies, that can be representative of the economy, and each is used separately as a predictor for both US federal government current expenditures and receipts. In order to deal with data sampled at different frequencies, we employ Mixed Data Sampling (MIDAS) type regression models, and specifically Augmented Distributed Lag–Mixed Data Sampling (ADL-MIDAS) regression models. 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 data for time periods up to that point in time. In-sample estimates are produced with the vintages of data restricted to be only those available at that time, while forecasts are obtained using the latest values from the most recent vintage.[2] After constructing the longest possible samples with the data available, individual forecasts with each single indicator are obtained, resulting in multiple forecasts of expenditures and receipts. Stock and Watson (2001), employing various different types of models and variables, find that combined forecasts generally outperform the forecast performance of the best individual model. Also, as was indicated by Timmermann (2006), it is not reasonable to think that the same individual model will dominate all of the others at all points in time, since forecasting models are considered to be local approximations. Therefore, in order to obtain more accurate forecasts by using evidence from

---

[1] As was shown by 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.

[2] 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.