Measure Thrice, Punctuate Once

Assessing Measures of Local Budgetary Change

Samuel Workman
Institute for Policy Research and Public Affairs, West Virginia University

Herschel F. Thomas
LBJ School of Public Affairs, University of Texas

August 31, 2024

Contact:
Samuel Workman, Director
Institute for Policy Research and Public Affairs
West Virginia University
Woodburn Hall, 221D
1515 University Ave
Morgantown, WV 26506
sam.workman@mail.wvu.edu

Acknowledgments: This research is based on work supported by the *Center for the Study of Federalism* and the *Institute for Policy Research and Public Affairs* at *West Virginia University* (WVU). Findings, recommendations, conclusions, or errors are those of the authors and do not reflect the views of the *Center for the Study of Federalism*. We thank E.J. Fagan, Chris Koski, and Christian Breunig for valuable feedback on this manuscript.

Abstract

A rich literature in public policy studies uses public budgeting to assess public and institutional priorities and analyze how institutional or policymaking systems exacerbate or smooth budget punctuations (Baumgartner et al. 2017; Jones et al. 2009; Breunig and Koski 2006). More recently, a debate has emerged over the best way to measure the punctuated or "tailedness" of budgetary change distributions. The traditional approach using l-kurtosis of budgetary change distributions (Jones, Sulkin, and Larsen 2003; Breunig and Jones 2011) has been challenged by proponents of t-scores, Gini coefficients, or generalized paretian fits as better measures of tail behavior (Fernández-i-Marín et al. 2019; Kaplaner and Steinebach 2022; Workman, Robinson, and Bark 2024). Do these measures behave differently when comparing governing or policy systems regarding budgetary change distributions? We use data on budgetary change in Appalachia to assess which measures are best or if they matter substantively when comparing counties and municipalities. After constructing each measure for counties and cities, we statistically assess how much change in ordering we observe among the local governments, comparing the measures. Our findings are important for assessing the link between theory and measurement in public policy, especially for punctuated equilibrium studies. The findings also matter for real-world governance and auditing. They offer insights into comparing counties and cities that are not attainable with the usual attention to geographic or population size.

Introduction

In this research note, we examine three measures of policy change found in the literature on Punctuated Equilibrium Theory (PET). Two emergent approaches seek to offer better conceptual fits to PET's core conceptual framework. We use data on local government expenditures in Appalachia, specifically West Virginia (WV), to assess the measures, address correlations, and ask if these measures are substantively different and lead to substantively different conclusions when comparing real-world governing systems.

While addressing conceptually distinct features of PET, we find that two of the measures correlate well and are not substantively different when comparing governing systems. The analysis also offers the opportunity for theoretical development in PET in understanding the non-linear relationship between government revenues and other resources and the nature of policy punctuations. Our study also highlights the important role of geospatial features in promoting extreme punctuations in policymaking and centers counties as important governing units for understanding policy change, especially in rural areas that are generally understudied in political science.

Background

Scholars working in PET have collectively been using l-kurtosis to measure policy punctuations for a long time (Breunig and Koski 2006; Breunig and Jones 2011; Jones et al. 2009; Breunig 2011). L-kurtosis is particularly useful because it provides a direct measure of the tails of a distribution (i.e., policy punctuations) that is robust to the small sample sizes policy scholars usually have at hand. Because l-kurtosis is scalable, it provides a comparable measure for institutions, governing, and decision systems across time and space. L-kurtosis tends to be high in the presence of empirical policy change distributions that have “fat tails” and peaked center mass. However, there has been ample debate lately about whether l-kurtosis is a good conceptual, if not empirical, fit to the features of data that PET wants to study.

What are these features? As a theory, PET describes policy change as occurring in small incremental bits punctuated by drastic, quick, and large departures from the incremental pattern (Baumgartner and Jones 2009; Jones et al. 2009). This pattern is evident the world over in a host of countries (Breunig 2011), governing systems that range from Western Democracies to authoritarian regimes (Baumgartner et al. 2017; Jones, Epp, and Baumgartner 2019), and

political and economic institutions and organizations (Jones, Sulkin, and Larsen 2003; Epp 2018; Workman, Robinson, and Bark 2024).

This ubiquitous pattern of policy change is caused by shifting attention (Jones and Baumgartner 2005), how governing systems process information (Fagan 2022; Workman, Shafran, and Bark 2017), and institutional and organizational features (Workman, Robinson, and Bark 2024; Andersen and Mortensen 2009), including subsystems (May, Sapotichne, and Workman 2009). The temporal dynamics of policy change also play heavily in the story PET tells about how, when, and why policy changes, and does so drastically more often than we would or should expect.

Recent debates abound concerning the conceptual fit of the classic measure, l-kurtosis, to what PET wants to measure. One challenger to the measurement hegemony of l-kurtosis is the t-distribution (Fernández-i-Marín et al. 2019). These scholars argue that incrementalism and punctuation are separate features of PET that are well represented by a t-distribution. In particular, this work suggests that l-kurtosis, as a measure of tails, underestimates the stasis or incrementalism in policymaking. Theoretically and conceptually, l-kurtosis is a poor measure of policy subsystems—the most potent cause of policy incrementalism or stasis.

The second challenger is the Gini coefficient proposed by Kaplaner and Steinebach (2022). They center PET's concern with temporal dynamics as the key to understanding policy change. In particular, they argue that the Gini coefficient better captures how drastic policy change is disproportionately evident in a few temporal periods. In other words, the influx of information creates an urgency problem, requiring governing and decision systems to act in haste and stretching their capacity to prioritize, leading to large departures from incrementalism. Gini coefficients measure inequality. In this case, punctuations occur unequally across periods.

Core Research Questions

These measurement approaches are powerful because they relate empirical measures to two of PET's core conceptual features: subsystems and temporal dynamics. Given this, we set out to compare the three measures—Gini coefficients, t-distributions, and l-kurtosis to a population of local government expenditures (the canonical type of data a PET analyst might use). Comparison over time and space are key features of any good research design. We ask four core questions of the measures:

1. Do these measures broadly correlate across the set of governing systems we examine?

2. Are these measures so different as to cause *substantive* differences in comparison across policy output distributions?
3. Do the measures lead to different substantive conclusions when comparing one governing system to another regarding punctuated policymaking?
4. How might these measures contribute to theory development in PET?

Research Design

Our research design uses comprehensive data on local government expenditures across time to compare each measure of punctuated policy change. In this case, we collect expenditure data from all 55 counties in WV for FY 2012–2023, comprising the whole population for the period. This prevents issues with sampling and over-reliance on larger, more populated areas, thus reducing the comparison to those features attributable to each measure. The data collection is a heavy lift and does much to control for variation not attributable to the nature of each measure.

After collecting data on county expenditures, we ensure this data is comparable across counties and years regarding substantive expenditure categories and accounting for inflation. From there, our methodological approach is a straightforward comparison of measures. We are specifically interested in whether the measures lead to different substantive comparisons among county expenditure mixes. As such, we are interested in the measures' correlation (or lack thereof) and how the substantive rankings of counties in terms of punctuatedness might change, moving from one measure to another.

Data

To assess these comparative measures for policy change, we collect expenditure data for counties in Appalachia, specifically West Virginia. The US Census Bureau's Census of Governments is a perennial source of data on revenues and expenditures in local governments. However, we opt for data collected from the WV State Auditor's office. These data are more complete and detailed and represent the entire population of local expenditures for the state. We collect data for all 55 counties in the state at the general and sub-category levels.

These data are available in digital format on the website of the State Auditor's Office's Local Government Division.[1] The expenditure tables, stored in PDF format, date back to FY 2012. Our

[1] See: https://www.wvsao.gov/LocalGovernment/Default.

data cover FY 2012–2023, with updates for FY 2024 underway. For each county, we collect the PDFs manually and transfer the data to CSV files via a macro deployed by the authors and research assistants.

Once these data were collected, manual manipulation was required to ensure compatibility and comparability across time and space. Following (Jones et al. 2009), we adjusted the budget documents to ensure the general and sub-categories were comparable across all years and counties. The budget documents contain expenditure and revenue categories from eight general categories (e.g., capital projects, culture and recreation, general government, public safety, revenues, etc.) and 223 specific sub-categories. The latter include but are not limited to, functions related to mental health, emergency services, libraries, rail trails, road maintenance, wastewater, public health programs, and many other functions.

To construct the overall data infrastructure, we parse and append each budget file for county expenditures as reported to the Auditor. Expenditure data in WV is provided in four variables, including a revised expenditure category, reconciling various tensions in the original reported data, and information on the state's coal severance tax revenues—an important source of revenue in the state for most years of our data. We use the revised general fund as our primary variable of interest and restrict the study to 164 sub-categories (Johnston, Pagano, and Russo 2000; Pagano and Johnston 2000). For counties in the data, this yields 2,358 sub-category panels with enough observations to calculate percentage changes across FY 2012–2023 (N = 18,834 percentage changes). Because our data must be comparable across time and space, we adjust for inflation using the GDP-based deflator the St. Louis Federal Reserve Bank provides.[2] We use the FY 2012 deflator for October of each FY. With data that is reliable and comparable across space among counties and over time among each county's spending, we can construct the three measures of expenditure punctuations identified in the literature that inform current measurement debates in the study of punctuated equilibrium.

Measures

We assess three measures emergent in the literature on PET and how they relate to policy punctuations, specifically expenditures in our case. Our endeavor is not to adjudicate among these measures and make prescriptions but to outline how they differ if they do, assess if these differences are substantive, and articulate how they help us understand the sources of policy

[2] See: https://fred.stlouisfed.org/series/GDPDEF

punctuations. These measures are so ubiquitous across the social sciences that we do not provide equations for them here but cite the key pieces in which they are developed. We encourage the reader to consult these for a technical explanation of each.

The first of these measures, l-kurtosis, has been the standard for almost two decades of PET research ((Breunig and Koski 2006; Hosking 1990). Kurtosis is the fourth moment around the mean of the distribution, giving a distribution's shape. Standard kurtosis measures are unstable in small sample sizes, which characterizes much social science research. L-kurtosis provides a measure of the shape of the distribution that is more robust to these small sample sizes. L-kurtosis specifically pertains to the fat tails and slender peaks of a policy change distribution—a point of much contention in the research cited above.

The second is the Gini coefficient. PET has maintained that temporal dynamics are key to understanding policy change (Baumgartner and Jones 2009). In arguing for using the Gini coefficient to capture punctuated equilibrium dynamics, Kaplaner and Steinebach (2022) rightly note that prioritization becomes harder as information ebbs and flows and creates urgency for decision-making. They argue that the Gini coefficient calculated over time better captures this temporal dynamic and the ebb and flow of information that makes prioritization difficult.

The third measure is the degrees of freedom of a t-distribution proposed by (Fernández-i-Marín et al. 2019). The benefit of estimating the t-distribution's degrees of freedom is that the t-distribution takes account of the entire shape of the distribution rather than particular features (e.g., tails for l-kurtosis or time for Gini coefficients). In this sense, the t distribution is argued to better model the underlying substantive features of policy change, resulting in both fat tails and peaked centers.

Methodological Approach

We contribute to the literature on the nature of policy change by comparing three proposed measures of policy punctuation. At the outset, our county expenditure categories have been made uniform across counties so that they are consistent and comparable. The expenditure values have also been adjusted for inflation to FY 2012 dollars.

Our approach begins by calculating the measures for l-kurtosis and Gini coefficients for all 55 counties pooling across time and sub-functions as is typical in the PET literature on budgetary

change.[3] For the t-distribution degrees of freedom, these are fits of a t-distribution to the same pooled data and are estimates rather than calculations. We should note here that fitting t-distributions of such low degrees of freedom to empirical rather than simulated data is difficult. It does not take many degrees of freedom before the t-distribution starts to approximate a normal distribution. In particular, the shoulders of the distribution contain a lot more probability, or cases, than is typical of the empirical expenditure distributions found in most of the literature on PET.

Once we have calculated or estimated the measures of expenditure punctuation characterizing the distributions, we first examine the simple correlations of the measures across counties. In a broad sense, correlation tells us that the measures are tapping the same broad features of the pooled expenditure changes across counties. Strong correlations might also indicate that the measurement debates are really about conceptual differences than substantive, empirical ones.

What really matters, however, is whether different measures would lead us to draw different substantive conclusions from comparing the counties. To see if measurement differences are substantive, we construct rank orders of the counties in terms of each measure and use a rank-order test to see if these differences are statistically significant. If the rank-order tests are significant, this lends credence to the notion that one might order the counties differently depending on which measure one uses to gauge policy punctuations.

Finally, we assess each county's average change in rank when pairing measures. This simple approach sheds light on how county ranking might change due to using one measure or another. Comparing change-in-rank also tells us whether one or more of the measures tap into fundamentally different features of the empirical expenditures and county comparisons drawn from them.

Finally, we map the comparative measures of expenditure punctuations to contextualize what the measures tell us about punctuations in county expenditures. We use this as a springboard to generate hypotheses about what drives policy punctuations generally beyond the current status of PET budgetary literature. The analysis also moves to conjecture about more parametric approaches to modeling policy punctuations generally.

---

[3] In calculating l-k at the county level we drop one extreme observation: Marion County Capital Projects Expenditure with 1.4m percentage change.

Findings

Table 1 displays the simple correlations of the measures across counties. Recall the data are percentage change expenditures from FY 2012–2023 pooled over expenditure sub-categories for each county. So, the correlations pertain to calculated or estimated measures across counties. Table 1 shows high correlation for l-kurtosis and the Gini coefficient for expenditures at almost 0.85. In a broad sense, l-kurtosis and Gini move similarly over the set of counties in our data.

Table 1. Correlation of Measures of Distributional Characteristics in County Expenditures

Measure	l-kurtosis	Gini Coefficient	t - distribution
l-kurtosis	1.000	0.848	-0.192
Gini Coefficient	0.848	1.000	-0.279
t - distribution	-0.192	-0.279	1.000

Source: Calculated by the authors. Underlying data collected by the *Institute for Policy Research and Public Affairs* from the West Virginia State Auditor's Office.

Table 1 also shows the estimated t-distribution stands out from the calculated measures. The correlation is small and negative at -0.19 for l-kurtosis and -0.28 for the Gini coefficient. In a broad sense, the t-distribution does not move with the l-kurtosis or Gini coefficient over the set of West Virginia counties. It is important to remember that the l-kurtosis and Gini coefficient are calculations that assess the tailedness of the policy distributions. However, when fitting the t-distribution, the entire shape of the distribution must be fitted to the empirical data—center mass, dispersion, and effectively the shape. As we note in prior research (Workman, Thomas, and Connor 2022), Appalachian county and municipal budgets exhibit degrees of punctuation on par with authoritarian regimes—meaning extremely narrow shoulders, peaked centers, and fat tails. Fitting a t-distribution to empirical data of this nature is difficult when the entire shape must be faithful to the data.

As noted above, these broad correlations say little substantively. Table 2 contains the results of Kendall's $\boldsymbol{\tau}$ rank-order correlation and test. Table 2 exhibits strong evidence that the strong correlation between l-kurtosis and the Gini coefficient is maintained in the rank ordering of the counties. All the paired measures reach conventional levels of statistical significance. The rank

order correlations between the calculated measures and the t-distribution remain weak, though statistically significant.

Table 2. Comparative Rank-Order Correlations for Paired Distributional Measures

Comparison	Kendall's τ	Z-Score	P-Value
l-kurtosis vs Gini	0.848	9.134	0.000
l-kurtosis vs t - df	0.192	2.069	0.039
Gini vs t-df	0.279	3.013	0.003

Source: Calculated by the authors. Underlying data collected by the *Institute for Policy Research and Public Affairs* from the West Virginia State Auditor's Office, Local Government Division.

After compiling the rank orders of counties and conducting the Kendall $\boldsymbol{\tau}$ pairings in Table 2, we calculated the average number of rank-order changes for each county in the paired measures. In other words, if we compare any two of the measures, on average, how many ranks does a typical county "move" in rank order?

Figure 1 displays each measure pair's average rank-order movement for a county. For l-kurtosis and the Gini coefficient, the average change in rank for a county is 3–4 ranks. When paired with the estimated fit of the t-distribution, both l-kurtosis and the Gini coefficient display an average change in rank of 14–15. Further investigation is needed to understand whether changing 3–4 ranks in our counties is substantive. However, an average change in rank of 14–15 is most certainly substantive among 55 counties. This does not mean that l-kurtosis or the Gini coefficient is necessarily a better measure. Still, it does mean that substantive comparisons using a t-distribution will lead to very different conclusions among the counties for this data.

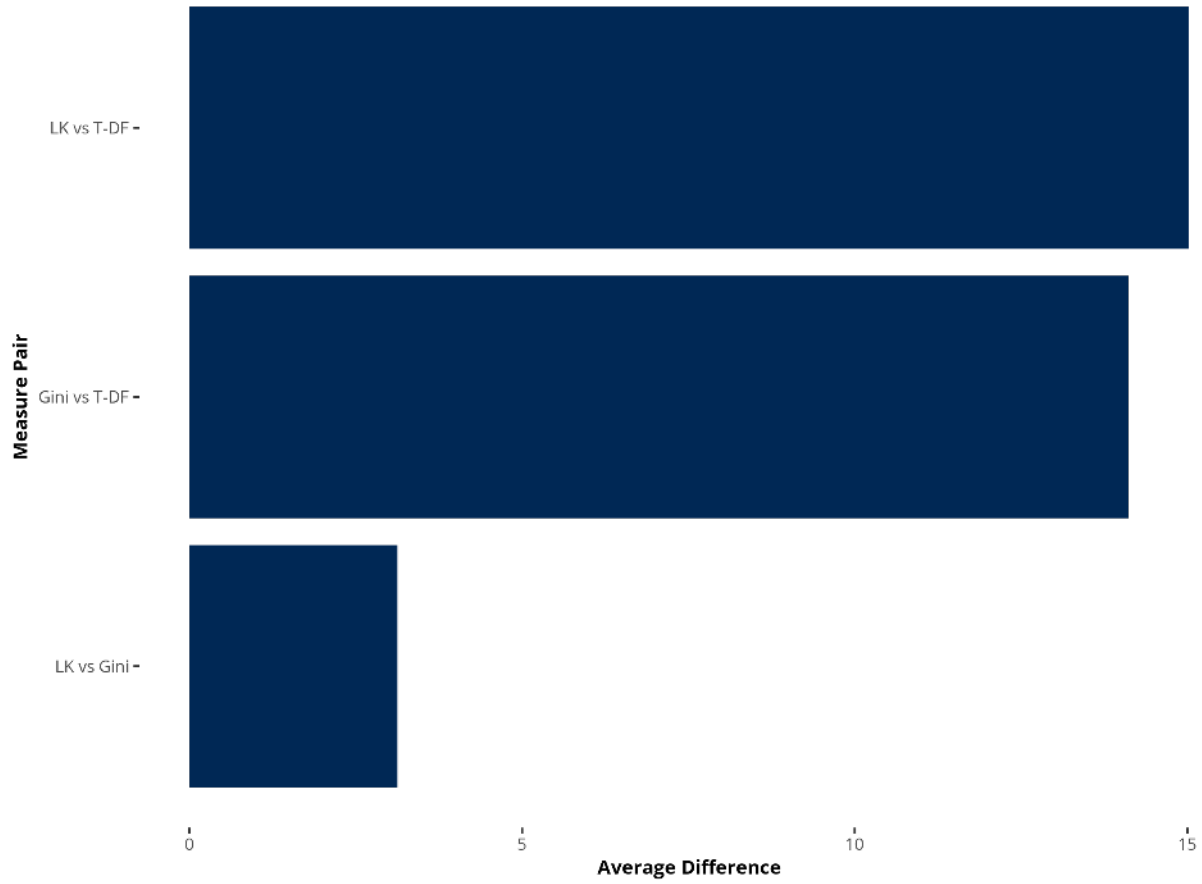

Figure 1. Average Difference in County Ranking for each Paired Measure.
Source: Calculated by the authors from data collected by the Institute for Policy Research and Public Affairs from the West Virginia State Auditor's Office, Local Government Division.

Figure 2 maps the expenditure punctuation measures for the West Virginia counties to contextualize these changes further. The left map displays l-kurtosis, the middle displays the Gini coefficient, and the right shows the t-distribution. The substantive similarity in comparing counties is evident in the first two maps. l-kurtosis and the Gini coefficient are nearly identical, substantively if not in magnitude when comparing counties. Comparing these measures to the t-distribution on the right shows that different conclusions emerge immediately. The l-kurtosis and Gini coefficient measures draw attention to the substantive sources of expenditure punctuations and their non-linearity.

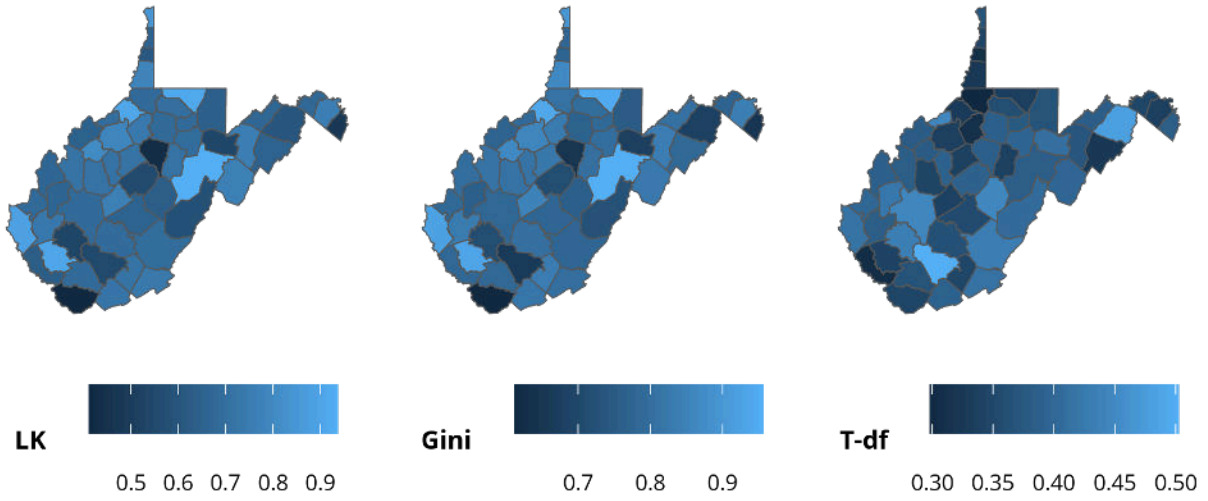

Figure 2. Maps of Each Measure by County in West Virginia. Source: Developed by the authors from data collected by the Institute for Policy Research and Public Affairs from the West Virginia State Auditor's Office, Local Government Division.

The maps show that extreme punctuations are found in two types of counties—well-developed, economically flourishing counties (e.g., Monongalia) and those that experience extreme resource strain (e.g., McDowell). How does one reconcile these data features? The literature on PET suggests that expenditure fluctuations result from difficulty in processing information about policy problems (Fagan 2022). Unlike cities, however, counties in West Virginia all have the same basic institutional features. Furthermore, no city in West Virginia is so large as to exceed it county boundaries in a way larger cities might.

When we observe expenditure volatility in distressed counties like McDowell, the institutional features of federalism loom large (López-Santana and Rocco 2021). Most of the time, along most subcategories of expenditure, the low inflow of revenues and resources means very little movement in expenditures—exacerbating already very peaked change distributions. In these circumstances, prioritization weighs less heavily as there are simply no resources to prioritize. However, in these places, resources tend to come sporadically from the federal and state levels, often in the form of grants. In these instances, counties are strained even more to prioritize. So, one might think of resource constraints as accelerating common prioritization problems. Given their common institutional structure, it also follows that counties are excellent cases in which to isolate and understand the influence of resources on prioritization and punctuation.

Conclusions

In this research note, we assessed the various measures of "tailed-ness" or punctuations in expenditure change distributions for local governments in Appalachia. Given the measures proposed in the existing literature, our task was straightforward. We calculated three measures of spending volatility from real-world expenditure data at the local level of government, specifically for counties and municipalities. Our choice here is important because counties, in particular, are vastly understudied in the literature on punctuated equilibrium, the study of policy change, and budgetary research. This is an important area for further inquiry and attention, given that counties are responsible for overlapping and unique services in the federal architecture of the United States. Education, emergency services, and public health are prime examples. One cannot understand these issues without accounting for county policymaking. In addition, political science generally understudies rural areas. Most literature is lodged at the federal, state, and municipal level. Yet, in rural areas especially, counties are the locus of much policymaking.

We find that the expenditure change distributions for counties in West Virginia approximate the most extreme distributions found in the literature for authoritarian regimes (Baumgartner et al. 2017; Sebők and Berki 2018; Jones, Epp, and Baumgartner 2019). We suspect this is a common finding for local governments in general and for local governments in rural areas in particular. In part, this is due to their smaller geographic jurisdictions. Smaller jurisdictions are less able to insulate themselves from perturbations in the problems they face, larger regional transitions, and shifting economic, social, and demographic features of their communities.

Our work also highlights spending punctuations resulting from two fundamentally different features of the counties we study. First, in distressed areas, it results from a lack of resources to deal with emergent and burgeoning policy problems. Second, in counties with largesse, volatility results from a fundamentally different problem - prioritization and whether decision systems are set up to deal with an influx of resources. In other words, volatility can result from a weak inflow of resources in the face of difficult circumstances or from swelling resources, pressuring policy and decision systems ill-equipped to prioritize how to deal with the influx.

In this research note, we asked whether three proposed measures for describing the punctuatedness of a policy change distribution correlated and, if not, whether the differences were significant both statistically and substantively. We find very little difference in l-kurtosis and the Gini coefficient as measures of spending volatility. The estimated degrees of freedom

for a t-distribution are significantly and substantively different from these other two measures. Further work is needed to delineate whether this stark difference is substantive and what features of these localities relate to these comparative differences. We also note that t-distributions represent technical difficulty in estimation for real-world, empirical spending changes. The degrees of freedom of the t-distribution are themselves likely distributed by a Gamma distribution.

Future Directions

Our analysis here leads straightforwardly to extensions along both methodological and substantive lines. We demonstrated above that the rank ordering of counties in terms of punctuation changes regarding the measure we use to describe the expenditure distributions. Our next steps involve investigating what features of counties relate to these changes. For example, the federal government and states classify counties regarding their level of distress, rurality, demographic and geospatial characteristics, and economic information. Concerning the shifting ranks, we measure which features or mixes of features best characterize these rank-order differences.

We are currently extending this analysis and investigation to cities in West Virginia. Like many states, West Virginia classifies cities into tiers based mostly on population. Future work that characterizes how spending volatility relates to city characteristics is a natural extension of our work here.

Ultimately, our work lays the foundation for moving beyond these descriptive measures to parametric models of spending fluctuations accounting for a host of local community, county, and city characteristics (Workman, Robinson, and Bark 2024). Moving to a framework of estimating shape parameters directly will bring studies of policy change back into the realm of general regression approaches so that we can assess the impact of these features on policy change.

Conceptually, we tend to think of expenditure volatility as resulting from resource constraints and or extreme changes in the problem environments facing local governments. Our analysis highlights resource expansions' role in spurring volatility in local governments where decision and policy systems become quickly overwhelmed (e.g., note the current rush to spend expiring ARPA funds three years later) and face difficulty prioritizing among demands. Large federal programs like ARPA, BBB, and the IRA strain the ability of local governments to prioritize.

Finally, there is also a revenue story underlying these fiscal decisions. At the federal level, where the government acts as an insurer of states and localities, the revenue stream is less important in the short run. Revenue at the local level looms large in influencing these patterns of expenditure.

References

Andersen, Simon C., and Peter B. Mortensen. 2009. "Policy Stability and Organizational Performance: Is There a Relationship?" *Journal of Public Administration Research and Theory* 20 (1): 1–22. https://doi.org/10.1093/jopart/mup005.

Baumgartner, Frank R., Marcello Carammia, Derek A. Epp, Ben Noble, Beatriz Rey, and Tevfik Murat Yildirim. 2017. "Budgetary Change in Authoritarian and Democratic Regimes." *Journal of European Public Policy* 24 (6): 792–808. https://doi.org/10.1080/13501763.2017.1296482.

Baumgartner, Frank R, and Bryan D Jones. 2009. *Agendas and Instability in American Politics*. 2nd ed. Chicago: University of Chicago Press.

Breunig, Christian. 2011. "Reduction, Stasis, and Expansion of Budgets in Advanced Democracies." *Comparative Political Studies* 44 (8): 1060–88.

Breunig, Christian, and Bryan D Jones. 2011. "Stochastic Process Methods with an Application to Budgetary Data." *Political Analysis* 19 (1): 103–17.

Breunig, Christian, and Chris Koski. 2006. "Punctuated Equilibria and Budgets in the American States." *Policy Studies Journal* 34 (3): 363–79. https://doi.org/10.1111/j.1541-0072.2006.00177.x.

Epp, Derek A. 2018. *The Structure of Policy Change*. Chicago: University of Chicago Press.

Fagan, E. J. 2022. "Political Institutions, Punctuated Equilibrium Theory, and Policy Disasters." *Policy Studies Journal*, February, psj.12460. https://doi.org/10.1111/psj.12460.

Fernández-i-Marín, Xavier, Steffen Hurka, Christoph Knill, and Yves Steinebach. 2019. "Systemic Dynamics of Policy Change: Overcoming Some Blind Spots of Punctuated Equilibrium Theory." *Policy Studies Journal*, December, psj.12379. https://doi.org/10.1111/psj.12379.

Hosking, J.R.M. 1990. "L-Moment Analysis and Estimation of Order Statistics." *Journal of the Royal Statistical Society Series B* 52:105–24.

Johnston, Jocelyn M., Michael A. Pagano, and Philip A. Russo. 2000. "State Limits and State Aid: An Exploratory Analysis of County Revenue Structure." *State and Local Government Review* 32 (2): 86–97. https://doi.org/10.1177/0160323X0003200201.

Jones, Bryan D., and Frank R. Baumgartner. 2005. *The Politics of Attention: How Government Prioritizes Problems*. Chicago: Chicago University Press.

Jones, Bryan D., Frank R. Baumgartner, Christian Breunig, Christopher Wlezien, Stuart Soroka, Martial Foucault, Abel François, et al. 2009. "A General Empirical Law of Public Budgets: A Comparative Analysis." *American Journal of Political Science* 53 (4): 855–73. https://doi.org/10.1111/j.1540-5907.2009.00405.x.

Jones, Bryan D., Derek A. Epp, and Frank R. Baumgartner. 2019. "Democracy, Authoritarianism, and Policy Punctuations." *International Review of Public Policy* 1 (1): 7–26. https://doi.org/10.4000/irpp.318.

Jones, Bryan D., Tracy Sulkin, and Heather A. Larsen. 2003. "Policy Punctuations in American Political Institutions." *American Political Science Review* 97 (1): 151–69. https://doi.org/10.1017/S0003055403000583.

Kaplaner, Constantin, and Yves Steinebach. 2022. "Why We Should Use the Gini Coefficient to Assess Punctuated Equilibrium Theory." *Political Analysis* 30 (3): 450–55. https://doi.org/10.1017/pan.2021.25.

López-Santana, Mariely, and Philip Rocco. 2021. "Fiscal Federalism and Economic Crises in the United States: Lessons from the COVID-19 Pandemic and Great Recession." *Publius: The Journal of Federalism* 51 (3): 365–95. https://doi.org/10.1093/publius/pjab015.

May, Peter J., Joshua Sapotichne, and Samuel Workman. 2009. "Widespread Policy Disruption: Terrorism, Public Risks, and Homeland Security." *Policy Studies Journal* 37 (2): 171–94. https://doi.org/10.1111/j.1541-0072.2009.00309.x.

Pagano, M. A., and J. M. Johnston. 2000. "Life at the Bottom of the Fiscal Food Chain: Examining City and County Revenue Decisions." *Publius: The Journal of Federalism* 30 (1): 159–70. https://doi.org/10.1093/oxfordjournals.pubjof.a030058.

Sebők, Miklós, and Tamás Berki. 2018. "Punctuated Equilibrium in Democracy and Autocracy: An Analysis of Hungarian Budgeting between 1868 and 2013." *European Political Science Review* 10 (4): 589–611. https://doi.org/10.1017/S1755773918000115.

Workman, Samuel, Scott E. Robinson, and Tracey Bark. 2024. "Organisation, Information Processing, and Policy Change in US Federal Bureaucracies." *Policy & Politics* 52 (2): 278–97. https://doi.org/10.1332/03055736Y2023D000000013.

Workman, Samuel, JoBeth Shafran, and Tracey Bark. 2017. "Problem Definition and Information Provision by Federal Bureaucrats." *Cognitive Systems Research* 43 (June): 140–52. https://doi.org/10.1016/j.cogsys.2016.11.002.

Workman, Samuel, Herschel Thomas, and Corinne Connor. 2022. "County Budgetary Dynamics: Findings from Appalachia." In. https://tinyurl.com/yyyve2wt.