Survey Article

The Status of Budget Forecasting

Daniel W. Williams – Baruch College
Thad D. Calabrese – New York University

This article examines the breadth of the current forecast literature as it relates to public budget making. It serves to provide summary information to decision-makers who otherwise do not have the resources to learn more than a small amount focused on much more narrowly defined areas of forecasting (such as the politics of forecast bias). Next, it serves those who perform forecasting related to budgeting by reviewing the current methods and practices commonly used in this domain. It also provides a ground level for future public budget forecasting research. Finally, this article identifies several areas in which the public forecasting literature needs additional development. Several of these areas, such as the effectiveness of nonregression-based forecasting techniques, are quite important to the majority of governments in the United States and other subnational jurisdictions, where budget offices are limited and resource investments in technology are scarce.

Keywords: Budget Forecast, Revenue, Expenditure, Balanced Budget,

Consider the forecasting practices of New York City: In May 2014, the New York City Independent Budget Office (IBO) questioned the mayor’s revenue projections, suggesting that the surplus for 2015 would be $1.8 billion, over $100 million more than the mayor’s proposed budget (Independent Budget Office, 2014; Katz, 2014). According to the New York State Comptroller, the actual surplus was $3 billion, almost twice the amount suggested by the IBO (Office of the New York Comptroller, 2015). In December 2015, the IBO predicted an $900 million surplus for fiscal year 2016, $800 million more than predicted by the mayor’s office (Durkin, 2015). In May, the state comptroller reported an expected surplus of $3.4 billion (Office of the New York Comptroller, 2016). These recent news items reflect the continuation of a decades-long practice: New York City recurrently underestimating revenues and, until recently, anticipating nonexistent shortfalls (D. W. Williams, 2012; D. W. Williams & Onochie, 2013).

While the magnitude of this uncertainty or bias is greater in New York City than in other local jurisdictions, this article reviews evidence that such forecasting practice is actually quite common. Such practice influences budgetary decisions, which can restrict or liberate policy making. D. W. Williams and Onochie (2013) show that, when funds are found after the year begins as a result of underforecasting during the budget process, decision-making authority may be shifted from legislative bodies to executives. Levine, Rubin, and Wolohojian (1981) identify similar shifting with respect to related revenue practices. Shifting decision authority is just one of many ways in which forecasting is as much a political function as it is a technical one.

This article examines the breadth of forecast literature as it relates to public budget making.¹ It can serve to provide information to decision makers who otherwise do not have the resources to learn more than a small amount focused on much more narrowly defined areas of forecasting (such as the politics of forecast bias). Next, it serves those who perform forecasting related to budgeting by reviewing the current methods and practices commonly used in this domain. It

¹ Because of the unique nature of capital budgeting, forecast-like practices for capital budgeting are not discussed.

also provides a ground level for future public budget forecasting research. Finally, this article provides suggestions for future research. Because the methodological side of forecasting is replete with technical terms, there is an appendix of definitions, which include “forecast” itself.

Forecasts are needed to enable planning. From the first decade of the twentieth century in the United States and earlier in Europe, budgeting has been a means of adding a planning stage to appropriating. Here we discuss the current state of forecasting as it specifically relates to public budgeting. While a substantial number of publications address budget-related forecasting, they are found in a wide array of journals and disciplines. Our goal is to summarize this literature in one place.

Commonly, budget forecasting is treated as synonymous with revenue forecasting (for subnational jurisdictions) or budget balance forecasting (for nations), which is conducted primarily for the budget year with an eye toward also predicting subsequent years. While the article addresses these topics, it also addresses numerous other matters that clarify budget-related forecasting and identify the current state of the practice. Specifically, it examines forecast bias resulting in systematic errors, the use of forecasting in expenditure planning, techniques and practices, and the risk of dynamic forecasting.

The following sections examine: Forecasting for budgeting; state and local government revenue and expenditure forecasts; national budget forecasts; forecast techniques; forecast practices; forecasting, predicting, estimating and dynamic estimation; and opportunities for future research. The article summarizes research and discusses topics that do not frequently appear in the literature.

**Forecasting for Budgeting**

It is useful to think about how forecasting is related to the budget process. It has different functions within three distinct budget periods. The most immediate period is the appropriated period, which is the remainder of the current fiscal year. The second period is the budget year, which begins the day after the current year ends. The third period begins the day after the budget year ends and is labeled “out years” here.

**The Current Year**

For the current year, the purpose of forecasting is to support tracking of revenue and spending during the fiscal year. Actual revenues and expenditures are compared with their budgeted values in the form of variance reports, which are typically then used by analysts to examine the causes of significant deviations from appropriations based on prior forecasts. As variance analyses are performed midyear, governments need predictions of how much revenue or expenditure to expect within the remaining part of the year. Because many revenue and expenditure lines are seasonal, the remaining part of the year cannot be treated as a simple straight line. D. W. Williams (2008) shows how forecast confidence intervals can be used to achieve precision for such tracking. However, any method that provides within-year periodic updates accounting for seasonality as appropriate can be used to estimate whether revenue and expenditures at year-end will match expectations.

---

2 Here the word “forecast” is used to refer to using some forecast method or practice, not simply preparing the financial component of a budget request. Further definition is in the appendix.

3 As most budgeting is for an annual period, we refer to fiscal periods as years.
Current year forecasts are also used for cash management, the choice among investing, holding, cashing out, and borrowing to pay for current expenses. Techniques for cash management can be relatively sophisticated (Stone & Wood, 1977; M. Williams, 2013) or fairly simple (Chen, Weikart, & Williams, 2015). For these purposes, the forecast needs to be sufficiently granular to be beneficial. During most of the fiscal year, this likely means updating the forecast with monthly data for revenue tracking and possibly weekly or daily for cash management. Near the end of the fiscal year, forecasts may need to be updated weekly or daily for both purposes.

The Budget Year

In budget making, forecasts are made for the upcoming fiscal period. Forecasts are used to predict resource constraints. Unlike forecasts for the current year, budget year forecasts only need to address the entire fiscal year (D. W. Williams & Kavanagh, 2016). For budget constraints, forecasts include predicting the availability of revenue from various taxes, fees, and transfers from superordinate governments and for predicting extrinsic sources that drive expenditures, such as school populations, health care users, or jail and prison inmates. For these forecasts, the primary objective is accuracy.

For subnational jurisdictions in the United States and a mixed set of jurisdictions elsewhere, there is a secondary objective of risk reduction, which may be labeled “prudence” or “asymmetric loss function.” Risk reduction means selecting a forecast that has a higher probability of favorable error than unfavorable error. For revenue this means underestimating the revenue; for expenditures this may mean overestimating expenses. However, for expenditures, appropriations are generally distributed to spending agencies, so overestimation can lead to overfunding of these agencies and creating unintended discretion. Consequently, the motive for overestimation of expenditures may be much weaker than the well-established tendency to underestimate revenue. While forecasts with granular data may perform better than annualized data, the forecast for the budget year need only be accurate for the entire year.

The Out Years

For the period subsequent to the budget year, the forecast predicts structural balance or structural imbalance, sometimes labeled structural deficit. A structural balance occurs when for the length of the forecast the revenue is adequate to meet the currently forecast expenditures, assuming that all obligations are being fully met with recurring revenue. Excess future revenue is generally treated as acceptable, as it allows room for either tax reductions or policy options. A variety of conditions can hide imbalance, such as (1) using debt or nonrecurring revenue to meet recurring obligations; (2) underforecasting the current share of future obligations, typically retirement commitments; (3) overforecasting future revenue; or (4) underforecasting future expenditures. No literature addressing the possible use of forecasts to hide structural imbalance for subnational governments has been identified. When risk-reducing forecasts of the budget period are extended into future periods, they can create a false belief in structural deficit (D. W. Williams, 2012). This misperception may be accidental, or it may more likely serve the strategic purpose of suppressing policymaking that leads to long-term expenditure commitments. For the US federal budget, there is evidence of optimistic bias over the out years (Kamlet, Mowery, & Su, 1987), which may mask structural deficits or create an appearance of available tax or policy options.

---

4 The article describes the policymaking uses in the section “Forecasting, Estimating, Predicting, and Dynamic Estimation” below.
State and Local Government Revenue and Expenditure Forecasts

This section examines research that has looked at the revenue side on its own and then examines research that focuses on the expenditure side of the budget on its own. This discussion is focused primarily on state and local governments within the United States and also includes material from other subnational governments as well as material that is related to national practices.

Revenue Forecasting

Much of the extant relevant literature focuses on the revenue side of forecasting. The literature is largely in agreement that forecast errors are not simply the result of technical shortcomings in how forecasts are established but also reflect political decisions, as discussed below. Beginning with Burkhead (1956), most of this literature finds that American state and local governments engage in underestimation bias. Hou (2006) demonstrates that the average revenue error rate is positive, meaning actual revenues exceed forecasts on average across the 50 states. D. W. Williams (2012) similarly shows that revenues as of year-end are above forecast for New York City from 2001–2011.

The practice of revenue underestimation serves as a rational hedge against future revenue uncertainty (Bretschneider & Gorr, 1992; Bretschneider, Gorr, Grizzle, & Klay, 1989; Rodgers & Joyce, 1996). Local public finance managers and elected officials use conservative revenue forecasting as a budget constraint. This practice has the effects of limiting expenditure growth and generating incidental reserves (Frank, 1993; Kelly, 2013). Frank and Zhao (2009), in fact, define the revenue constraint as the key political factor in the budget process and find that approximately 90% of surveyed cities underestimate revenues by 1% to 7% annually. Frank and McCollough (1992) identify conservative revenue forecasting as a means to constrain expenditure growth from political pressure to increase particular spending categories. Tyer (1993) notes that conservative revenue forecasting is also one strategy employed to accumulate fund balances and other reserves. Only about one-half of the states have statutes or constitutions that legally bind budgets by revenue forecasts Morozov (2013), yet even the optics of forecasting expenditures in excess of revenues may be enough to limit spending.

Revenue underestimation bias may lead to year-end budget surpluses, meaning this bias can result in de facto stabilization funds or funds for other uses (Anessi-Pessina, Sicilia, & Stecolini, 2012; Dougherty, Klase, & Song, 2003; D. W. Williams & Onochie, 2013). Hou (2003) and Marlowe (2005) find that states and localities use funds accumulated during earlier fiscal periods to address fiscal stress, which can result from structural, managerial, or cyclical sources (Hou, 2006). Hence, revenue underestimation bias can be one of the ways by which decision makers increase savings during good economic times, so that, during lean times (whether self-inflicted or external), expenditure reductions or tax increases are not necessarily required. Alternatively, legislators might use such surpluses to fund tax cuts or to add programs. When these surpluses are partly resultant from cyclical upturns, such decisions become difficult to sustain over time (Nelson A. Rockefeller Institute of Government & Pew Center on the States, 2011). Hou (2006) notes that multiyear budgeting rather than single-year budgeting might better address revenue and expenditure swings; D. W. Williams (2012), however, finds evidence that longer forecast are associated with severe underestimation of revenue.

Beyond the budget stabilization function, revenue underestimation can fund within-year budget changes, which may serve political or managerial purposes. Forrester and Mullins (1992) note that jurisdictions frequently re-budget money during the year. Dougherty et al. (2003) and D.
W. Williams and Onochie (2013) find evidence that municipalities do in fact use revenue underestimation for such purposes and are able to use revenue underestimation to generate a budget stabilization fund even when explicitly prohibited (by statute, constitution, or process). Rodgers and Joyce (1996) also note that conservative revenue forecasts reduce pork barrel politics because these lower revenue estimates reduce discretionary funds available for ingratiating politicians with constituents. Choate and Thompson (1988, 1990) hypothesize that the source of conservative revenue estimation derives from the political decision maker rather than the technical forecaster. While their work is consistent with other analyses that find conservative revenue forecasting in governments, the authors argue that the goal of this behavior is not risk aversion but rather tax minimization.

Related to this literature on revenue forecast bias, others examine the source of the forecast to explain revenue forecasting behavior. Bland (2007) notes that some forecasters are “revenue conservers,” that is—those forecasters who are biased toward more pessimistic forecasts. This might be expected in executive budget offices. Such bias not only serves as a hedge but also maintains a lower target for public agencies as they prepare budget requests. Forecasters from legislative and agency budget offices, therefore, may estimate less biased (that is, more accurate) revenue forecasts because their goal is to fund expenditures (Bretscher Schneider, Straussman, & Mullins, 1988). On the other hand, Krause and Douglas (2006) find evidence of herding behavior between various forecasters, in which forecast differences are minimized between parties.

While the typical state and local practice in the United States is underforecasting, Rubin (1987) notes that accurate revenue forecasts or even overforecasting revenues might be a sign of fiscal stress because these estimated revenues are needed to cover immediate spending. Regardless of the direction of a systematic error, the literature consistently finds evidence of political motivation for these biases. Somewhat relatedly, there is substantial overforecasting at the onset of cyclical downturns. A report by the Nelson A. Rockefeller Institute of Government and the Pew Center on the States (2011) finds that states tended to overestimate revenues for one to two fiscal years following the recession, which began in 2007–2008. More than 70% of states overestimated their revenue in fiscal year 2009, compared with just 45% in the prior recession (2001–2003). These errors may not reflect deliberate political decisions.

There is literature that shows that some subnational jurisdictions in other countries behave similarly to US state and local governments. Imbeau and Tellier (2012) detail the literature on conservative revenue forecasting by Canadian provinces. Chatagny and Soguel (2012) find underforecasting of revenues by Swiss cantons from 1980–2002, which leads to reductions in actual expenditures, and Chatagny and Siliverstovs (2013) similarly find conservative revenue forecasts over a longer time period (1944–2010) but increasingly less so (that is, forecasting became more accurate over time). Czech municipalities similarly underforecast revenues (although smaller cities underforecast less), and longer budget processes lead to increased revenue underforecasting (Sedmihradská, 2013; Sedmihradská & Čabla, 2013; Sedmihradská & Klazar, 2011), which is consistent with the American context (see D. W. Williams, 2012). Benito, Guillamón, and Bastida (2015) find opportunistic behavior by Australian politicians, who overestimate revenues during election years. Anessi-Pessina et al. (2012) find revenue underforecasting leads to more re-budgeting in Italian municipalities as well. The international context reveals that forecast bias is not merely an American phenomenon. In fact, the literature shows that regional biases differ, which suggests they are deliberate.

Whether underforecasting or overforecasting, the frequent appearance of bias shows that revenue forecasting is not simply a technical activity. This literature shows that top decision-
maker preferences, whether managerial or political, influence the point estimates creating a systematic error.

Expenditure Forecasts

While there is relatively rich and consistent literature on revenue forecasting and bias, little exists on the expenditure side. Hou (2006) finds evidence of expenditure overestimation in all 50 states. In principle, accurate expenditure-related forecasts are desirable because expenditure authority is commonly appropriated to the diverse agencies of government. As a result, excess expenditure authority may create unintended discretion for agency heads. However, insufficient expenditure authority prevents agencies from accomplishing their assigned responsibilities. There can be conflicting interests concerning expenditure overforecasting, as the implicit discretion created is potentially desired by public managers or decision makers.

The lack of attention to expenditure forecasting may reveal a belief that expenditures are fundamentally a choice of government and therefore do not need forecasting. In this sense, government expenditures represent the willingness to supply public goods and services to meet demand. Further, because public budgets are almost entirely on the cash or modified accrual basis of accounting, delaying or deferring payments can alter annual expenditures. For example, governments may choose to defer contributions to pension funds during times of fiscal stress (that is, when revenues fail to materialize as expected) to bring expenditures in line with available resources.

While governments might have control over certain expenditures, they certainly cannot control all expenditures. Public schools require a minimum number of teachers; Medicaid must pay service providers; and employee health insurance is usually established through multiyear contract negotiations with municipal labor unions, are but a few examples. While public officials may be unable to control these expenditures, they do suggest that a government could forecast the underlying causes of expenditures.

Although the budget and public finance literature does not frequently address expenditure forecasting, literature within subject matter domains discusses forecasting of underlying factors that lead to expenditures. Astolfi, Lorenzoni, and Oderkirk (2012) review 25 models used to forecast health care expenditures in OECD countries. Barnett (1987) forecasts prison population using demographic and sentencing policy variables. Such information is useful for forecasting justice system expenditures. Similarly, Campolieti (2015) forecasts applications for a disability program in Canada, which would inform projected expenditures as well. Deschamps (2004) discusses the consensus forecasting procedure for Medicaid forecasts in Washington state. A significant driver of state and local expenditures is education. Ploughman, Darnton, and Heuser (1968) evaluate forecasting of school age children for capital planning purposes and also for drawing district boundaries. Johnstone (1974) notes the need to forecast for education spending; almost paradoxically, Johnstone (1974) finds that, as forecast models become more complex, they tend to perform less accurately. Ferland and Guénette (1990) notes that decision makers need not just the total number of school age children but also the types of children. These forecasts give public officials data to assess reorganizations and resource distribution.

Overall, the public budgeting literature leaves expenditure forecasting underdeveloped. While there are diverse articles found in a variety of domains, this lack of focus within public budgeting may result in unidentified risks within budget making or in a lack of coordination with established research findings found within other disciplines.
National Budget Forecasts

The study of national budget forecasts includes four overlapping components. The first component focuses on the forecasts of developing nations; the second focuses on forecasts related to the United States; the third focuses on forecasts within the European Union (EU); and the fourth focuses on budget and forecasting in the Organization for Economic Cooperation and Development (OECD) and other countries.

Developing Nations

Caiden and Wildavsky (1974) and Caiden (1980) long ago determined that poor countries and countries experiencing economic or fiscal distress engage in repetitive budgeting (also called continuous budgeting or re-budgeting), which typically means that budget plans, including forecasts, made before the beginning of the fiscal year are materially revised after the budget is approved. Consequently, this sort of budget may be insufficiently useful as a fiscal or accounting device. This literature has expanded over the years (Bird, 1982; Gollwitzer, 2011; LeLoup, Ferfila, & Herzog, 2000; Martinez-Varquez & Boex, 2001; Patto, 1975; Peterson, 1994; Schick, 1998; Sharkansky, 1984; Vanagunas, 1995) focusing on a variety of less-developed countries and providing mixed evidence that countries with distressed economies may overestimate their revenue or underestimate expenditures, that is, make optimistic forecasts.

This behavior may allow decision makers to promise a richer package of public benefits than is supported by their revenue. Rubin (1987) and Levine et al. (1981) find evidence of similar behavior with distressed local governments in the United States. Such behavior is consistent with the more political characteristic of forecasting, as will be found in the following sections, which show that many countries reveal similar motivations resulting in biases.

The United States

Generally, after the 1974 creation of the Congressional Budget office and primarily after the passage of the Balanced Budget and Emergency Deficit Control Act of 1985 (Gramm-Rudman-Hollings) researchers have been interested in the accuracy, efficiency, and possible biasedness of forecasts related to the federal budget. This interest may partly reflect a reaction to David Stockman’s cynical claim that supply-side economics were really a Trojan horse to achieve a trickle-down tax policy (Greider, 1981, 1982).

Federal budget forecasts are generally associated with macroeconomic data, particularly unemployment, inflation, and the change in gross domestic product (GDP). In addition, there are many private forecasts of these variables and of federal revenue, expenditure, deficit, and debt. At least six federal government entities make some or all of these forecasts:

1. The Council of Economic Advisors (CEA)
2. The Office of Management and Budget (OMB)
3. The Federal Reserve Board (FRB)
4. The Congressional Budget Office (CBO)
5. The Social Security Administration (SSA)
6. The Bureau of Economic Analysis (BEA)

---

5 In earlier periods, the gross national product (GNP) was the preferred variable. This list should not be taken as exhaustive.
Numerous studies of the economic, budget, and deficit forecasts are made by these organizations, sometimes including private forecasts (Auerbach, 1994, 1997; Belongia, 1988; Blackley & DeBoer, 1993; Booth, Timmerhoff, & Weiner, 2015; Campbell & Ghysels, 1995; Cohen & Follette, 2003; Corder, 2005; Ericsson, 2013; Frendreis & Tatalovich, 2000; Howard, 1987; Huntley & Miller, 2009; Kamlet et al., 1987; Kliesen & Thornton, 2001, 2012; Kowalewski & Edelberg, 2015; Krol, 2014; Lipford, 2001; Martinez, 2011, 2015; McNees, 1975, 1976, 1978, 1981, 1990, 1995; McNees & Ries, 1983; Penner, 2001; Plesko, 1988). These studies examine a variety of forecasted variables, such as budget balance, deficit or debt, revenue, outlays, and macroeconomic variables that are associated with these governmental variables. They examine:

1. Are the forecasts accurate, efficient, rational, or unbiased in the budget year? In out-years?
2. If biased, what is the bias?
3. Are some better than others?
4. Are they better or worse than private sector forecasts?

With the wide variety of variables and forecasters examined, there are only a few relatively consistent results. Most studies find that the budget year forecasts are relatively accurate. However, there is some evidence of optimism, particularly in the OMB forecast. Optimism can be defined as a forecast that leads to an underestimated deficit. This error may be associated with underestimation of unemployment or inflation or an overestimation of GDP growth. Underestimation of unemployment leads to overestimation of revenue and, simultaneously, leads to underestimation of expenditures.

For the budget year, these errors are typically small. However, out-year errors and out-year bias are larger, which is consistent with other revenue forecast research (D. W. Williams, 2012). For some but not all periods, an OMB forecast is more biased than CBO’s. In particular OMB’s out-year forecasts are significantly more biased than CBO’s. There is little difference in error between CBO and other federal agency forecasts. For variables in some periods, private forecasts may be marginally better than government forecasts, but typically these differences are small. Some OMB forecast errors may reflect a failure of Congress to adopt proposed presidential policies, or they may reflect other policy adjustments such as subsequent year changes in tax policy.

Overall, the studies suggest accurate and unbiased, or nearly so, forecasts for the budget year. However, there is rapid deterioration in forecast accuracy in the out-years. This deterioration should provoke users to question the reliability of assertions about the structural (multiyear) budget balance. While the pattern is not completely consistent, anti-tax political affiliation is sometimes associated with optimistic bias (for example, the Economic Growth and Tax Relief Reconciliation Act of 2001 championed by the Bush administration).

---

6 Martinez (2015, p. 19) summarizes many of these studies in Table 1. The CBO has produced 15 reports between 1999 and 2015.
7 The legislative portion of the federal budget begins in January and, when on time, ends by September for the budget year that begins in October. Forecasts for this budget are made in advance of the legislative process and are updated through the legislative process. While not all studies are clear on the exact timing, the results typically imply that forecasts made for the federal budget have relatively small errors for the budget year.
The 1992 Maastricht Treaty, which created the European Union, sets out objectives for national economic performance, including limits on deficits (3% of a nation’s GDP) and debt (60% of GDP). In 1997 these rules were strengthened through the Stability and Growth Pact (SGP). The SGP was strengthened in 1999 with preventive rules and in 2005 with corrective rules (European Commission, 2016). Since the advent of the SGP, there has been considerable concern about forecasting practices of EU member states.

The full scope of EU forecast practice research is immense. Here, the article briefly summarizes findings. There is general agreement that there has been an optimistic bias in forecasts of revenue or budget balance since the implementation of the SGP (Barberi, 2014; Beetsma, Bluhm, Giuliodori, & Wierts, 2013; Bluhm, 2009; J. Frankel, 2011; J. A. Frankel & Schreger, 2013a, 2013b; Giuriato, Cepparulo, & Barberi, 2016; Jonung, Larch, Favero, & Martin, 2006; Milesi-Ferretti & Moriyama, 2006; Moulin & Wierts, 2006; Rülke & Pierdzioch, 2014). This bias may be more pronounced during the run up to elections, reflecting the political business cycle. Optimistic bias may allow for the appearance of compliance with the Maastricht Treaty and the SGP during budget development, while end-of-year performance may no longer be in compliance. Use of practices such as an independent forecast entity may ameliorate bias. In Europe, optimism may be associated with liberal political affiliation and with the electoral cycle (election years). As with the United States, bias becomes more severe over longer horizons. Rülke and Pierdzioch (2014) suggest that this apparent bias reflects an asymmetric loss function (the penalty for error differs depending on the direction of error).

Jón Blöndal and co-authors have examined the budget practices of many countries (Blöndal, 2001a, 2001b, 2003a, 2003b, 2005, 2006, 2010; Blöndal & Bergvall, 2008; Blöndal, Bergvall, Hawkesworth, & Deighton-Smith, 2008; Blöndal & Curristine, 2004; Blöndal, Goretti, & Kirstensen, 2003; Blöndal, Hawkesworth, & Choi, 2009; Blöndal & Kim, 2006; Blöndal, Kraan, & Ruffner, 2003; Blöndal & Kirstensen, 2002; Blöndal, Kirstensen, & Ruffner, 2003; Blöndal & Ruffner, 2004; Blöndal, von Trapp, & Hammer, 2016). Countries examined by Blöndal can be found in all the categories previously discussed and a few are not members of the OECD. While these studies do not examine forecast effectiveness, they include descriptive information about the use of the forecasts in budget development. Blöndal et al. typically frame the forecast discussion as a brief review of economic assumptions included in the budget. In most they address either optimism or prudence (pessimism). Prudence may be achieved either of two ways: either within the forecast itself (bias or an asymmetric loss function) or through some overt form of reserves. Blöndal et al. describe prudence or underestimation in Australia, Netherlands, Sweeden, Canada, Indonesia, and Thailand. In contrast, the United States (inconsistently), the Philippines, and Brazil overestimate revenue. For Brazil, Blöndal et al. (2003, p. 112) say, “[I]n most cases these actions do not reflect the early adoption of unrealistic economic assumptions.” Finland and Austria are said to not use deliberate prudence; likewise, they are not reported to exhibit optimism. While independence and use of consensus forecasting are remarked on for a few countries, no clear pattern is identified.

In other research, Calitz, Siebrits, and Stuart (2013a, 2013b) show that, in South Africa, revenue forecasts are optimistic and recommend increased legislative oversight. As with other findings, Parkyn (2010) finds that, for 1995 through 2009, New Zealand overestimated revenue, with increasing error over longer horizons. Posner and Blöndal (2012) and Debrun and Kinda (2014) discuss the beneficial use of fiscal councils or other independent entities to improve forecast
accuracy and reduce bias. In similar work Krause and Corder (2007); Krause and Douglas (2005, 2006, 2013); and Krause, Lewis, and Douglas (2006, 2013) have examined the relationship between institutional designs and organizational structure to identify elements that may affect forecast accuracy and bias. Some of their findings are that organizations that produce competing forecasts may obtain similar results and be associated with less effective forecasts (Krause & Douglas, 2006) and that consensus group forecasting, in which forecasters representing different stakeholders or points of view (usually from the executive and legislative branches of government) are assembled to arrive at a joint forecast, may improve forecast accuracy (Krause & Douglas, 2013). This second finding is similar to those of Mikesell and Ross (2014).

Forecasting Techniques

The discussion in the previous sections has focused on forecast results. The next sections turn to how budget-related forecasts are made. This section discusses quantitative techniques. The next section discusses additional forecast practices including some qualitative techniques. Then, there is discussion of related matters involving estimation when it is not forecasting. Generally speaking, quantitative budget forecasters use either time series or causal/causal-like methods, the latter of which can be divided into simulation9 and econometrics. Within each class are techniques of varying degrees of complexity. This section also addresses decomposition, mixed approaches, and the use of intentions.

Time Series Methods

Data for which observations repeat periodically are frequently labeled time series. Time series are typically autoregressive, that is, two sequential observations will be correlated, so that the earlier observation contributes to predicting the next. Autocorrelation is the theoretical justification for the use of time-series methods, which can be either simple or complex. Simple time-series methods include moving average, simple exponential smoothing, Holt exponential smoothing, and damped trend exponential smoothing (see D. W. Williams and Kavanagh [2016] for a complete description of these methods and the formulas by which they are produced). Frank and Zhao (2009) suggest that most quantitative forecasting at the local government level is likely simple moving averages or trend analysis. Additionally, other ad hoc simple techniques used may include using the last period’s observed value for the next period’s value, an average of past data, the rate of change, the growth rate (expressed as a percentage), and time-index regression (D. W. Williams & Kavanagh, 2016). These ad hoc techniques may be appealing to forecasters with moderate sophistication because of ease in learning how to use these techniques; however, they are generally inaccurate and should be avoided (Armstrong, 2001a; D. W. Williams & Kavanagh, 2016). Further discussion of exponential smoothing methods can be found in Gardner (1985), Gardner (2006), De Gooijer and Hyndman (2006), and Hyndman, Koehler, Snyder, and Grose (2002).

8 This section uses a number of technical terms, which are defined in the Appendix. For those interested in the equations used for many of these techniques see D. W. Williams and Kavanagh (2016) for simpler methods or Makridakis, Wheelwright, and Hyndman (1998) for an extensive treatment.

9 We use “simulation” to refer to any approach that uses math to imitate real world processes. These can be deterministic, which are sometimes labeled algorithms, or they can involve statistical modelling, such as Monte Carlo simulations.
There are many complex time series methods. Some of the more common ones\(^{10}\) are autoregressive integrated moving average (ARIMA), which is sometimes labeled Box–Jenkins following their text (Box & Jenkins, 1970) and is intended to be a systematic way of selecting the optimal univariate time series model; X-11/X-12/X-13 (Findley, Monsell, Bell, Otto, & Chen, 1998; Monsell, 2007, 2009), which is used to optimally determine seasonal factors; Kalman filtering (Morrison & Pike, 1977), which provides time-variant parameter fitting; vector autoregressive techniques (Clements & Galvão, 2013; Sinclair, Stekler, & Carnow, 2015), a multivariate time-series method; empirical Bayesian techniques (Carriero, Clements, & Galvão, 2015; Miller & Williams, 2003), which typically correct for excessive variance; or neural networks (Voorhees, 2006), which are borrowed from neurology. Sometimes these methods are used in combination.

These complex methods generally perform well; however, they may require specific statistical software or the ability to implement complex mathematical procedures, knowledge on how to build proper models, and how to interpret output correctly for forecasting purposes. Thus, they are more appropriate for jurisdictions with a substantial budget for forecasting.

Makridakis et al. (1982) compared the accuracy of many of the forecasting techniques available at the time and concluded that simple methods outperformed complex. In fact, they found de-seasonalized simple exponential smoothing (SES) is the most accurate forecasting method available. Makridakis et al. (1993) and Makridakis and Hibon (2000) further test different quantitative methods against each other and find that increasingly sophisticated techniques do not universally improve forecast accuracy or errors. The most recent literature suggests damped trend, a modified form of Holt exponential smoothing, is likely the most accurate (Makridakis & Hibon, 2000; D. W. Williams & Kavanagh, 2016).

Causal and Causal-Like Methods

Reddick (2004a, 2004b) labels both simulations, which he calls deterministic methods, and econometric methods as causal.\(^{11}\) This article labels deterministic methods as “causal-like.” This usage may be imprecise; however, it usefully distinguishes these methods from time-series methods. Some deterministic methods are likely insufficiently complex to truly reflect even weak causal theories and are better treated as simple or simplistic time-series methods. Frank and Wang (1994) compared a simple deterministic approach for two revenues for one locality with several other methods. The authors find that these methods may be no worse than the other methods they consider. No other studies of simple causal methods have been identified, although some research does note that some econometric models take advantage of correlation without clear causality (McDonald, 2013, 2015).

Simulation Models. Simulation models show the relationships between variables (Chen et al., 2015), and forecasters work through them to determine what the consequences of specific decisions are. Some nations and large subnational jurisdictions use simulations or systems of statistical models to forecast their economies and related budgetary data (Congressional Budget Office, 2011; New York City Office of Management and Budget, 2016). There is no identified research into the marginal benefit of using these complex methods despite the importance of these models to many governments’ budgets.

\(^{10}\) These methods are defined in the Appendix.

\(^{11}\) Some simulations, especially those used by larger governments, are econometric models.
There are also many simpler simulation models that are often deterministic. These deterministic simulation models include payroll simulations, in which salary and benefit levels are forecast with great accuracy based on several factors such as cost of living adjustments, efficiency pay, longevity, and performance incentive pay. In addition to salary, employees earn fringe benefits such as Social Security, Medicare tax, unemployment compensation, health insurance, accrued leave, and retirement benefits. These simulations incorporate pay lags, vacancies, and increment timing (that is, when specific “steps” are incorporated into payroll). Using several years of data to determine the average, these factors can then be used to estimate future payrolls reasonably well. For a complete example of this type of simulation, see Chen et al. (2015).

Another example of a deterministic simulation model used in forecasting relates to the property tax. Forecasters know how much property is located within a jurisdiction from prior year assessments and also have estimates of this property’s taxable value from market changes and physical changes. The property tax is forecast by simply applying the approved tax rate to this taxable value. Some minor estimates of tax delinquency, tax-exempt property being purchased or sold, and other tax abatements or expenditures are made, but the bulk of the property tax forecast is determined by relatively known factors.

Only a few studies that examine the effectiveness of deterministic simulation (Brown & Harding, 2002; Smith, Pearce, & Harland, 2011) have been identified. None of these empirically examines simulation in the context of revenue or expenditure forecasts. This is peculiar given that these deterministic simulations account for the largest expenditure line item (personnel) and revenue source (property tax) for most governments and are important tools for forecasters.

**Econometric Models.** The state of the art of econometric modeling, particularly as it applies to national macroeconomic variables, is beyond the scope of this article. However, the basics of these models, when used for forecasting of individual variables, is relatively simple.¹² A forecast is produced by associating a dependent variable (the revenue or expenditure item to be forecast), with a set of independent variables through regression. Typically, these models are causal in that the independent variables are precursors of the dependent variable. A typical model may predict sales tax revenue through various measures of commerce and possibly demographic data. If tax rates vary over time, they also may be treated as an independent variable. A condition that is required is that the independent variables have known or reliably predicted values for the time period for which the forecast of the dependent variable is desired. This is achieved either through lagging (associating the dependent variable with a temporally earlier instance of the independent variable) or through additional forecasting of the independent variables. Because both the dependent and independent variables are found in time series, the regression residual is subject to autocorrelation, which can be measured with the Durbin Watson statistic. There are various techniques for correcting autocorrelation of errors, most of which require sophisticated software. Makridakis et al. (1998) provide instructions for forecasting with regression, and Kavanagh and Williams (2016) provide relatively simple guidance for use with revenue forecasting.

---

¹² These basics do not closely approximate complex macroeconomic models. There is controversy (a small part of which is cited here) in the macroeconomic literature largely having to do with the relationship between models and economic theory, which is, therefore, outside the scope of this article (Adolfson, Linde, & Villani, 2007; Diebold, 1997; Edge, Kiley, & Laforte, 2010; Negro, Schorfheide, Smets, & Wouters, 2007; Smets & Wouters, 2004). Despite this controversy, the macroeconomic forecasts actually in use are relatively accurate.
Not all econometric models are causal. Regression can only establish correlation. Correlation is not adequate to establish causality. There are substantial conditions (see “causal/causal-like” in the appendix) and sophisticated tests for causality (Granger, 1988a, 1988b); however, the most basic tests are whether there is a theoretical reason for causality and whether the supposed cause precedes in time the effect. The second condition, alone, is inadequate because there can be some other reason for this temporal relationship, including the possibility of a mutual prior cause or simple accidental relationship. If the accidental relationship is ruled out, it is possible to use non-causal (or indirect causal) relationships to forecast, which may result in improved forecast accuracy (McDonald, 2013, 2015).

**Decomposition and Mixed Approaches**

It can be beneficial to decompose a time series by its causal elements before forecasting the components (Armstrong, 1985, 2001b; Armstrong, Collopy, & Yokum, 2005; Green & Armstrong, 2015). To decompose revenue, a government separately estimates each type of tax. The total forecast is the deterministic sum of the taxes. At a deeper level, a single tax may be decomposed by units (for example property transfers) and the value of those units (the recorded sales price of the property transfers). Each may be forecast independently; then the value may be computed deterministically as the multiplication of forecasted units times forecasted value, which would then be further multiplied by the tax rate. These represent mixed approaches that bear some resemblance to a deterministic simulation but may rely primarily on other forms of forecasting. The method for forecasting each element may be distinct from the method for other elements. For example, the tax rate may be set by law and not require forecasting. Decomposition may also help incorporate anticipated policy changes that may affect particular elements of the forecast.

For some governmental revenue problems, particularly concerning intergovernmental transfers, especially where there is high variability from year to year, it is likely that determining the intentions of policymakers is the best method of forecasting. There has been substantial research into the use of intentions in other contexts (for a few examples see Morwitz [2001] and Armstrong and Green [2005]). There is no identified research regarding the use of intentions in budget forecasting. Where intentions are relevant but cannot be accessed, exponential smoothing methods or moving averages that reduce year-to-year variability may be the most effective. More research in this area may be beneficial.

**Forecasting Practices**

Governments may use quantitative techniques mixed with other practices to arrive at final forecasts. Forecasters, for example, frequently employ judgments to adjust quantitative output before finalizing forecasts. This technique is quite common in local governments (Frank & Zhao, 2009) and relies upon forecaster insight or expertise on particular taxes, spending categories, or determinants of revenue or spending categories. Some research identifies judgment as antithetical to accuracy (Hogarth & Makridakis, 1981). Others, however, identify judgment as benefiting forecast accuracy even with the existence of forecast bias (Lawrence, Goodwin, O’Connor, & Önkal, 2006).

While Makridakis et al. (1982) compared quantitative models with each other, Lawrence, Edmundson, and O’Connor (1985, 1986) were the first to compare quantitative to judgmental forecasting. In this study, the researchers found that judgmental forecasting was nearly as accurate as statistical techniques and, in certain cases, was much more accurate. Perhaps more
importantly for budgeting purposes, Lawrence et al. (1985) found that the standard deviations of the judgmental forecast errors were smaller than the statistical techniques, suggesting more consistent accuracy. Later studies found that the accuracy of judgmental forecasting depends upon the biases of the forecaster, and these can lead to less accurate forecasts than statistical methods (Moore, Kurtzberg, Fox, & Bazerman, 1999). Sanders and Manrodt (2003), on the other hand, find quantitative models result in lower forecast errors than judgmental forecasts. Overall, then, Lawrence et al. (2006) note that judgmental forecasting may be as good and accurate as statistical techniques, but it is highly dependent upon the forecaster’s biases.

For the moderately skilled forecaster, an appealing approach is the use of forecasting software (D. W. Williams & Kavanagh, 2016), which provide less sophisticated forecasters with more advanced methods. Hyndman and Khandakar (2007), for example, document two automatic forecasting methods implemented in R, which is a free statistical software package: the exponential smoothing method and an ARIMA model. R, however, requires a considerable learning curve. D. W. Williams and Kavanagh (2016) find that Forecast Pro and Autobox generally produce results that modestly outperform typical spreadsheet approaches. Both of these products implement ARIMA and other sophisticated techniques that the modestly skilled forecaster is unlikely to successfully use without assistance. They provide that assistance through artificial-intelligence-style automation. Frank and Zhao (2009) note that, despite the availability and cost of these software, few local governments actually employ them. This finding is similar to a survey of businesses, which found low uptake of software usage for forecasting, mainly because the software is not easy to use, and results are not easy to understand (Sanders & Manrodt, 2003). Because these studies are seven to 13 years old, and because software usage changes rapidly, these results may be dated.

A third practice that is prevalent in the public sector is consensus forecasting. This practice is used in part to remove politics from the forecasting process and to prevent politicians from increasing these revenue estimates especially in election years (Nelson A. Rockefeller Institute of Government & Pew Center on the States, 2011). Mikesell and Ross (2014) note that, in government, political acceptance of revenue forecasts (a hard budget constraint) is critical. Even when simple techniques are more accurate, political actors might reject the forecasts for political gain. Further, Mikesell and Ross (2014) identify significant heterogeneity in the actors involved in the consensus forecast between states. Relatedly, Krause and Douglas (2013) and Krause et al. (2013) find that too much or too little political inclusion in the consensus forecasting exponentially increases forecast error. Consensus forecasting may serve an alternate purpose of obtaining forecast acceptance, particularly among included participants.

A fourth practice involves combining forecasts (Armstrong, 2001c; Clemen & Winkler, 1986; Gardner, 1985; Timmermann, 2006). Some literature aims to identify optimal weights for combining forecasts, but some evidence suggests simple forecast averaging achieves most of the benefit. Averaging forecasts obtained through sharply different methods may achieve the greatest benefit.

**Forecasting, Estimating, Predicting, and Dynamic Estimation**

The word “forecast” is ambiguous in that it can refer to future values or it can refer to all output of a statistical model built to predict the future. From a model-building point of view, the forecast is all of the model’s output data, which may range from before the beginning of the
actual data through the end of the predicted future periods. Segments of the data may be labeled “backcast” (predicted values for the period before the temporal beginning of the input data), “nowcast” (predicted values for times roughly contemporaneous with the end of the series), and “forecast” (predicted values beyond the end of the series). Typically, the common use of “forecast” refers to predictions of the future. However, even predictions of the future are conditional in that they may predict the future assuming current practices continue or, instead, the future hypothesizing alternates. For governmental revenue and expenditures, these alternatives typically reflect policy changes.

We label a prediction with respect to the result of a deliberate policy change as an “estimate.” Estimation reflects the prediction of future values where there is limited or no directly relevant historical data series, or where such data is incorporated through analogy or computation not contained within the core historical series. Estimation may be accomplished through a variety of methods:

1. It is commonly taught in colleges and universities that economic policy analysis is the best method for making an estimate. While there are many specific forms of this method, the general approach is to build a regression model that correctly captures the a priori rationally selected independent causal variables and uses these to predict the dependent variable of interest, such as the revenue produced. Econometric causal modeling can be used to estimate the effect of changes in current practices or, when proposed practices occur in other locations, they may be used to impute the effect of those practices if implemented in a different locality. For this use, simple association is inadequate because the statistical model is used to impute the effect, which is a causal construct, to the proposed policy. If it is assumed that revenue is the product of rate times base, econometric modeling also can produce reasonable estimates of the tax base even where there has been no prior practice of taxing such a base in any locality. An estimate made with econometric models may be similar to a forecast, particularly if the causal model is also used for forecasting; however, care should be taken to understand how the future values of the independent variables are determined. Causal modeling requires access to data sources, skilled users of statistical methods, and a sound basis for causal modeling. Some of these resources may not be available for many estimates.

2. If there is a regression-style forecast model, then, by adjusting the future value of independent variables to reflect a policy change and comparing those to the future value of those independent variables under current policy, one can determine the difference resultant from the policy change. For example, a simple regression of a sales tax revenue series may include the tax rate as an independent variable. An estimation of the consequence of changing the rate could be produced by substituting the proposed new rate for a continuation of the old rate within the statistical model for temporal periods beginning with the change. By comparing the output with the substitute data with the output of the original data, one estimates the value of the change in the tax rate. However, even sophisticated use of this method may be controversial because methods for modeling may assume continuation of historical

---

13 For regression modeling this would be the temporal beginning of the dependent variable.
14 There is no definite border between forecast and estimate. When predicting a single proposed change or determining the effect of some highly probable future event, the likely term is estimate; however, a forecast that incorporates one or more estimates remains a forecast.
relationships that may not be continue under new policy (Leamer, 1985; Sargent, 1979, 1984; Sims, 1986; Sims, Goldfeld, & Sachs, 1982).

3. Where estimators do not have the benefit of the resources implied with the previous methods (Grizzle & Klay, 1994), estimates can be made through deterministic calculations. For example, if an econometric model produces an estimate of a tax base (as with the first method), that estimate is converted to a revenue by deterministically multiplying it by the tax rate. In other circumstances, both the base and rate may be determined through less robust methods and be combined to compute a revenue.

4. Some estimates or values of variables for estimates may be determined through expert judgment.

Few identified studies have examined the effectiveness of deterministic policy estimates (Brown & Harding, 2002; Smith et al., 2011), and none that empirically examine the relative effectiveness of a variety of approaches to estimation.

One form of estimation related to econometric causal modeling is dynamic estimation, which is also called dynamic forecasting or dynamic scoring. As with all estimation, dynamic estimation is not forecasting. However, it is often treated as a form of forecasting, and, to be performed at all, it requires effective causal models. The basic idea of dynamic estimation is to include behavioral change within the policy change estimation. For example, if property taxes were increased, citizens may “vote with their feet” and move to another locality where property taxes are lower. The dynamic effect would be the gradual decline in property value leading to lower than expected revenue with the higher rate. Dynamic scoring is associated with anti-tax advocates and with the Laffer curve (Laffer, 1981, 2004; Oudheusden, 2016), a theoretical view that, if tax rates are too high, then reducing the rate will produce more, not less, revenue because the current taxes excessively discourage economic behavior. Consequently, a tax reduction may pay for itself. The empirical evidence does not support the view that taxes in the United States or Europe exceed this theoretical limit, although they may be close in Europe (Trabandt & Uhlig, 2009, 2011). The recent experience of the state of Kansas, where tax changes have not produced the expected revenue effect, should serve as a warning that ideological commitment to dynamic scoring is risky for politicians (Fox, 2016; Stapleton English, Løppenthin, & Roca Diaz, 2015). Generally, states have been disappointed when expecting dynamic results (Bluestone & Bourdeaux, 2015).

While not universal, the typical national practice includes some inaccuracy and forecast optimism. Consequently, this overforecasting may interact with dynamic scoring. If actual dynamic effects are small or absent and forecasts are already optimistic, the consequence of adjusting forecasts for dynamic effects to include tax policy changes may exacerbate the optimistic bias.

Opportunities for Future Research

Throughout this article, we have discussed budget-related forecasting and identified topics that would benefit from future research.

1. There is little empirical research into the relative effectiveness of various approaches to estimation. As a practical matter many jurisdictions may not have access to regression-
The Status of Budget Forecasting

2. While there is a vast literature of sophisticated forecasting techniques that generally use econometric and time-series methods, a fair amount of actual forecasting, particularly among smaller governments, is completed using deterministic techniques (Frank & Zhao, 2009). We classify a technique as deterministic when the forecast is made using equations that are not fit by minimizing a statistic (excluding equations that combine multiple individually fit forecasts); some cited literature may use alternate definitions. These methods can be as simple as increasing an expenditure by an anticipated inflation rate and can be as complex as accounting for all of the elements of a payroll system and computing expected future expenditures. Frank and Wang (1994) compared a simple deterministic approach for two revenues for one locality with several other methods obtaining mixed results and Reddick (2004b, p. 46) examines deterministic forecasting somewhat more broadly, finding “very modest support” for deterministic forecasting. We are otherwise unaware of any research within the past 30 years that examines the effectiveness or relative effectiveness of deterministic techniques used by governments for budget forecasting. The largest single source of local government revenue is property tax, and the largest local jurisdiction in the United States, New York City, uses a deterministic quasi-simulation to forecast property tax revenue (New York City Office of Management and Budget, 2016). Reddick (2004a) finds that 53% of local governments use deterministic methods, including 27% using them for property taxes and nearly 30% for other fees. Forrester (1991) finds that 44% of cities use deterministic approaches for property taxes.15

3. There is an intriguing inconsistency in revenue forecast bias for different jurisdictions. Some, such as state and local governments in the United States, prefer underforecasting as a strategy that likely minimizes risk, while others, such as members of the European Union, prefer overforecasting, which as a strategy that likely provides maximum short-term benefits to the public. While these differing practices clearly reflect institutional or cultural differences, it is not clear what the differences are. It is particularly intriguing that, with the United States federal government, overforecasting revenue is a conservative strategy, but the same practice in Europe is a liberal strategy. These phenomena require further examination.

4. Subnational jurisdictions in the United States exhibit strategic forecasting of revenue through underestimation. This behavior does not appear to be matched with strategic overestimation of expenditures possibly because of the way appropriation works. However, we are unaware of any research into whether there is or is not such strategic overestimation.

5. With respect to the United States’ federal economy, there are numerous forecasters in at least three classes: (a) the government itself, (b) forecasters associated with private firms, and (b) forecasters associated with public-policy-oriented nonprofits. In Europe, there are also publicly sponsored nongovernmental entities that forecast. Finally, international organizations such as the World Bank engage in forecasting. While sometimes two or three forecasters, sometimes from different sectors, are compared, there is no apparent literature that addresses the entire domain. Questions that might be

---

15 Forrester’s table labels are vague, “Deterministic & others” where the “& others” occurs for every category except “Unspecified.”
addressed from this array of forecasters include: (a) Do more forecasters from more sectors increase or decrease bias? (b) Would some aggregate of multiple forecasts be more accurate and less biased than any single forecast?

6. Multiyear forecasting appears to have substantial risk, yet multiyear budgeting remains a likely best practice. Consequently, more research is needed on the effectiveness of multiyear budgeting and the availability of low-bias multiyear methods.

7. When it is addressed in this research, out-year forecasts quickly exhibit substantial bias. What are the consequences of this bias? In the United States, where the bias may lead to a perception of severe debt in the mid-term future, does it contribute to current state of political gridlock? Are there other consequences when bias misleading suggests surplus or even just balance? Has the proliferation of forecasters contributed to bias and possibly to the inability to achieve political consensus?

8. There are numerous studies within several components of budget-related forecasting. Over the last 40 years, there have been 28 identified studies of the US federal budget related to forecasting. There are a similarly large number related to the European Union. There are numerous studies of state and local budget forecasting in the United States. While there is some diversity in precisely what is studied, it is likely that these component domain collections are ripe for aggregate evaluation through meta-analysis, which may lead to a better understanding of forecast practice.

9. There is no identified research into forecasting where intentions are relevant, particularly with respect to intergovernmental revenue transfers. Such research would be beneficial.

Conclusion

In all organizations, budgets are plans that reveal objectives established by decision makers, how the organization will obtain resources, and how it will use these resources to reach goals. Organizations typically budget not just for the upcoming fiscal year, but for several years beyond as well so that the organization may strategically move toward its goals over time. Because of the prospective nature of financial planning, forecasting these resources or the underlying causes of revenues or expenditures is a core function of budgeting.

In public organizations, however, forecasting is not merely an attempt to accurately predict future values; that is, forecasting is not merely a technical exercise in which budget analysts aim to predict with minimal error revenue and expenditure line items. Instead, the extant literature consistently reveals that forecasting serves other ends that are valuable to managers and decision makers. Most importantly, most US subnational governments use conservatively biased revenue and expenditure forecasts. Given widespread balanced budget requirements, these biases make it easier to meet statutory and constitutional financial requirements. These pessimistic biases also provide funds for midyear re-budgeting/budget modifications, which are potentially valuable for politicians (either to curry favor with other elected officials or the public at large). International studies confirm the political nature of forecasting within different government systems as well. Of particular importance, out-year budget forecasting indicates large errors, suggesting that such out-year budgeting is of limited value for decision making.
Beyond the politics of budget forecasting, budget analysts also face the question about the best techniques for forecasting. With the proliferation of data and computing power, the costs of complex forecasting are increasingly minimized. However, most assessments of forecasting techniques find that simple methods work as well or better when compared with complex methods, or, in the absence of skilled forecasters, the use of forecast software.

Forecasting practice also reveals heterogeneity across the United States as well as internationally. The literature quite consistently shows that forecasters apply judgments to their own technical forecasts. Some experts even eschew statistical models and forecasts based on their own knowledge. In other cases, consensus forecasting is found to result in more accurate forecasts. The literature overall, therefore, does not dismiss judgmental forecasts, assuming knowledgeable people are the forecasters.

Finally, we identify a number of areas in which the public forecasting literature needs additional development. Several of these areas, such as the effectiveness of nonregression-based forecasting techniques, are quite important to the majority of governments in the United States and other subnational jurisdictions, where budget offices are limited and resource investments in technology are scarce.

**Disclosure Statement**

The authors declare that there are no conflicts of interest that relate to the research, authorship, or publication of this article.

**References**


**Author Biographies**

**Daniel W. Williams** has taught budgeting and related topics at Baruch College for 20 years. His research encompasses budgeting, forecasting, history of public administration, performance measurement, and related topics.

**Thad D. Calabrese** teaches and researches in the field of public and nonprofit financial management. He is especially interested in the areas of employee benefits, the financial implications of collaborative governance and contracting, and capital structure decisions in public service organizations.
Terms as used in this discussion include:

**Accuracy** – A measure of how close predicted values are to actual values. For forecasting, the two most common measures of accuracy are RMSE and MAPE. Smaller values are more accurate. MSE is commonly used when examining a single series. MAPE is commonly used when examining accuracy across multiple series.

**ARIMA** – Autoregressive Integrated Moving Average model; a statistical technique in which a lagged variable is used to predict current values, and incorporates past error terms.

**Autoregression** – The correlation between sequential observations.

**Asymmetric Loss Function** – The penalty for an error differs depending on the direction of the error. Similar to bias. Also see confidence interval.

**Bias** – To systematically over-predict or under-predict. When ME is positive, the forecast is systematically under-predicting, and in reverse it is over-predicting.

**Cause/Causal/Causal-Like** – Hill (1965) asserts that two variables are causally related when a change in the variable labeled “cause” is temporally prior or simultaneous with a change in the variable labeled “effect,” where there is a plausible reason why the cause leads to the effect, the relationship is consistent, and there is a dose effect (the size of the change in cause is related to the size of change in effect). He includes four additional or alternative criteria (strength of relationship, specificity of relationship, subject to experimental modification, and reasoning by analogy) and one criterion (coherence) that is at the level of epistemology. Granger (1988a, 1988b) adds a complex test for causality when performing statistical modeling. With statistical models the change/change relationship can be established by correlation. When plausible causal variables are included in a statistical model, it is widely understood that plausible alternative causal variables – representing alternative hypotheses – should be excluded (Newton & Rudestam, 2012), which may be achieved through the relative strength of correlation diagnostics. Because correlation relates to only one of the conditions of causation, there is a widely known principle that “correlation is not causation.” This principle can be too broadly applied in that sometimes correlation is disparaged as irrelevant to causation. In this article “causal” refers primarily to the temporal, plausibility, alternate, consistency, and dose criteria, particularly when established through correlation and possibly meeting the Granger criterion.

**Central Tendency** – The estimated middle of a set of observations. In its simplest form, the average of some observations. For more complex methods, the predicted location on the center line. For forecasting, this value may be labeled a point estimate.

---

16 Many of the terms here are defined in Makridakis et al. (1998) Statistical terms are defined broadly in most statistics textbooks. Terms related to bias are defined diffusely through the forecasting literature. Formulae for many of these methods can be found in Makridakis et al. (1998) or in D. W. Williams and Kavanagh (2016).
Confidence Interval – An estimate of the range of values surrounding the point estimate for within which there is a specified large probability of finding the actual observation.

Consensus Group Forecasting – A system in which forecasters representing different stakeholders or points of view are assembled to arrive at a joint forecast.

Conservative Forecast – See “pessimism”.

Damped Trend – A variation on Holt’s exponential smoothing in which the trend element of forecast is reduced by a small percentage over each successive period eventually leading to zero trend.

Decompose – To break a time series into smaller component time series.

Dependent Variable – The variable that is predicted in an econometric or other model.

Deterministic Model – A tool in which all independent variables are treated as known with certainty.

Deviation – See “error”.

Dynamic Scoring/Forecasting – A method for analyzing policy options in which economic actors’ behavior changes are forecasted based on the policy’s incentives. These forecasted behavioral changes are then used to forecast tax or expenditure changes.

Econometric Model – The use of sophisticated statistics to model and predict a variable. Typically, econometric modeling relies on regression or closely related correlation based techniques.

Effective/Effectiveness – A measure of how much a forecast influences decisions.

Efficient/Efficiency – A relative measure of whether a forecast can be improved by using more information.

Empirical Bayesian Analysis – Empirical Bayesian methods use information from similar data to adjust mean values and narrow confidence intervals. They are especially useful with small samples.

Error – Actual results minus forecast. Also labeled deviation in some statistical literature.

Estimate – A prediction of the consequences of some change, either deliberate or anticipated. For budgeting, an estimate is expressed in dollars.

Exponential Smoothing – Technique in which a weighted average between older and more recent observations in a time series is determined (usually where more recent observations are more heavily weighted than older ones), and this mean is then used as a forecast.

Favorable Error – Error in which actual expenditures were less than forecast, or actual revenues were greater than forecast.

Flexibility – See “policy options”.

The Status of Budget Forecasting
**Fiscal Year** – A twelve-month period that can begin with any date during which revenue and expenditures are appropriated (authorized). The most common fiscal years begin on July 1 and end on June 30 (46 states and many localities) or begin on October 1 and end on September 30 (the federal government and 2 states). Smaller localities may have their fiscal years coincide with the calendar year. Commonly appropriations are for years, however they may be for other periods such as two sequential years (biennial budgets).

**Forecast** – A prediction, commonly of the future, resulting from a technique or method that is intended by its user to produce future values. A forecast is not merely the preparation of the financial component of a budget. As discussed in the article, forecasts are distinct from estimates, where the future value is contingent on a future decision; however, it is not uncommon for estimates to be treated as forecasts. For revenue and expenditures, forecasts are of dollars or units, such as individuals or transactions that contribute to dollar values. Forecast can also refer to any output of a forecast model, whether of the future, present, or past.

**Holt Exponential Smoothing** – A type of exponential smoothing in which separate equations estimate the time series level of observations and the time series trend (or change) of observations.

**Horizon** – The length of time between the production of a forecast (or its release) and the time period to which the forecast applies.

**Independent Variables** – The variables that are thought to contribute to a prediction.

**Lags** – Associating a dependent variable with a temporally earlier instance of the independent variable.

**Mean Absolute Percent Error (MAPE)** – The average of the absolute errors divided individually by the actual values times 100. MAPE treats errors proportional to their size. Because it is expressed in percentage, it is not sensitive to the magnitude of the data.

**Mean Error (ME)** – The average of the errors.

**Mean Squared Error (MSE)** – The average of the errors after each error has been squared.

**Model** – (1) The estimated central tendency and variance of a set of observations. For example, the mean, also called the average, is the unadjusted central tendency of a set of observations with a standard deviation that is the square root of the variance. In general, models are more sophisticated than the simple average. A model is providing a method for predicting the value of a dependent variable. See econometric model and time series methods. (2) An algorithm that is used to produce a predicted value. See deterministic model.

**Moving Average** – Time series forecasting technique in which values are averaged over some time period and then this average is used to forecast.

**Neural Networks** – A forecasting technique that, unlike ARIMA methods, does not assume a linear relationship in the data.

**Optimism** – Bias that over-predicts revenue, under-predicts expenditures, or both.

**Pessimism** – Bias that over-predicts expenditures, under-predicts revenue, or both.
Policy options – The perception that funding is available allowing for voluntary choices such as increasing expenditures or reducing taxes.

Point Estimate – See “central tendency”.

Predict – Specify an unknown value.

Prudence – Either pessimism or deliberately not spending all expected revenue.

Rational/Rationality – For a forecast, efficient and unbiased.

Repetitive Budgeting – A system in which a budget for the upcoming and current period(s) is modified on an ongoing basis, diminishing the value of the budget for planning and control purposes.

Root Mean Squared Error (RMSE) – The square root of MSE. RMSE values large errors much more than small errors as a result of squaring. It is sensitive to the magnitude of the data, thus it is not good when comparing series that are of different size. This statistic is nearly identical to the common form of the standard deviation.

Seasonal/Seasonality – Time series data with an underlying predictable variation during the fiscal year.

Shortfall – More expenditures than revenue during a fiscal year.

Simulation - Any approach that uses math to imitate real world processes. These can be deterministic, which are sometimes labeled algorithms, or they can involve statistical modeling, such as Monte Carlo simulations.

Standard Deviation – For the simplest statistics, RMSE. Otherwise comparable values determined through statistical theory.

Stochastic Model – A tool in which random variation exists in the independent variables.

Structural Deficit – Over the foreseeable horizon recurrent revenue is less than recurrent expenditure requirements.

Surplus – Revenues in excess of expenditures during a fiscal year.

Systematic Error – Bias.

Time-Index Regression – A model in which a variable is predicted using time as the chief or only independent variable.

Time Series – A variable that takes on alternate values demarked by time units.

Time Series Methods – Techniques that implicitly rely on an expectation that change in a time series is gradual.
**Trend** – Tendency for a time series to increase or decrease from observed point to observed point.

**Uncertainty** – The degree to which a forecast, which is yet to be actualized, may ultimately be in error.

**Unfavorable Error** - Error in which more actual expenditures were spent than forecast, or actual revenues were less than forecast.

**Variable** – An object or characteristic that can take on values when observed.

**Variance** – For the statistics included in forecasting, MSE.

**X-11/X-12/X-13** – A complex nonparametric procedure used to determine seasonal factors. These methods are closely associated with the United States Census Bureau. The X-12 version integrates older approaches with older multi-level moving averages. The X-13 version integrates the method with more complex statistical procedures.