Working Paper Series
Congressional Budget Office
Washington, D.C.

AN EVALUATION OF CBO FORECASTS

Jonathan Huntley (E-mail: jonathan.huntley@cbo.gov)
Congressional Budget Office
Washington, D.C.

Eric Miller (E-mail: eamiller@artsci.wustl.edu)
Formerly of Congressional Budget Office
Washington, D.C.

August 2009
2009-02

Abstract

We compare the performance of a subset of CBO's economic forecasts against that of an unrestricted vector autoregression (VAR) model. We evaluate forecasts of real economic indicators as well as budget-related nominal statistics. We find that under most specifications, the VAR performs competitively with, if slightly worse than, the corresponding CBO forecasts at up to 20 quarters. Therefore, a simple VAR is unlikely to be able to contribute directly to improving budget forecasts. The only series for which the VAR outperforms the CBO forecast is the growth in real consumption.

1. Introduction

The Congressional Budget Office's (CBO) economic forecasts play an important role in the federal budgetary process. Inaccurate forecasts of key economic variables can distort the overall budgetary picture because projected tax revenues are heavily influenced by future macroeconomic conditions.[1] The variables with the greatest effect on future revenues are nominal gross domestic product (GDP), the capital and labor income shares of GDP, the unemployment rate, and various interest rates.

Every year, CBO evaluates the accuracy of a subset of the approximately 60 variables in its economic forecasts relative to those of the Office of Management and Budget (OMB) and the Blue Chip Consensus, an average of private-sector forecasts.[2] CBO forecasts were generally in line with these two benchmarks over the period from 1982 to 2005. Although this comparison is useful, it is just one of many tools that can be used to evaluate CBO's forecasts. To this end, we propose using a vector autoregression (VAR) as an alternative method for evaluating both nominal variables that directly affect the budget and real variables that indirectly influence budget forecasts. We find that by our measures, the CBO forecast performs about as well as or slightly better than the VAR forecast except with regard to real consumption, for which the CBO forecast shows a large downward bias. Although this suggests a need to revisit the procedures by which real consumption is forecast, this finding has little impact on the variables that directly affect budget predictions.

Direct comparison between forecasts may be misleading because forecasts are rarely made in an informational vacuum. In the absence of explicit structural models, forecasters may rely on one another to inform their forecasts. If a forecaster includes other institutions' forecasts in his or her information set, then correlated outcomes are inevitable.

[1] Projected government spending is influenced by macroeconomic variables as well, but to a lesser extent.

[2] For a further discussion of CBO's forecast accuracy see CBO (2007a)

Furthermore, forecasters may use other institutions' forecasts as implicit parameters in their loss functions. In other words, accurate forecasts that differ from the consensus may yield a very small reward compared with the costs of inaccurate forecasts that differ from the consensus. As a result, forecasters may be more inclined to make forecasts that are similar with those of their competitors. Instead of comparing CBO's forecasts with those of other forecasters, this paper compares the CBO forecasts with those generated mechanically by a VAR. This avoids two potential problems: that CBO is comparing its forecast with those of institutions with different objectives, and that CBO may be making the same systematic errors as other forecasters.

To this end, we develop our own naïve forecasting model in the form of a simple VAR. This widely used forecasting method has been shown by Litterman (1986) and Edge and others (2008) to compare favorably with forecasts made by both the Federal Reserve and private-sector forecasters. Sims (1980) has also argued in favor of small time-series models. VARs are also attractive because they force modelers to make their underlying assumptions explicit. This makes for good science because it allows the model to be easily reproduced and analyzed by outside observers (Todd 1990).

We compare the results of our VAR model with 20 five-year CBO forecasts spanning a roughly 10-year window between August 1993 and August 2003. During this period CBO's forecasts were made in a wide variety of economic environments. Consistent deviations from the VAR are thus unlikely to be the product of conditions specific to a particular year or economic episode, but are instead likely due to systematic differences between the two forecasting methods. One of these differences is already well known: CBO's projections of government spending are not true forecasts, but rather baseline estimates that assume current laws and policies remain in place.[3] This rule is useful for policymaking because it

[3]This is in accordance with provisions set forth in the Balanced Budget and Emergency Deficit Control Act of 1985, the Budget Enforcement Act of 1990, and the Congressional Budget and Impoundment Control Act of 1974. For a further discussion of the baseline concept in CBO's economic forecasts see CBO (2008).

serves as a benchmark against which proposed policy changes can be measured. However, it also means that the VAR may have a forecasting advantage because it is not forced to abide by this constraint.

Our experiment provides a useful point of comparison because it can help to highlight biases and systematic errors that may not be obvious when CBO's forecasts are compared with those of other institutions. Although our experiment is far from perfect, we believe it is more transparent and informative than comparisons against competing entities with unknown objective functions. Whereas we have very little insight into the motivations and formulations of other forecasters' problems, we know the objective function that characterizes the VAR, and we have some insight into how the VAR's objective function compares with that of CBO. We find that under a variety of specifications, the VAR generally performs as well as the CBO forecasts at up to 10 quarters ahead and is still competitive at twenty quarters out. We also find that within the period of study, CBO has consistently underestimated real consumption, whereas VAR forecasts show much less downward bias in this series.

The remainder of the paper is organized as follows. Section 2 formalizes the problems inherent in comparing CBO's forecasts with those of other institutions. Section 3 briefly outlines the two forecast models used in our analysis. Section 4 describes our data set and the procedure we used to compare forecasts. We discuss the results of our experiment in section 5. Section 6 concludes.

2. The Comparison Problem

Using mean squared error (MSE) to compare two forecasts may not be appropriate if they are formed using two dissimilar objective functions. Consider the following generic loss function for the forecast of a particular economic indicator $\mathbf{X}$ beginning in period p and

evaluated over the next t periods:

$$L(p, t, \mathbf{X}_{p+t}, \Psi_p, \widehat{\mathbf{X}(\Psi_p)}_{p+t}).$$

Here we denote the forecast period as p, the time from the date when the forecast was made as t, the realized observation at that point as $\mathbf{X}_{t+p}$, and the forecast value at that time period as $\widehat{\mathbf{X}(\Psi_p)}_{t+p}$, where Ψ_p is the information set available to the forecaster at time p. This loss function takes the time period, the forecast horizon, the information set, the forecast, and the actual data as parameters. Presumably, the forecaster is attempting to minimize the expected loss over all possible outcomes.

In the past, CBO has evaluated its forecasts using three simple loss functions: the mean error, the absolute mean error, and the root MSE. CBO then uses these metrics to compare the quality of its forecasts against those of other professional forecasters. Any forecast comparison that uses the root MSE assumes a loss function that is some transformation of

$$L(.) = (\mathbf{X}_{p+t} - \widehat{\mathbf{X}(\Psi_p)}_{p+t})^2.$$

However, we consider a more generic loss function that can accommodate the possibility that these other forecasters have different objectives in mind:

$$L(.) = W(\mathbf{X}_{p+t})M(\mathbf{X}_{p+t} - \widehat{\mathbf{X}(\Psi_p)}_{p+t}),$$

where W is some generic weight function, in this case of the realized state, and M is some nonlinear function of the difference between the realized and forecast states. This formulation suggests that different weights may be applied to different outcomes. For example, forecasters may place a different weight on missing a recession than on missing

a rapid economic expansion. This hypothesis is advanced by a large literature, particularly Bennett, Geoum, and Laster (1999), who find that if an institution is more likely to benefit from publicity than its peers, its forecasts tend to exhibit greater deviation from the Blue Chip Consensus at the cost of forecast accuracy.

A simple comparison of MSEs implicitly assumes that these errors are independent across institutions. However, this is highly unlikely because forecasts are rarely made in an informational vacuum. If institutions rely on one another to inform their forecasts, even informally, then outcomes are bound to be correlated. In this case the loss function that is used when maximizing a forecaster's objective function may look like

$$L = W(\Psi_p, \widehat{\mathbf{X}(\Psi_p)}^{i}_{p+t}, \widehat{\mathbf{X}(\Psi_p)}^{-i}_{p+t}) M(\mathbf{X}_{p+t} - \widehat{\mathbf{X}^i(\Psi_p)}_{p+t})$$

where i refers to one particular forecast, W is the generic weight function, in this case a function of the information set and the set of forecasts; $-i$ refers to all other forecasts; and the information set Ψ_p can be thought of as containing some set of publicly available forecasts.

Bauer and others (2003) find evidence that the Blue Chip Consensus is unbiased and tends to outperform any individual forecaster. This consensus forecast is updated monthly and is used as an input in CBO's forecasts. Because CBO uses this information to weight the plausibility of its forecasts, one should not be surprised to find that the two forecasts tend to show similar patterns of MSE. Only when the MSEs of the two forecasts differ significantly does the observer gain any insight into whether CBO's forecasts add value.

By estimating the posterior density of a VAR, we provide a different metric by which to judge the accuracy of the CBO forecast. The objective function for the VAR is well defined and is not subject to the biases mentioned above.

3. CBO's Forecasts and a Simple VAR

3..1 CBO's Forecasts

CBO makes economic projections for 10 years from the year in which the forecast is formulated (for example, the January 2008 projection extends to 2018). CBO generally emphasizes two different forecast horizons in its reports: the near term (the "forecast period"), which constitutes the first two years, and the medium term (the "projection period"). The forecasts over these two horizons are formulated using separate but interrelated processes.

The near-term forecast is based on CBO's assessment of the various factors that influence the economy in the short run. Econometric models are used to investigate issues such as the effect of housing wealth on personal consumption spending or the effect of import prices on inflation. An econometric model is also used to ensure consistency within the forecast, but CBO does not rely on an explicitly defined macroeconomic model when preparing its forecast. The CBO forecast incorporates various sources of information such as private-sector forecasts, private-sector issue-specific analyses, futures markets, short-run indicators and surveys, the federal budget outlook, and recent monetary policy changes. In addition, CBO staff present a preliminary version of the forecast to an outside panel of economic advisers for discussion.

The medium-term projections are driven largely by CBO's estimates of long-run trends in the economy. The projections are more rule-based and do not take the business cycle into account. For example, there is no attempt to forecast year-to-year changes in inflation rates. Therefore, CBO characterizes the projected values as averages for the projection period. Some of the long-run economic trends that are used to anchor the medium-term projections are CBO's assessments of potential GDP, the natural rate of unemployment, the natural rate of interest, the average labor income share of GDP, and

various relationships among interest rates and inflation measures. Real GDP is projected to converge to CBO's estimate of potential GDP during the medium term, the labor share of GDP to converge to its long-run average, and the unemployment and interest rates to converge to their natural rates. Rules are used to ensure that some longer-run relationships are maintained in the projection. For example, the projection of private investment (an important component of the estimate of potential GDP) is affected by estimates of depreciation, the federal deficit, and the growth of the labor force. Lastly, the projection for inflation assumes that the Federal Reserve will be able, on average, to achieve an inflation rate for the price index for personal consumption expenditures within the putative "target band" of 1 to 2 percent. Federal government spending is an exogenous input in the model.[4]

3..2 The VAR

We consider a very simple VAR model that features a variety of macroeconomic variables. The standard form we use is

$$\mathbf{y}_s = \mathbf{C} + \mathbf{\Phi}_1 \mathbf{y}_{s-1} + \mathbf{\Phi}_2 \mathbf{y}_{s-2}, \cdots + \mathbf{\Phi}_n \mathbf{y}_{s-n} + \boldsymbol{\epsilon}_s \tag{1}$$

where $\mathbf{y}_s$ is an (n x 1) vector of variables at date s, $\mathbf{C}$ is a vector of constants, and $\boldsymbol{\epsilon}_t \sim$ i.i.d. $N(\mathbf{0}, \mathbf{\Omega})$.[5]

There is a possibility that unanticipated structural breaks can severely erode forecast accuracy. Such breaks are endemic in macroeconomic time series (Stock and Watson 1996). However, Sims and Zha (2006) and Primiceri (2005) show that for monetary

[4]For a further discussion of the methodology underlying CBO's forecasts see CBO (2006) and CBO (2007b).

[5]We also construct a version of the model that includes a linear time trend. In general, this specification yields little or no improvement in forecast accuracy and in some cases produces forecasts that are significantly less accurate than the standard VAR. In light of these findings, we do not report these results.

structural models, the change tends to manifest itself in the decomposition of the error-term structure, which does not affect the VAR forecasts. Consistent with past experience such as that of Litterman (1986), we first-difference the log values of our data. This greatly improves the accuracy of our forecasts, a feature also noted by Hendry and Clements (2003).

We assume flat priors[6] and apply Gibbs sampling to generate posterior densities and expectations for the forecast in each period.[7] We use the Akaike information criteria on ordinary least squares (OLS) estimates to set the number of lags, thereby removing one layer of subjective decision making from a process that could result in more favorable results for the VAR forecast.

4. Data and Comparison Procedure

We analyzed 20 CBO forecasts made between August 1993 and August 2003 and compared the total forecast error against the error produced by the VAR over a five-year forecast horizon (20 quarters).[8] We repeated this experiment for a variety of specifications. The optimal lag length, on the basis of the Akaike information criterion, was found to be one quarter in all cases.

In order to make fair comparisons, the VAR was fed the same unrevised data that were available to the CBO modelers when the forecast was initially released. We account for data revisions by normalizing the difference of the log of the forecasts and historical data to zero in the initial forecast period. For example, if we want to examine the series $lnGDP_{p+t,p}$, where p is the forecast period and t is a quarter subsequent to the forecast

[6] See Bauwens, Lubrano, and Richard (1999) for the precise formulation.

[7] We also constructed a Bayesian version of the model that makes use of the Minnesota prior. This specification yielded no apparent improvement in forecast accuracy. We do not report these results.

[8] We selected these dates based on data availability. CBO produces economic forecasts twice a year, usually in January and August. CBO did not produce a forecast in August 1996.

period, we generate the series

$$\overline{lnGDP_{p+t,p}} = lnGDP_{p+t,p} - lnGDP_{p,p} \tag{2}$$

We evaluate the performance of the VAR and CBO forecasts by taking the MSE of each individual economic indicator at up to 20 quarters subsequent to the forecast period p. There are J distinct forecasts, and the function $p(j)$ is the time period in which forecast number j was constructed. For example, the MSE of the GDP series, t periods beyond forecast j is defined in the two models as:

$$MSE_t^{VAR}(GDP) = \frac{\sum_{j=1}^{J}\left(\overline{lnGDP^{VAR}_{p(j)+t,p(j)}} - \overline{lnGDP_{p(j)+t,p(j)}}\right)^2}{J} \tag{3}$$

$$MSE_t^{CBO}(GDP) = \frac{\sum_{j=1}^{J}\left(\overline{lnGDP^{CBO}_{p(j)+t,p(j)}} - \overline{lnGDP_{p(j)+t,p(j)}}\right)^2}{J} \tag{4}$$

where $lnGDP^{VAR}_{p(j)+t,p(j)}$ is the VAR forecast and $lnGDP^{CBO}_{p(j)+t,p(j)}$ is the CBO forecast. We construct the distribution for this statistic for $t = 1, 2, ..., 20$, and we plot the expected value and the 90 percent interval for this statistic at all of the aforementioned time periods.

5. Results

We performed a number of experiments to evaluate the accuracy of CBO forecasts. The results for each experiment are presented in the Figures section. Unless otherwise noted, all of the results are MSEs. We simulated each experiment with 2,500 draws from the sample, dropping the first 500 to allow for convergence. In all graphs, the dashed line corresponds to the CBO forecast, the dash-dotted line is the expected value of the VAR forecast, and the shaded area represents the 90 percent bounds on the marginal distribution of the forecast at each point in the forecast horizon.

Due to problems inherent in forecasting nominal series using VARs, we begin by forecasting several real variables and price indices that are of interest to CBO: GDP, profits, wages, nonwage income, unemployment, and both the GDP and personal consumption expenditure (PCE) price indices.[9] The results, in terms of MSEs, are shown in Figure 1. At up to 10 quarters the VAR forecasts about as well as the CBO for all the variables of interest. At up to 20 quarters out the VAR is still competitive, and it does a better job of forecasting real GDP and produces noticeably better estimates for the future path of the PCE price index.[10] Assuming that the errors are uniformly distributed across the forecasts, the difference in MSE suggests that except in the case of unemployment, there is typically less than a 1 percent average difference between the forecasts at the five-year horizon.

Since CBO budget projections are constructed using nominal indicators whereas the VAR is used to predict real variables, we inflate the real variables in this experiment using the corresponding set of price indices: we inflate the real GDP forecast using the forecast value of the GDP deflator, and we inflate the components of income forecasts using the forecast value of the PCE deflator. We construct the marginal distributions of the nominal variables and report them in Figure 2. The upper-left panel demonstrates that the two forecasts of nominal GDP are now very close, with the VAR performing slightly better in the third and fourth years. For the rest of the variables, the VAR forecasts are competitive to about 10 quarters and then, except for the forecast of nominal wages, slightly worse on average. CBO does a noticeably better job of predicting profits. This error can be largely attributed to the VAR greatly underestimating the growth of profits during the late 1990s and the recovery after the 2001 recession. Aside from this discrepancy, there is little statistical or economic significance to the differences between the CBO and the

[9] Nonwage income consists of proprietor's income, rental income, personal interest income, and dividends.

[10] We repeated this experiment using two lags and observed a very similar outcome; we omit this specification from the results.

VAR forecasts.

The more important nominal forecasts show that although the VAR does a marginally better job of forecasting real GDP, this does not translate into a superior forecast of nominal GDP. A similar result is observed in the other time series, and from a budgeting perspective, nominal series are the ones that matter.

To determine how well the CBO forecast predicts real economic activity, we consider a specification for the main components of the national accounts: real investment, consumption, exports, imports, and government spending. We disaggregate investment into its nonresidential and residential components because these two series are often thought to move separately–as noted by Edelberg, Eichenbaum, and Fisher (1999)–and aggregating them may remove valuable information. Furthermore, we disaggregate government spending into state and local as well as federal spending, because the forecast of state and local consumption is not subject to the same rules as the projection of federal consumption.[11] This specification is subordinate to the other one in which we forecast budget-related variables because it does not directly affect budget projections.

The results from this experiment with seven variables are presented in Figure 3.[12] In most cases the two forecasts are close to each other, as measured by the MSE. Assuming that the errors are uniformly distributed across forecasts, and with the exception of real consumption, the forecasts do not appear to be more than 3 percent different at the five-year horizon. The forecasts of nonresidential investment are very close. The CBO forecast outperforms the VAR by a narrow margin, a result which is exactly flipped for residential investment. Both results are well within the 90 percent bounds on the marginal posterior distribution. CBO's projection of federal spending has a substantially higher MSE than

[11] Federal spending is considered an exogenous input in CBO's economic forecasts.

[12] Here we considered two specifications, one with real GDP and one without. Although chain-weighted real GDP is not a static linear combination of the real components of GDP, it is close enough that it adds very little information to the forecast. This specification looked nearly identical to the specification that did not include real GDP. Therefore, we omit the specification that includes real GDP and report only the version described in the main text.

the expected MSE for the VAR, but still remains within the 90 percent interval. The VAR does objectively worse at predicting state and local expenditures, with an MSE that is nearly twice that of CBO's forecast in the final forecast quarter. Forecast imports and exports are nearly identical, except for a small and unimportant deviation at the end of the forecast horizon.

The largest and most noticeable difference between the forecasts is in real consumption, which is not used explicitly in producing budget projections as it does not directly generate federal tax revenue. Here, the VAR greatly outperforms the CBO's forecast, particularly for horizons greater than two years. Personal consumption was unusually stable during the years in which the VAR forecasts were constructed, and this may make the MSE appear to be particularly low.[13] In order to determine whether or not this error is due to a higher-volatility forecast, we plot the average errors at every horizon in Figure 4.[14] Using this same measure, we plot the error generated by the VAR for purposes of comparison.

The average error in the VAR forecast at the five-year horizon is slightly more than 1 percent. The 90 percent marginal distribution of the forecast errors includes zero, which is well in line with expectations. The CBO forecast, however, has a bias exceeding 5 percent at the five-year horizon. Given that the MSE presented in Figure 3 is not that large and consumption is not that volatile, this suggests a consistent underestimate of private consumption. Although the CBO forecast shows little bias in the short run (one year or less), the principles that guide the longer-term estimates may have biased the forecast downward. The VAR predictions for consumption are extremely stable; the variation from forecast period to forecast period is negligible in the unreported VAR forecasts.[15] Given the minor differences in the other components of GDP, it may be worth putting more

[13]Before 2008, the last time year-over-year growth in personal consumption dropped below zero was in 1992.

[14]The actual CBO forecasts for this variable are unpublished; therefore we omit them from the paper.

[15]The VAR does predict a substantial decline in consumption based on data through the third quarter of 2008.

weight on historical consumption patterns rather than trying to predict an inflection point in this series. Furthermore, it may be worth reexamining the procedure used to generate forecasts of personal consumption.

In short, aside from real consumption, CBO's forecasts compare very favorably with the VAR. Most variables generated endogenously by macroeconomic forecasters at the agency have MSEs that are roughly comparable to or better than the VAR counterparts.

6. Conclusion

If one compares the MSE of CBO's forecasts against that of other forecasting institutions, the agency seems to be doing quite well. However, there is some concern that such comparisons are of limited usefulness. We aimed to develop an experiment that would avoid these problems and make evaluations of CBO's forecasts more transparent and informative. To achieve this, we created a simple VAR model under several different specifications and used a well-defined method for choosing the number of lags to help avoid the criticism of "data mining," or selecting models that appear to bias the results against CBO.

Following some of the recommendations set forth by Hendry and Clements (2003), we found that the VAR forecast method was competitive with CBO forecasts for most of the variables. These simple suggestions helped guide the construction of systems of equations that, with one exception, were roughly comparable to the CBO forecasts. The agency does as well as or better than the VAR at forecasting nominal budget variables directly used in formulating budget forecasts. As for real variables, the agency does as well or slightly worse among the set of components of GDP, except for personal consumption, which CBO has consistently underpredicted. The VAR shows that choosing a consistent estimate for real consumption has historically been a far better approach than attempting

to pick points of inflection in what is a low-variance series. This result suggests that CBO should revisit any rules it has in determining personal consumption.

Although the VAR may offer some guidance on the growth of consumption, it has a few disadvantages compared with the existing CBO forecast. Compared with the rest of the data used to make these forecasts, much of the recent U.S. experience has been outside of the space of the regressors on which the VAR is estimated. The large-scale banking crisis that characterized 2008 and 2009 is unique to the data, in this environment the VAR may do a poor job of extrapolating an accurate forecast from historical experience.

Furthermore, whereas the VAR is a rule-based forecast, CBO's current method allows the flexibility to impose a more realistic response to ongoing economic conditions. Our VAR, as it is formulated, is based on quarterly data. The CBO, by virtue of having access to higher frequency data, may have access to information that is unavailable when the VAR forecasting procedure is used. Lastly, the VAR performs slightly worse when it comes to predicting the budget-related variables that are of greatest interest in CBO's budget forecasts. Therefore, as a forecasting tool, the VAR offers few clear advantages over CBO's normal forecasting procedure in the current economic environment. It does, however, provide a reliable benchmark against which to compare CBO's forecasts over the medium term and serves as a reminder that sometimes a stable, uneventful forecast can be nearly as accurate as a more elaborately justified prediction.

References

Bauer, A., R.A. Eisenbeis, D.F. Waggoner, and T. Zha, "Forecast Evaluation with Cross-Sectional Data: The Blue Chip Surveys," *Economic Review - Federal Reserve Bank of Atlanta*, 2003, *88* (2), 17–32.

Bauwens, L., M. Lubrano, and J.F. Richard, *Bayesian Inference in Dynamic Econometric Models*, Oxford University Press, 1999.

Bennett, P., I.S. Geoum, and D. Laster, "Rational Bias in Macroeconomic Forecasts," *Quarterly Journal of Economics*, 1999, *114* (1), 293–318.

Congressional Budget Office, "CBO's Method for Estimating Potential Output: An Update," 2001.

_ , "How CBO Forecasts Income," 2006.

_ , "CBO's Economic Forecasting Record: 2007 Update," 2007a.

_ , "How CBO Projects the Real Rate of Interest on 10-Year Treasury Notes," 2007b.

_ , "The Budget and Economic Outlook: Fiscal Years 2008 to 2018," 2008.

Edelberg, W., M. Eichenbaum, and J.D.M. Fisher, "Understanding the Effects of a Shock to Government Purchases," *Review of Economic Dynamics*, 1999, *2* (1), 166–206.

Edge, R.M., M.T. Kiley, and J.P. Laforte, "A Comparison of Forecast Performance Between Federal Reserve Staff Forecasts, Simple Reduced-Form Models, and a DSGE Model," *Finance and Economics Discussion Series, Federal Reserve Board*, 2009.

Gelman, A., J.B. Carlin, H.S. Stern, and D.B. Rubin, "Bayesian Data Analysis. Texts in Statistical Science," *Chapman & Hall/CRC, Boca Raton, FL*, 2004, *25*, 668.

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

Hendry, D.F. and M.P. Clements, "Economic Forecasting: Some Lessons from Recent Research," *Economic Modelling*, 2003, *20* (2), 301–329.

Landefeld, R. and R. Parker, "BEA's Chain Indexes, Time Series, and Measures of Long-Term Economic Growth," *Survey of Current Business*, 1997.

Litterman, R.B., "Forecasting with Bayesian Vector Autoregressions: Five Years of Experience," *Journal of Business & Economic Statistics*, 1986, *4* (1), 25–38.

Primiceri, G.E., "Time Varying Structural Vector Autoregressions and Monetary Policy," *Review of Economic Studies*, 2005, *72* (3), 821–852.

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

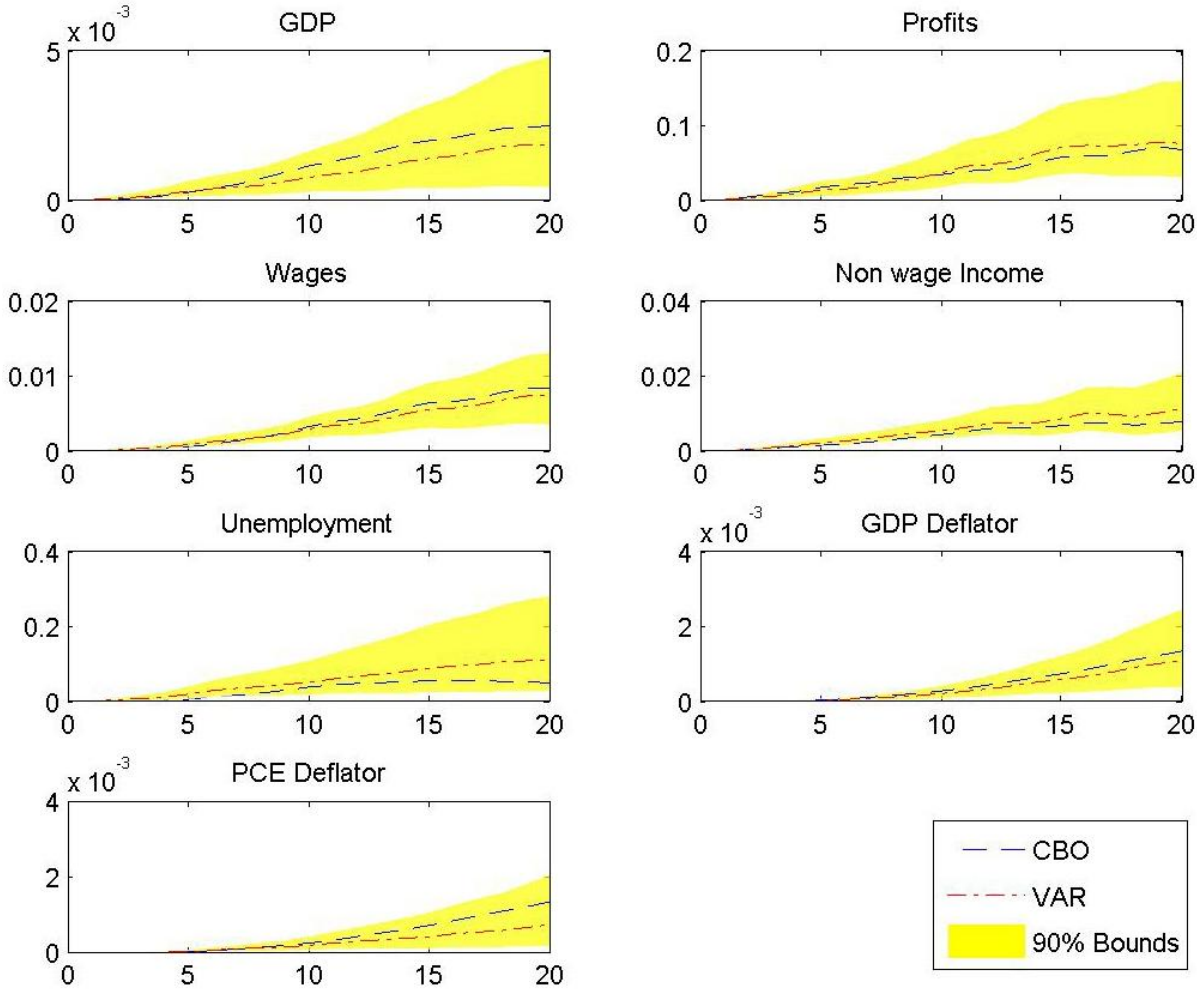
_ **and T. Zha**, "Were There Regime Switches in US Monetary Policy?," *American Economic Review*, 2006, *96* (1), 54–81.

Stock, J.H. and M.W. Watson, "Evidence on Structural Instability in Macroeconomic Time Series Relations," *Journal of Business & Economic Statistics*, 1996, *14* (1), 11–30.

Todd, R.M., "Improving Economic Foreasting with Bayesian Vector Autoregression," in C.W.J. Granger, ed., *Modelling Economic Series: Readings in Econometric Methodology*, Oxford: Oxford University Press, 1991, chapter 10, pp. 214–34.

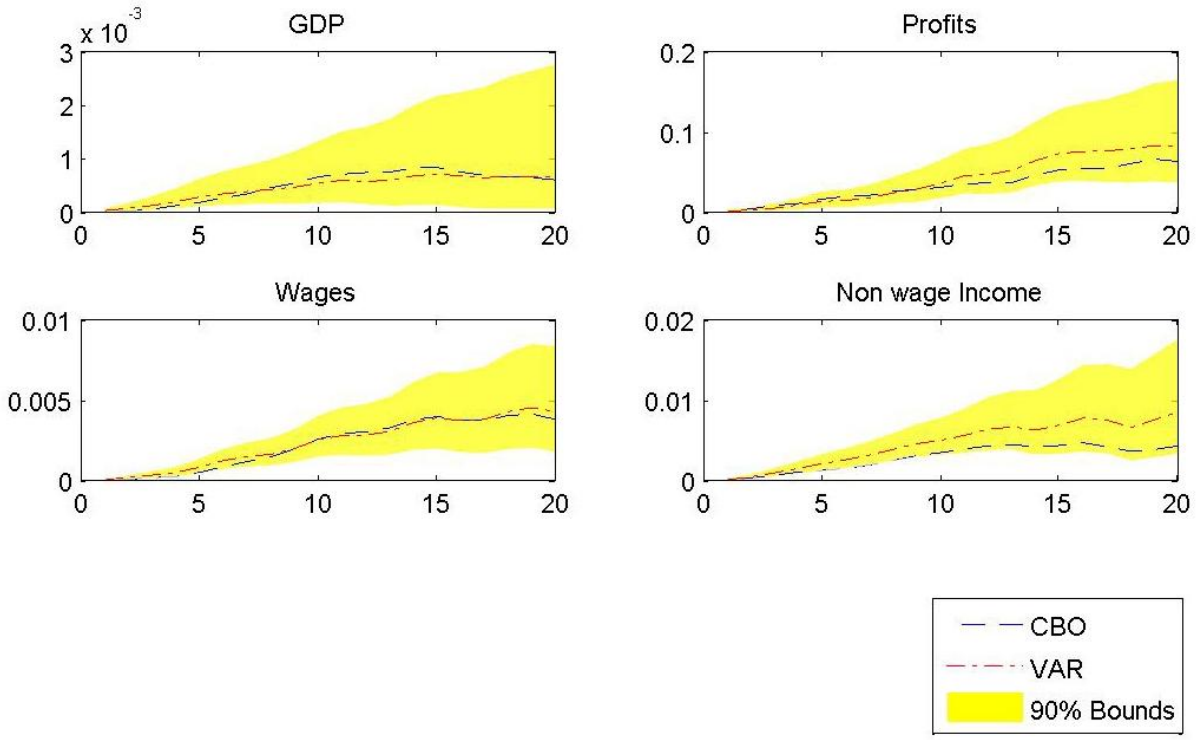
A Figures

Figure 1: MSEs of CBO and VAR Forecast of Real Budget Variables



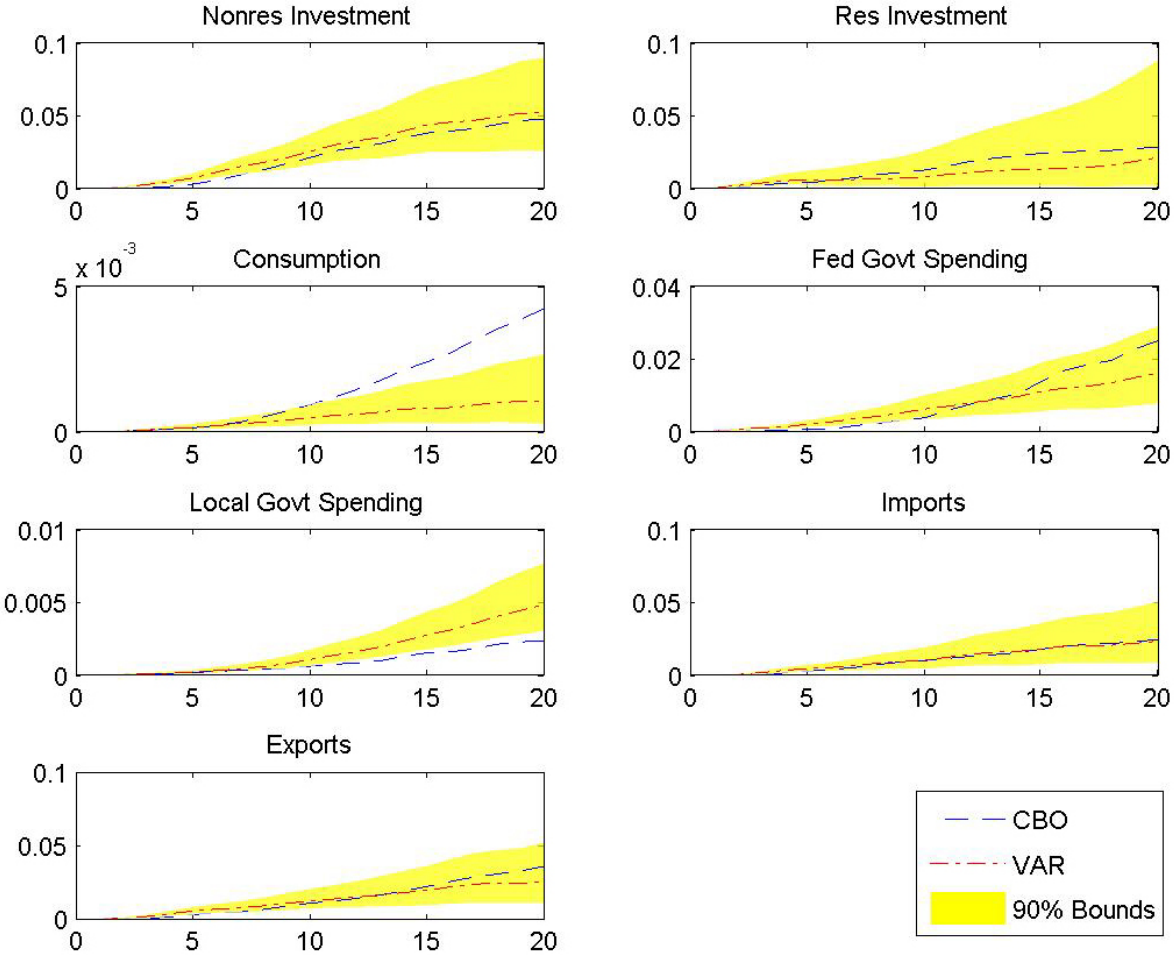
Note: Values on the x axis represent forecast quarters. Values on the y axis represent MSE.

Figure 2: MSEs of CBO and VAR Forecast of Nominal Budget Variables



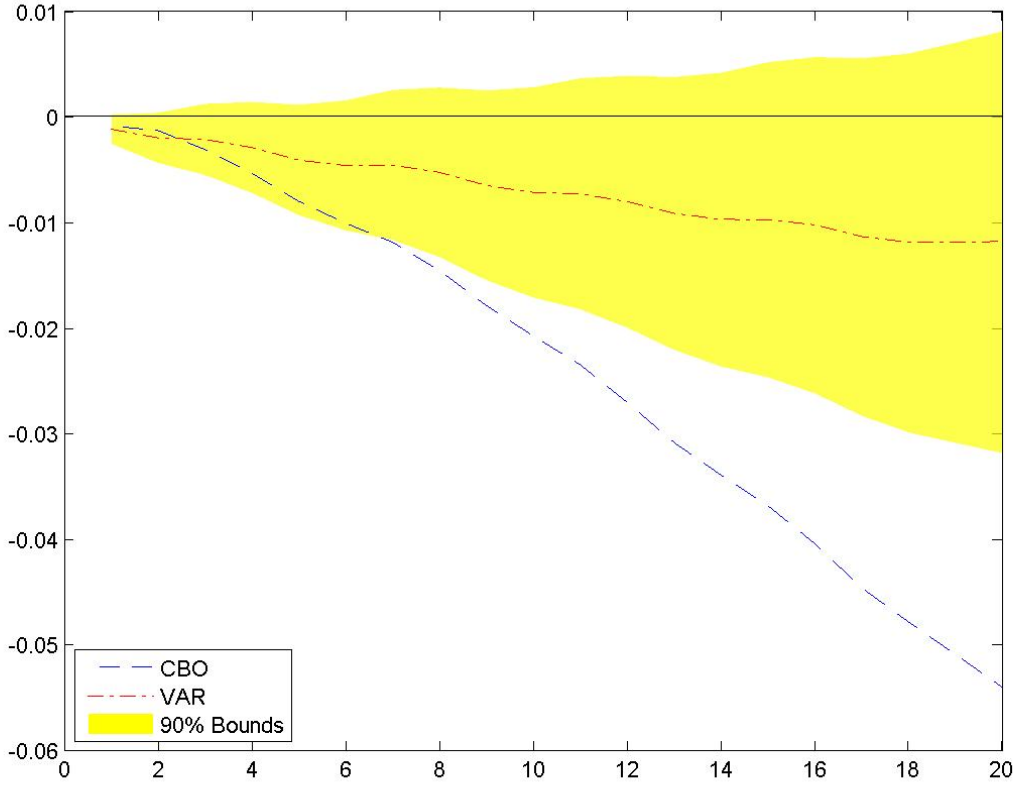
Note: Values on the x axis represent forecast quarters. Values on the y axis represent MSE.

Figure 3: MSEs of CBO and VAR Forecast of Components of Real GDP



Note: Values on the x axis represent forecast quarters. Values on the y axis represent MSE.

Figure 4: Average Error in CBO and VAR Forecasts of Real Consumption



Note: Values on the x axis represent forecast quarters. Values on the y axis represent the average error. The y axis values multiplied by 100 give the percent error.