REVENUE FORECASTING

STATE TAX REVENUE FORECASTING ACCURACY

Technical Report

September 2014

Donald J. Boyd and Lucy Dadayan

Rockefeller Institute of Government

State University of New York

Support for this project was provided by

The Pew Charitable Trusts

THE NELSON A. ROCKEFELLER INSTITUTE OF GOVERNMENT

SUNY The State University of New York

The Public Policy Research Arm of the State University of New York

411 State Street
Albany, NY 12203-1003
(518) 443-5522

www.rockinst.org

Acknowledgements

We thank revenue forecasters and analysts in the states for responding to the survey on revenue forecasting conducted in the course of this project.

We graciously thank The Pew Charitable Trusts for providing funding for the research and release of this report.

The views expressed herein are those of the authors and do not necessarily reflect the views of The Pew Charitable Trusts.

A Note on the Term "Forecast Error"

Throughout this report we refer to the difference between a forecast and actual results as a "forecast error." This term is common in analyses of forecasts, whether they be forecasts of the economy, or of the weather, or of state tax revenue. It does not imply that the forecaster made an avoidable mistake or that the forecaster was somehow unprofessional. All forecasts of economic activity will be wrong to some extent, regardless of the expertise of the forecaster or the quality of the tools used. Forecasting errors are inevitable because the task is so difficult: Revenue forecasts are based in part upon economic forecasts, and economic forecasts prepared by professional forecasting firms often are subject to substantial error; tax revenue is volatile and dependent upon idiosyncratic behavior of individual taxpayers, which is notoriously difficult to predict; and tax revenue is subject to legislative and administrative changes that also are difficult to predict.

Contents

State Tax Revenue Forecasting Accuracy

Technical Report

September 2014

Executive Summary

This report updates data on state revenue forecasting errors that was initially presented in "Cracks in the Crystal Ball," a 2011 report on revenue forecasting in the states produced in collaboration with The Pew Charitable Trusts. It supplements these data with additional data, including the results of a survey of state forecasting officials conducted by the Rockefeller Institute. In addition to examining how revenue forecasting errors have changed since 2009, which was the last data point in the prior project, we examine the relationship between revenue forecasting accuracy and:

- Tax revenue volatility;
- Timing and frequency of forecasts; and
- Forecasting institutions and processes.

Our main conclusions and recommendations follow.

Description and Summary of Data Used in This Study

Our analysis for this report is based on four main sources. First, as with the 2011 report, we computerized data on revenue estimates and revenue collections for the personal income, sales, and corporate income taxes from the Fall *Fiscal Survey of the States* from the National Association of State Budget Officers (NASBO) for each year from 1987 through 2013, covering a total of twenty-seven years. We define the forecasting error as actual minus the forecast, and thus include both overforecasting and underforecasting. In other words, a positive number is an underestimate of revenue (actual revenue is greater than the forecast), and a negative number is an overestimate.

Second, we developed a new measure of forecast difficulty that we used both in descriptive analyses of the data, and in statistical analyses to allow more accurate estimates of the influence of other factors on forecast error, after controlling for the difficulty of a forecast. The measure of forecast difficulty is the error that results from a uniformly applied simple or "naïve" forecasting model, which we then compare to forecasting errors that result from states' idiosyncratic and more elaborate revenue forecasting processes.

Third, as before, we included other secondary data in our analyses, including measures of year-over-year change in state tax revenue from the Census Bureau, and measures of the national and state economies from the Bureau of Economic Analysis. In addition, several other researchers provided us with data that they had developed on the institutional and political environment in which revenue forecasts are made, and we incorporated some of these data into our analyses.

Fourth, we surveyed state government officials involved in state revenue forecasting. We did this to gain a more detailed understanding of the NASBO data on forecasts and results, and also to learn more about the institutional forecasting arrangements in each state, and how forecasts are used in the budget process. Among other things, this survey allowed us to develop a measure of the typical lag in each state between the time a revenue forecast is prepared and the start of the fiscal year, which we used in our analysis of the impact of the timing and frequency of forecasts on forecast accuracy.

Our main conclusions from our initial descriptive analyses of these data sources are:

- Corporate income tax forecasting errors are much larger than errors for other taxes, followed by the personal income tax and then the sales tax. The median absolute percentage error was 11.8 percent for the corporate income tax, 4.4 percent for the personal income tax, and 2.3 percent for the sales tax.
- Smaller states and states dependent on a few sectors of the economy (particularly states reliant on oil or natural gas, or gambling) — such as Alaska, Montana, Nevada, New Hampshire, and North Dakota — tend to have larger errors. Those states' errors also tend to be more variable. (Among these states, the taxes reported to NASBO by Alaska, Delaware, Montana, New Hampshire, Vermont and Wyoming are a relatively small share of total taxes as measured by the Census Bureau. Their errors for total taxes in their budget, which also include other less volatile taxes, appear to be smaller than those examined here. However, these states also have large errors from naïve forecasting models using full Census data.)
- When taxes are particularly difficult to forecast, states tend to be more likely to underforecast revenue, suggesting that they may try to do so in an effort to avoid large shortfalls. Thus, there is a pattern to the apparent bias in state revenue forecasts. By contrast, our naïve forecasting model does not become more likely to underforecast when forecasting becomes more difficult, suggesting that this phenomenon may reflect the behavior of forecasters rather than underlying factors in the economy or tax revenue structures.
- Errors become particularly large in and after recessions.
- Variation in errors across states also increases in and after recessions.
- Errors near the 2001 and 2007 recessions were much worse than in the recession of the early 1990s.
- States have returned to "more normal" forecasting errors with the 2007 recession now behind us.

Revenue Forecasting Accuracy and Revenue Volatility

Increases in forecasting errors have been driven by increases in revenue volatility, which in turn have been driven in large part by volatile capital gains, which have grown as a share of adjusted gross income over the last several decades.

States have relatively little opportunity to reduce volatility simply by restructuring their tax portfolios. Only by virtually eliminating the corporate income tax and significantly increasing reliance on the sales tax relative to the personal income tax could the typical state reduce revenue forecasting errors, and even then most tax combinations would not reduce forecast errors very much. Changing the structure of individual taxes may be more promising, but it can raise difficult tax policy issues. For example, making income taxes less progressive might also make them less volatile and easier to forecast, but that may run counter to distributional objectives that some policymakers hope for.

Because it is so hard to reduce forecasting errors by changing tax structure, it is especially important for states to be able to manage the effects of volatility. Rainy day funds are one important tool states can use. However, the data suggest that states with larger forecast errors and more difficult forecasting tasks do not have larger rainy day funds; perhaps they should.

Revenue Forecasting Accuracy and the Timing and Frequency of Forecasts

Further-ahead forecasts are more error prone, whether we examine the forecasts for the second year of a biennium, which generally are prepared further in advance, or use a measure of the lag between forecast preparation and the start of the forecast period that we developed based on the information collected from our survey. In both cases the effect is meaningful: the data suggest that accuracy worsens by well over a percentage point in the forecast for the second year of a biennium, and worsens by about a half a percentage point for every ten weeks of lag between the time a forecast is prepared and the start of the forecast period.

There is no evidence from our data that frequent forecast updates lead to greater accuracy. This is consistent with past research.

The policy implications of the first part of our analysis are clear: States should minimize unnecessary lags between forecast preparation and the start of the fiscal year, and should update those forecasts as close as possible to the start of the fiscal year, as many states do. Even though there is no evidence that more frequent forecast updates during the forecast period will lead to more accurate forecasting, it is good management practice to update forecasts regularly as the year progresses so that finance officials are managing the budget using the best available information.

Revenue Forecasting Accuracy and Institutional Arrangements

Our reading of the literature is that there is very little relationship between consensus forecasting and forecast accuracy. That is consistent with our examination of these data, also. However, as we have noted before, the evidence in favor of examining and combining forecasts is overwhelming: Combining forecasts tends to lead to lower errors. Processes that encourage this to happen may also lead to lower errors.

Beyond that, it is good practice to try to insulate forecasting from the political process and consensus forecasting can help to achieve that.

STATE TAX REVENUE FORECASTING ACCURACY

Technical Report

THE NELSON A. ROCKEFELLER INSTITUTE OF GOVERNMENT

State University of New York
411 State Street
Albany, New York 12203
(518) 443-5522
www.rockinst.org

Carl Hayden
Chair, Board of Overseers

Thomas Gais
Director

Robert Bullock
Deputy Director for Operations

Patricia Strach
Deputy Director for Research

Michael Cooper
Director of Publications

Michele Charbonneau
Staff Assistant for Publications

SUNY The State University of New York

Nancy L. Zimpher
Chancellor

Introduction

This report is based upon work conducted in 2013 and 2014 and updates and extends analysis by the Rockefeller Institute of Government for "Cracks in the Crystal Ball," a 2011 report on revenue forecasting in the states produced in collaboration with The Pew Charitable Trusts.[1] Among other things, the 2011 report concluded:

> … errors in the annual revenue estimates have worsened. Revenues have become more difficult to predict accurately…. [R]evenue overestimates during the nation's past three recessions grew progressively larger as did the underestimates in the past two periods of economic growth.

This analysis updates data on state revenue forecasting errors developed during that project and supplements it with data from a survey of state forecasting officials conducted by the Rockefeller Institute. In addition, this report examines the relationship between revenue forecasting accuracy and:

- Tax revenue volatility
- Timing and frequency of forecasts
- Forecasting institutions and processes

Finally, we offer policy recommendations for how to improve revenue forecasting and to manage outcomes of the process.

Description and Summary of Data Used in This Study

Description of Data

Data on Forecasting Errors From NASBO

Each fall the National Association of State Budget Officers (NASBO) ask state budget officers for revenue estimates used at the time of budget enactment and for preliminary actual revenue collections.[2] They publish these results in the fall *Fiscal Survey of States*. The fall *Survey* typically is conducted in the months of July through September and published in December.

Below is an example of the question as it was asked in August 2013 for fiscal year 2013 preliminary actuals and fiscal year 2014 estimates.

> Please complete the following table with respect to actual revenue collections for FY 2012, estimates used at FY 2013 budget enactment, preliminary actual revenue collections for FY 2013 & estimates used at budget enactment for FY 2014 ($ in millions).

Taxes	Actual Collections FY 2012	Estimates Used When Budget was Adopted FY 2013	Preliminary Actuals FY 2013	Estimates Used When Budget Adopted FY 2014
Sales Tax Collections				
Personal Income Tax Collections				
Corporate Income Tax Collections				

We computerized data on revenue estimates and revenue collections for the personal income, sales, and corporate income taxes from the NASBO Fall *Fiscal Survey of States* for each year from 1987 through 2013, covering a total of twenty-seven years. In the survey, the "original estimates" are intended to be the forecasts on which the adopted budget was based, and the "current estimates" are the preliminary actual estimates for the fall after the year was closed (e.g., the estimate in fall 2013 for the fiscal year that ended in June 2013). In 1987 and 1988 the survey covered personal income and sales taxes only: from 1989 forward it included personal income, sales, and corporate income taxes. Because NASBO data do not include the District of Columbia, we did not incorporate the District in this analysis.

The NASBO data held several great advantages for the purposes of our analysis: They are self-reported by states and thus reflect states' own assessment of appropriate data; they are collected by a single source; they are intended to be collected under a common set of definitions; they are collected for all fifty states in most years; and they go back more than twenty-five years, covering all or part of three recessions (in 1990, 2001, and 2007). It would be impractical to assemble such a data set from scratch, collecting historical forecast and actual revenue data from individual states over nearly three decades; records disappear, memories fade, and staff move on, making it difficult to reconstruct these data.

As with any self-reported numbers, there were some anomalies in the NASBO data, which we were diligent in cleaning. We eliminated data in the following situations: cases in which only the original estimate was reported but not the current, and vice versa; cases in which the original and current estimates were identical for two or more taxes, suggesting a possible reporting error; and cases in which we believed the estimating errors were too large to be plausible (the top 1 percent of cases with the highest absolute value of forecast error). California was missing from the NASBO data in 2001 and 2009. Because it was such a large state, we contacted state officials in California directly and supplemented the NASBO data with estimates of what would have been submitted to NASBO in those years. Data were also missing for Texas in 1996, 1997, and 1999. However, as noted below, it was not practical to obtain comparable data for missing values for Texas.

After these adjustments, we have 3,347 observations of forecast errors for individual taxes. In addition, where data allowed, we computed forecast errors for the sum of the three taxes, for each state and year, for an additional 1,311 values.[3]

Survey of State Officials

We supplemented forecast-error data from NASBO with a survey we conducted of state government officials.[4] Our intent was to strengthen our ability to use the NASBO data in regression analyses, and to gain a richer understanding of state forecasting processes more generally. The survey had two main purposes:

1. To gain a clearer understanding of what, exactly, the forecast data provided by states to NASBO represent; to gather information on when those forecasts are prepared; and to understand how they are used in the budget process in the states.

2. To gain a richer understanding of the variety of forecasting arrangements in the states and how these processes relate to the budgeting process.

We conducted the survey in two rounds. During the first round, we asked questions of state forecasting officials related to the data reported to NASBO. The survey questions were designed to help us get a better understanding of when estimates that underlie state budgets are developed, as that influences the difficulty of the forecasting job, and to be sure we understand how the forecasts are used in the budget process.

The main purpose of the second round was to get a better understanding of general procedures and practices of revenue estimating processes in the states. The survey questions were targeted at getting more information about the timing of forecasts, frequency of forecast updates, the long-term forecasting practices, parties involved in the forecasting processes, and whether there is a formal group that plays a major role in the revenue estimating process.

Both surveys were conducted online, using Google Docs. The first round was conducted during July-August of 2013, while the second round was conducted during August-September of 2013.

For the first round of the survey, we sent an online survey questionnaire to state budget officers in all fifty states, who were also the respondents for the NASBO's survey. The list of respondents was retrieved from the *Fiscal Survey of the States* report. The response rate for the first round of the survey was 68 percent (34 usable responses were received).

For the second round of the survey, we sent an online survey questionnaire to both legislative and executive branch officials in all fifty states who were likely to be involved in the forecasting process. The response rate for this round was much higher: 90 percent (forty-five responses) for the executive branch and 92 percent (forty-six responses) for the legislative branch. In aggregate (executive and legislative branches combined), we have received responses from all fifty states.

The responses from the first round indicate that in most states the estimates provided to NASBO are the final estimates used for the adopted budget. States vary widely in how far in advance of the budget they prepare their estimates. Some states prepare revenue estimates as early as thirty-five weeks ahead of the start of the fiscal year (e.g., Alabama), while other states use revenue estimates updated about two weeks before the start of the state fiscal year (i.e., Delaware, Pennsylvania, and Tennessee). The survey results also revealed that, in general, states with a biennial budget cycle provide updated estimates to NASBO for the second year of the biennium in order to account for legislative changes and updated economic forecast conditions. Nonetheless, on average, estimates for the second year of a biennium are prepared further in advance of the second year than are estimates for the first year.

The responses from the second round indicate wide variation among the states in the timing of the forecasts prepared for the adopted budget as well as in frequency of revenue estimate updates. In terms of long-term forecasting, states indicated that they typically generate revenue estimates for one or more years beyond the upcoming budget period; most states release revenue estimates for two to five years beyond the budget period.

The state officials were also asked to indicate who is involved in the revenue estimating process. The survey results indicate that in most states one or more agencies of the executive branch are involved in the revenue estimating process and in a handful of states the legislative branch also plays a major role in revenue estimating. However, only very few states also involve academicians and/or experts from the private sector in the revenue estimating process. Finally, in a few states the revenue estimating process is under the jurisdiction of the nonpartisan independent group/agency, the members of which are normally appointed by the governor and/or legislative branch. According to the survey responses, in approximately two-thirds of the states the participants in the revenue estimating process come to an agreement on a single number.

In general, the survey results indicate that there is a wide range of practices and processes involved among the fifty states in terms of revenue estimates. There are no universally agreed upon standards and practices among the fifty states. The survey responses, as well as various discussions with the state government officials, indicate that it is highly desirable to have more frequent forecasts as well as to have a minimum time gap between the forecast and the adoption of the budget so that the states can take into account the latest economic conditions as well as the impact of any legislative and/or administrative changes. Overall, it is also desirable to have a consensus revenue estimating practice in the state, so that various parties involved in the forecasting process can come into an agreement.

Other Data

In addition to the NASBO data, we used U.S. Census state tax collection figures in estimating the size of errors across years or across the type of tax. We also used data on national price changes (the gross domestic product price index), gross domestic product by state, and state personal income from the U.S. Bureau of Economic Analysis. We also obtained data from several other researchers, and describe as needed below.[5] Finally, we constructed a measure of forecast difficulty from U.S. Census state tax collection data, which we describe in the section, *Constructing a Forecast Difficulty Measure*.

How do the revenue estimating data relate to Census tax revenue data? Not all states rely heavily on the three taxes that NASBO collects forecast data for. Six states relied on these three taxes for less than half of their tax revenue in the typical year: Alaska, Delaware, Montana, New Hampshire, Vermont, and Wyoming (see Table 1).

Measures of Forecasting Error

There are several common ways of measuring forecast error. Some of these measures do not distinguish positive from negative errors and are useful for gauging magnitude of error, and others maintain the sign and are useful for gauging bias and asymmetries. One common measure is the root-mean-squared error (RMSE), or more appropriate when comparing across states, the root-mean-squared percent error (RMSPE).

Table 1. Personal Income, Sales, and Corporate Income Taxes as Percent of Total Tax Revenue

Median, 1987-2013			
US average: 72%			
State	Percent	State	Percent
AK	20.5	OR	73.4
NH	27.3	RI	73.4
WY	32.1	IA	73.4
MT	41.2	ME	74.2
DE	46.6	NJ	75.1
VT	46.7	NC	75.1
TX	50.4	MN	75.3
NV	51.0	MI	75.8
ND	51.9	AZ	75.8
SD	57.7	ID	75.8
WA	60.1	CT	75.9
LA	61.1	VA	76.0
AL	62.4	KS	77.5
OK	62.6	MO	77.7
FL	63.4	NE	77.9
WV	63.5	WI	77.9
NM	66.2	SC	78.1
PA	66.3	CO	78.9
KY	66.5	NY	80.2
TN	67.7	HI	81.2
IL	70.8	CA	81.3
MD	71.7	IN	81.4
OH	72.9	UT	81.7
AR	73.0	MA	83.4
MS	73.0	GA	85.0

Source: Rockefeller Institute analysis of data from Census Bureau.

One very common measure in the literature examining state revenue forecasting is the absolute value of the error as a percentage of the actual result (known as absolute percentage error), or variants on this (Voorhees 2004). Summary measures of estimating error across states or years often use the mean or median of the absolute value of the percentage error (mean absolute percentage error is quite common and is known as the MAPE). These measures, which treat positive and negative errors the same, may not be the most appropriate in all situations. In particular, other measures are needed to examine questions of whether forecasts are biased (e.g., more likely to underestimate revenue than to overestimate) or whether forecasters believe the costs of overestimates are greater than the costs of underestimates.

As many observers have noted, the revenue estimating error is quite large relative to what policymakers might find acceptable. In 2002, when the previous fiscal crisis hit sharply, twelve states had revenue estimating errors of 10 percent or more (Willoughby and Guo 2008). To provide some perspective, a 4 percentage point error in the state of California would be more than $4 billion — an amount that policymakers would find quite disconcerting.

We define the forecasting error as actual minus the forecast. Thus, a positive number is an underestimate of revenue (actual revenue is greater than the forecast), and a negative number is an overestimate. Forecasters often view underestimates as better than overestimates. Our primary measure of error is the error as a percentage of the actual revenue. Thus, if a forecast of revenue was $90 million and the actual revenue was $100 million, then the error was $10 million (an underestimate), and the percentage error was 10 percent.

We use two main versions of this percentage — as is, where it can be either positive or negative, and the absolute value, which is always positive. In general, the absolute value is useful when we are interested in the concept of accuracy, without regard to whether revenue is above target or below, although the as-is value can be useful, too, and we use that in much of the analysis below. When we want to examine bias (not a significant focus of this report), the absolute value measure is not useful, because we care about the direction of error, and we focus solely on the as-is value.

When we want to summarize error across states, years, taxes, or other groupings, we almost always use the median as a measure of central tendency. The reason is that our data are fairly noisy and we don't generally place great faith in any single data point, although we do believe the data tell truthful stories in aggregate. Because any single data point might be wrong or an outlier, it could have an undue impact on the mean, pulling it away from what we think is the true central tendency of the data.

When we want to measure how much the errors vary across states, years, taxes, or other groupings we use either the standard deviation or the interquartile range.

In the next section, we provide simple descriptive statistics of errors, first using the percentage error "as is" and then the absolute value. There are some useful insights, which we summarize in the concluding subsection of this section.

Forecasting Errors as Percentage of Actual Revenue

Below we provide summary tables of errors by tax, by fiscal year, and by state, and explore patterns with several plots.

Table 2 displays selected summary statistics for forecasting errors by tax for 1987-2013. The corporate income tax median forecasting error, at 2.8 percent, was much larger than the median for other taxes. The personal income tax median forecasting error was 1.8 percent, while the sales tax forecasting error was 0.3 percent. The standard deviation of the error — a measure of variability in forecast error and a potential indication of forecasting difficulty — was largest by far for the corporate income tax, nearly three times as large as for the personal income tax, and five times as large as for the sales tax.

Table 2. State Revenue Forecasting Percentage Errors by Tax, 1987-2013

Tax	*Forecast errors as percentage of actual revenue*					
	# of observations	Mean error	Standard deviation	25%	Median (50%)	75%
Personal income tax	1,100	1.1	7.7	(2.3)	1.8	5.5
Sales tax	1,189	0.2	4.6	(2.0)	0.3	2.6
Corporate income tax	1,058	1.3	22.6	(9.3)	2.8	14.0
Sum of PIT, sales & CIT	***1,311***	***1.0***	***7.6***	***(2.2)***	***1.5***	***4.4***

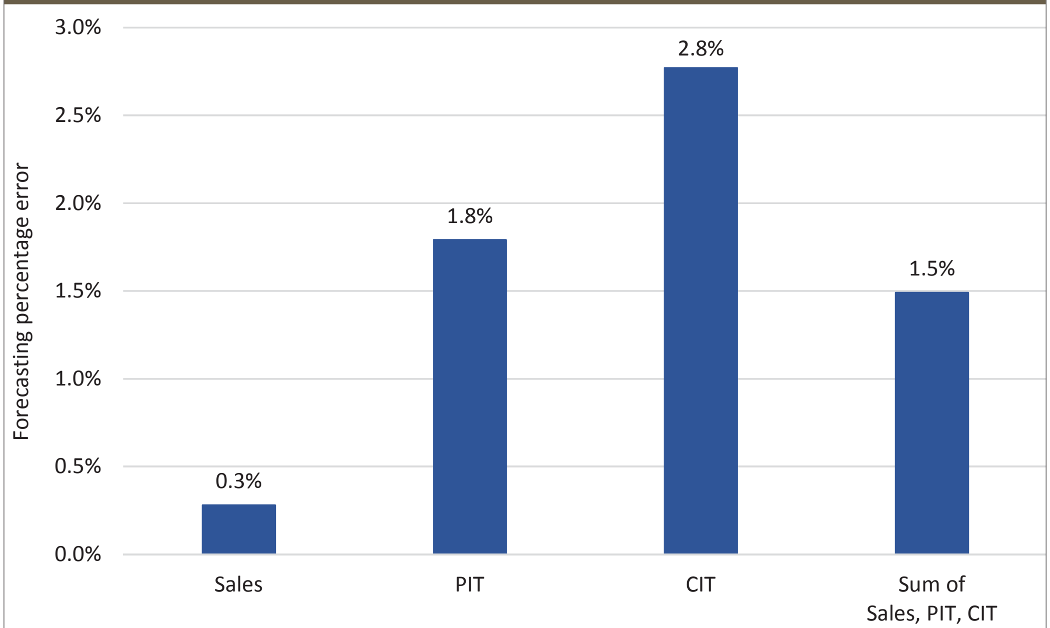
Source: Rockefeller Institute analysis of NASBO forecast error data.

Figure 1 on the following page displays the median forecast error by tax.

There is some evidence that forecasting errors are "serially correlated" — that is, that a revenue shortfall in one year is more likely to be followed by a shortfall in the next than by an overage, and vice versa. Rudolph Penner, a former director of the Congressional Budget Office, concluded this about federal government forecasts: "… if CBO makes an error in an overly optimistic or pessimistic direction one year, it is highly probable that it will make an error in the same direction the following year" (Penner 2008). Inspection of NASBO data in time periods surrounding recessions shows that, at least for the states in aggregate, revenue overestimates tended to be followed by additional overestimates.

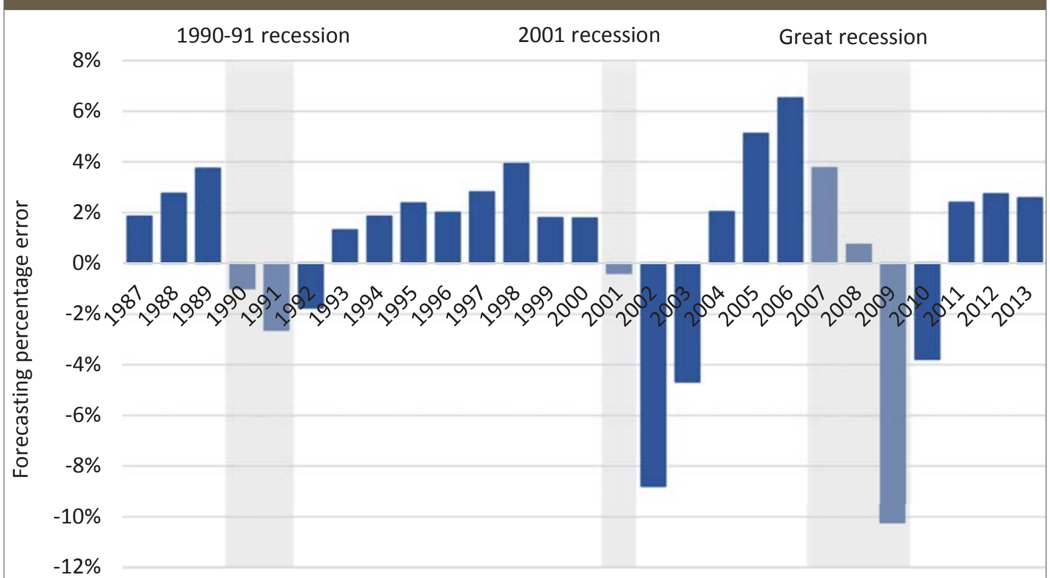
Figure 2 provides median forecasting errors by fiscal year for twenty-seven years, spanning from 1987 to 2013. As depicted on Figure 2, states normally underestimate during expansions and overestimate during the downturns. Moreover, overestimating

Figure 1. State Revenue Forecast Errors as Percent of Actual Tax, Median for 1987-2013



Source: Rockefeller Institute analysis of data from NASBO Fall Fiscal Surveys.

Figure 2. Median State Revenue Forecast Errors as Percent of Actual Tax, by Fiscal Year



Source: Rockefeller Institute analysis of data from NASBO Fall Fiscal Surveys.
Note: Forecasting error is for the sum of personal income, sales and corporate income taxes.

errors were particularly large near the 2001 and 2007 recessions. States now returned to their pattern of underestimation after the worst of the fiscal crisis was behind them.

Table 3 on the following page provides summary statistics of forecasting errors by state from 1987 to 2013. Overall, smaller states and states dependent on a few sectors of the economy (particularly states reliant on oil or natural gas, or gambling), such as Alaska, Montana, Nevada, New Hampshire, and North Dakota, tend to have larger errors. Those states' errors also tend to be more variable, as measured by the standard deviation or the interquartile range. As already mentioned above, personal income, sales, and corporate income taxes in Alaska, Delaware, Montana, New Hampshire, Vermont, and Wyoming represent less than half of total taxes as measured by the Census Bureau. Their errors for total taxes in their budget, which also include other less volatile taxes, appear to be smaller than those examined here. We strongly caution not to read too much into the errors for any single state because the data are noisy and influenced by many forces, and larger errors need not indicate less effective revenue forecasting.

Table 3. Forecast Errors as Percent of Actual Tax, by State

State	Mean	Standard deviation	25%	Median	75%
US	*0.1*	*4.3*	*(1.1)*	*1.9*	*2.5*
AK*	7.6	22.2	(9.0)	10.3	15.2
AL	0.5	6.1	(1.3)	1.5	3.8
AR	2.1	4.1	(0.2)	2.1	3.5
AZ	(0.4)	9.7	(1.6)	1.4	4.4
CA	(0.3)	7.4	(4.6)	2.5	4.3
CO	0.9	8.5	(0.2)	1.6	5.5
CT	0.5	7.1	(2.4)	2.8	5.3
DE*	1.0	6.8	(2.8)	(0.1)	5.6
FL	(1.0)	5.2	(2.3)	(0.4)	2.1
GA	(0.3)	5.0	(4.1)	1.5	3.3
HI	1.2	6.5	(2.7)	1.5	4.9
IA	1.9	3.7	(0.3)	1.8	4.8
ID	1.9	6.8	0.1	1.5	5.1
IL	0.7	5.3	(2.3)	1.9	2.8
IN	(0.9)	4.9	(2.5)	0.9	2.0
KS	1.8	4.5	(0.0)	1.1	2.0
KY	(1.1)	5.1	(3.6)	0.5	2.6
LA	0.6	6.6	(3.8)	1.7	4.1
MA	1.2	7.5	(2.2)	2.3	5.7
MD	0.4	5.6	(2.9)	1.6	4.0
ME	1.7	4.6	(0.1)	1.4	3.9
MI	(1.7)	5.2	(4.2)	(1.0)	2.3
MN	1.7	5.1	(0.6)	1.8	5.5
MO	(0.2)	5.4	(1.3)	1.1	3.6
MS	1.0	5.2	(2.4)	1.8	4.3
MT*	4.2	11.2	(1.0)	6.6	9.3
NC	0.5	5.9	(1.2)	2.5	4.2
ND	9.3	18.9	0.7	3.0	13.8
NE	1.0	5.2	(1.5)	1.7	4.4
NH*	(0.4)	16.1	(3.4)	2.1	7.0
NJ	(0.7)	5.6	(3.6)	1.0	2.5
NM	0.7	6.0	(2.5)	0.6	4.5
NV	2.2	7.0	(3.9)	3.9	7.1
NY	(2.5)	3.8	(5.2)	(2.1)	0.3
OH	(0.2)	5.3	(1.9)	1.3	2.6
OK	0.3	7.0	(1.5)	1.1	4.3
OR	0.3	10.0	(1.7)	2.1	4.9
PA	(0.2)	3.7	(1.9)	0.2	3.2
RI	0.3	5.7	(2.8)	(0.4)	4.3
SC	(0.6)	8.6	(5.1)	1.4	4.6
SD	1.1	2.5	(0.2)	0.8	2.4
TN	0.3	4.5	(1.4)	0.6	3.4
TX	3.3	6.4	0.0	2.6	6.7
UT	1.5	6.1	0.2	2.9	5.3
VA	(0.1)	6.7	(2.3)	1.5	2.6
VT*	2.9	6.3	(0.8)	4.2	7.5
WA	0.3	7.0	(1.0)	1.3	3.3
WI	0.2	3.3	(0.0)	0.9	1.8
WV	1.6	3.4	(1.1)	1.6	3.2
WY*	3.4	7.1	0.3	2.0	7.1

Source: Rockefeller Institute analysis of data from NASBO.
Note: States with * relied on PIT, CIT & sales taxes for less than half of their tax revenue in a typical year.

Figure 3 on the following page maps the distribution over time of state forecasting errors for the sum of the three taxes (personal income tax, sales tax, and corporate income tax), and also shows the extent of missing data. We are missing data for Texas in 1996, 1997, and 1999, but other than that no large state is missing in any year. It is difficult to obtain data from state documents that is comparable to data in the NASBO survey, and as noted earlier we were unable to obtain such data with reasonable effort for Texas.

The figure shows how forecasting errors have rolled out over time across the country. The revenue shortfalls in the middle of the country in 1987 likely are related to the oil price bust that hit Texas and other oil-extraction states in the mid-1980s. The effects of the 1990 recession, which lingered for four years, appear to have led to significant revenue shortfalls on the coasts and in the north but not as much in the center of the country. The 2001 recession and the effects of the September 11th attacks are evidenced in revenue shortfalls that began on the east coast but spread to the rest of the country; revenue shortfalls of more than 8 percent were far more common than in the 1990 recession. The 2005-07 years were a golden period of revenue overages in most of the country. The 2007 recession appears to have hit early in the housing bust states of Arizona, Florida, and Nevada, as well as Michigan. It spread in 2008 and hit the rest of the country with force in 2009 and 2010, before state forecasters were once again on more-solid ground in 2011 through 2013.

Forecasts may be intentionally biased — more likely to be off in one direction than another — because forecasters believe there is an "asymmetric risk" and that the most accurate estimate is not always the best estimate. Many executive-branch forecasters are assumed to believe that the costs of overestimating revenue are greater than the costs of underestimating revenue, because the political and managerial complications of managing deficits are believed to be worse than the consequences of managing surpluses. These forecasters, sometimes explicitly and other times unknowingly, may choose conservative forecasts of the revenue that is available to finance spending. By contrast, legislative forecasters or their principals, not generally responsible for financial management, might not be bothered in the least by revenue shortfalls that the governor has to manage, but may be troubled to learn of "extra" revenue that they could have spent on their priorities — that is, they may be more willing to risk revenue overestimates and revenue shortfalls than is the executive.

Figure 3. Percentage Forecast Errors for Sum of PIT, CIT, and Sales Taxes, 1987-2013

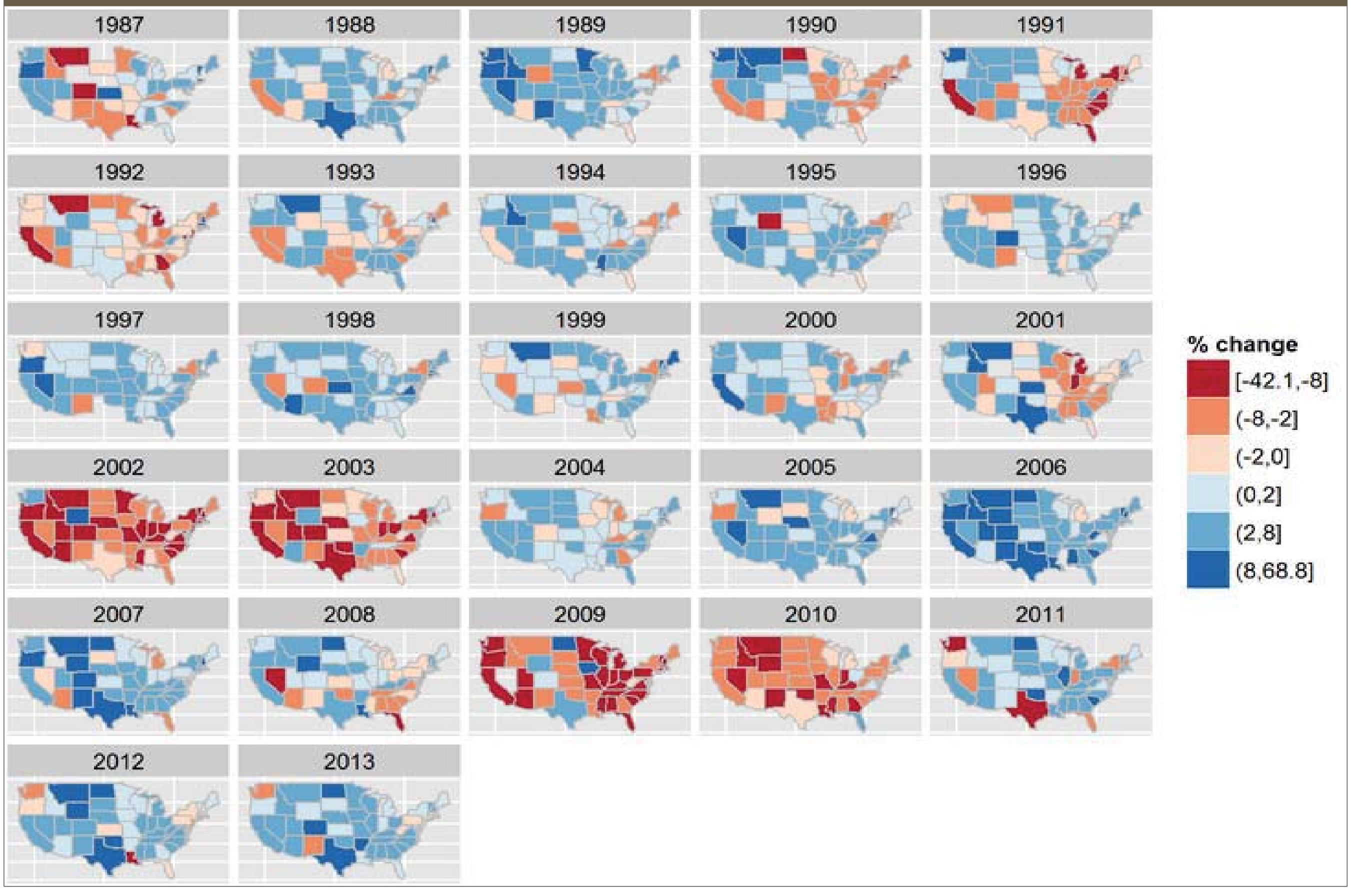
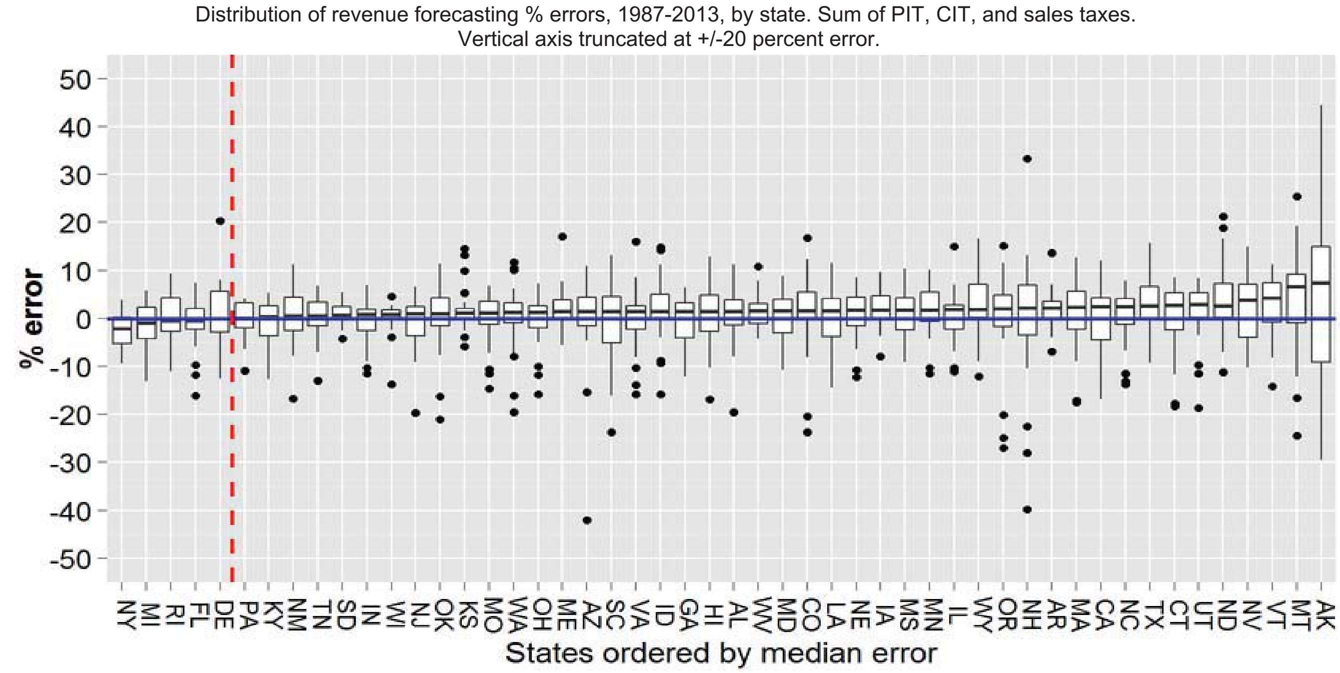


Figure 4 on the following page displays boxplots of each state's percentage forecasting error for the sum of personal income tax, sales tax, and corporate income tax. The top horizontal line of each box marks the 75th percentile of forecast errors, and the lower horizontal line marks the 25th percentile. Each box also has vertical lines (sometimes called "whiskers") extending from its top and bottom, indicating variability in the top and bottom 25 percent of the data (longer lines indicate more variability in this part of the data), and dots marking extreme outliers.[6] The taller the box (the greater the difference between the 75th and 25th percentiles), and the longer the lines, the greater the variability in forecast errors for a given state.

Only five states have a negative median forecasting error suggesting again that states may be "trying" to have small positive forecasting errors — underforecasts of revenue. However, in a common formal but rather strict test of unbiasedness known as the Mincer-Zarnowitz test we reject unbiasedness at the 5 percent level for only eight states: Arkansas, Iowa, Montana, New York, North Dakota, Texas, Vermont, and Wyoming.[7] For each of these states, virtually all of the box in Figure 4 is above the line (about three quarters of the errors are positive), or else below the line (in the case of New York).

Figure 4. Forty-Five States Underestimated Revenue in the Typical Year



Source: Rockefeller Institute Analysis of Data From NASBO Fall Fiscal Surveys.

Judging by the graph, there appears to be a slight tendency for states with greater variability in the middle half of their data, measured by vertical length of the box, to have slightly more-positive errors (more underforecasting) than those with lesser variability, suggesting that when the possibility of large negative errors is greater, states may be more likely to protect against downside risk by aiming for small positive errors — i.e., bias.

Absolute Value of Percentage Forecasting Errors

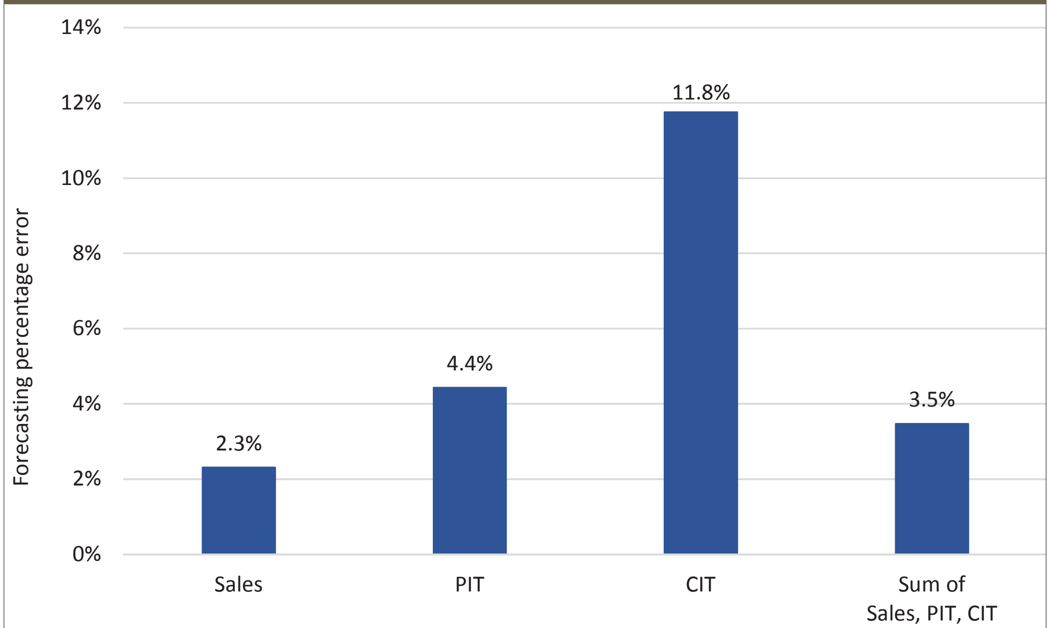
As noted earlier, the absolute value of the percentage error is useful in examining accuracy. Figure 5 on the following page, which shows median errors by tax, reinforces the idea that the corporate income tax has been the hardest tax to forecast, by far.

Table 4 on the following page provides another view, showing the size distribution (across states) of the median absolute percentage error, by tax. It is clear that large errors are extremely common for the corporate income tax, and much more common for the personal income tax than the sales tax.[8]

Figure 6 on page 13 shows median absolute percentage errors by fiscal year. The trends shown on indicate that:

- Errors become particularly large in and after recessions.
- Variation in errors across states also increases in and after recessions.
- Errors near the 2001 and 2007 recessions were much worse than in the recession of the early 1990s.

Figure 5. Absolute Value of Percentage Forecasting Errors by Tax, Median for 1987-2013



Source: Rockefeller Institute analysis of data from NASBO Fall Fiscal Surveys.

- States have returned to "more normal" forecasting errors with the 2007 recession now behind us.

Table 5 on page 14 shows the median absolute percentage errors by state. As before, the pattern of larger errors among smaller states and states heavily reliant on one or two large industrial sectors (such as oil and gas, or gambling) is clear: for example, median absolute percentage errors are particularly large for Alaska, Montana, Nevada, New Hampshire, Texas, and Vermont. In general these states also have greater variation in their errors over time, too. (Among these states, the taxes reported to NASBO for Alaska, Montana, New Hampshire, and Vermont are a relatively small share of total taxes as measured by the Census Bureau. Their errors for total taxes in their budget, which also include other less volatile taxes, appear to be smaller than those examined here.)

Table 4. Very Large Errors Are Much Rarer for the Sales Tax Than for Personal Income or Corporate Income Taxes

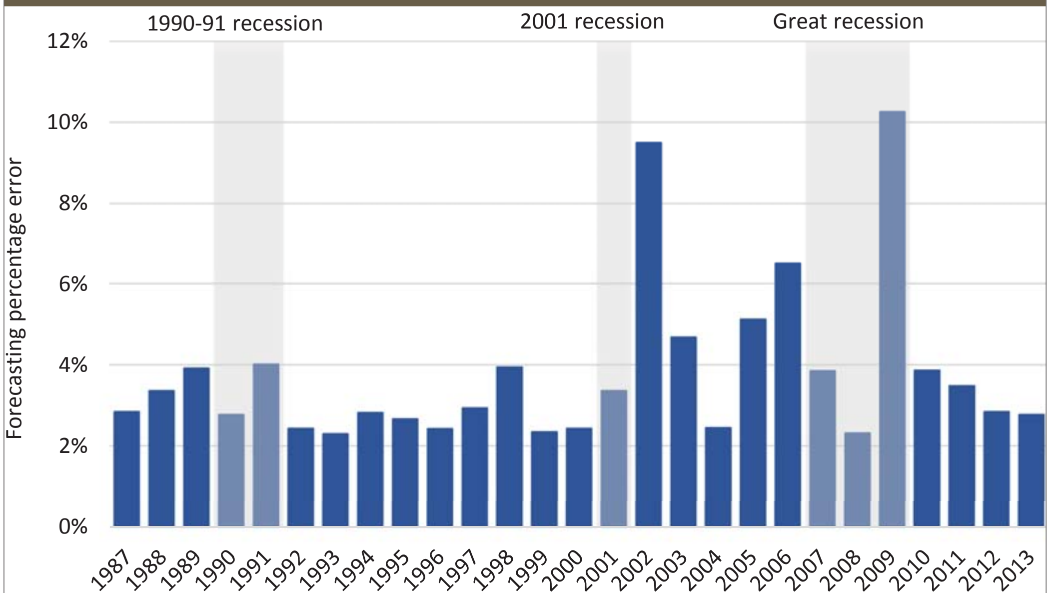
Number of states by median absolute percentage error and tax, 1987-2013				
Size of median absolute error	**Sales**	**PIT**	**CIT**	**Sum of Sales, PIT, CIT**
< 2%	15			4
2% to 3%	18	8		13
3% to 4%	8	10	1	18
4% to 5%	2	9		8
5% to 10%	3	15	17	6
10% to 20%			26	1
> 20%			3	
Total	**46**	**42**	**47**	**50**

Source: Rockefeller Institute analysis of data from NASBO Fall Fiscal Surveys.

Constructing a Forecast Difficulty Measure

One of the problems in comparing forecast errors across states or over time is that it is more difficult to be accurate with some forecasts than others. Some taxes are more volatile than others, some periods are more uncertain than others, some states have

Figure 6. Absolute Value of Percentage Forecasting Errors by Fiscal Year, 1987-2013



Source: Rockefeller Institute analysis of data from NASBO Fall Fiscal Surveys.

more volatile economies than others, and so on. If forecast errors are greater in one year or state than in another, it may not mean the forecaster was less adept, but rather may mean that the forecasting task was more difficult.

Many researchers who examine revenue forecasting errors attempt to control for this, implicitly, by including measures of economic conditions in regressions such as unemployment rates and changes in personal income. While this is useful, tax revenue depends on many things other than broad measures of the economy. For example, capital gains can vary widely even when the economy is stable, leading to forecast errors not easily tied to measures of economic volatility. It would be useful to have a measure of forecast difficulty that does a better job of capturing whether a particular forecast is difficult or not. It is an important technical issue because it is not possible to draw many conclusions in regression analyses about factors affecting forecast accuracy unless we take into account how difficult a particular forecast is. Beyond that, having a strong measure of forecast difficulty is useful in its own right because it allows much greater insights, even when examining forecast errors in a descriptive sense rather than in a statistical sense.

Our contribution to this issue is to develop a more-sophisticated measure of forecast difficulty.[9] The way we have done this is by constructing a uniformly applied “naïve” model that forecasts each tax in each state and each year, using data that generally would have been available to forecasters at the time. We then compute the error in the naïve model’s forecast and use that as our measure of forecast difficulty. If a model has trouble forecasting an observation, that particular forecast probably is difficult.

The particular model we use is a common and very useful one: an exponential smoothing model with a trend, which uses ten previous years of data to forecast each tax in each state and each year. The model uses Census Bureau data on taxes for each state and tax (so that it can use historical data for years before our data start, and for other technical reasons), looks back ten years, and forecasts ahead one. To do this we need to estimate and

Table 5. Absolute Value of Percentage Forecasting Errors, By State

State	Mean	Standard deviation	25%	Median	75%
US	*3.3*	*2.7*	*2.0*	*2.3*	*4.0*
AK*	18.4	14.2	9.7	14.6	25.2
AL	4.3	4.2	1.5	3.6	5.2
AR	3.4	2.9	1.8	3.0	3.8
AZ	5.2	8.1	1.8	2.7	5.4
CA	5.9	4.3	2.7	4.3	8.2
CO	5.9	6.2	1.4	3.8	7.7
CT	5.4	4.5	2.6	4.4	6.8
DE*	5.3	4.3	2.6	4.7	7.0
FL	3.7	3.8	1.3	2.2	4.7
GA	4.1	2.6	2.2	3.7	5.2
HI	5.1	4.1	2.1	3.9	6.7
IA	3.4	2.4	1.6	2.9	4.8
ID	5.0	4.9	1.2	2.7	8.6
IL	3.9	3.5	2.0	2.7	4.7
IN	3.7	3.3	1.1	2.4	5.4
KS	3.1	3.8	1.0	1.7	3.2
KY	4.2	3.1	2.4	3.4	5.0
LA	5.2	3.9	2.3	4.1	6.5
MA	5.8	4.7	2.6	5.1	8.1
MD	4.6	3.0	2.2	3.7	7.0
ME	3.5	3.4	1.3	2.2	4.6
MI	4.2	3.5	1.3	3.6	5.6
MN	4.2	3.3	1.8	3.4	6.2
MO	4.0	3.7	1.3	3.1	5.4
MS	4.4	2.9	2.3	3.8	6.2
MT*	9.9	6.4	6.0	8.0	13.0
NC	4.6	3.6	2.2	3.1	6.1
ND	11.7	17.5	2.7	4.8	14.6
NE	4.1	3.2	1.6	3.2	6.2
NH*	10.9	11.6	2.2	6.6	12.6
NJ	4.0	3.9	1.8	3.0	4.8
NM	4.7	3.8	1.9	3.7	7.1
NV	6.3	3.6	3.9	6.2	7.9
NY	3.6	2.8	1.6	2.7	5.2
OH	3.8	3.7	1.5	2.5	4.2
OK	4.8	5.1	1.5	2.7	6.3
OR	6.8	7.2	2.0	4.4	9.2
PA	2.8	2.4	1.2	2.8	3.9
RI	4.6	3.3	2.2	3.5	7.7
SC	6.6	5.5	2.6	4.6	7.9
SD	2.1	1.7	0.7	1.7	3.0
TN	3.5	2.9	0.9	3.0	5.0
TX	5.6	4.4	2.3	5.2	9.0
UT	4.9	4.0	2.3	4.2	6.1
VA	4.6	4.7	1.5	2.7	4.4
VT*	6.0	3.4	3.1	5.7	8.0
WA	4.8	5.1	1.2	3.0	6.8
WI	2.0	2.6	0.8	1.6	2.2
WV	2.9	2.4	1.4	1.9	3.4
WY*	5.6	5.4	1.2	3.4	9.5

Source: Rockefeller Institute analysis of data from NASBO.
Note: States with * relied on PIT, CIT & sales taxes for less than half of their tax revenue in a typical year.

forecast from a separate model for every single data point. For example, to forecast Massachusetts's income tax in 1998, we would use data for the Massachusetts income tax from 1988 to 1997, estimate an exponential smoothing model, and forecast one year ahead. To forecast Massachusetts's income tax in 1999, we would estimate a new model using data from 1990 to 1998 and then forecast one year ahead. Since we need a separate model for every data point, we had to estimate and forecast from approximately 5,000 models (one for each state, tax, and year, and one for the sum of the taxes in each state and year).

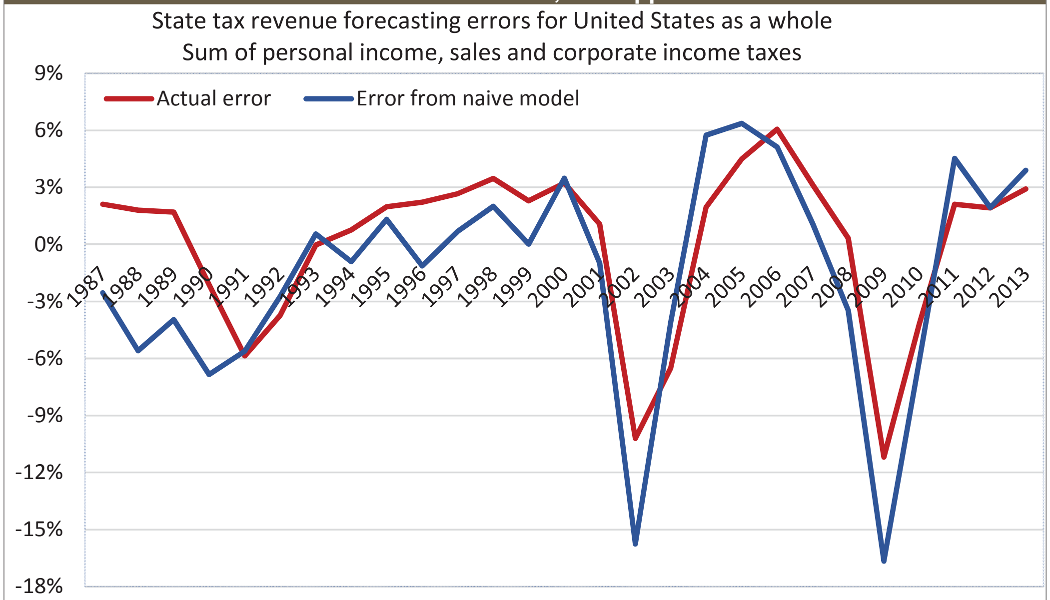
The great advantages of this approach are that it is uniformly applied to each data point, it relies largely (but not perfectly) on data that forecasters would have had available at the time, and it works well. The main disadvantages are that it is a very naïve model — exponential smoothing models forecast a variable using data only from that variable's past, and they cannot call turning points and are likely to be particularly wrong when the economy changes direction. Because revenue forecasters read the newspapers, talk to economic forecasters, and make use of a wide variety of information sources that cannot easily fit into uniform models, they are likely to be more accurate than our naïve model.

Figure 7 on page 15 shows the percentage error from our naïve model and the actual forecasting error by states for the U.S. as a whole, over time. It is clear that the model does a fairly good job of capturing the difficulty state forecasters face. We'll say more about this below.[10]

Figure 8 on page 16 shows histograms with the distribution of state forecasters' errors (the top panel) and the distribution of errors from our naïve model. Note that the forecasters' errors are **not** clustered around zero but rather tend to be positive on average. As we have discussed, this is an indication of *possible* bias. Of course, it is also possible that this is simply a function of our data, and that for the years we have good actual revenue was likely to be higher than estimates from unbiased models. However, note that the errors from the naïve models **are** clustered around zero and are quite symmetric (although not normal). Given that the naïve model's errors tend to be clustered around zero, but the naïve model's errors are not, this suggests that economic and revenue patterns were **not** the cause of positive errors, further reinforcing the idea that state forecasters were biased.

Finally, Figure 9 on page 17 shows the geographic distribution of the median absolute percentage forecast error from our naïve model. Forecast difficulty appears to be greater in smaller states

Figure 7. Actual Forecasting Errors Follow Same Pattern as Naïve Model's Errors But Are Smaller, and Appear to be More Positive



Sources: Rockefeller Institute analysis of data from NASBO Fall Fiscal Surveys & Census Bureau.

and in resource rich states, suggesting that the larger actual errors seen in those states is not simply an indication of bad forecasting, but that there is something about those states that has made revenue harder to forecast in those states.

Conclusions and Policy Implications From Descriptive Analysis

The main conclusions from our descriptive analysis of the data are:

- Corporate income tax forecasting errors are much larger than for other taxes, followed by the personal income tax. The median absolute percentage error was 11.8 percent for the corporate income tax, 4.4 percent for the personal income tax, and 2.3 percent for the sales tax.
- Smaller states and states dependent on a few sectors of the economy (particularly states reliant on oil or natural gas, or gambling), such as Alaska, Montana, Nevada, New Hampshire, and North Dakota, tend to have larger errors. Those states' errors also tend to be more variable. (Among these states, the taxes reported to NASBO by Alaska, Montana, and New Hampshire are a relatively small share of total taxes as measured by the Census Bureau. Their errors for total taxes in their budget, which also include other less volatile taxes, appear to be smaller than those examined here. However, these states also have large errors from naïve forecasting models using full Census data.)
- When taxes are particularly difficult to forecast states tend to be more likely to underforecast revenue, suggesting that they may try to do so in an effort to avoid large shortfalls. Thus, there is a pattern to the apparent bias in state revenue forecasts. By contrast, our naïve forecasting model does not become more likely to underforecast when forecasting becomes more difficult, suggesting that this phenomenon may reflect the behavior of forecasters rather than underlying factors in the economy or tax revenue structures.

Figure 8. Errors from Naïve Forecasting Model Are More-Centered Around Zero Than Forecasters' Actual Errors and Tend to be Larger

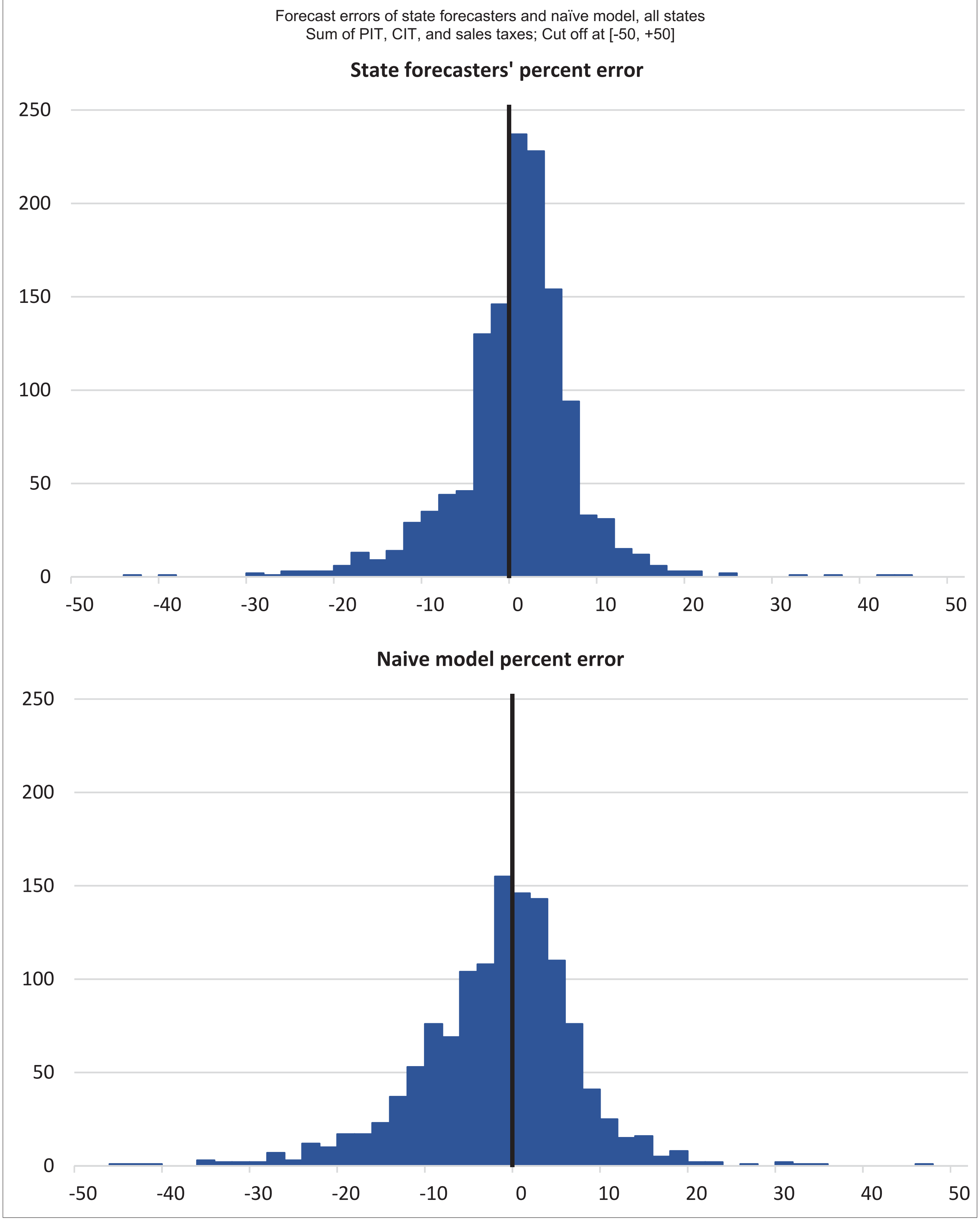
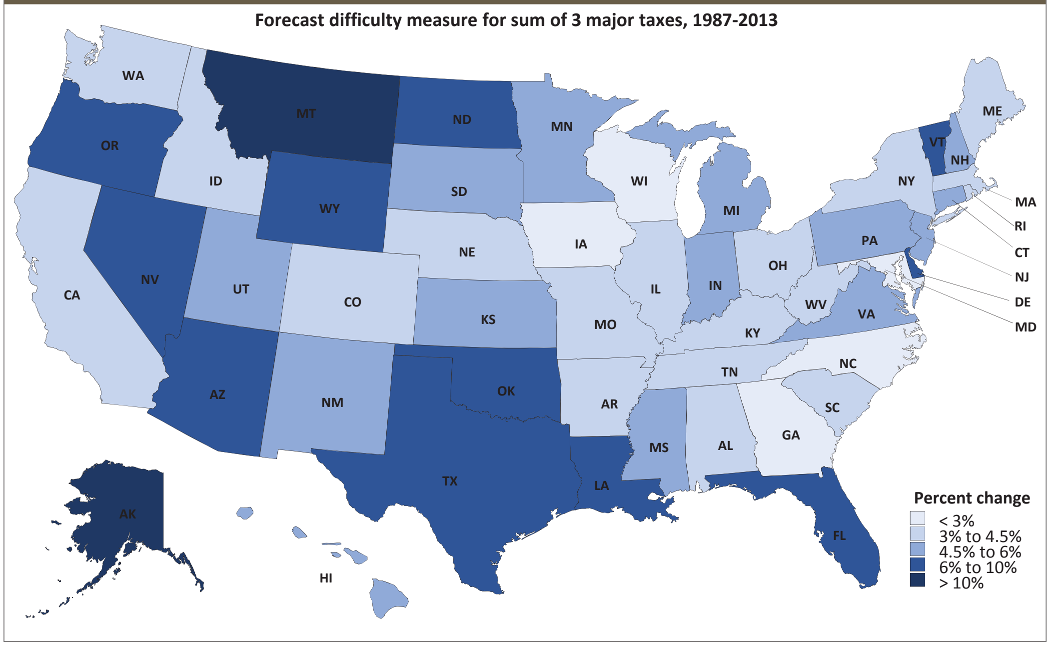


Figure 9. Forecasting Appears to Have Been More Difficult in Smaller States and Resource-Rich States



- Errors become particularly large in and after recessions.
- Variation in errors across states also increases in and after recessions.
- Errors near the 2001 and 2007 recessions were much worse than in the recession of the early 1990s.
- States have returned to "more normal" forecasting errors with the 2007 recession now behind us.
- One limitation of this study is that we do not have detailed and comprehensive data on total revenue forecasts and our analysis is limited to the sum of personal income, corporate income, and sales taxes. Thus, these data do not tell whether and by what magnitude an overall state forecast was missed.

Revenue Forecasting Accuracy and Revenue Volatility

Revenue volatility plays a significant role in revenue forecasting accuracy and, as noted in the prior report, tax revenue has become more volatile and difficult to forecast. States do have options for reducing volatility, such as diversifying the set of taxes they rely upon or broadening the bases of those taxes. They also have options for strengthening their ability to manage volatility, through use of rainy day funds and other mechanisms.

Prior Research on Revenue Volatility and Revenue Forecasting

A number of factors influence the volatility of state tax revenues, including the structure of each state's tax system, the uniqueness of each state's economy, and the severity of economic cycles (Felix 2008; Cornia and Nelson 2010).

Several studies have examined strategies that might reduce budget uncertainty caused by revenue volatility, or that would allow better management of the effects of volatility. One such strategy is diversifying tax revenue to reduce volatility, and another is establishing a rainy day fund to manage the effects of volatility.

Crain (2003) used panel data for all fifty states from 1968 to 1998 to examine the volatility of sales and income tax revenues in each state, as well as the impact of tax structure diversification on overall revenue volatility. Crain defined revenue volatility as "the standard deviation of the deviation in revenue from its long-run trend line" (Crain 2003). According to the study results, "Among the states that levy both types of taxes, 62 percent (twenty-three out of thirty-seven states) experience less volatility in individual income tax revenues than in sales tax revenues" (Crain 2003). This conclusion may seem inconsistent with the permanent income hypothesis and consumption smoothing, where consumption patterns are based on long-term expected income (not immediate income) and therefore might be less volatile than annual income. However, state income and sales tax bases differ substantially from broader economic measures of income and consumption, and the conclusions need not be inconsistent. (However, unlike Crain, we conclude in this report that income taxes are more volatile than sales taxes, in part because we include more-recent data.)

In addition, Crain also examined the impact of diversified tax structure on revenue volatility. Crain found that in thirteen out of thirty-seven states that have both sales and income taxes, the volatility in one tax source is large enough to offset any potential benefits from tax diversification. The author concludes, "that in nearly two-thirds of the states, tax diversification effectively reduces revenue volatility, as standard portfolio theory would predict. In one-third of the states, relying on a single-tax instrument would predictably reduce revenue volatility relative to the status quo level of tax diversification" (Crain 2003). Interestingly, only half of the states levying both forms of taxes used the *less volatile tax* to collect the majority of revenue.

Crain's work preceded dramatic growth in capital gains and other nonwage income, and income tax volatility is a much greater problem now than when he wrote. Two researchers at the Chicago Federal Reserve Bank concluded in a 2008 paper, "Our findings suggest that increasing income cyclicality, in particular of capital gains, have made state revenues more responsive to the business cycle since the mid-1990s" (Mattoon and McGranahan 2012).

Mattoon (2003) suggested creating a national rainy day fund as a strategy for improving state fiscal performance. He argues, "States seem reluctant to 'fix' their tax structures to better manage volatility. In addition, it is unclear that revenue volatility is necessarily a bad thing if states are willing to create budget stabilization tools. Efforts to broaden major tax bases, such as subjecting services to sales taxation, have seen little progress" (Mattoon 2003). The author finds that if states had access to the national rainy day fund he proposed, for fiscal years 2002 to 2004, 74 percent of the aggregate state budget deficit would have been covered by the fund, where in practice only one of the fifty states avoided a budget shortfall over these three years (Wyoming). However the fund would have required a 15 percent savings rate for each state, which forty-eight of fifty states had not achieved already, thus inflating the rainy fund's success rate (states could have performed better simply by saving 15 pecent).

Elder and Wagner (2013) also suggest establishing a national rainy day fund, but argue a different rationale. The authors contend that since state business cycles are not perfectly synchronized — economic booms and busts do not always occur concurrently for all states — disparate state business cycles can counteract each other as a collective insurance pool. The study finds that states are synchronized with the national business cycle only 83.6 percent of the time, on average. Ten states were synchronized with the national economy 90 percent of the time, while seven states were synchronized less than 75 percent of the time. Given this lack of perfect synchronization, a national rainy day fund could be created for less than the cost of each state insuring downturns independently, as states with a surplus can support the states in a downturn (Elder and Wagner 2013).[11]

Thompson and Gates (2007) provide several recommendations for reducing state revenue volatility and improving state revenue forecasting. Based on premises of modern financial economics, the authors emphasize the following set of tools that would help policy and budget makers to reduce revenue volatility or manage its effects, and improve revenue forecasting accuracy: substituting more progressive taxes with a less progressive tax; having a well-designed portfolio of tax types; investing in rainy day fund pools; and balancing budgets in a present-value sense by using savings and debt to smooth consumption (Thompson and Gates 2007). Thompson and Gates conclude that a state government *cannot* significantly reduce tax volatility by simply switching from one tax-type to another, unless *also* replacing a more progressive tax structure with a less progressive tax structure. The authors maintain that a diversified portfolio of taxes can significantly reduce revenue volatility, but caveat that "Even the best-designed tax portfolio would not eliminate all volatility in revenue growth" (Thompson and Gates 2007).

In a recent study, Cornia and Nelson (2010) explore the impact of state economies and tax portfolios on revenue growth and volatility. According to their results, states with a combination of taxes would likely reduce the volatility of tax revenues. By the time of their study, the income tax had become more volatile than previously, and more volatile than the sales tax. In contradiction to Crain's study (and consistent with ours), the authors conclude, "The sales tax offers stability but at the cost of a lower growth rate. The individual income tax offers growth but at the cost of increased volatility" (Cornia and Nelson 2010).

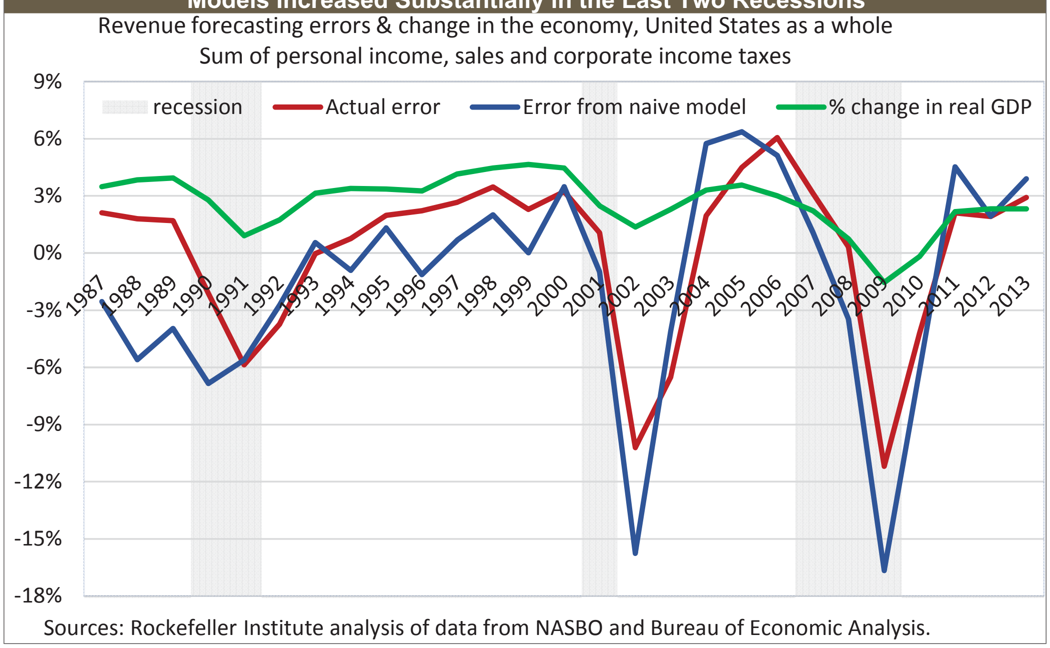
Trends in Revenue Volatility and How It Relates to Forecasting Accuracy

Tax revenue volatility has increased in recent decades, making forecasting more difficult. Figure 10 shows states' forecasting errors as a percentage of actual results, the error from our naïve forecasting model, and the year-over-year percentage change in real gross domestic product, in each case for the span of our data, i.e., 1987-2013 and for the United States as a whole.

The figure shows three periods in which revenue forecasting errors were sharply negative, with states overestimating how much would be available: in and after 1990, in and after 2001, and in and after 2008, in each case associated with a national recession. The green line shows growth in real gross domestic product (GDP), and in each of those periods it slowed substantially or fell.[12] Several points are evident from the graph:

- Revenue forecasting errors were more than twice as large in the two most recent recessions compared to the 1990 recession;
- Those errors were much greater than the

Figure 10. Forecasting Errors of State Revenue Forecasters and of "Näive" Forecasting Models Increased Substantially in the Last Two Recessions



Sources: Rockefeller Institute analysis of data from NASBO and Bureau of Economic Analysis.

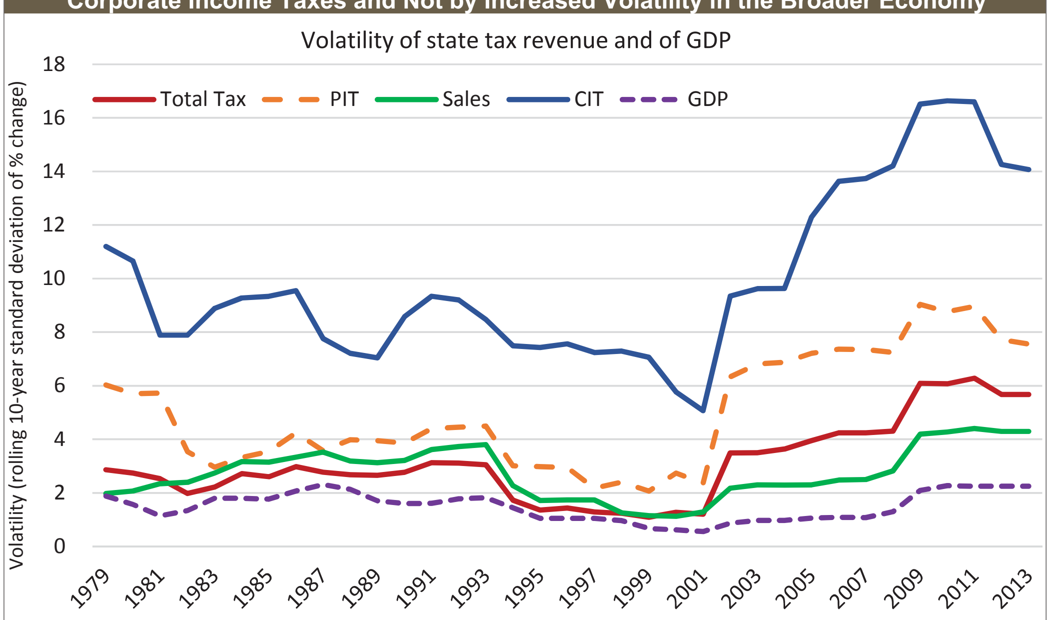
change in real GDP, suggesting that even relatively small changes in a broad measure of the economy can lead to large tax revenue forecasting errors;

- Errors from our naïve forecasting model in the two most recent recessions also were more than twice as large as earlier errors, suggesting that the task of revenue forecasting became more difficult rather than that forecasters became less adept; and
- Revenue forecasters' errors generally were smaller on the downside than errors from our naïve model, but about the same as or even slightly more positive than those from our naïve model, when errors were positive. This suggests that revenue forecasters have been more accurate than our naïve model, and perhaps biased toward positive errors (underestimates of revenue).

One common measure of volatility is the standard deviation of the periodic percentage change in a variable.[13] Applied to tax revenue, it is the standard deviation of the annual percentage change in tax revenue.[14] It is always a positive number, and the larger it is, the more volatile the revenue. The intuition is this: suppose a state's tax revenue grew by 5 percent each year, come rain or shine. Its mean percentage change would be five and because it never varies its standard deviation would be zero — revenue changes from year to year, but it is not volatile. And it is easy to predict. However, suppose the mean remains 5 percent but revenue grows in some years by as much as 8, 10, or even 20 percent, while in other years it grows by 2 percent or even falls by 5 or 15 percent. In this case, the standard deviation will be much larger — the tax is more volatile. And revenue probably will be difficult to predict. The standard deviation of percent change also can be used to measure the volatility of economic variables such as real GDP. By examining changes in the standard deviation over time, we can see whether a variable is becoming more volatile or less volatile over time.

Figure 11 shows the rolling ten-year standard deviation of the percentage change in the Census Bureau's data on state tax revenue for

Figure 11. Increasing Volatility of State Tax Revenue Has Been Driven by Personal and Corporate Income Taxes and Not by Increased Volatility in the Broader Economy



Sources: Rockefeller Institute analysis of data from Census Bureau & Bureau of Economic Analysis.

the personal income tax, sales tax, and total taxes, and of the percentage change in nominal gross domestic product, for the United States as a whole. Each point on a line represents the standard deviation of percentage changes over the ten previous years. For example, the first point for taxes is 1979, and it measures the standard deviation of the annual percentage changes in each year from 1970 through 1979. As we move forward in time, we maintain the ten-year window, so that, for example, a data point for 2009 is the standard deviation of annual percentage changes for each year from 2000 through 2009. By comparing standard deviations across variables, we can see whether one variable is more volatile than another.[15] By examining how the standard deviation has changed over time, we can see whether volatility has been increasing or decreasing.

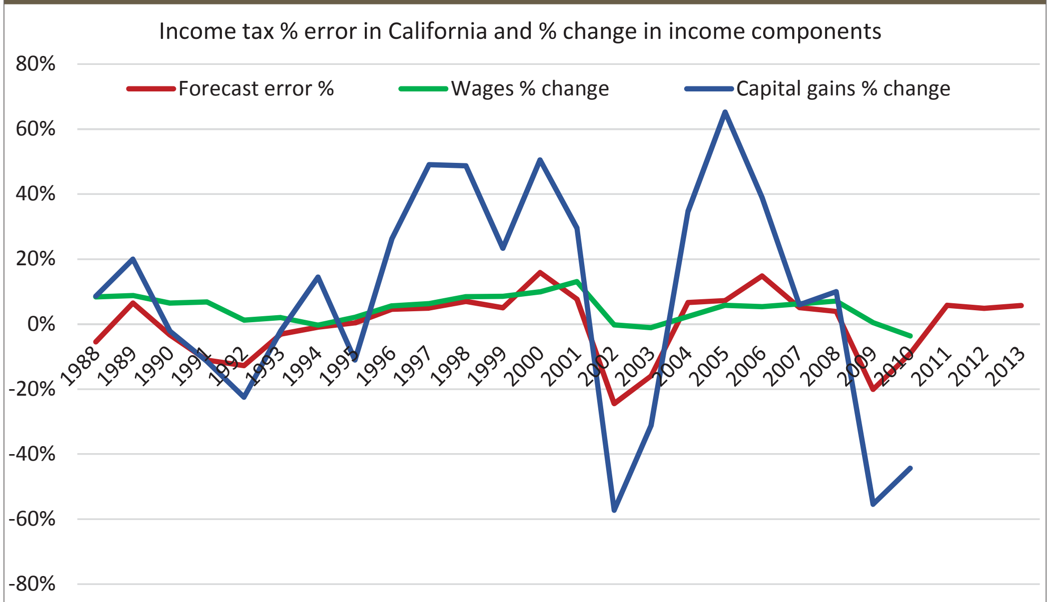
From this figure we can see that:

- The rolling standard deviation of the percent change in total taxes was below 3.5 percent in every year before 2001, after which it rose sharply and exceeded 5.5 percent in 2009-13. (It's important to remember that each point summarizes ten years of data, so the increase in volatility seen beginning in 2001 reflects percentage changes from 1992 through 2001 and how they differed from previous periods.)
- Corporate income taxes are by far the most volatile.
- Personal income tax volatility increased much more than did sales tax volatility, although the sales tax became somewhat more volatile.
- Throughout the period, the income tax was more volatile than the sales tax, and volatility of total taxes generally was between the two.
- Taxes are more volatile than gross domestic product, a broad measure of the economy.
- While the economy became more volatile between 2001 and 2010, the increase was much smaller than the increase in tax volatility.[16] Put differently, the increase in tax revenue volatility appears to have been caused by forces other than just an increase in the volatility of the economy, broadly measured.

Increases in tax revenue volatility have been widespread: Our analysis of Census tax data shows that forty-two states had an increase in revenue volatility between 2000 and 2013.

We have noted elsewhere that the increase in tax revenue volatility is related to increasing volatility of the income tax, in turn influenced by increasing reliance on capital gains and other nonwage income, and increasing volatility of those sources. Figure 12 illustrates this with data from California: annual percentage forecast errors, and the year-over-year changes in income from wages and from capital gains.[17] Capital gains clearly are

Figure 12. Capital Gains Are Far More Volatile Than Wage Income and Are Associated With Increased Forecast Error



Sources: Rockefeller Institute analysis of data from NASBO & Internal Revenue Service.

more volatile than wages, and forecasting errors generally were much larger in years when capital gains changed significantly.

In general, volatility in the overall economy makes it harder to produce accurate forecasts and states often revisit their revenue estimates during the business cycles. For example, an official from Vermont said: "During periods of economic volatility especially downturns, updates and revisions have been made more frequently to allow timely executive and legislative response to fiscal events." Volatility can be compounded by federal policies that have unintended and hard to measure consequences at the state level. For example, the fiscal cliff not only contributed to the volatility of income tax receipts, but also made the job of the forecasters in many states much harder.

Reducing Volatility by Changing Tax Structure

How much could tax revenue volatility and associated revenue forecasting errors be reduced if states chose different portfolios of taxes? We examined this by recomputing the total forecast error that each states would have had in each year of our data if they had had different mixes of personal income, sales, and corporate income taxes. We did this for each state that used all three taxes in our data, for each year, recomputing the percentage error for the sum of the three taxes while varying the percentage share of each tax from zero to 100 percent, in 10 percentage-point increments. We then summarized the error for each portfolio by computing the median absolute percentage error for the sum of the three taxes, across all states and years.

Table 6 on the following page shows the results of this analysis. Each cell represents a combination of personal income, sales, and corporate income taxes. For example, the fourth row and second column represents a portfolio in which 30 percent of the revenue is from the income tax and 10 percent of the revenue is from the sales tax and therefore, by subtraction, 60 percent of the revenue is from the corporate income tax. It has a median absolute percentage error, across all states and years, of 7.8 percent. This is

Table 6. States Have Relatively Little Opportunity to Reduce Forecasting Error by Changing Their "Tax Portfolios"

Absolute percentage error, median across states and years, different tax portfolios
Compare to median absolute % error of actual portfolios for these data: 3.2%

Personal income tax as share of total taxes	Sales tax as share of total taxes										
ps	**0%**	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**	**80%**	**90%**	**100%**
0%	11.6	10.5	9.6	8.6	7.5	6.4	5.4	4.4	3.4	2.6	2.1
10%	10.6	9.7	8.7	7.5	6.5	5.4	4.6	3.5	2.7	2.1	
20%	9.9	8.9	7.8	6.6	5.5	4.6	3.6	2.8	2.3		
30%	8.9	7.8	6.8	5.8	4.7	3.7	3.0	2.4			
40%	8.0	6.9	5.8	4.8	3.9	3.2	2.5				
50%	7.1	6.0	5.0	4.1	3.3	2.8					
60%	6.2	5.3	4.4	3.5	3.0						
70%	5.5	4.6	3.8	3.2							
80%	4.9	4.0	3.5								
90%	4.3	3.8									
100%	4.2										

Notes: (1) Corporate income tax is the residual. For example, when sales tax is 70% and income tax is 10%, the corporate income tax is 20%.
(2) Limited to the 905 observations for which a state had all three taxes (personal income, sales, and corporate income).
(3) Cells with lines around them have zero corporate income tax.
(4) Green-shaded portfolios have lower total error than historical experience.
(5) Blue-shaded portfolios have total error above 9 percent.

considerably greater than the 3.2 absolute percentage error that states with all three taxes actually experienced (noted in the table subtitle). Cells in the table that are shaded pink have median errors that are larger than 9 percent, cells that are shaded green are lower than the actual error of 3.2 percent, and cells that have solid border lines include no corporate income tax. As we move to the northwest, the portfolios have increasing corporate income tax shares.

Several conclusions are evident:

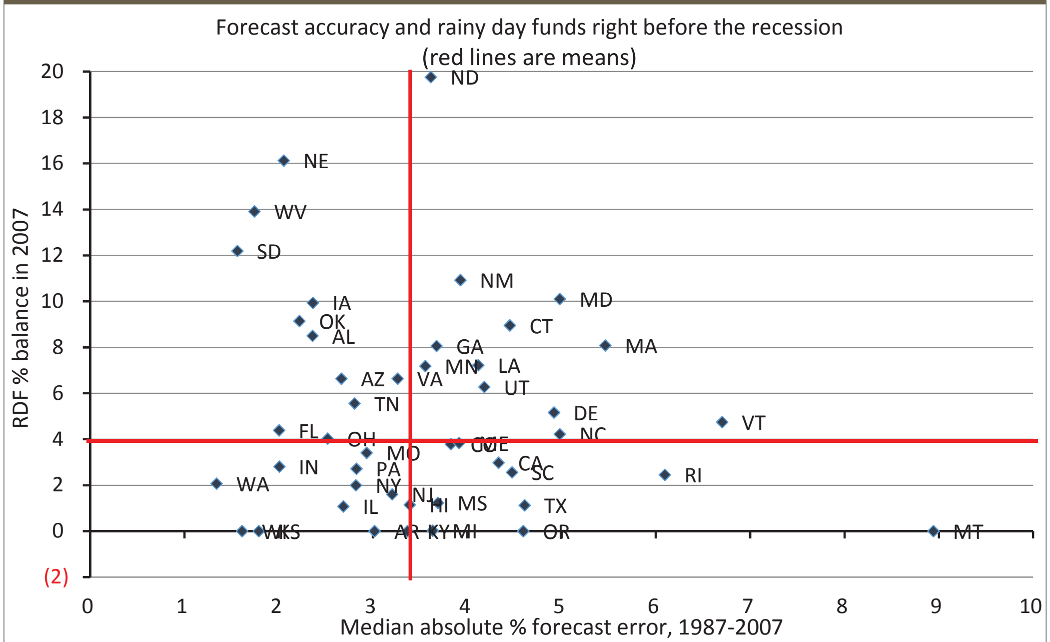
- Increasing reliance on the corporate income tax increases errors dramatically;
- Increasing reliance on the personal income tax relative to the sales tax tends to increase overall forecast errors.
- Relatively few combinations of revenue would reduce the median absolute percent error from what states currently have, and they would not reduce forecast error by very much. Only by virtually eliminating the corporate income tax and significantly increasing reliance on the sales tax relative to the personal income tax could the typical state reduce revenue forecasting errors, and even then most tax combinations would not reduce forecast errors very much.

In summary, states are unlikely to solve the problem of revenue forecasting error by changing the mix of taxes that they have.

Managing Volatility and Errors

States have many possible ways of managing the effects of revenue volatility and forecasting errors. For example, Delaware budgets less than 100 percent of forecasted revenue. Another

Figure 13. Size of Rainy Day Funds Does Not Appear to be Related to Historical Forecast Error



major tool is rainy day funds, which can help to smooth out volatility. We might expect states that have the most volatile and difficult to predict tax revenue sources to have larger rainy day funds to help them manage the consequences of tax revenue shortfalls. As Figure 13 shows, that is not true. There is little apparent relationship between rainy day fund size right before the recession hit and the forecast errors states had experienced to that point. We are missing rainy day fund data for five states: Alaska, Idaho, Nevada, New Hampshire, and Wyoming. Thus, those five states are not shown on Figure 13. As mentioned previously, we strongly urge caution when interpreting data for individual states due to various tax structures and reliance on various sectors of economy.

Conclusions and Policy Implications Regarding Forecasting Accuracy and Revenue Volatility

We draw several conclusions from this analysis:

- Increases in forecasting errors have been driven by increases in revenue volatility, which in turn have been driven in large part by increased reliance on increasingly volatile capital gains.
- States have relatively little opportunity to reduce volatility simply by restructuring their tax portfolios. Only by virtually eliminating the corporate income tax and significantly increasing reliance on the sales tax relative to the personal income tax could the typical state reduce revenue forecasting errors, and even then most tax combinations would not reduce forecast errors very much. Changing the structure of individual taxes may be more promising, but it can raise difficult tax policy issues. For example, making income taxes less progressive might also make them less volatile and easier to forecast, but that may run counter to distributional objectives that some policymakers hope for.

- Because it is so hard to reduce forecasting errors by changing tax structure, it is especially important for states to be able to manage the effects of volatility. Rainy day funds are one important tool states can use. However, the data suggest that states with larger forecast errors and more difficult forecasting tasks do not have larger rainy day funds; perhaps they should.

Revenue Forecasting Accuracy and the Timing and Frequency of Forecasts

Prior Research on Forecasting Accuracy and Forecast Timing and Frequency

Many factors affect forecasting accuracy, including errors in national and state economic forecasts, tax revenue volatility, and other factors. In this section we focus on the relationship between forecast accuracy and forecast timing and frequency.

One important determinant of forecast accuracy is the length of time between forecast preparation and the start of the forecast period, since in general the further ahead the forecast is, the less accurate it is likely to be. There are big differences across states in this lead time. Bretschneider (1989) noted that 68 percent of states prepared their forecasts within six months of the start of the budget year, 20 percent prepared forecasts more than six months ahead and less than a year ahead, and that 12 percent of states prepared forecasts a year or more before start of a budget year, as a consequence of biennial budgeting. Mikesell and Ross (2014) in a working paper examined the implications of biennial budgeting for forecasting error in Indiana. According to the authors, in Indiana the revenue forecast is prepared in December of each even year, and there is an April update prepared in odd years.

Because of widely varying differences across states and within states over time, it is important to take the lead time between a forecast and its time period into account when examining differences in forecast accuracy across states.

Another important factor to take into consideration is the frequency of forecasts. Some scholars argued that more frequent forecasts actually decrease forecast accuracy. For example, Deschamps (2004) writes, “Though it seems intuitive that revising a forecast more often would improve accuracy, such a practice may increase the risk of adjusting a forecast to random error.” Voorhees (2004) found a negative correlation between forecasting frequency and forecast accuracy: as the forecast frequency increases, forecast accuracy decreases. According to Voorhees, “While frequent forecasts will provide an early warning of a downturn, it may be difficult to determine if the downturn is indeed a trend or just randomness” (Voorhees 2004). These conclusions might in part reflect reverse causality — states with harder revenue forecasting challenges may choose to update those forecasts more frequently.

Measuring Timing of Forecasts

We have developed two measures that allow us to examine the implications of time lags between when a forecast is prepared and when it is used.

The first such measure is a simple indicator variable signifying when a state’s forecast is for the second year of a biennium. The idea is that the lag between forecast preparation and forecast use generally is longer for the second year of a biennium. This is a crude measure because the forecast data states provide to NASBO may not always reflect this lag — in some cases states may provide updated estimates to NASBO — and because some second-year-of-biennium states will have longer lags than other states in the same position. On the other hand, the variable has the great advantage that it has some variation within states over time, because several states have shifted from annual budgeting to biennial budgeting, and others have shifted in the other direction. We capture this variation in our measure.

Our second measure is based on the survey we conducted, wherein we asked states explicitly when they prepared their forecasts. The survey results indicate that there is a wide variation among the states in terms of timing of the forecasts prepared for the adopted budget. For example, officials in Arkansas reported that they prepare the preliminary estimates as early as in October, while officials in Pennsylvania indicated that they prepare preliminary estimates as late as in April or May.

From these questions we developed an estimate of the number of weeks between when a forecast was prepared and when the associated fiscal year began. This has the advantage of capturing more subtle variation across states as it is measured in number of weeks rather than by a single indicator. It has the disadvantage, however, of not varying within states over time: We only have one number per state, and we have to assume that the lag is constant within states over time, which is not always the case.

Accuracy and the Timing of Forecast

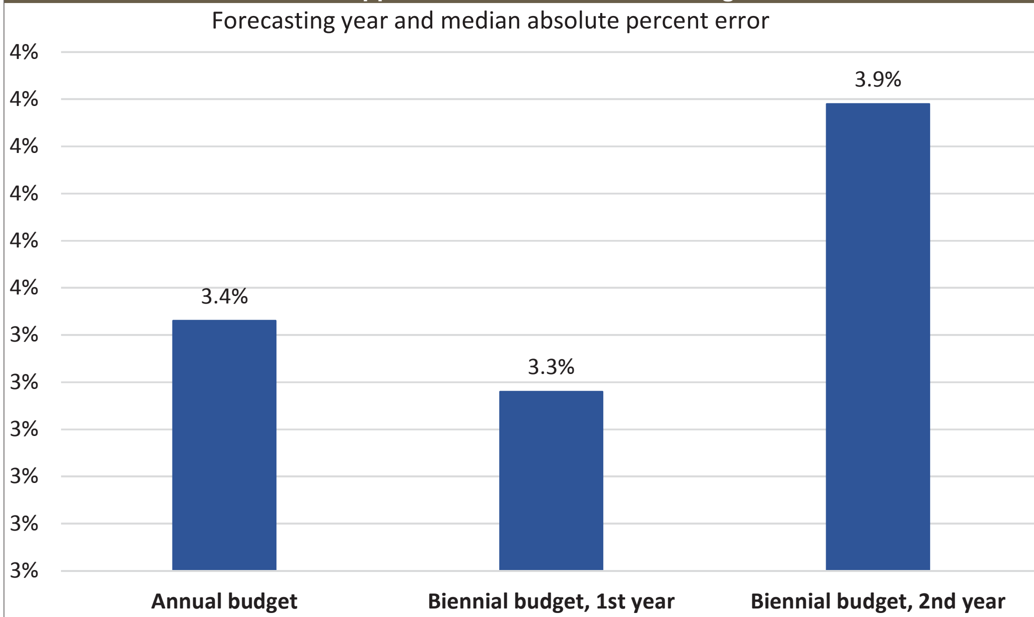
Descriptive Analysis

Figure 14 on the following page shows the median absolute percentage error and the median percentage error for the sum of the three major taxes, categorized by whether the forecast is for an annual budget, the first year of a biennium, or the second year of a biennium. It is not at all apparent from this that this measure of timing matters, although there is a hint that absolute errors may be larger in the second year of a biennium. However, this does not take into account the fact that forecast difficulty might have varied by these groups. We need statistical analysis to take that into account.

Table 7 on the following page shows that the lag between forecast preparation and use (our second measure) tends to be greater in the second year of a biennium (our first measure), and so we would expect greater errors for those forecasts, given that they must be produced further ahead of time.

The survey results indicate that in some states there are longer lags between forecast preparation and the start of the fiscal year, which means states provide forecasts further ahead. Presumably, forecasting is a much harder task for those states that not only don't prepare multiple forecasts but also have to prepare forecasts far in advance. Therefore, forecasting is likely to be more error

Figure 14. Forecast Errors, Without Controlling for Forecast Difficulty, Do Not Appear Related to the Forecasting Year



Source: Rockefeller Institute analysis of data from survey of state forecasting officials.

Table 7. Revenue Forecasts for the Second Year of a Biennium Typically Are Prepared Well Before the Fiscal Year Starts

Lag between when revenue forecast is prepared and when it is used

Budget cycle	# of states	# of weeks from preparation to budget adoption		
		25th percentile	Median	75th percentile
Annual budget	31	7.5	13.5	20.5
Biennial budget, 1st year	19	8.0	12.0	19.0
Biennial budget, 2nd year	19	10.0	22.0	29.5

Source: Rockefeller Institute analysis of data from survey of state forecasting officials.

prone. This is particularly relevant in some biennial budgeting states where they have to forecast the second year of the biennium before the first year has even begun.

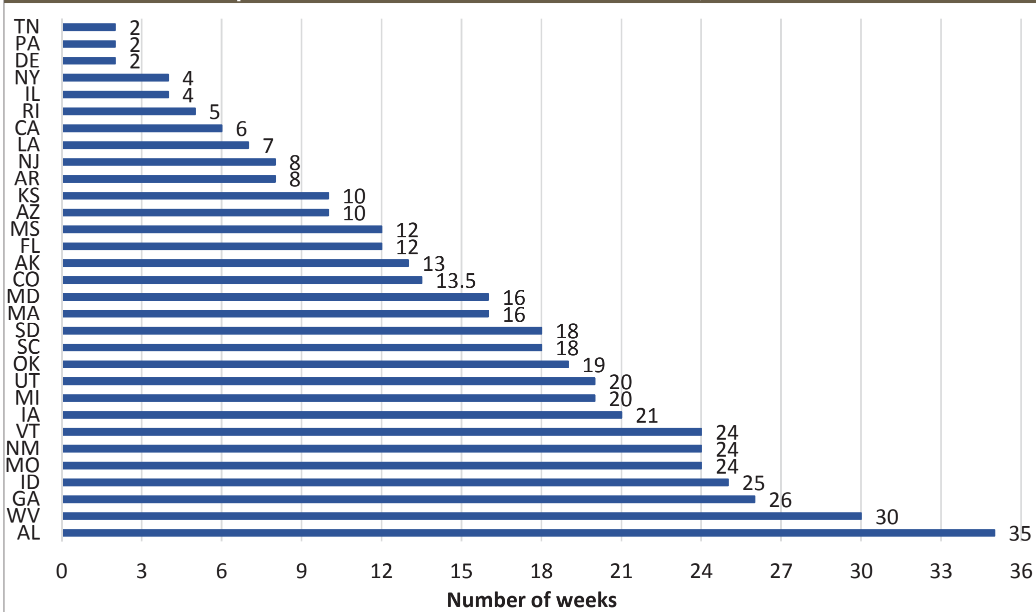
The survey results reveal that there is a wide variation among the states both in terms of timing of the forecast and frequency of forecasts.

In the survey we asked the states to indicate approximately how many weeks ahead of the new state fiscal year the estimates in the adopted budget are prepared in a typical year. We tried to fill in the data for any missing states based on publicly available information on state government websites. In some cases, states indicated the number of weeks between the initial forecast and the start of the state fiscal year. In such cases we have revised the number of weeks to include the lag between the latest forecast used in the adopted budget and the start of the fiscal year. There is a wide variation among the states in terms of how far in advance each state prepares revenue estimates. For example, among the states with an annual budget cycle, Alabama prepares revenue forecasts as early as thirty-five weeks before the start of the state fiscal year on October 1st. On the other hand, states such as Delaware, Pennsylvania, and Tennessee use the revenue estimates updated about two weeks before the start of the state fiscal year (see Figure 15 on the next page).

Among the states with a biennial budget cycle, Texas prepares revenue forecasts as early as thirty-four weeks before the start of the first year of the biennium while Ohio uses updated forecasts just about two weeks before the start of the first year of the biennium (see Figure 16 on the next page).

Figure 17 on page 31 shows the average lag between forecast preparation and use for the second year of a biennium. In general, the lag between forecast preparation/update and the start of the second year of a biennium is greater than the lag between forecast

Figure 15. Average Number of Weeks the Forecast Is Completed Before the Start of the Fiscal Year, Annual States



Source: Rockefeller Institute analysis of data from survey of state forecasting officials.

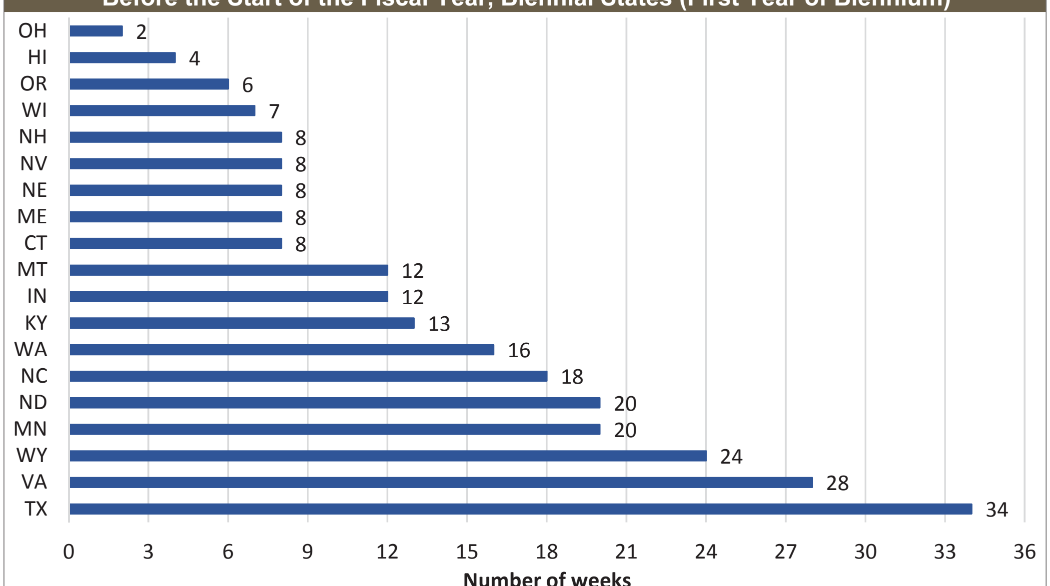
preparation and the start of the first year of the biennium. For example, in Kentucky the lag between forecast preparation and the first year of the biennium is thirteen weeks, while the lag between forecast preparation and the second year of the biennium is sixty-five weeks. (Note: Based on survey results and based on the information retrieved from the publicly available documents, it appears that Texas does not update the forecasts for the second year of the biennium.)

Regression Analysis

Table 8 on the next page reports the results of regression models that examine how forecast accuracy (the absolute percentage error) varies depending on whether the state is forecasting for the second year of a biennium, after controlling for forecast difficulty and other important factors. The different formulations of the model are designed to ensure that we examine this robustly.

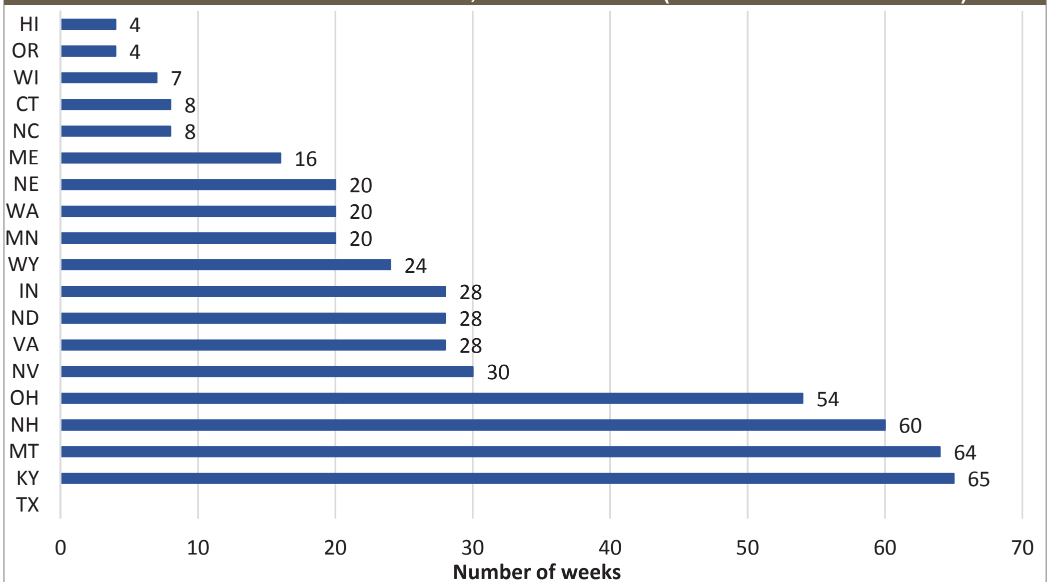
The results of the analysis are very clear: After controlling for forecast difficulty, forecast errors for the second year of a biennium are considerably larger than for other forecast periods,

Figure 16. Average Number of Weeks the Forecast Is Completed Before the Start of the Fiscal Year, Biennial States (First Year of Biennium)



Source: Rockefeller Institute analysis of data from survey of state forecasting officials.

Figure 17. Average Number of Weeks the Forecast Is Completed Before the Start of the Fiscal Year, Biennial States (Second Year of Biennium)



Source: Rockefeller Institute analysis of data from survey of state forecasting officials.

about 1.3 to 1.7 percentage points, depending on the model.

Table 9 on the next page summarizes a similar analysis in which the time-lag variable of interest is our measure of the number of weeks between forecast preparation and use, from the survey. Again, we report on several models (we do not report on a model with state fixed effects because this measure does not vary across years within states). Again the variable is statistically significant at the 5 percent level. The coefficients suggest that each week of lag is associated with increased error of about a one-twentieth of a percentage point (0.05) – for example, ten weeks of lag would be associated with about a half a percentage point greater error, which seems meaningful.

The lag for an annual budgeting state at the 25th percentile is about 7.5 weeks, whereas at the 75th percentile it is 20.5 weeks. That thirteen week difference appears to cost about 0.6 to 0.7 percentage points in additional error, on average.

Table 8. Forecasts for the Second Year of a Biennium Are Less Accurate Than Other Forecasts

Relationship between forecast accuracy and whether the forecast is for 2nd year of biennium

Dependent Variable: Forecast error as % of actual, absolute value	Coefficients and significance levels			
	Model 1	Model 2	Model 3	Model 4
Key variable of interest				
2nd year of biennium, indicator variable	1.337 **	1.649 **	1.676 **	1.650 **
Controls that vary across models				
State fixed effects	No	No	No	Yes
Year fixed effects	No	No	Yes	Yes
Biennial budgeting	NA	(0.396)	(0.409)	(1.985)
Controls for forecast difficulty				
Naïve forecast % error, absolute value	0.235 ***	0.235 ***	0.189 ***	0.185 ***
Square of naïve forecast % error	(0.001) *	(0.001) *	(0.001)	0.000
% change in tax revenue, absolute value	0.217 ***	0.217 ***	0.224 ***	0.213 ***
Square of % change in tax revenue	(0.001) ***	(0.001) ***	(0.001) ***	(0.001) ***
% change in personal income, absolute value	(1.539) ***	(1.546) ***	(0.461)	(0.398)
Square of % change in personal income	0.130 ***	0.131 ***	0.075 .	0.060
Tax-type controls (relative to CIT)				
PIT	(6.299) ***	(6.300) ***	(6.626) ***	(6.921) ***
Sales Tax	(7.987) ***	(7.984) ***	(8.414) ***	(8.454) ***
Adjusted r-squared	0.398	0.398	0.423	0.447
Number of observations	3,256	3,256	3,256	3,256

Significance codes: 0.1% level ***; 1% level **; 5% level *, 10% level .

Data: NASBO forecast errors for states, by year and tax.

Accuracy and the Frequency of Forecast Updates

We also examined the relationship between forecast accuracy and the frequency of forecast updates. One might argue that more frequent forecasts are a sign of greater professionalism, and perhaps greater

Table 9. Forecasts Become Less Accurate the Longer the Lag Is Between Preparation and Use

Relationship between forecast accuracy and lag between forecast preparation and use

Dependent Variable: Forecast error as % of actual, absolute value	Coefficients and significance levels			
	Model 1	Model 2	Model 3	Model 4
Key variable of interest				
# of weeks between forecast preparation & use	0.060 ***	0.058 ***	0.052 ***	0.055 *
Controls that vary across models				
Year fixed effects	No	No	Yes	Yes
Biennial budgeting	NA	0.226	0.288	3.196
Controls for forecast difficulty				
Naïve forecast % error, absolute value	0.240 ***	0.240 ***	0.193 ***	0.189 ***
Square of naïve forecast % error	(0.001) *	(0.001) *	(0.001)	(0.001)
% change in tax revenue, absolute value	0.216 ***	0.216 ***	0.222 ***	0.211 ***
Square of % change in tax revenue	(0.001) ***	(0.001) ***	(0.001) ***	(0.001) ***
% change in personal income, absolute value	(1.583) ***	(1.572) ***	(0.509)	(0.431)
Square of % change in personal income	0.133 ***	0.132 ***	0.077 *	0.063
Tax-type controls (relative to CIT)				
PIT	(6.258) ***	(6.259) ***	(6.608) ***	(6.900) ***
Sales Tax	(8.006) ***	(8.008) ***	(8.456) ***	(8.440) ***
Adjusted r-squared	0.401	0.401	0.425	0.448
Number of observations	3,166	3,166	3,166	3,166

Significance codes: 0.1% level ***; 1% level **; 5% level *, 10% level .

Data: NASBO forecast errors for states, by year and tax.

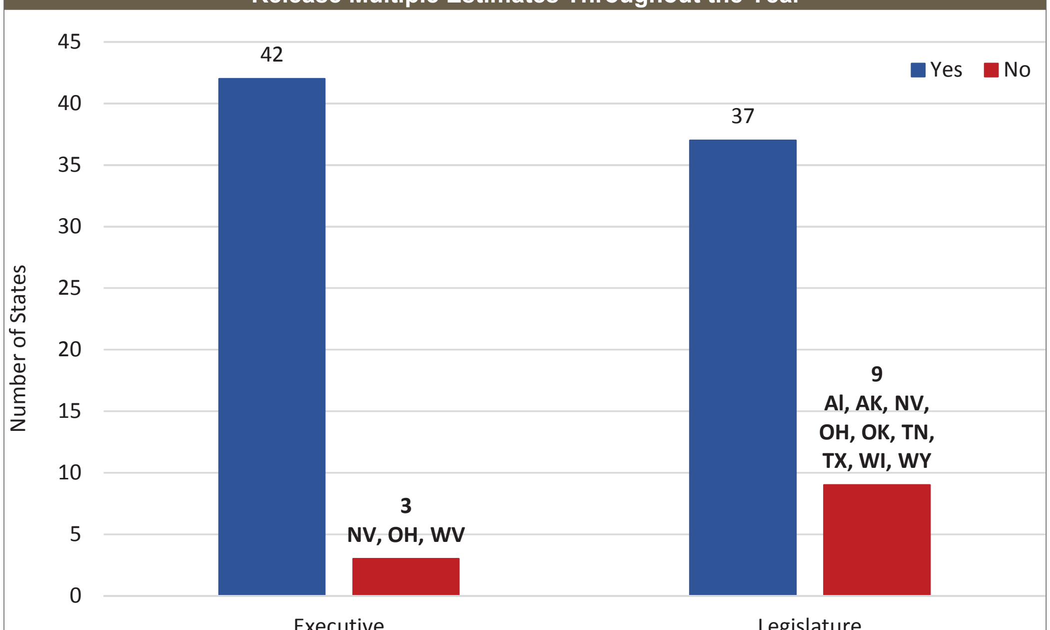
Note: Data for # weeks between forecast preparation and use do not vary over time within states.

professionalism is associated with smaller errors. However, as noted earlier the research on this is ambiguous.

In the survey, state forecasting officials were asked whether they revise revenue estimates or release multiple estimates throughout the fiscal year. Most state officials both in the executive and legislative branches indicated that they do usually provide multiple revenue estimates throughout the fiscal. Figure 18 below captures the list of states where officials in executive and legislative branches indicated absence of revised or multiple revenue estimates.

Officials in both executive and legislative branches of the government indicated various timelines for the updates of the revenue estimates. In some states the updates are done on an ad hoc basis, while in others updates are done bimonthly, quarterly, semiannually, or annually. In very few states revenue estimates are updated more frequently during the legislative session.

Figure 18. States That Revise Revenue Estimates or Release Multiple Estimates Throughout the Year



Source: Rockefeller Institute analysis of data from survey of state forecasting officials.

Table 10 on the next page shows the median absolute percentage error for each grouping of number of forecasts in a year. It does not look from this like more frequent forecasts are associated with smaller errors.

The table, however, does not control for forecast difficulty, and that may be different across the states. We

Table 10. Frequency of Forecast Updates Does Not Appear Related to Forecast Accuracy (Before Adjusting for Forecast Difficulty)

Frequency of forecast updates and forecast accuracy				
# of forecasts in a year	# of states	Median absolute percentage error		
		PIT	Sales tax	CIT
1	22	4.3	2.3	11.8
2	17	4.7	2.4	11.1
3	3	3.5	1.7	10.8
4	6	4.0	2.7	12.5
6	1	4.1	NA	17.9
12	1	6.5	1.8	10.5

Source: Rockefeller Institute analysis of data from survey of state forecasting officials and NASBO.

ran regression models to examine this question, but none of them suggest that forecast accuracy is associated with frequency of forecast updates. This is consistent with past research.

There are other good reasons for states to update forecasts regularly. Forecasters understand that forecasts will be wrong, hence the dictum cited by Paul Samuelson and others, "If you must forecast, forecast often" (Samuelson, 1985). The Government Finance Officers Association has long recommended monitoring, measuring, and evaluating budgetary performance, and adjusting as necessary (National Advisory Council on State and Local Budgeting, GFOA, 1998).

Conclusions and Policy Implications Regarding Forecasting Accuracy and the Timing and Frequency of Forecasts

Further-ahead forecasts are more error prone, whether we examine the forecasts for the second year of a biennium, which generally are prepared further in advance, or use a measure of the lag between forecast preparation and the start of the forecast period that we developed from our survey. In both cases, the effect is meaningful: The data suggest that accuracy worsens by well over a percentage point in the forecast for the second year of a biennium, and worsens by nearly a half a percentage point for every ten weeks of lag between the time a forecast is prepared and the start of the forecast period.

There is no evidence from our limited data that frequent forecast updates lead to greater accuracy, and this is consistent with past research.

The policy implications of the first part of our analysis are clear: States should minimize unnecessary lags between forecast preparation and the start of the fiscal year, and should update those forecasts as close as possible to the start of the fiscal year, as many states do. Even though there is no evidence that more frequent forecast updates during the forecast period will lead to more accurate forecasting, it is good management practice to update forecasts regularly as the year progresses so that finance officials are managing the budget using the best available information.

Revenue Forecasting Accuracy and Institutional Arrangements

Prior Research on Forecasting Accuracy and Institutional Arrangements

Many factors contribute to revenue forecasting accuracy including forecast methodology (such as utilization of expert opinion, nominal groups, Delphi methods, ARIMA, exponential smoothing, moving average, simple regression, multiple regression, multiple equation regression, and simulation methods), political factors (such as political party composition of the government), economic factors (such as the economic condition of the state, unemployment rate, or per capita income) and institutional factors (such as tax and expenditure limits, budget cycles, parties involved in revenue forecasting process, frequency of the forecast, whether the budget is bound by the forecast, the use of the university faculty in the forecast preparation, the presence of an economic advisory council) (Rose and Smith, 2011; Voorhees, 2004).

One institutional arrangement of great interest is consensus forecasting, but it turns out that it is quite difficult to define and measure. NASBO defines a consensus forecast as a "revenue projection developed in agreement through an official forecasting group representing both the executive and legislative branches" (NASBO 2008).

The definition used in the Government Performance Project survey for 2005 was:

> A process that requires a panel of experts brought together for purposes of generating the requested forecast. Experts may include officials from the executive and legislative branches of the state, as well as external researchers or officials from universities, private consultants or citizens.[18]

According to a study published by the Federation of Tax Administrators (FTA), "There are two types of consensus forecasting. In the first type, members of the consensus group are composed of professionals from the different agencies responsible for revenue forecasting (i.e., Budget Office, Revenue Department, or Legislative Agencies). Each agency brings its own forecast and assumptions to the table. By discussing its differences, the group comes to an agreement on one single forecast for the state.… The second type of consensus forecasting includes independent groups whose members are appointed by the executive and legislature. These groups typically turn to the resources of other agencies to make their economic and revenue forecast" (FTA 1993).

Before discussing the research on consensus forecasting, it is important to touch on a related area: combining separate forecasts (regardless of whether there is a formal consensus body or not).

This research goes well beyond revenue forecasting and pertains to forecasts in general, including business forecasts, economic forecasts, and other kinds of forecasts. It concludes emphatically that combining multiple forecasts leads to increased forecast accuracy. (For summaries of this research, see Batchelor and Dua 1988; Clemen 1989; Voorhees 2004; Weltman 1995-96).

The research on combining economic and business forecasts is convincing, and certainly suggests that consensus state revenue forecasting — or at least combining forecasts as part of consensus-like institutional arrangements — could increase accuracy. However, research bearing directly on institutional arrangements of consensus state revenue forecasting is less convincing, although several researchers have argued that the data support the hypothesis that consensus forecasting is more accurate than other institutional arrangements. The evidence for this so far is slim, although that does not necessarily mean the conclusion is wrong.

Some analysts assume that consensus forecasting diminishes forecast bias and increases forecast accuracy as it "takes the politics out of the revenue forecast accuracy" (Qiao 2008). But depending on how it is structured, consensus revenue forecasting may not remove politics from the process, it faces its own challenges such as being an extremely time-consuming process, or involving genuine differences among the members of forecasting group (Qiao 2008). The time-consuming nature of consensus forecasting can increase lags between forecast preparation and forecast use, in turn potentially reducing accuracy, a topic we discuss later in this report.

Qiao (2008) used data from the Government Performance Project[19] and analyzed revenue forecast errors for each state for fiscal years 2002-04, using 5 percent as a threshold for forecast accuracy. The results indicated that eleven out of twenty-seven states that utilize consensus forecasting made errors less than 5 percent in each study year and only one state made errors greater than 5 percent in all the three years. Qiao argues that while the consensus forecasting process certainly contributes to revenue forecast accuracy, there are other important variables that are essential for improving revenue forecasting accuracy. Such variables include state revenue structures, the duration of time between the development of a revenue forecast and its implementation, forecast frequency, the use of economic advisors and university consultation, economic stability of the nation and the state, political pressures, and political ideology (Qiao 2008).

Voorhees (2004) examined the impact of consensus forecasting on state revenue estimating error. He argued that errors in state revenue forecasts are not solely caused by the fluctuations in the economy, but also are attributable to the state institutional structures and the degree of consensus required for the state forecast. Voorhees conducted a survey of forecasters in forty-four states and collected institutional and methodology data pertaining to state revenue forecasting for the years 1990-97. In addition, Voorhees relied on secondary data from NASBO's Fiscal Survey of States for the years 1989-98 for calculating the forecast errors. Voorhees classified forecasting processes in the states into one of the following six categories: (1) single executive agency, (2) single legislative agency, (3) two separate forecasts (executive and legislative), (4) consensual multiple executive agencies, (5) consensual executive and legislative agencies, and (6) consensual independent conference. Voorhees used the locus of forecast formulation and the degree of diversity in participation at that locus as a proxy for consensus forecasting. He concluded that an increase of consensus in the formulation of the forecast leads to decrease in forecast errors. More specifically, "for each unit increase in the consensus variable, the forecast error is expected to decrease by approximately 0.1075 and 0.1084 standard errors respectively" (Voorhees 2004).

In a recent study, Krause and Douglas (2013) examined the relationship between the organizational structure of the consensus groups (both in terms of the size and the diversity of the consensus groups) and revenue forecasting accuracy. The authors used data for fiscal years 1987 to 2008 across twenty-five consensus group commissions. The central theme of their study was to empirically test the following panel size-diversity trade-off hypothesis: "Increases (Decreases) in CG [consensus group] commission diversity will improve revenue-forecast accuracy when CG commission size is decreasing (increasing)" (Krause and Douglas 2013). The authors operationalize the organizational diversity based on the presence of the following three types of players: appointee (partisan, nonpartisan); institutional (legislature, governor, independent elected agency heads); and partisan (Democrat, Republican, third party). Based on empirical results the authors argue that organizational diversity in smaller consensus groups (with three or four members) yields a marginal improvement in revenue forecasting accuracy. However, organizational diversity yields an increasingly adverse marginal impact on revenue forecast accuracy for larger consensus groups. "While organizational diversity has salutary benefits for collective decision making, it can reduce the accuracy of collective decisions in larger panels since coordination of heterogeneous expertise becomes problematic. This is because the marginal coordination losses from inefficient pooling of expertise within the group will exceed its marginal information gains regarding accurate collective decisions" (Krause and Douglas 2013).

There are quite a few lessons from this literature review about the impact of institutional arrangements on revenue forecasting:

- There is overwhelming evidence that combining forecasts, even in simple ways, tends to reduce the magnitudes of errors on average, and reduce instances of extreme error. This may be one of the reasons that some studies have found consensus forecasting of state tax revenue to be more effective than nonconsensus methods.

- Nonetheless, the evidence supporting consensus forecasting of state revenue is decidedly less overwhelming than that in favor of combining forecasts (with combining being a more mechanical task).
- In evaluating consensus forecasting, it is important to realize that there are important differences in how consensus forecasting is defined, and important differences across states in how consensus forecasting is implemented.

Observations From the Survey

There is wide variation across states in terms of budgeting processes and rules and regulations governing revenue forecasting. Some of these forecasts are prepared by a single entity alone, such as either executive or legislative branch or both independently, and some are a joint effort of both executive and legislature. Some are prepared by an independent formal group generally formed by a mix of executive, legislature, university economists, and experts from private or public sectors. It is generally believed that a depoliticized forecast is more objective.

In the survey by the Rockefeller Institute, state forecasting officials were asked to indicate who is involved in the revenue estimating process. The state forecasting officials were specifically asked to specify which agencies are involved from the executive and legislative branches as well as whether there are other parties involved such as academicians, experts from the private sector, or any other parties. According to the survey responses, in most states one or more agencies of the executive branch are involved in the revenue estimating process. The executive branch does not take any role in producing revenue estimates only in very few states including Hawaii, Montana, Nevada, and Washington.

In Hawaii, the revenue estimates are prepared by the Council on Revenues, which is attached to the Department of Taxation for administrative purposes. The Council on Revenues consists of seven members, three of whom are appointed by the governor for four-year terms and two each of whom are appointed by the president of the Senate and the speaker of the House of Representatives for two-year terms.

In Montana, the preparation of the revenue estimates is under the jurisdiction of the legislative branch and is carried out by the Legislative Fiscal Division. The Legislative Fiscal Division prepares and presents the revenue estimates to the Revenue and Transportation Interim Committee (RTIC). The RTIC, which is a twelve-member joint bipartisan committee of the legislature, uses the information provided by the Legislative Fiscal Division to estimate the amount of revenue that the state will receive during the next biennium.

In Nevada, the revenue forecast is provided by the state's Economic Forum, which is a panel of five economic and taxation experts from the private sector, all appointed by the governor. All agencies of the state, including the governor and Nevada legislature, are required to use the Forum's forecast, which is provided shortly before the beginning of a new legislative session.

In Washington, revenue estimates are provided by the Economic and Revenue Forecast Council, which is an independent agency and is not a part of the executive or legislative branches. The Council was created in 1984 and consists of two members appointed by the governor and four members from the legislature.

There is some discrepancy in the answers provided by the officials from the executive versus legislative branch for some states in terms of the role of other parties, such as the legislative branch, academicians, and private sector experts, in the revenues forecasting process. Thirty of the forty-five respondents from the executive branch and thirty-three of the forty-six respondents from the legislative branch indicated that the legislative branch is involved in the revenue estimating process in their respective states. According to survey responses, only a few states involve academicians and private sector experts in the revenue estimating process.

State officials were also asked to indicate if there is a formal group or organization tasked with developing, or signing off, on revenue estimates. Once again, there is some discrepancy in the responses provided by the officials from the executive branch versus by the officials from the legislative branch. Fifteen of forty-six respondents from the legislative branch indicated that there is no formal group and only nine respondents from the executive branch indicated the absence of any formal group for revenue estimates. In most states where officials indicated existence of a formal group tasked with the preparation of revenue estimates, they also indicated that those groups hold formal meetings throughout the year. Again, wide variation exists among the states in terms of frequency of the meetings by the groups/organizations tasked with the revenues estimation. For example, in some states such as Alaska, Mississippi, Nevada, North Dakota, or Virginia, the group meets only once or twice in the given fiscal year. In other states, such as Washington, the group meets more frequently, eight times per fiscal year.

According to the survey responses, in approximately two-thirds of the states the participants in the revenue estimating process come to an agreement on a single number. Moreover, in most of these states, both the governor and the legislature are required to use the agreed upon estimates in the adopted budget.

Due to ambiguity of the definition, it is hard to identify the exact number of states that have a consensus forecasting process. For example, in some states there is no official consensus forecasting group, but there is an informal process for reaching agreement between the legislature and executive.

The survey results indicate that states use various institutional arrangements for revenue forecasting. According to survey responses, there is a formal group tasked with the preparation of revenue estimates in at least two-thirds of the states. Moreover, such groups hold formal meetings throughout the year. In some states, the revenue estimating groups have more clear and routine responsibilities and processes in place. For example, a state official from Connecticut indicated: "The consensus process occurs three times per year; in November, January, and April. The January estimate binds the Governor's budget proposal and the April estimate binds the Legislature's adopted budget." On the other hand, the consensus revenue estimating process is more informal process in North Carolina. "The consensus revenue forecast is an informal process that does not require either the executive or legislative branch to use the revenue estimate," said a state official from North Carolina.

In very few states, not only there is no consensus revenue estimating process in place, but also revenue estimating is under the jurisdiction of a sole entity. For example, in New Jersey, official revenue estimates are prepared by the executive branch only and there are no formal processes in place.

Descriptive Analysis

We incorporated data on consensus forecasting and other forecasting mechanisms into our analysis, relying primarily on data collected by Rose and Smith (2011). Table 11 shows the median absolute percentage error under each of several forecasting arrangements. It is apparent that consensus forecasting, as measured by Rose and Smith, was not associated with more accurate forecasts (absent controls for forecast difficulty). Both expert forecasts and separate legislative forecasts were associated with less accurate forecasts, although the difference may not be significant.

Table 11. Summary Data, Without Controlling for Forecast Difficulty, Do Not Suggest Consensus Forecasts Are More Accurate

Median absolute percentage forecast error under different arrangements to forecasting process
Sum of PIT, sales taxes, and CIT, 1987-2007 by type of forecast

	States with this process	States with a different process
Consensus forecast	3.4	3.5
Executive forecast	3.4	3.5
Expert forecast	3.7	3.3
Separate legislative forecast	3.7	3.4

Sources: Rockefeller Institute analysis of NASBO forecast error data, and data from Rose, S., & Smith, D. L. (2011). Budget Slack, Institutions, and Transparency. *Public Administration Review, 72* (2), 187-195.

We also examined the potential impact of these variables econometrically, repeating our earlier models that controlled for forecast difficulty. In none of the specifications did the variables indicating the type of institutional arrangement have a statistically significant association with forecast accuracy at the 0.05 significance level. We repeated the analysis with the forecasting arrangement variables we collected in the survey, and none of them had a statistically significant association with accuracy either.

Conclusions and Policy Recommendations Regarding Forecasting Accuracy and Institutional Arrangements

Our reading of the literature is that there is very little relationship between consensus forecasting and forecast accuracy. That is consistent with our examination of these data, also. However, as we have noted before, the evidence in favor of examining and combining forecasts is overwhelming: Combining forecasts tends to lead to lower errors. Processes that encourage this to happen may lead to lower errors.

Beyond that, it is good practice to try to insulate forecasting from the political process and consensus forecasting can help achieve that.

Appendix A: Survey of State Officials

The Rockefeller Institute of Government conducted two rounds of surveys. Both surveys were conducted online, using Google Docs.

The first round of the survey was conducted in July-August of 2013. The purpose of the first round of the survey was to ensure that we understand as precisely as possible the information that state budget officials provide to NASBO on revenue forecasts and preliminary actuals, as published in the *Fiscal Survey of States*. The survey questions were designed to help us get a better understanding of when estimates that underlie state budgets are developed, as that influences the difficulty of the forecasting job, and to be sure we understand how the forecasts are used in the budget process. For the first round of survey, we sent out the survey questionnaire via online survey tool to state budget officers in all fifty states. We developed two survey instruments: one for the states with an annual budget cycle and another one for the states with a biennial budget cycle. For the biennial states, we asked two additional questions to understand what estimates they provide to NASBO for the second year of the biennium. We received thirty-four usable responses from the states, of which twenty-three responding states were on annual budget cycles, while eleven states were on biennial budget cycles.

The second round of the study was conducted in the months of August and September of 2013, with a few follow-ups in October 2013. The purpose of the second round of the survey was to collect additional information from all the states to get further information related to procedures and practices of revenue forecasting process in the states. We sent out the survey questionnaire to state budget officers representing both executive and legislative branches of government in all fifty states. We received fort-five responses from the state officials representing the executive branch and forty-six responses from state officials representing the legislative branch. On the executive side, we did not receive responses from Georgia, Missouri, Montana, North Carolina, and Oregon and on the legislative side no response were received from Massachusetts, Michigan, New Hampshire, and South Carolina.

The survey questions asked during the first round of the survey administration were targeted on getting better understanding of what is being reported to NASBO. The most important questions asked of the states were focused on whether the revenue estimates being reported to NASBO are based on the final estimates for the adopted budget, whether the revenue estimates take into account legislative and administrative changes, approximately how many weeks ahead of the start of the new state fiscal year are the estimates in the adopted budget prepared, and whether the biennial states report estimates for the first and second years of biennium separately or combined.

Among the respondent states, thirty-one states indicated that the estimates provided to NASBO are the final estimates used for the adopted budget. Alabama, Arizona, and Idaho were the only three states indicating that the revenue estimates provided to NASBO are not the final estimates used for the adopted budget. Twenty-six states also indicated that the revenue estimates bind the budget. In most of the twenty-six states where the revenue estimates are binding the budget, such binding is defined either in the constitution or by statute. And only in a handful of states are the revenue estimates binding based on customary practice.

We also asked the states with the biennium budget cycle whether they submit updated estimates to NASBO for the second year of the biennium. Ten of the eleven responding states with the biennial budget cycle indicated that they do provide updated estimates to NASBO for the second year of the biennium in order to account for legislative changes and updated economic forecast conditions.

The main purpose of the second round of the survey was to get a better understanding about general procedures and practices of revenue estimating processes in the states. The survey questions were targeted into getting more information about the timing of forecasts, frequency of forecast updates, the long-term forecasting practices, parties involved in the forecasting processes, and whether there is a formal group that plays a major role in the revenue estimating process.

Officials in at least half of the states indicated that they typically generate revenue estimates for one or more years beyond the upcoming budget period. Overall, most states release revenue estimates for another two to five years. Only in very few states are the forecasts provided for more than five years beyond the budget cycle. One example is Alaska, where officials in the legislative branch indicated generating revenue estimates for six years while officials in the executive branch said that the revenue estimates are provided for nine additional years beyond the budget cycle.

Endnotes

1 *States' Revenue Estimating: Cracks in the Crystal Ball* (Philadelphia, PA, and Albany, NY: Pew Charitable Trusts and Rockefeller Institute of Government, March 2011), http://www.rockinst.org/pdf/government_finance/2011-03-01-States_Revenue_Estimating_Report.pdf.

2 Before the fall of 2013, NASBO conducted the *Fiscal Survey of States* jointly with the National Governors Association.

3 These observations are not independent of those for the individual taxes, of course.

4 For more information, please see Appendix A: Survey of State Officials, at the end of this report.

5 Richard Mattoon and Leslie McGranahan, *Revenue Bubbles and Structural Deficits: What's a state to do?,* Working Paper (Chicago, IL: Federal Reserve Bank of Chicago, April 2012), https://www.chicagofed.org/digital_assets/publications/working_papers/2008/wp2008_15.pdf.

6 The upper whisker extends from the hinge (the 75th percentile) to the highest value that is within 1.5 times the interquartile range, and the lower whisker extends similarly downward from the lower hinge. Data beyond the end of the whiskers are outliers and plotted as points.

7 This test involves regressing the forecast on the actual and testing the joint hypothesis that the intercept is zero and the slope is one. See Jacob A. Mincer and Victor Zarnowitz, "The Evaluation of Economic Forecasts," in *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance* (New York, NY: National Bureau of Economic Research, Distributed by Columbia University Press, 1969): 1-46, http://www.nber.org/chapters/c1214.

8 The table shows forty-two states with a personal income tax. There are only forty-one states with a broad-based personal income tax. Tennessee, which has a narrow tax on interest and dividends, is included in the NASBO data. This tax is a relatively small share of total Tennessee tax revenue.

9 This is similar to an approach developed contemporaneously in a working paper by John Mikesell and Justin Ross (see Bibliography) regarding forecasting in Indiana.

10 Our naïve forecasting model ends in 2013 because we do not yet have full data on actual tax revenue for 2014.

11 Synchronization was calculated by using a monthly coincident index at the state level, relying on: (a) nonfarm payroll employment; (b) average hours worked in manufacturing; (c) the unemployment rate; and (d) wage and salary disbursements. Given the idiosyncrasies and data limitations between states, these measures are not a perfect barometer for each state's economy, but are a good start. Perhaps further measures could be incorporated, for example, measures of tourism and travel, as Hawaii was found to synchronize with the national economy only 57.4 percent of the time.

12 While recessions typically include one or more quarters in which real GDP declines, these are annual data, and GDP declines on an annual basis are less common.

13 For example, analysts often use this to measure volatility of daily stock prices.

14 This is a different measure of volatility than used by Crain (see Bibliography), discussed previously. For technical reasons, we believe this is a better measure of volatility of tax revenue than his measure, which examines deviations from a trend line.

15 The tax data have not been adjusted for legislative changes because it is impractical to do so. A small portion of observed changes in state tax volatility may be attributable to legislative changes. We do not believe this would affect conclusions materially.

16 For several years before the 2007 recession, a strand of economic research identified what came to be known as the "Great Moderation" — the idea that the United States economy was becoming less volatile — and examined its potential causes. While the data clearly supported declining economic volatility over longer periods than shown in this figure, the idea that we have entered a new period of stability is now in tatters.

17 The data on wages and capital gains are from the Statistics of Income branch of the Internal Revenue Service, and were provided to us by Leslie McGranahan of the Chicago Federal Reserve Bank.

18 States are listed in the spreadsheet "GPP_Revenue Estimate Methods FY 2004.xls"worksheet "Measurement Methods 2005."

19 See "Archived Project: Government Performance Project," Pew Charitable Trusts, http://www.pewtrusts.org/en/archived-projects/government-performance-project.

Bibliography

Batchelor, Roy and Pam Dua. "Forecaster Ideology, Forecasting Technique, and the Accuracy of Economic Forecasts." Eighth International Symposium on Forecasting. Amsterdam, 1988.

Bretschneider, Stuart I., Wilpen L. Gorr, Gloria Grizzle, and Earle Klay. (1989). "Political and organizational influences on the accuracy of forecasting state government revenues." *International Journal of Forecasting* 5, 3 (1989): 307-19.

Clemen, Robert T. "Combining forecasts: A review and annotated bibliography." *International Journal of Forecasting* 5, 4 (1989): 559-83.

Cornia, Gary C. and Ray D. Nelson. "State Tax Revenue Growth and Volatility." *Regional Economic Development* 6, 1 (2010): 23-58.

Crain, William Mark. "Reliability of Revenues from Alternative Tax Instruments." In *Volatile States: Institutions, Policy, and the Performance of American State Economies.* Ann Arbor, MI: University of Michigan Press, 2003: 73-81.

Deschamps, Elaine. "The impact of institutional change on forecast accuracy: A case study of budget forecasting in Washington State." *International Journal of Forecasting* 20, 4 (2004): 647– 57.

Elder, Erick M. and Gary A. Wagner. "Revenue Cycles and Risk-Sharing in Local Governments: An Analysis of State Rainy Day Funds." *National Tax Journal* 66, 4 (2013): 939-54.

Felix, R. Alison. "The Growth and Volatility of State Tax Revenue Sources in the Tenth District." *Economic Review* 93, 3 (2008): 63-88.

FTA. *State Revenue Forecasting and Estimation Practices*. Research Report No. 139. Washington, DC: Federation of Tax Administrators, March 1993.

Krause, George A. and James W. Douglas. "Organizational Structure and the Optimal Design of Policymaking Panels: Evidence from Consensus Group Commissions' Revenue Forecasts in the American States." *American Journal of Political Science* 57, 1 (2013): 135-49.

Mattoon, Rick. *Creating a National State Rainy Day Fund: A Modest Proposal to Improve State Fiscal Performance*. Working Paper No. 2003-20. Chicago, IL: Federal Reserve Bank of Chicago, 2003.

Mattoon, Rick and Leslie McGranahan. *Revenue Bubbles and Structural Deficits: What's a State to Do?* Working Paper 2008-15. Chicago, IL: Federal Reserve Bank of Chicago, 2008.

Mikesell, John L. and Justin M. Ross. "State Revenue Forecasts and Political Acceptance: The Value of Consensus Forecasting in the Budget Process." *Public Administration Review* 74, 2 (2014): 188-203.

NASBO. *Budget Processes in the States*. Washington, DC: National Association of State Budget Officers, 2008.

National Advisory Council on State and Local Budgeting, GFOA. *Recommended Budget Practices: A Framework for Improved State and Local Government Budgeting*. Chicago, IL: Government Finance Officers Association (GFOA), 1998.

Penner, Rudolph G. "Federal Revenue Forecasting." In *Government Budget Forecasting: Theory and Practice*, ed. Jinping Sun and Thomas D. Lynch, 11-25. Boca Raton, FL: Auerbach Publications, 2008.

Pew Center on the States and Rockefeller Institute of Government. *States' Revenue Estimating: Cracks in the Crystal Ball*. Philadelphia, PA, and Albany, NY: Pew Charitable Trusts and Rockefeller Institute of Government, March 2011.

Qiao, Yuhua. "Use of Consensus Revenue Forecasting in U.S. State Government." In *Government Budget Forecasting: Theory and Practice*, ed. Jinping Sun and Thomas D. Lynch, 393-413. Boca Raton, FL: Auerbach Publications, 2008.

Rose, Shanna and Daniel L. Smith. "Budget Slack, Institutions, and Transparency." *Public Administration Review* 72, 2 (2011): 187-195.

Samuelson, Paul "Quotes from Paul Samuelson's February 1985 lecture at Trinity University, February 6, 1985, http://www.trinity.edu/nobel/Samuelson_files/Samuelson%20web%20quotes.htm.

Thompson, Fred and Bruce L. Gates. "Betting on the Future with a Cloudy Crystal Ball? How Financial Theory Can Improve Revenue Forecasting and Budgets in the States." *Public Administration Review* 67, 5 (2007): 825-36.

Voorhees, William R. "More Is Better: Consensual Forecasting and State Revenue Forecast Error." *International Journal of Public Administration* 27, 8-9 (2004): 651-71.

Weltman, Jeremy C. "Using Consensus Forecasts in Business Planning." *The Journal of Business Forecasting* 14, 4 (Winter 1995-96): 13-6.

Willoughby, Katherine G. and Hai Guo. "The State of the Art: Revenue Forecasting in U.S. State Governments." In *Government Budget Forecasting: Theory and Practice*, ed. Jinping Sun and Thomas D. Lynch, 393-413. Boca Raton, FL: Auerbach Publications, 2008.