



Institute for Illinois Public Finance

UNIVERSITY OF ILLINOIS SPRINGFIELD

MEASURING FISCAL SUSTAINABILITY OF LOCAL GOVERNMENTS: A STRESS TESTING APPROACH USING ILLINOIS MUNICIPALITIES

Kenneth A. Kriz

University Distinguished Professor of Public Administration, University of Illinois
at Springfield

Richard Funderburg

Associate Professor, University of Illinois at Springfield

Working Paper 19-01

April 2019

BACKGROUND

Fiscal sustainability is a term that is used frequently in the academic literature on applied public finance. The concept itself is easily understood. We say that something is sustainable if it can be maintained over time. An organization (or jurisdiction) is unsustainable by contrast if it will not survive forces that act on all such organizations. In that sense, the term sustainability in a fiscal sense is somewhat overstated. Even in the worst case of bankruptcy, as in Stockton or Vallejo, California, Puerto Rico, or Detroit, Michigan, the corpus of the organization that manages the jurisdiction survives even if it declared financial insolvent. However, the very nature of those jurisdictions changed dramatically as they went through the bankruptcy process (Lewis, 2011). Thus, perhaps sustainability in a fiscal context means that a jurisdiction will be able to maintain its current profile of services and leadership structure.

The importance of fiscal sustainability can be illustrated by the “Meredith Whitney” effect. When former Wall Street analyst Meredith Whitney went on 60 minutes in December 2010 to predict that 50 to 100 municipalities would default on debt payments totaling “hundreds of billions of dollars” and that the problem would become evident within the next year, investors in the municipal bond market reacted strongly, with municipal bond fund outflows reaching a then record amount shortly after the announcement (Wiesenthal, 2011).

Fiscal sustainability has been widely studied under many different names (fiscal health and fiscal space most recently) over the past three decades. Numerous studies have presented systems that purport to measure fiscal sustainability. And numerous tools have been invented to compare and contrast jurisdictions in terms of their ability to sustain operations in the future. Most of these studies and tools share a common logic and approach. A few to many indicators of fiscal

sustainability are gathered on a jurisdiction, largely consisting of single measures (such as fund balance) or ratios of measures (debt service per capita) and then a summary of those measures is generated, accompanied by an analytical assessment of how those measures may affect the sustainability of a jurisdiction. However, these studies most often have failed to try to quantify the risk posed by a given indicator or provide a summary measure of the likelihood of a jurisdiction being at risk of fiscal insolvency. What the reader of these studies is left with is that it is the “professional judgment” of the author that the items in question are important in determining fiscal sustainability and that a given jurisdiction is better or worse on these key measures. But the reader is left without guidance as to why and how much risk the jurisdiction faces. Further, most of the studies have failed to incorporate dynamic analysis, looking at only past and current values of measures or ratios. This is despite the essential definition of sustainability as forward looking. Indeed, one could and should ask that if a particular indicator is deemed to be negative for a jurisdiction why this has not caused insolvency in the past.

In this paper, we seek to outline the use of a different tool for assessing the risk of fiscal sustainability in jurisdictions. Drawing inspiration from the financial crisis of 2008-09 and the approach used by banking regulators in its wake, we propose to use stress testing to illustrate the risks of “exceptional yet plausible” risks to jurisdictions (Sorge, 2004). Stress testing involves the use of computer simulations to detect the likelihood that a community faces risk at a certain level over the next several years. For this paper we choose five years as the timeframe for analysis and employ a model of the fiscal future of jurisdictions developed by Kriz (2015) to assess the level of budgetary reserves necessary to withstand economic or fiscal shocks of a given magnitude. Employing the stress tests on a population-stratified random sample of Illinois municipalities, we find that most of the jurisdictions face significant risks from even a one-standard deviation shortfall in forecasted revenues

or increase in expenditures above forecast, with some jurisdictions facing nearly certain and ruinous outcomes from a two-standard deviation event. We then discuss implications for the jurisdictions and for their stakeholders.

LITERATURE REVIEW

FISCAL SUSTAINABILITY/HEALTH/STRESS

The concept of fiscal sustainability has actually been referred to in the literature by three terms. During the 1980s and 1990s, the focus was on fiscal stress and the effects of that stress on public organizations, started by the publication of Levine's seminal works (e.g., Levine, Rubin and Wolohojian, 1981). As governments struggled to adapt to changing economic and fiscal situations, many encountered significant financial difficulties. The one-two punch of a deep industrial recession in the early 1980s followed by an agricultural recession, along with stresses brought on by the "taxpayer revolt" of the 1970s and 80s led many subnational governments to the brink. Most of the academic work on fiscal stress focused on broad forces that impact the amount of fiscal stress that governments were facing. The works from this period had a marked political science approach and lacked applicability to the measurement of fiscal sustainability. It simply was not the focus of the work.

The focus on government finances shifted gradually during the 1990s to the concept of fiscal health. This concept emphasized the development of systems of indicators that officials could use to assess the sustainability of the jurisdiction's finances (there is another stream of literature that emphasizes using these measures in a broader discussion of factors that affect fiscal health, for this see Honadle, Cigler and Costa (2004) and McDonald (2015)). One of the earliest efforts in this area was the development of a "10-point" test for smaller cities to assess their financial condition (Brown, 1993). Over the years this literature went in two directions. One direction emphasized the development of a better set of indicators for assessing

fiscal health. Examples of this include Maher and Nollenberger's (2009) updating of Brown's test, Honadle and Lloyd-Jones' (1998) demonstration of three systems of indicators for analyzing rural government fiscal health, Hendrick's (2004) illustration of the use of fiscal condition measures in the Chicago area, and Groves and Valente (1986), who describe the International City/County Management Association's Financial Trend Monitoring System.

The most recent effort in this vein is from Gorina, Maher and Joffe (2018). They point out that most of the systems of indicators that had been developed lacked an empirical element and were not assessed for predictive ability. The most that users of a system of indicators could hope to do is to calculate some ratios or include some measures but could not assess how much a given indicator would contribute to fiscal health. Further, fiscal health measurement is almost always retrospective. There is no focus on whether governments will encounter difficulties in the future, which is the crux of the sustainability question. They develop a model where fiscal distress (measured in many different ways) is regressed on various indicators of health. They find that several indicators have predictive power and provide estimates of the marginal effect of various indicators on fiscal distress. However, their model has low predictive power and more importantly, they only analyze contemporaneous values or a single lag of the fiscal health measures. Given reporting lag, the most that their model can offer in terms of predictive ability is an ability to measure whether the jurisdiction is likely to be in fiscal distress in the current year. This is hardly a forward-looking measure of fiscal sustainability.

A related literature on fiscal sustainability developed in the 2000s, mainly driven by a perceived need to assess the ability of national governments to sustain growing debt levels. Earlier work on budget balance had emphasized principles and techniques for determining whether budgets were balanced over time (see, e.g.,

Hamilton and Flavin (1986) and Quintos (1995)), however, there was little sustained work on the concept of sustainability. Then during the early 2000s, researchers at the World Bank expanded on these ideas and produced a series of working papers setting forth concepts of sustainability, culminating in the publication of a set of principles and documenting their application (Burnside, 2005). The problem of sustainability in this literature has been framed as one of cointegration, whether over the long-term revenues and expenditures shared a common-trend, indicating that they moved together over the long-term and if a deviation from trend occurred that there was an adjustment back to the long-term trend.

Much of the work in this area has been focused on national government finances. There are at least three reasons for this. First is the general paucity of time-series data of sufficient length for using in a cointegration framework. The detection of long-run trends takes a data series long enough to analyze for unit roots and cointegration. This cannot be done with less than 20-30 observations at a minimum. And most local governments do not have that much data readily available. The second reason is the changing nature of state and local government accounting. Analyzing federal government revenues and expenditures is greatly aided by a more or less consistent accounting over time. However, government accounting standards for state and local governments have changed dramatically at important points during the last several decades. Finally, there is simply too much diversity in state and local reporting requirements. In some states, the majority of smaller governments are on cash or modified-cash accounting bases. Attempting to compare governments reporting in this way to those reporting on a modified accrual basis would entail crosswalks that may or may not exist.

For these reasons, the literature on the use of fiscal sustainability in a cointegration sense is sparse. Two exceptions are Mahdavi and Westerlund (2011) and Ji, Ahn and Chapman (2016). The former paper uses a panel of data whereas the latter uses summary data from the Census Survey of Local Government Finances. They find broad patterns of sustainability in financial aggregates. However, they do not attempt to disaggregate the data fully to investigate the sustainability of individual units of government. Therefore, as a practical tool to assess sustainability, it falls short.

STRESS TESTING

Stress testing developed in the 1990s as concerns over increased financial instability in many countries created a sense that policy makers did not fully understand the magnitude or prevalence of risks in financial markets. Work on the development of a unified framework for implementing stress tests on financial institutions began in the late 1990s through the development of the Financial Sector Assessment Program (FSAP), a joint program of the International Monetary Fund and World Bank (Blaschke et.al., 2001). Since then it has been used in numerous domains ranging from banks (Schuermann, 2014) to household debt levels (Bhutta et.al., 2019) and public pension systems (Mennis, Banta and Draine, 2018).

At the most basic level, stress testing addresses two of the issues that have plagued the literature on fiscal sustainability/fiscal health. First, it allows for the development of richer models that go beyond a few indicators to examine a full model of variables of interest. In much of the fiscal sustainability literature, revenues and expenditures are treated as aggregates. This aggregation may hide volatility in underlying revenue sources and expenditure demands that may signal future problems of sustainability. This relates to the second benefit of using stress

testing. Previous attempts to measure sustainability have largely ignored the temporal dimension and treated measures as static. Sustainability has an inherent temporal dimension. Sustainability at a given point in time is not based on current measures but rather what those measures may or may not say about how they will change over time. Gorina, Maher and Joffe (2018) attempt to get at this, but as stated above are only able to provide current estimates of sustainability. Stress testing allows us to look into the future and assess what current trends in revenues and expenditures and uncertainty about them says about the likelihood that a jurisdiction will encounter fiscal distress in the future.

MODELING APPROACH

Blaschke, et.al. (2001) define the essential components of a stress test as shown in the left column of Table 1. We choose a Financial-Operational Risk Model (Decision Item 1) through building a historical record of a jurisdiction’s finances and then building a forecast model based on that data and economic data from the area surrounding a jurisdiction. We choose to use county-level economic data for a practical and theoretical reason. In practical terms, counties are the smallest geographies with many of the variables we consider to be measures of economic activity in the cities. Theoretically, cities should benefit not only from their own economic activity, but activity in surrounding jurisdictions. This is especially true of larger cities, that exhibit a “pull” on local income through being centers of retail sales, but it is true for most cities of even modest size.

TABLE 1 GOES HERE

We employ an “Other” Type of Stress Test (Decision Item 2 in Table 1) by estimating the effects of a shortfall in revenues below an expected level (as defined by a forecast model) and/or higher than expected city expenditures. Our forecasting

model generates standard errors for forecasts based on model error but also on historical volatility of revenues and expenditures. Thus the “risk” of unexpected increases or decreases can be measured directly. In this way, the Type of Shock (Decision Item 3) is Underlying Volatilities, derived from the forecast and based on standard errors. With these measures, we could develop a Monte Carlo simulation (as in Kriz, 2015). However, we rather choose for simplicity of presentation to employ a Hypothetical Scenario approach (Decision Item 4).

Our Core Assets to be Shocked (Decision Item 5) is a measure of financial situation we term “Operating Balance”. We build the variable through subtracting Total Governmental Fund Expenditures from Total Governmental Fund Revenues (the revenue forecasts are done on individual sources of revenue except for an “other” source that aggregates smaller sources of revenue). We use Governmental Fund revenues and expenditures as this definition should incorporate the bulk of government operations. We then adjust this measure by subtracting Capital Outlays (Expenditures) to isolate it from the inherent lumpiness of capital spending and since capital expenditures are frequently financed using debt issuance, which is other source of financing. If the city gets significant resources from Other Financing Sources, we incorporate this into the analysis and note its effects. So our final measure of the Core Asset is:

$$\text{OPERATING BALANCE} = \text{TOTAL GOVERNMENTAL FUND REVENUES} - (\text{TOTAL GOVERNMENTAL FUND EXPENDITURES} - \text{GOVERNMENTAL FUNDS CAPITAL OUTLAY}) \quad (1)$$

In constructing our Operating Balance measure, we use historical data from a jurisdiction’s (for this paper we use data on Illinois municipal governments) Comprehensive Annual Financial Report to build a history of the city’s finances. Following Kriz (2015) we then build a forecasting model on various revenue sources in the Governmental Funds (taxes by major category if available or aggregate taxes,

intergovernmental revenue, charges for services, and other revenues) and operating expenditures (Governmental Fund total expenditures – capital outlays) using various time-series methodologies (ranging from naïve models such as smoothing models and autoregressive models to vector autoregressive models).¹ The model produces point estimates and standard errors for key financial variables. We then use equation (1) to calculate point estimates of operating balance. Then we estimate the effects of a 1 standard error negative shock on the financial outcome by adding the standard error to operating expenditures and subtracting it from the various revenue sources. Then we recalculate the operating balance, as in:

$$\text{OPERATING BALANCE} - 1 \text{ SE} = \text{TOTAL GOVERNMENTAL FUND REVENUES} - 1 \text{ SE} - (\text{TOTAL GOVERNMENTAL FUND EXPENDITURES} - \text{GOVERNMENTAL FUNDS CAPITAL OUTLAY}) - 1 \text{ SE}$$

Finally, we do calculations for a 2 standard error negative shock from forecasted values. In this way, the Peripheral Assets to be Shocked are revenues and expenditures of the city (Decision Item 6 in Table 1) and the Size of Shocks (Decision Item 7) is 1 and 2 standard error unexpected deviations in the peripheral assets. Finally, we develop the forecasts for the current year (FY 2020) as well as four future years (FY 2021-24) for a total time horizon of 5 years. We also implicitly assume that the city will not employ any specific risk management techniques to offset revenue losses or extra expenditures (Decision Item 9). This is a modeling choice that produces estimates of future risk. Our hope is that cities with particularly high risk of encountering fiscal distress would develop risk management techniques (such as fully funding a robust reserve fund). However, as

¹ As part of the forecasting process, we also develop forecasts of key economic variables such as county taxable retail sales, county per capita personal income, county wage and salary income, and county salary and wage employment. Those variables become inputs into the forecasting models as predictor variables as described in Table 3.

many studies have indicated that many jurisdictions fall short of funding their reserves or fail to fully employ them during even serious financial distress periods, we can only conclude that cities are passively accepting financial risks. We will highlight cities that do have a robust reserve fund in place to offset initial losses.

SAMPLE CITIES

Our goal with the sampling strategy was to try to get cities in as many circumstances as possible. Our sampling frame is municipalities in Illinois. Therefore, we first stratified the list of municipalities by population into six population groups: (1) greater than 100,000; (2) 50,000 – 100,000; (3) 20,000 – 50,000; (4) 5,000 – 20,000; (5) 500 – 5,000; and (6) under 500. We then did a random sample using a random number generator. This produced a list of cities in each category from which we could search for CAFRs. We then attempted to get a sample of cities in each group proportional to the relative size of the group. For example, there are only eight cities in Illinois with populations greater than 100,000 while there are 212 in the 5,000 – 20,000 population size category. We excluded Chicago as that city is likely more affected by national rather than regional economies, producing a category size of seven cities. So we included only one city in that category. Ultimately, we had to abandon a strict adherence to the sampling strategy due to the relative lack of information on the smaller cities (many smaller cities do not prepare CAFRs) and an added goal to have some geographic diversity (many of the smaller cities are in Northeastern Illinois).

Table 2 shows the cities that are in our sample. We were able to get information on cities in all income categories except for the smallest category. We were also able to get some geographic diversity. The data availability is mixed. It was somewhat surprising that the third smallest city in the sample had the most years of financial data readily available. Due to data availability, we were not able to analyze all of

the cities in the sample. We decided to try to get geographic diversity, so analysis was not completed on Park Forest, Prospect Heights, Hawthorn Woods, Hillside, and Oakbrook Terrace. In all, analyses were conducted on six cities.

TABLE 2 GOES HERE

RESULTS

FORECAST MODELS

As discussed above, we first developed economic models on data from counties in the area of the cities (Table 3²). In all cases, the economic models produced forecasts for taxable retail sales (variable TAXABLE), per capita personal income (PCPI), wage and salary income (WAGE) and wage and salary employment (EMP).³ Data on taxable retail sales were obtained from the Illinois Department of Revenue website (<https://www.revenue.state.il.us/app/kob/KOBReport?r=TotalMenu>). Other variables were obtained from the U.S. Bureau of Economic Analysis Regional Data Interactive Data Tool (Tables CAINC1 and CAINC4 from <https://apps.bea.gov/itable/iTable.cfm?ReqID=70&step=1>).

TABLE 3 GOES HERE

A variety of models were used. The criteria for model selection was the minimization of the Mean Absolute Percentage Error (MAPE) statistic. Best fitting models were also checked for remaining serial correlation using the Box-Ljung “Q” Portmanteau statistic and for normality of residuals using a Doornik-Hansen test. The results of the models are generally good. All of the MAPEs are less than 4.00,

² Full results of all forecast models are available from the authors.

³ For Springfield, the variable TAXABLE was better fit in the analysis of financial variables.

indicating our average model predictions were less than four percent from the actual realized value, and the majority were less than 2.00.

The results of the economic forecasts were then used to forecast financial variables over the sample period. Table 4 shows the models used and the respective fit for each financial variable. Due to inconsistent financial reporting, different variables had to be used in each jurisdiction. We attempted to develop individual forecasts for revenue variables accounting for greater than 10 percent of total revenues, then create an “other revenue” category (variable OTHERREV) for less important revenue sources.

For Springfield, we forecast total property valuation (TOTVALUATION) and then apply the current effective property tax rate to create a forecast for property taxes. We also forecast the sales tax base (TAXABLE) separately from the economic variables through including it in the vector autoregression with the financial variables. We then apply the current local sales tax rates to arrive at a forecast for sales tax revenue. Along with these variables we forecast OTHERREV, our operating expenditure measure from equation (1) (OPERATINGEXP), and finally net transfers in to the Governmental Funds (NETTRANSFERS). This is necessary in Springfield because of the recent use of this mechanism to collect a payment in lieu of tax (PILOT) on the city-owned electric and water utility. Rather than listing it as a tax, it is recorded as a transfer in to the General Fund. The fact that the forecast does not fit the data particularly well for this variable can be explained by the irregular use of transfers until 2016 and then a large structural break in the data at that time as the PILOT was enacted (we entered a dummy variable for this year into the model). The other area with less than acceptable forecast error is intergovernmental revenue. This source is challenging to forecast separately in all of our cities due to its inherent volatility. The coefficient of variation for the annual

change in Waukegan's intergovernmental revenue is 12.55, indicating that the standard deviation of the annual change is more than 12 times greater than the average value. This type of volatility makes forecasting extremely difficult but represents real risk to a city if it becomes too reliant (as reflected in most of the ratio measures of fiscal health).

In Waukegan, we forecast total taxes (TAXES) and intergovernmental revenues (IGREV), along with other revenues and operating expenditures. The model does well with the largest categories (TAXES, OPERATINGEXP) but not well with the smaller items. Other revenue also difficult to forecast with the exception of Springfield and Carbondale, where the largest component of other revenues (service charges, permits and fees) has shown steady growth over time.

Normal has revenue sources similar to Waukegan. However, their charges for services revenue (CHARGES) accounts for more than 10 percent of total revenues and is forecast separately. Our forecasting models here work fairly well, except for the aforementioned intergovernmental revenue forecasting challenge and total expenditures. Here we use total expenditures (TOTALEXP) instead of operating expenditures because the city does not report capital outlays separately in their CAFR historical data schedule. Therefore, the forecasts reflect the lumpiness of capital expenditures and the associated volatility reduces forecasting accuracy.

Carbondale breaks out property taxes from sales taxes so we forecast TOTVALUATION and use the current effective property tax rate to forecast property taxes. This base forecasting is unfortunately not available for sales taxes. This is because the city forecasts utility taxes and other types of revenue in this category (SALESSERVICEUTIL) so we forecast it separately. The models here are mostly sufficient, again with the notable exception of intergovernmental revenue.

In Geneseo, we did have some models with fairly significant errors (the obvious outlier here is intergovernmental revenue, however, Theil's U is less than 1.0 indicating the model does have some predictive power). However, in this community taxes account for over 80 percent of revenues and our model here has an acceptable MAPE. We did violate our 10 percent rule for separate forecasts. Intergovernmental revenue and charges each account for around five percent of total revenue. One additional thing to note about these estimates is that something very strange happened in the finances of Geneseo in FY 2013. In the charges and operating expenditure categories, there is an obvious outlier. As with Springfield's structural break, we corrected the forecasts by entering a dummy variable for 2013 into the model.

Somewhat surprising to us, Thornton had data adequate for using a vector autoregressive model. Errors for taxes and operating expenditures here are acceptable. Once again, taxes account for a large share (75%) of total revenue, therefore the errors in intergovernmental and other revenues (each 12-13% of total) are less important in forecasting overall balance.

TABLE 4 GOES HERE

STRESS TESTS

As discussed in the Approach section, forecasts were generated for each year in the sample period along with standard errors in order to develop the standard error based shocks for stress tests. We next used equations (1) and (2) to build forward

looking stress tests of operating balance. Table 5 and Figures 1 – 6 summarize the main results for each city.⁴

Examining the results, we can divide the cities into roughly three categories. Two cities, Geneseo and Thornton, are most likely to be sustainable over the near-to-medium term. Geneseo has very little chance of running an operating deficit, even in extreme situations (shortfall in revenue or increased expenditures of two standard errors. Thornton has slightly more risk of running a large deficit, at one standard error from forecast they have a five-to-seven percent deficit as percent of operating revenues, which would entail some cuts in expenditures or revenue increases. And at two standard errors they would face an operating deficit of around 20 percent.

The second group of cities, Waukegan and Normal, face more risk over the next five years. At one standard error from forecast, Waukegan will run a deficit of 15 to 25 percent of operating revenues, requiring large cuts or revenue increases to maintain solvency. And at two standard errors their average loss increases to over 50 percent, a potentially catastrophic loss. Normal is in a slightly better position, with losses in the 40 percent range at two standard error deviations from forecast.

The third group of cities is in the most fiscal trouble. Even if Springfield and Carbondale hit their forecasts, they will still have significant operating deficits. Carbondale's situation might be slightly exaggerated by our analysis as they do not report capital outlays separately. This might give them room to make cuts in expenditures. However, an average loss of 60 percent at two standard errors is unlikely to be made up by suspending capital outlays alone. Springfield is arguably

⁴ Details available from the authors.

in the direst situation. Their situation in terms of average losses is similar to Carbondale's. However, our calculations already include net transfers, which most recently have been driven by the PILOT on the city-owned utility. This could be increased, but since it increases utility bills, there is a limit to how much it can be raised.

CONCLUSIONS AND LIMITATIONS

In this paper we have developed a model of fiscal sustainability based on the concept of stress testing, analyzing what is likely to happen when jurisdictions are exposed to an unexpected fiscal shock. We provide evidence that it does provide a way to analyze the future prospects for a city's financial situation and that it does provide divergent validity in the sense that it makes predictions that vary by certain characteristics of a city.

In terms of what causes these results, one can only speculate. We can observe past behavior and make some tentative associations. Springfield has been running significant operating deficits since around 2012 and has had to cover it by transfers. Revenue continued to grow at around the same pace as earlier, but expenditures increased dramatically. At the other end of the spectrum, Geneseo's growth in revenues has matched its growth in operating expenditures, producing a consistent operating surplus. Whether this is reflective of a fiscal management philosophy or exogenous factors that affect Springfield more than Geneseo, one can only speculate. But with this tool we can provide an assessment of risk that should help managers and can be the basis for further study into the causes of what we have measured.

There are some limitations of our model that should be discussed. First, although we were able to develop our data, it is still not sufficient to use many of the

forecasting tools we wished to use, such as vector error-correction models for cointegrated data. And the VAR models we used likely were somewhat underpowered in the sense that with more data we could examine longer lag structures. The Schwarz BIC indicated the maximum possible lag with the available data was the one necessary to extract the maximum information out of the data. This means that if we had more data, it is possible that even more lags would improve predictive ability. We were also constrained in building a holdout sample for model testing. This would provide more preferred “pseudo-out of sample” measures of goodness of fit that could be employed in model building.

Another limitation of the model is the timeframe of the analysis. Five years is medium-term at best. However, once we go beyond five years we reach a point where the number of data points pre-forecast is nearly the same as the forecast period. This does not bode well for forecast accuracy, and standard errors would grow dramatically in magnitude. Finally, although by incorporating more revenue sources and expenditure categories into our model we gain more detail, the sacrifice is the possibility of generating more noise. It could be possible that aggregation, while potentially masking changes at lower levels of detail, may generate more stability. This would come out as greater uncertainties in our forecasts. We choose to not address this directly but instead not get too focused on specific breakpoints in assessing sustainability. The differences in potential outcomes across our categories is likely to be more informative than our differences within categories.

Within these limitations, we have developed and documented a candidate for use by practitioners to assess financial sustainability and by academics to examine factors impacting true sustainability. We will continue to develop and use these models and hope that other models are developed and tested.

REFERENCES

- Bhutta, N., Bricker, J., Dettling, L., Kelliher, J., & Laufer, S. (2019). *Stress Testing Household Debt* (Finance and Economics Discussion Series No. 2019-008). Washington, DC.
- Blaschke, W., Jones, M. T., Majnoni, G., & Peria, S. M. (2001). *Stress Testing of Financial Systems: An Overview of Issues, Methodologies, and FSAP Experiences* (No. WP/01/88). Washington, DC.
- Brown, K. (1993). The 10-point test of financial condition: Toward an easy-to-use assessment of tool for smaller governments. *Government Finance Review*, 9(6), 21.
- Burnside, C. (2005). *Fiscal Sustainability in Theory and Practice: A Handbook*. Washington, DC: The International Bank for Reconstruction and Development/The World Bank.
- Gorina, E., Maher, C., & Joffe, M. (2018). Local fiscal distress: Measurement and prediction. *Public Budgeting & Finance*, 38(1), 72–94.
- Groves, S. M., & Valente, M. G. (1986). *Evaluating Financial Condition: A Handbook for Local Government*. Washington, DC: International City Management Association.
- Hamilton, J. D., & Flavin, M. A. (1986). On the limitations of government borrowing: A framework for empirical testing. *American Economic Review*, 76(4), 808.
- Harvey, A. C. (1981). *The Econometric Analysis of Time Series*. New York: Wiley.
- Hendrick, R. (2004). Assessing and measuring the fiscal health of local governments. *Urban Affairs Review*, 40(1), 78–114.
- Honadle, B. W., Cigler, B. A., & Costa, J. M. (2004). *Fiscal Health for Local Governments: An Introduction to Concepts, Practical Analysis and Strategies*. Amsterdam: Elsevier Academic Press.
- Honadle, B. W., & Lloyd-Jones, M. (1998). Analyzing rural local governments' financial condition: An exploratory application of three tools. *Public Budgeting & Finance*, 18(2), 69.
- Ji, H., Ahn, J., & Chapman, J. (2016). The role of intergovernmental aid in defining fiscal sustainability at the sub-national level. *Urban Studies*, 53(14), 3063–3081.

- Kriz, K. A. (2015). Is there an optimal size of fiscal reserves for local governments? In *Local Government Budget Stabilization: Explorations and Evidence* (pp. 171–190). Cham: Springer International Publishing.
- Levine, C. H., Rubin, I. S., & Wolohjian, G. G. (1981). Resource scarcity and the reform model: The management of retrenchment in Cincinnati and Oakland. *Public Administration Review*, 41(6), 619–628.
- Lewis, M. (2011). *Boomerang: Travels in the New Third World*. New York: W.W. Norton & Co.
- Mahdavi, S., & Westerlund, J. (2011). Fiscal stringency and fiscal sustainability: Panel evidence from the American state and local governments. *Journal of Policy Modeling*, 33(6), 953–969.
- Maher, C. S., & Nollenberger, K. (2009). Revisiting Kenneth Brown’s “10-point test”. *Government Finance Review*, 25(5), 61–66.
- McDonald, B. D. (2015). Does the charter form improve the fiscal health of counties? *Public Administration Review*, 75(4), 609–618.
- Mennis, G., Banta, S., & Draine, D. (2018). *Assessing the Risk of Fiscal Distress for Public Pensions: State Stress Test Analysis* (M-RCBG Associate Working Paper Series No. 92). Boston, MA.
- Mills, T. C. (1990). *Time Series Techniques for Economists*. Cambridge, UK: Cambridge University Press.
- Quintos, C. E. (1995). Sustainability of the deficit process with structural shifts. *Journal of Business & Economic Statistics*, 13(4), 409–417.
- Schuermann, T. (2014). Stress testing banks. *International Journal of Forecasting*, 30(3), 717–728.
- Sorge, M. (2004). *Stress-Testing Financial Systems: An Overview of Current Methodologies* (BIS Working Papers No. 165). Basel, CH.
- Wiesenthal, J. (2011). The Meredith Whitney effect in one big, beautiful chart. *Business Insider*. Retrieved April 19, 2019, from <https://www.businessinsider.com/the-meredith-whitney-effect-2011-1>

Table 1. Decision Items for a Stress Testing Model and Our Model Definitions.

Decision Item	Our Model
1. Type of Risk Model	Financial – Operational Risk Model
2. Type of Stress Test	Other – Effects of an Unexpected Shortfall in Revenues or Increase in Expenditures
3. Type of Shock	Underlying Volatilities
4. Type of Scenario	Hypothetical – Based on Standard Errors of Forecasts
5. Core Assets to be Shocked	Operating Balance (equation (1))
6. Peripheral Assets to be Shocked	Revenues of Various Types, Operating Expenditures
7. Size of Shocks	One and Two Standard Error Deviations from Forecasted Values
8. Time Horizon	Current and Four Future Years (2019-2024)
9. Risk Management Techniques	None - Passive Acceptance of Risk

Source: Decision items adapted from Blaschke, et.al. (2001, pp. 4-6).

Table 2. Cities in the Sample.

City	Population (2017 est.)	Data Availability	Region
Springfield	114,868	2003-2018	Central
Waukegan	87,729	2000-2018	Northeast
Normal	54,284	2005-2018	Central
Carbondale	25,899	2004-2018	South
Park Forest	21,682	2001-2018	Northeast
Prospect Heights	16,180	2002-2018	Northeast
Hawthorn Woods	8,412	1999-2017	Northeast
Hillside	8,043	2001-2018	Northeast
Geneseo	6,533	1999-2018	West Central
Thornton	2,488	2003-2018	Northeast
Oakbrook Terrace	2,161	2005-2018	Northeast

Table 3. Forecasting Models and Fit, Economic Variables.

City	Variable and Final Model	MAPE (In-Sample)
Springfield	PCPI: VAR(1)	0.92
	WAGES: VAR(1)	1.18
	EMP: VAR(1)	0.79
Waukegan	TAXABLE: AR(1) w/PCPI	1.86
	PCPI: ARIMA (0,1,0) w/WAGES	2.49
	WAGES: ARIMA (0,1,0) w/EMP	2.16
	EMP: ARIMA (1,0,1)	1.06
Normal	TAXABLE: AR(1) w/PCPI, EMP	3.99
	PCPI: ARIMA (1,1,0)	1.50
	WAGES: ARIMA (0,1,0) w/EMP	0.94
	EMP: MA(2)	0.61
Carbondale	TAXABLE: VAR(2)	0.91
	PCPI: VAR(2)	0.62
	WAGES: VAR(2)	0.68
	EMP: VAR(2)	0.26
Geneseo	TAXABLE: ARIMA (1,0,[4]) w/PCPI, WAGES	2.39
	PCPI: AR(1) w/EMP	3.02
	WAGES: ARIMA ([4],0,0) w/PCPI	1.82
	EMP: AR(1)	1.79
Thornton	TAXABLE: VAR(2)	2.14
	PCPI: VAR(2)	1.28
	WAGES: VAR(2)	0.71
	EMP: VAR(2)	1.43

Notes: (1) VAR(X) is a vector autoregression with lag length X . Optimal lag lengths determined by the Schwarz Bayesian Information Criteria; (2) AR(1) models indicate a first-order autoregressive model, estimated by the Kalman filter except in the case where independent variables were included (designated by “w/YYYY” where YYYY is the independent variable included). Those models were estimated by the Cochrane-Orcutt iterative procedure; (3) ARIMA(P,D,Q) indicates an Autoregressive Integrated Moving Average model, estimated by the Kalman filter, where P are the number of autoregressive lag terms in the model, D is the order of differencing necessary to induce stationarity, and Q are the number of moving average terms in the model. A bracket around a term indicates that only that lag was entered into the model (therefore, an ARIMA(2,0,0) would be have autoregressive lag terms 1 and 2 in the model whereas an ARIMA([2],0,0) would have only autoregressive lag term 2; (4) Methodologies discussed in Harvey (1990) and Mills (1990).

Table 4. Forecasting Models and Fit, Financial Variables.

City	Variable and Final Model	MAPE (In-Sample)
Springfield	TOTVALUATION: ARIMA(0,1,0) w/PCPI	1.07
	TAXABLE: ARIMA(0,1,0) w/PCPI	1.93
	IGREV: OLS w/PCPI, WAGES, EMP	10.09
	OTHERREV: ARIMA([2],0,0) w/PCPI, EMP	7.44
	OPERATINGEXP: AR(1) w/PCPI	3.13
	NETTRANSFER: ARIMA(0,1,1) w/Y2016	45.55
Waukegan	TAXES: VAR(2) w/PCPI	3.57
	IGREV: VAR(2) w/PCPI	34.79
	OTHERREV: VAR(2) w/PCPI	10.53
	TOTEXP: VAR(2) w/PCPI	4.04
Normal	TAXES: AR(1) w/PCPI, TAXABLE	3.50
	IGREV: AR(1) w/PCPI, TAXABLE	11.93
	CHARGES: AR(2) w/PCPI	1.14
	OTHERREV: AR(1) w/PCPI, TAXABLE, WAGES	5.34
	OPERATINGEXP: OLS w/PCPI, TAXABLE, WAGES	22.93
Carbondale	TOTVALUATION: ARIMA(2,1,0)	1.27
	SALESSERVICEUTIL: AR(1)	8.14
	IGREV: AR(1) w/TAXABLE, EMP	10.52
	OTHERREV: ARIMA(0,0,[2]) w/TAXABLE, EMP	3.02
	TOTALEXP: AR(1) w/PCPI, EMP	1.40
Geneseo	TAXES: ARIMA(0,1,0) w/TAXABLE, EMP	3.58
	IGREV: AR(1) w/TAXABLE, EMP, WAGES	48.88
	CHARGES: OLS w/ TAXABLE, EMP, WAGES, Y2013	8.62
		15.52
	OTHERREV: ARIMA([1,4],0,0) w/PCPI	2.94
	OPERATINGEXP: ARIMA(2,0,[3]) w/PCPI, WAGES, Y2013	
Thornton	TAXES: VAR(1) w/PCPI	3.86
	IGREV: VAR(1) w/PCPI	10.57
	OTHERREV: VAR(1) w/PCPI	11.56
	OPERATINGEXP: VAR(1) w/PCPI	5.70

Notes: Notation and methodology sources are as Table 3, except OLS which is Ordinary Least Squares regression.

Table 5. Stress Test Results, Sample Cities, 2019-2024.

City	Average Operating Balance (% of Operating Revenue)		Low Operating Balance (Year)	
Springfield	FORECAST:	-17.24%	FORECAST:	-9.62% (2019)
	FORECAST -1 SD:	-32.22%	FORECAST -1 SD:	-21.32% (2019)
	FORECAST -2 SD:	-50.02%	FORECAST -2 SD:	-34.74% (2019)
Waukegan	FORECAST:	4.1%	FORECAST:	2.12% (2023)
	FORECAST -1 SD:	-20.92%	FORECAST -1 SD:	-44.47% (2023)
	FORECAST -2 SD:	-54.86%	FORECAST -2 SD:	-100.00% (2023)
Normal	FORECAST:	6.66%	FORECAST:	-0.97% (2019)
	FORECAST -1 SD:	-18.90%	FORECAST -1 SD:	-25.67% (2019)
	FORECAST -2 SD:	-42.43%	FORECAST -2 SD:	-49.79% (2019)
Carbondale	FORECAST:	-12.89%	FORECAST:	-19.85% (2022)
	FORECAST -1 SD:	-33.20%	FORECAST -1 SD:	-42.53% (2022)
	FORECAST -2 SD:	-60.10%	FORECAST -2 SD:	-73.23% (2022)
Geneseo	FORECAST:	23.13%	FORECAST:	22.03% (2019)
	FORECAST -1 SD:	13.28%	FORECAST -1 SD:	11.17% (2023)
	FORECAST -2 SD:	1.19%	FORECAST -2 SD:	-3.13% (2023)
Thornton	FORECAST:	6.10%	FORECAST:	5.39% (2021)
	FORECAST -1 SD:	-6.47%	FORECAST -1 SD:	-7.77% (2022)
	FORECAST -2 SD:	-21.34%	FORECAST -2 SD:	-23.74% (2022)

Figure 1. Operating Balance as Percent of Operating Revenue, Springfield.

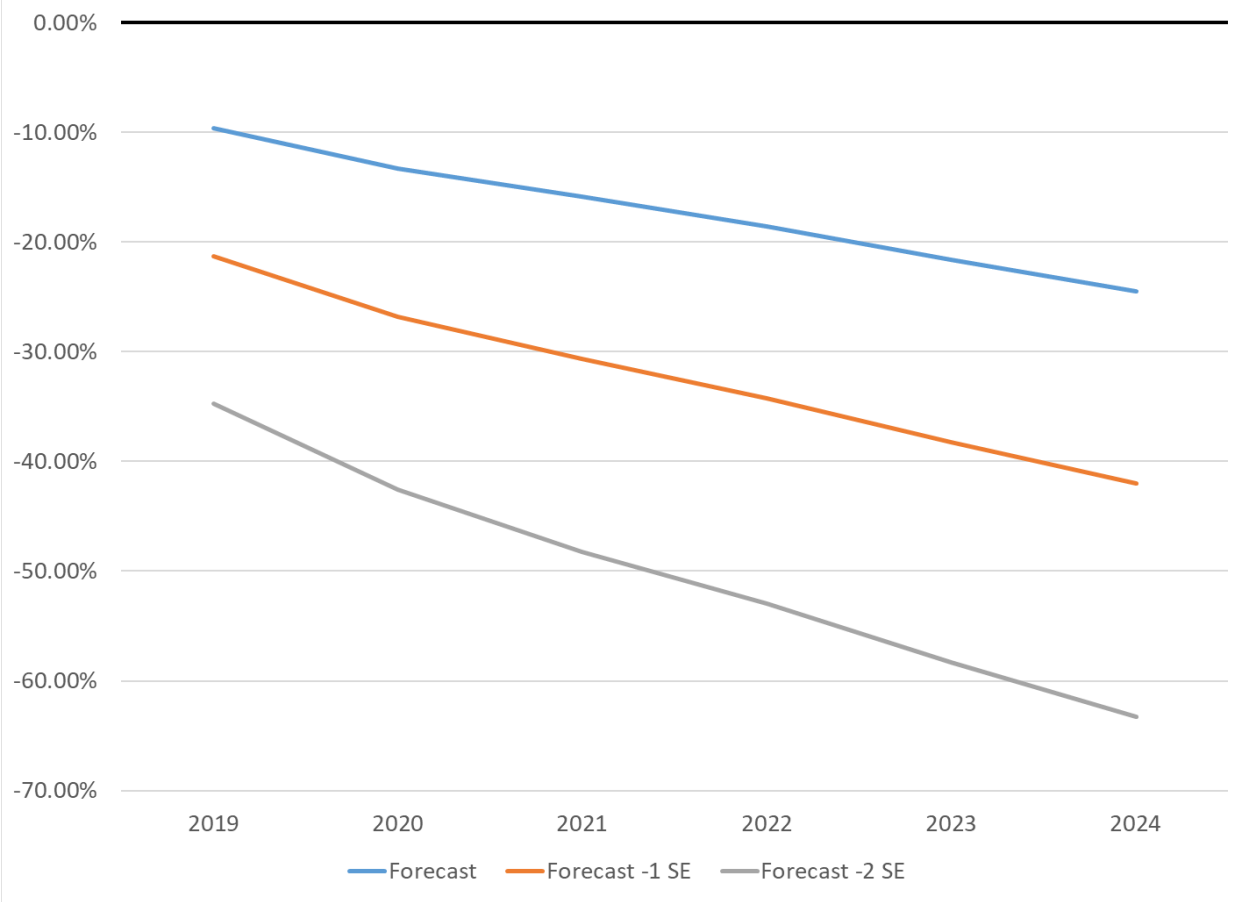


Figure 2. Operating Balance as Percent of Operating Revenue, Waukegan.

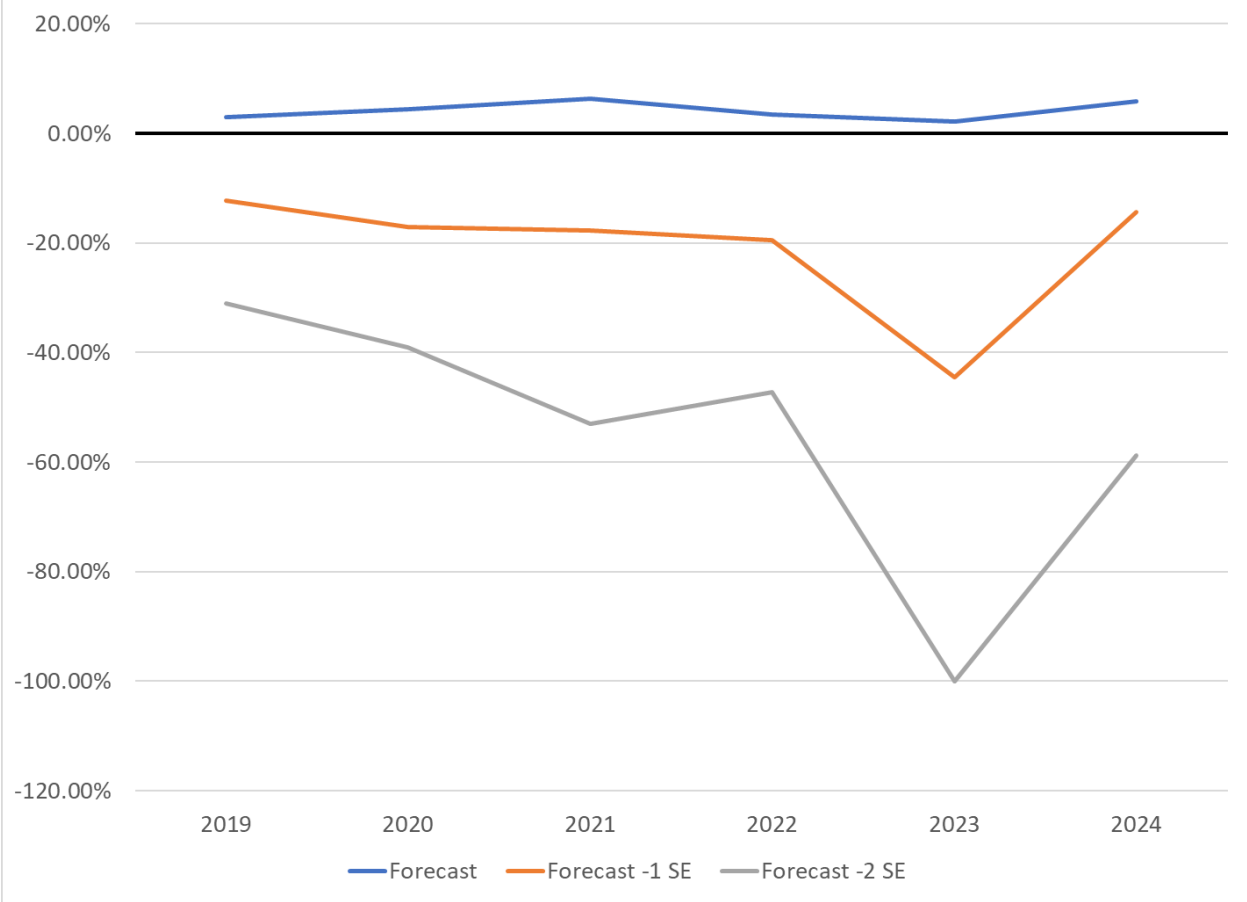


Figure 3. Operating Balance as Percent of Operating Revenue, Normal.

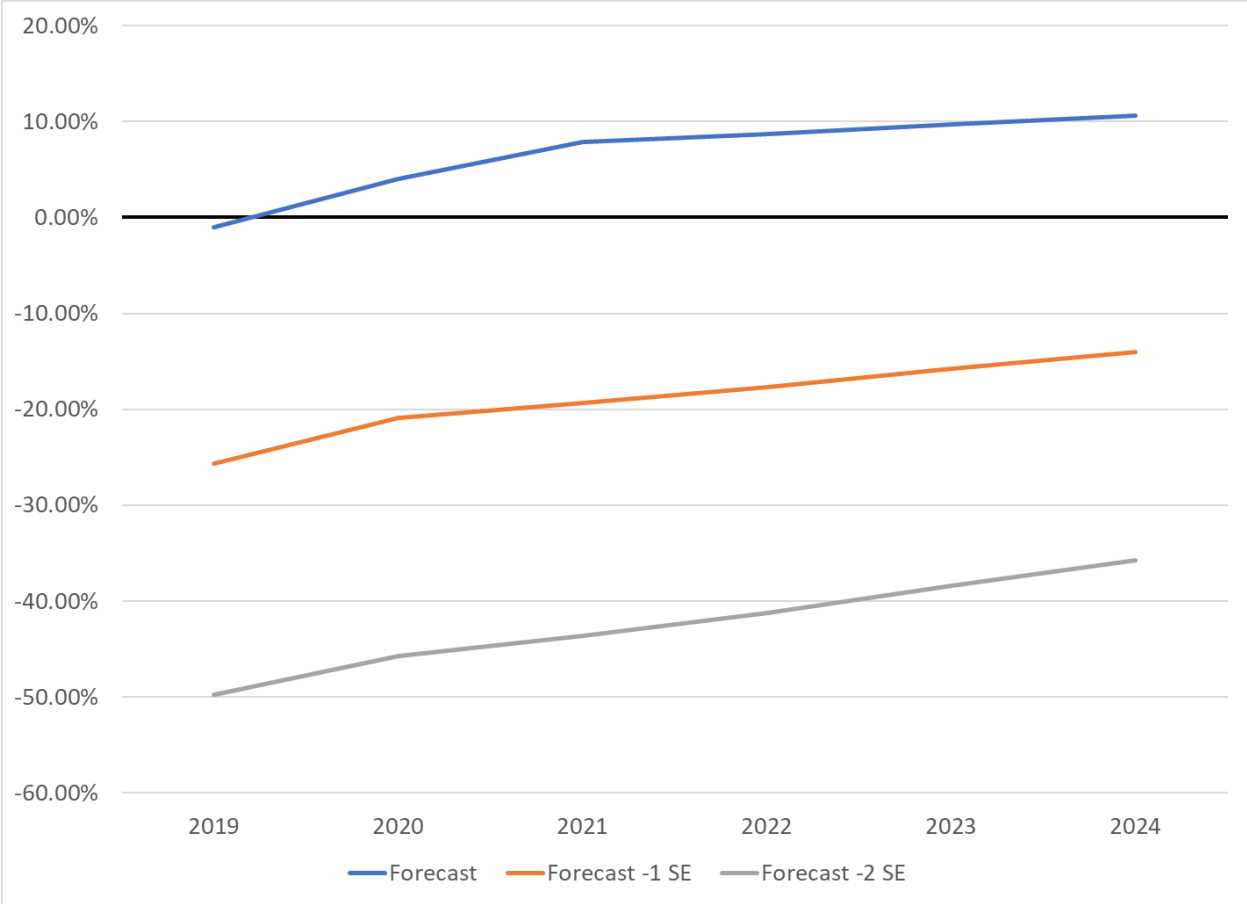


Figure 4. Operating Balance as Percent of Operating Revenue, Carbondale.

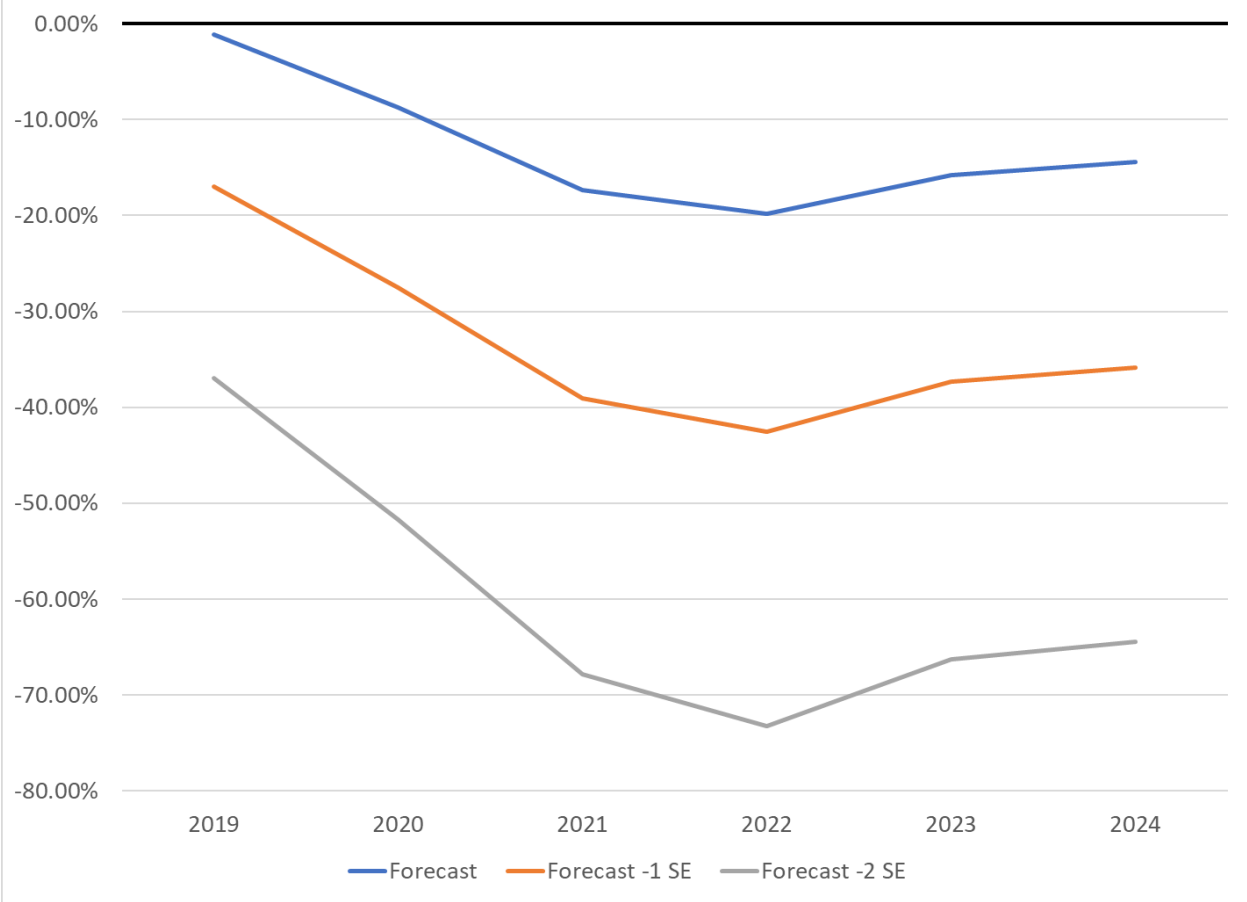


Figure 5. Operating Balance as Percent of Operating Revenue, Geneseo.

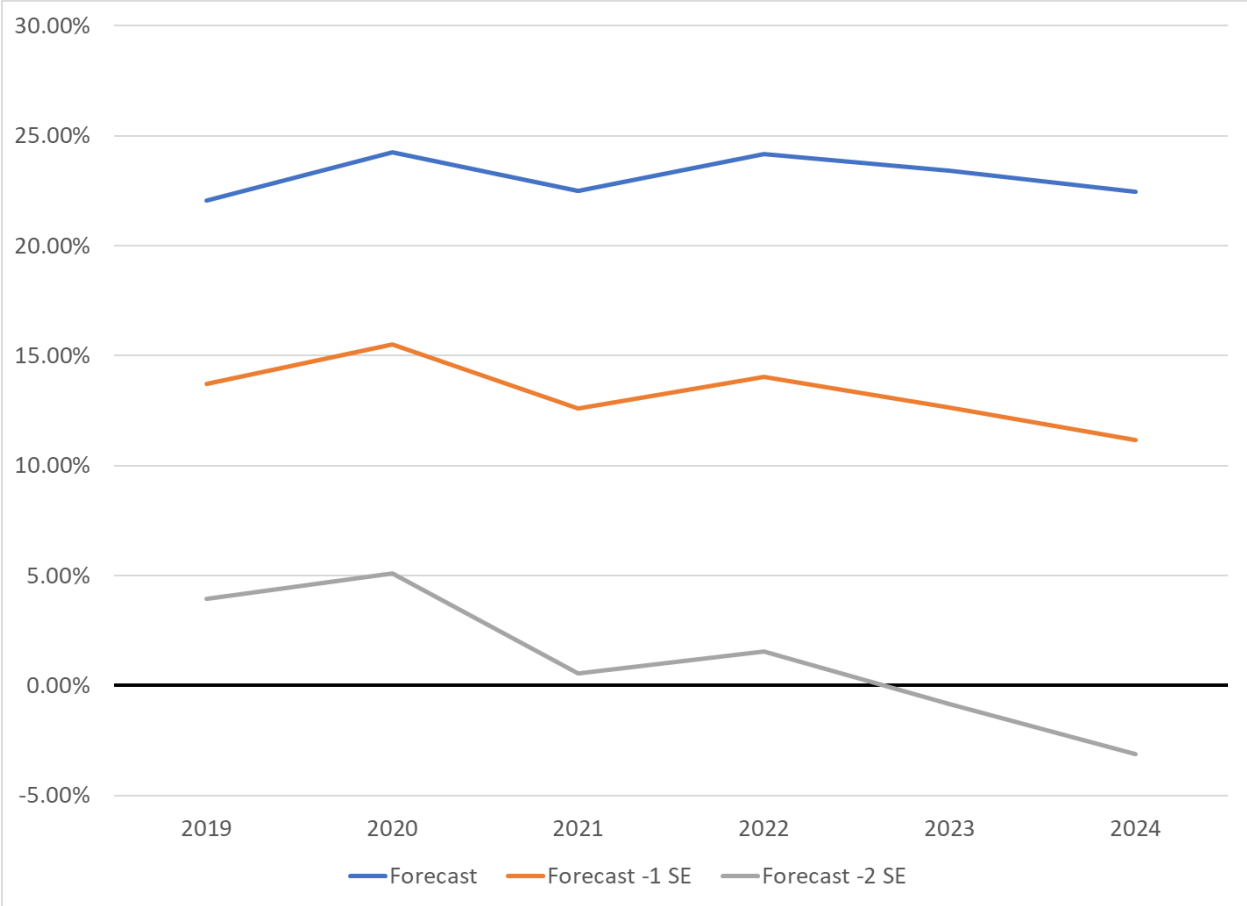


Figure 6. Operating Balance as Percent of Operating Revenue, Thornton.

