Forecasting State Tax Revenue: A Bayesian Vector Autoregression Approach

By

Robert Krol*

Professor
Department of Economics
California State University, Northridge
Northridge, CA 91330-8374
Robert.krol@csun.edu

August 2010

Abstract

This paper compares alternative time-series models to forecast state tax revenues. Forecast accuracy is compared to a benchmark random walk forecast. Quarterly data for California is used to forecast total tax revenue along with its three largest components, sales, income, and corporate tax revenue. For one- and four-quarter-ahead forecasts from 2004 to 2009, Bayesian vector autoregressions generally forecast best based on root mean squared errors compared to standard vector autoregressions or a random walk model. Similar to the macroeconomic forecasting experience, Bayesian vector autoregressions should be consider as a useful and cost effective revenue forecasting model for state governments.

*This paper was written while on sabbatical leave during the Fall 2009 semester. The author thanks the College of Business and Economics for financial support. I thank Shirley Svorny for helpful comments.

Forecasting State Tax Revenue: A Bayesian Vector Autoregression Approach

INTRODUCTION

This paper examines the state tax revenue forecasting performance of alternative time-series models. There is a small economics literature that examines state tax revenue forecasting.[1] These studies have examined the tax revenue forecasting performance of standard regression models and vector autoregressions. No papers directly compare the tax revenue forecasting performance of alternative time-series models.

Vector autoregressions (VARs) and Bayesian-vector autoregressions (BVARs) have become standard models in macroeconomic analysis and forecasting.[2] They have also been used as effective tools in regional forecasting.[3] Bayesian vector autoregressions generally outperform standard vector autoregressions and simple univariate models in forecasting macroeconomic variables. This paper will determine whether the superiority of BVARs holds when applied to forecasting state tax revenues.

Unit root tests suggest the tax revenue data used in this paper follow a random walk. As a result, I use the random walk forecast as a benchmark forecast to be compared to the VARs and BVARs.[4] Comparing alternative forecasting models to a simple univariate model is common in the forecasting literature (Stock and Watson, 2003).

---

[1] Previous papers using time-series methods include Litterman and Supel (1983) for Minnesota, Otrok and Whiteman (1997) for Iowa, and Rich et. al. (2005) for New York. At the national level Baghestani and McNown (1992) use time-series methods to forecast the federal budget.

[2] See Litterman (1986), Favero and Marcellino (2005), and Banbura, Giannone, and Reiculin (2010).

[3] See Shoesmith (1995) or Banerji, Dua, and Miller (2006) as examples.

[4] This approach has a long history in the exchange rate forecasting literature. Meese and Rogoff (1983) found most structural and time series models of exchange rates fail to beat a random walk forecast. Since that time, most exchange rate model forecasts are compared to a random walk forecast. For a recent review of this literature see Rogoff and Stavrakeva (2008).

Standard VARs and BVARs are estimated and are compared based on root mean squared errors of one-quarter and four-quarters-ahead forecasts. Bivariate Granger causality tests between national (state) economic variables and state tax revenues are conducted to help determine the variables to be included in the multivariate models.

The models are applied to California.[5] The tax revenue data used is taken from the U.S. Census beginning in the first quarter of 1979 and ending in the first quarter of 2009. Total tax revenues along with the three largest components, sales, income, and corporate tax revenues are examined.

The models are first estimated for the period 1979 to 2003. Then a short run (one quarter) and long run (four quarters) forecast is made. One quarter of data is added to the sample period and the models are re-estimated to produce new forecasts. This rolling regression approach continues through 2009.

A comparison of the forecasts and actual revenue data for 2004 to 2009 allows the root mean squared error to be computed for each model. In most cases, the BVAR models have the smallest root mean squared error compared to the other models examined. As a practical issue, tax revenue forecasters should consider using Bayesian vector autoregressions when producing revenue forecasts.

The paper is organized in the following manner. The first section briefly discusses the models that are used in the forecasting exercises. Section two summarizes data issues. Section three presents the model specification tests. Section four provides an evaluation of the forecast results. The paper ends with a summary and conclusions.

FORECASTING MODELS

[5] See Williams, Ingenito, and Vasche (1999) for a discussion of tax revenue forecasting in California.

The two multivariate models estimated and used in forecasting are a standard VAR and a BVAR (Litterman, 1986 and Enders, 1995). Equation (2) represents a standard VAR model with each variable expressed in levels.

$$X_t = \sum_{j=1}^{s} \Psi_j X_{t-j} + \mu_t \quad (1)$$

Where X_t is an n x 1 vector of variables and Ψ_j is an n by n matrix of coefficients that are estimated. Because of the strong seasonal movement in tax revenue, quarterly dummy variables are included when the model is estimated.

A well known issue associated with unrestricted VARs is the over-parameterization of the model. In most cases, the number of coefficients estimated is large relative to the sample size. You end up estimating less important relationships in the data that are random and do not help in forecasting. This problem results in large out-of-sample forecast errors (Litterman, 1986 and Doan, 2007). Rather than imposing strict zero restrictions on coefficients, Litterman (1986) proposed placing weaker restrictions on the coefficients to help solve the over-parameterization problem. These restrictions imply that the coefficients on longer lags are more likely to be zero than shorter lags. However, absolute zero restrictions on long lags are not imposed. If the data indicate that a non-zero coefficient is appropriate, then the coefficient on the long lag is non zero. A normal prior distribution is assumed with a mean of zero with a small standard deviation. The mean on a variable's first own lag is one with a larger standard deviation. The coefficients are estimated using Theil's mixed estimation approach (Doan, 2007).

This standard prior has three characteristics. First, the prior on deterministic variables such as seasonal dummy variables is flat or non-informative. Second, the prior distribution is independent normal. Third, the mean of the distribution is zero except for the first lag of the dependent variable of the equation which is equal to one. With these characteristics, we need only specify the standard deviation of the prior distribution, which is determined by the standard deviation function provided in equation two (Doan, 2007).

$$S(i,j,l) = [\{\gamma\, g(l)\, f(i,j)\}\, s_i] / s_j\,; \qquad f(i,i) = g(l) = 1.0 \qquad (2)$$

Here s_i is the standard error of a regression of variable i on lags of itself. The formula is scaled by the ratio of the standard errors to adjust for the different magnitudes of the variables in the model. $\{\gamma\, g(l)\, f(i,j)\}$ is the tightness of the prior distribution on coefficient i in equation j for lag l. γ = overall tightness of the standard deviation of the prior distribution. A smaller value for γ results in a smaller S(i,j,l) and a tighter standard deviation for the prior. The default value is .2. $g(l) = d^{l-1}$ is a geometric lag structure. This controls how the standard deviation changes as the number of lags increase. The parameter d can take on any value but the default value is set to 1.

f(i,j) = 1 if i = j and w otherwise. The parameter w is the weight placed on variable j relative to variable i in equation i. A smaller value for w reduces the standard deviation of the distribution prior. The default value is set to 1.

TAX REVENUE DATA

The tax revenue data used in this paper come from the U.S. Census.[6] The data is quarterly, starting with the first quarter of 1979 and ending with the first quarter of 2009. The categories of data examined are total tax revenue and the three largest components, sales tax, income tax, and corporate tax revenue.

---

[6] The data was downloaded from the U.S. Census *Quarterly Summary of State and Local Government Tax Revenue* at http://census.gov/govs/www/qtax.html.

Figure 1 plots of total tax revenue, sales tax revenue, corporate tax revenue, and income tax revenue over the sample period. Three observations can be made from Figure 1. First, there is a strong seasonal pattern in the data. Income tax revenue spikes in the second quarter of each year reflecting the income tax filing deadline. Second, the variability of the data increases with the level of tax revenue. Third, there are some missing observations in the early 1990s. Figure 2 plots the logarithm of the four tax revenue series. While the seasonality remains in the transformed data, the variability appears to be less linked to the level of the tax revenue. For this reason, I will use the logarithm of the tax revenue data and include seasonal dummy variables in the models for forecasting.

The remaining data used in this paper can be organized into national and state variables. California's economy and tax revenues are influenced by the performance of the overall U.S. economy. We measure overall economic activity using real GDP, real personal income (U.S. and California), and real personal income in the Far West census region. I also use an aggregate coincident business cycle index constructed by the Federal Reserve Bank of Philadelphia (Crone and Clayton-Matthews, 2005).

Other U.S. variables include the consumer price index, real price of West Texas Intermediate crude oil, real defense expenditures, the interest rate spread, and a tech sector index. The consumer price index captures movements in the overall cost of living. The crude oil price represents an important possible business cycle shock. Defense contractors have historically been an important sector of the California economy. Changes in defense expenditures may help predict the performance of the California economy and tax revenues. The spread between the ten-year Treasury bond rate and the three-month Treasury bill rate has been shown to be a good predictor of economic activity (Estrella and Trubin, 2006). Inclusion of this variable may help predict the economy and tax revenues. Finally, a national index of economic activity in the tech sector constructed by the Federal Reserve Bank of San Francisco is used to capture the performance of the tech sector in California (Hobijn, 2009). California-specific variables include the real price of housing, real personal income, a state coincident index, and state unemployment rate. All four of these variables should influence the California economy and tax revenues.[7]

MODEL SPECIFICATION TESTS

Specification tests were used to examine the trend properties of the data. Augmented Dickey-Fuller tests were used to test the null hypothesis that each data series contains a unit root and is non-stationary. These results are reported in Table 1.[8]

For all of the tax revenue data, the tests fail to reject the null hypothesis of a unit root implying they follow a random walk. This result allows me to use the random walk forecast as a benchmark forecast to be compared to the other models estimated in the paper. The tests also fail to reject the null hypothesis of a unit root for real defense spending, the tech index, both coincident indices, real GDP, all three real personal incomes, the interest-rate spread, and the California unemployment rate. The null

[7] The data on the coincident indices was downloaded from the Federal Reserve Bank of Philadelphia at http://www.phil.frb.org/index.cfm. The tech-sector index was downloaded from the Federal Reserve Bank of San Francisco at http://www.frbsf.org/csip/. All other data was downloaded from the Federal Reserve Bank of St. Louis at http://www.stlouisfed.org/fred2/.

[8] The lag length for the Augmented Dickey-Fuller test was determined using the Akaike information criterion (Stock, 1994).

hypothesis of a unit root is rejected for California real housing prices (5 percent level), real crude oil prices (5 percent level), and the consumer price index (1 percent level).[9]

Bi-variate Granger causality tests are conducted to help determine which national and state variables help predict tax revenues. If a variable Granger causes state tax revenue, it is included in the VAR and BVAR forecasting models. These tests are conducted for each tax revenue category using both a deterministic time trend and first differences. The tests are run with a lag length of three and five quarters. The results from these tests are reported in Table 2.

The null hypothesis for these tests is that the coefficients on the lagged explanatory variables are jointly zero. In other words, the lagged explanatory variables do not help predict tax revenues. Table 2 reports the p-values from this test.

Real GDP and all three measures of real personal income have p-values below 5% in most cases except for corporate tax revenues. I choose to use U.S. and California real personal income in the models for a more consistent comparison. Real defense, real California housing prices, real crude oil prices, and the consumer price index are consistently significant at the 10% level or less. This does not always hold in the case of corporate tax revenue. Real defense and the consumer price index have greater predictive ability when specified in levels. The technology index, U.S. and California coincident indices, California unemployment rate, and the interest rate spread appear to have little ability to predict state tax revenues.

Based on these tests, six variables are included in the VARs and BVARs. They include real defense spending, real crude oil prices, real U.S. personal income, the

---

[9] Because the VAR and BVAR models will contain variables with unit roots and deterministic time trends, I do test for cointegration (Doan, 2007).

consumer price index, real California housing prices, and real California personal income. To compare the impact of trends on forecast accuracy, each model is estimated in levels with a deterministic time trend and in first differences for the variables that contain unit roots.

FORECAST RESULTS

Each model's forecasting performance is evaluated by calculating the root mean squared error for the out-of-sample forecasts between 2004 and 2009.[10] Each VAR and BVAR is estimated in levels with a deterministic time trend and using the trend for each variable based on individual unit root tests. This way I can compare the forecast performance of models where the de-trending method is based on Dickey-Fuller tests and when it is not. Seasonal dummy variables are included to help predict the seasonal patterns in the tax revenue data. With quarterly data, I include five lags for each variable.

The parameter values used in specifying the BVARs follow the literature. The weight placed on non-revenue variables w and the decay parameter d are both set at the default value one. I experiment with two values for overall tightness γ with a loose value of .2 (denoted BVARa) and a tighter value of .1 (denoted BVARb).

The forecasting performance of the different models is first evaluated using the root mean squared error based on a series of one-quarter and four-quarters-ahead rolling forecasts for the period starting from the first quarter of 2004 to the first quarter of 2009. The forecast period begins with a growing economy that moves into a deep recession beginning in December 2007 through the end of the sample period. The one-quarter ahead forecast results are reported in Table 3. The BVAR in levels with tightness

---

[10] The root mean squared error equals the square root of the sum of forecast errors divided by the sample size T. The forecast error equals actual tax revenue in a quarter minus the forecasted tax revenue.

parameter value of .2 (BVARa levels) has the smallest root mean squared errors of all the multivariate models examined. BVARb levels with the tightness parameter set at a tighter .1 has the second smallest root mean squared errors. In 26 of the 28 cases, VARs and BVARs beat a random walk forecast.

The results of the four-quarters-ahead forecasts are reported in Table 4. As expected, the root mean squared errors are larger for the longer multivariate forecasts than for the shorter forecasts. The BVARs have the smallest root mean squared errors in each case except income tax revenue. However, a different specific model generates the lowest root mean squared error for each revenue category. The lowest root mean squared error for total and sales tax revenues is the BVAR in levels with a tightness parameter value of .1 (BVARb levels). The preferred models for income and corporate tax revenues are a standard VAR in first differences (VAR 1st diff) and a BVAR in levels with a tightness parameter of .2 (BVARa levels) respectively. The only multivariate case where a BVAR did not produce the lowest root mean squared error was for income tax revenue forecast. Interestingly, the random walk forecast does better at the longer horizon with smaller root mean squared errors. The random walk forecast beats the VAR and BVAR forecasts 13 of the 28 comparisons.

CONCLUSIONS

State governments can use a variety of methods to forecast tax revenues including structural econometric models, time-series models, and judgemental forecasts based on experience. This paper examined a number of alternative time series tax revenue forecasting models for California. Using a rolling regression approach for the period

2004 to 2009, a series of one- and four-quarters-ahead forecasts were made. Root mean squared errors were computed.

Based on root mean squared error, the BVAR models generally outperformed the VAR and random walk models. Tax revenue forecasters should consider Bayesian vector autoregressions as a cost effective method to forecast state tax revenues.

Figure 1

Tax Revenues

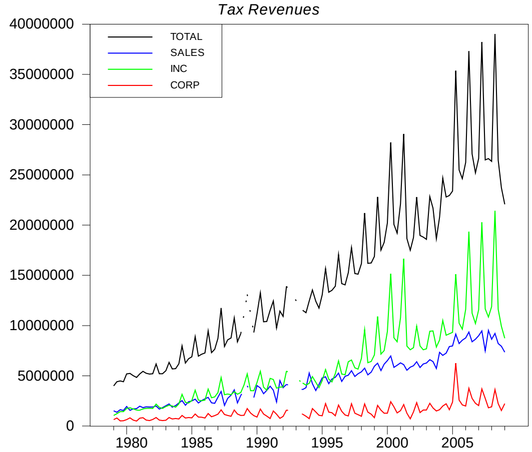

Figure 2

Log Tax Revenues

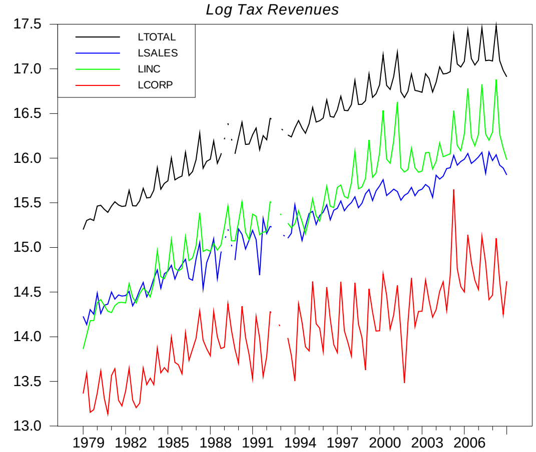

Table 1
Unit Roots Test Results

Variable	Dickey-Fuller Test Statistic	Number of Lags
Total Tax Revenue	-2.33	5
Sales Tax Revenue	-2.31	5
Income Tax Revenue	-2.35	4
Corporate Tax Revenue	-2.32	5
Real Defense Spending	-2.32	6
Real CA Housing Prices	-3.77[a]	3
Tech Pulse Index	-1.76	6
CA Coincident Index	-2.91	2
U.S. Coincident Index	-2.32	3
Real Price WTI	-3.88[a]	2
CPI	-4.11[b]	1
Interest Rate Spread	-1.65	6

Superscripts a and b indicate rejecting the null hypothesis at the five and one percent levels respectively.

Table 2
Bi-variate Causality Tests P-Values

Variable	Total		Sales		Income		Corporate	
Lags	3	5	3	5	3	5	3	5
Defense								
Level	.001	.040	.001	.022	.058	.294	.000	.124
1st Dif	.137	.173	.262	.173	.342	.149	.891	.547
CAHP								
Level	.105	.013	.163	.016	.081	.058	.146	.029
1st Dif	.005	.037	.005	.027	.033	.180	.177	.082
Tech								
Level	.038	.288	.071	.359	.012	.150	.714	.544
1st Dif	.382	.626	.473	.775	.126	.149	.439	.516
CACI								
Level	.028	.752	.142	.558	.091	.480	.288	.820
1st Dif	.842	.979	.593	.751	.734	.770	.934	.894
USCI								
Level	.026	.826	.131	.670	.145	.890	.650	.999
1st Dif	.933	.985	.862	.942	.967	.969	.879	.995
WTI								
Level	.076	.008	.020	.097	.340	.010	.260	.119
1st Dif	.056	.007	.293	.084	.248	.019	.378	.151
CPI								
Level	.001	.111	.052	.711	.046	.246	.100	.090
1st Dif	.123	.292	.679	.663	.039	.547	.714	.445
Spread								
Level	.147	.393	.006	.043	.730	.441	.836	.950
1st Dif	.786	.712	.147	.159	.739	.709	.956	.640
CAU								
Level	.034	.794	.189	.603	.101	.628	.240	.793
1st Dif	.522	.936	.561	.752	.440	.668	.899	.618
GDP								
Level	.004	.005	.009	.003	.001	.002	.440	.464
1st Dif	.001	.026	.005	.012	.015	.013	.192	.631
USPI								
Level	.000	.009	.000	.022	.000	.001	.039	.565
1st Dif	.000	.111	.006	.041	.000	.049	.507	.483
CAPI								
Level	.000	.001	.000	.012	.000	.000	.018	.443
1st Dif	.000	.027	.001	.017	.000	.007	.453	.381
FWPI								
Level	.000	.005	.000	.105	.016	.029	.070	.005
1st Dif	.000	.018	.008	.296	.001	.072	.059	.011

Where Defense equals real defense spending, CAUP equals real housing prices in California, Tech equals the tech pulse index, CACI equals the California coincident index, USCI equals the U.S. coincident index, WTI equals the real price of West Texas Intermediate crude oil, CPI equals the U.S. consumer price index, Spread equals the interest rate spread, CAU equals the unemployment rate in California, GDP equals real U.S. gdp, USPI equals real U.S. personal income, CAPI equals California real personal income, and FWPI equals far west census region personal income.

Table 3
Root Mean Squared Errors for One-Quarter-Ahead Forecasts
(2004:1 to 2009:1)

Model	Total Tax Revenues	Sales Tax Revenue	Income Tax Revenue	Corporate Tax Revenue
Random Walk	.23432	.11356	.34348	.44405
VAR Levels	.10966	.13533	.16772	.26007
BVARa Levels	.08347	.07282	.15243	.22367
BVARb Levels	.08796	.07495	.16343	.23424
VAR 1st Dif	.13858	.15445	.17883	.32600
BVARa 1st Dif	.09365	.08680	.16633	.26907
BVARb 1st Dif	.10572	.09386	.20582	.30579

Table 4
Root Mean Squared Errors for Four-Quarters-Ahead Forecasts
(2004:1 to 2009:1)

Model	Total Tax Revenues	Sales Tax Revenue	Income Tax Revenue	Corporate Tax Revenue
Random Walk	.11850	.12845	.14519	.31315
VAR Levels	.15912	.14658	.19295	.30526
BVARa Levels	.11141	.09185	.17322	.24435
BVARb Levels	.09903	.08774	.16222	.25249
VAR 1st Dif	.16646	.67777	.15459	.37368
BVARa 1st Dif	.09295	.09064	.16578	.26078
BVARb 1st Dif	.21019	.08819	.16623	.25921

References

Baghestani H McNown R. 1992. Forecasting the Federal Budget with Time-series Models. *Journal of Forecasting* 11: 127-139.

Banbura, M, Giannone D, and Reichlin, L. 2010. Large Bayesian Vector Auto Regressions. *Journal of Applied Econometrics* 25: 71-92.

Banerji, A, Dua, P, and Miller, SM. 2006. Performance Evaluation of the New Connecticut Leading Employment Index Using Lead Profiles and BVAR Models. *Journal of Forecasting* 25: 415-437.

Crone T M, Clayton-Matthews A. 2005. Consistent Economic Indexes for the 50 States. *The Review of Economics and Statistics* 87: 593-603.

Diebold FX. 2007. *Elements of Forecasting*. Thomson: Mason, OH.

Doan TA, 2007. *Rats User's Guide Version 7*. Estima: Evanston, IL.

Enders W, 1995. *Applied Econometric Time Series*. John Wiley & Sons: New York, NY.

Estrella A., Trubin, MR. 2006. The Yield Curve as a Leading Indicator: Some Practical Issues. Federal Reserve Bank of New York *Current Issues in Economics and Finance* 12: 1-7.

Hobijn B. 2009. The Tech Pulse Index: Recent Trends in Tech-Sector Activity. Federal Reserve Bank of San Francisco *Economic Letter* 3: 1-3.

Litterman RB. 1986. Forecasting with Bayesian Vector Autoregressions: Five Years of Experience *Journal of Business and Economic* Statistics 4: 25-38.

Litterman RB., Supel TA. 1983. Using Vector Autoregressions to Measure the Uncertainty in Minnesota's Revenue Forecasts. Federal Reserve Bank of Minneapolis *Quarterly Review* 7.

Meese, R. A. and Rogoff, K. 1983. Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample? *Journal of International Economics* 14: 2-24.

Newey WK, West KD. 1987. A Simple Positive Semi-Definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica* 55:703-708.

Otrok C., Whiteman CH. 1997. What to Do When the Crystal Ball is Cloudy: Conditional and Unconditional Forecasting in Iowa. Paper presented at the National Tax Association 90^{th} Annual Meeting on Taxation.

Rich R. Bram J. Haughwout A. Orr J. Rosen R. and Sela R. 2005. Using Regional Economic Indexes to Forecast Tax Bases: Evidence From New York *Review of Economics and Statistics* 87: 627-634.

Rogoff, K. and Stavrakeva, V. 2008. “The Continuing Puzzle of Short Horizon Exchange Rate Forecasting. National Bureau of Economic Research working Paper # 14071.

Shoesmith GL. 1995. Multiple Cointegrating Vectors, Error Correction, and Forecasting with Litterman’s Model *International Journal of Forecasting* 11: 557-567.

Stock JH. 1994. Unit Roots, Structural Breaks, and Trends. In *Handbook of Econometrics*. Robert Engle and Daniel McFadden (eds.) Elsevier: Amsterdam.

Stock, JH. And Watson , MW. 2003 Forecasting Output and Inflation: the Role of Asset Prices. *Journal of Economic Literature* 41: 788-829.

Williams B, Ingenito R, and Vasche JD. 1999. Forecasting Revenue Receipts in the States: Current Challenges in California. *National Tax Journal* LII: 361-371.