A General Empirical Law of Public Budgets: A Comparative Analysis

Bryan D. Jones  University of Texas at Austin
Frank R. Baumgartner  University of North Carolina at Chapel Hill
Christian Breunig  University of Toronto
Christopher Wlezien  Temple University
Stuart Soroka  McGill University
Martial Foucault  University of Montreal
Abel François  University of Strasbourg
Christoffer Green-Pedersen  University of Aarhus
Chris Koski  James Madison University
Peter John  University of Manchester
Peter B. Mortensen  University of Aarhus
Frédéric Varone  University of Geneva
Stefaan Walgrave  University of Antwerp

We examine regularities and differences in public budgeting in comparative perspective. Budgets quantify collective political decisions made in response to incoming information, the preferences of decision makers, and the institutions that structure how decisions are made. We first establish that the distribution of budget changes in many Western democracies follows a non-Gaussian distribution, the power function. This implies that budgets are highly incremental, yet occasionally are punctuated by large changes. This pattern holds regardless of the type of political system—parliamentary or presidential—and for level of government. By studying the power function’s exponents we find systematic differences for budgetary increases...
I...
the hypothesis of institutional friction, i.e., they correspond with differences in government structures. While this may not confirm the friction model, it establishes plausibility on a firm empirical base.

The Role of Empirical Generalizations in Science

An empirical generalization is a finding that can be rigorously described and which holds across different circumstances and locations that themselves can be described. In the best case, the empirical generalization can be described mathematically, in which case we refer to an empirical law. Empirical laws do not explain; they describe precisely. The scientific puzzle then becomes how to explain the empirical law—that is, what models might best account for the now-established mathematical relationship among variables. Invariably some models are ruled out as a consequence of the empirical law, but usually an investigator cannot confirm a single explanatory model for the generalization.

Nevertheless the investigator is not helpless at this point. The next step is to examine the conditions under which the law holds and those in which it does not, or to explore where the magnitude of the parameters describing the law varies according to systematic differences in conditions. This is possible only because the empirical law has been firmly established. Trying this approach without sound evidence is like building a skyscraper on a weak foundation.

Social scientists have lagged behind natural scientists in the appreciation for building models on firm generalizations. Indeed, there is often a rush to establish a theory not infrequently based on (or perhaps a better term, illustrated by) “stylized facts.” But some of the most productive interchanges in social science stem from empirical generalizations—the “democratic peace” comes to mind. Is it empirically true that democracies do not fight one another? If it is, then why? Is it because of the institutions of democracy, or because democracies share some other characteristics, such as cultural similarities? The finding stimulated both arguments about the robustness of the generalization and the causal mechanisms responsible for generating it.

An example directly relevant to our budgetary study is the Gutenberg-Richter Law of earthquake magnitudes, which was established long before the exact causal mechanisms were traced (Gutenberg and Richter 1949). The law states that the relationship between the number of earthquakes and the magnitude of them is a power function. Many earthquakes are very small, and a very few are exceptionally large; but there is a paucity of moderate-sized ones. The careful empirical work of these two scientists helped lead to the discovery of plate tectonics. To this day, many aspects of the complex causal dynamics underlying the Gutenberg-Richter Law are poorly understood. Yet it was an incredibly important generalization that has stood the test of time and formed one of the central “facts” that must be explained by any model of the dynamics of earthquakes (Saichev and Sornette 2004). Any model that cannot generate the Gutenberg-Richter pattern must be eliminated from consideration.

Change Distributions

Power functions are leptokurtic, in that they, among many other distributions, generate slender peaks and fat tails, implying an excess of cases in the center (with little or no change) and in the tails (with large changes). Frequency distributions of public budget changes are leptokurtic for all cases studied thus far, and hence rule out the standard incremental model (True, Jones, and Baumgartner 2007). But these studies are based on specific cases and most of the existing studies have not explored the specific underlying probability distribution. Can this invariance be extended to the comparative context of many countries and locations? Does a common probability density function, the Paretoian or power function, characterize all political systems? If so, can we describe different political systems systematically in terms of the parameters of the distribution?

Budgets set public priorities; they are the outcome of complex policy processes involving the nature of the decision-making institutions, the preferences of decision makers (organized by political parties), and informational signals from a changing environment. In many real-world information-processing situations, we do not have the luxury of observing the actual informational input, because we observe only whether the decision maker attends to that information and what action he or she subsequently takes. Nevertheless, we can make some inferences about the structure of the incoming information.

4In a dynamic growth process such as characterizes budgets, the mode would be a positive increment.

5Cases include U.S. Budget Authority, U.S. states, U.S. municipalities, Danish municipalities, U.K. national government, and the national budgets of France, Germany, Belgium, and Denmark.
The Central Limit Theorem (CLT) guarantees that in any situation where a decision maker combines numerous sources of information in an implicit index, the limit of the distribution of that information will be Gaussian, so long as any one stream is not too disproportionately weighted and the streams are not highly correlated (Jones and Baumgartner 2005b). In making budget decisions, when decision makers incrementally adjust this year’s budget from a starting point of last year’s budgets, annual changes will be Gaussian. This is but a special case of the index-construction model (Jones and Baumgartner 2005b) and leads to outputs that are proportionate to the strengths of input signals. Moreover, it can be shown that the incremental model, which Padgett (1980) showed must generate a Gaussian distribution of changes, is a special case of the proportionate updating model (Jones and Baumgartner 2005b).  

It is important to distinguish between information signals, detectable changes in the environment that are potentially relevant for policymaking, and the news, which is that part of the set of signals that decision makers (including newspaper editors) attend to. The Central Limit Theorem can be sensibly assumed to apply to signals, but does not necessarily characterize the distribution of attention or news. Several social processes come in between the signals received from the environment and how those are measured, translated into politically relevant understandings, and brought to the attention of decision makers. But we know from the CLT that the distribution of changes in the underlying signals must be Gaussian (in the limit) because it is based on a large number of independent processes. If priorities are changed moderately, in proportionate response to incoming signals, then budgetary outputs will approximate a Gaussian distribution. The Gaussian, unlike either the power or exponential family of distributions, has strong shoulders; moderate changes from the status quo are the norm.  

In real-world situations, decision makers prioritize information in a manner that invariably leads to deviations from this proportionate processing of information (Jones and Baumgartner 2005a). They prioritize, and prioritization leads to non-Gaussian dynamics. Indeed, setting priorities causes bursts of activity characterized by fat-tailed distributions, such as the Paretoian. Studying email communications, Barabasi (2005) shows that waiting-time models of processing information, which follow Poisson distributions if inputs are not weighted by their importance (such as FIFO inventory control systems or random processing), will follow power distributions if people prioritize the inputs based on the perceived urgencies of incoming messages.  

In more complex decision-making situations, decision makers often do not update the set of indicators that guide their behavior and prioritize their policy actions. Then a sense of urgency will occasionally lead to overcoming the built-in friction that occurs in all human institutions. Then we get such comments from policy makers as “that was not even on the radar screen” or “nobody could have foreseen this.” But almost always the problem was not in the signals, which were always there, but in their detection and interpretation. This implies that even if inputs are Gaussian (which indeed many are), outputs from governments and other complex institutions will not be, but are likely to be characterized by fat-tailed dynamics.  

Nevertheless, as numerous studies of policymaking suggest, prioritization of input streams leads to leptokurtic distributions, and Paretonian distributions are leptokurtic. There are a number of different processes that can lead to power functions, but generally speaking these models involve systems occasionally getting into critical states in which large-scale punctuations are much more probable than in subcritical states. Critical regimes are poised between ordered, incremental change and rapid, discontinuous change (Bak 1997; Mandelbrot and Hudson 2004; Sornette 2003, 2006a). In human terms, these are associated with crises and panics, on the one hand, and notions of incredible opportunities on the other. The general characteristic is a strong sense of urgency to act.  

Major policy changes are often associated with electoral replacement, so that electoral changes can set the conditions for critical states in the policymaking process (Peterson et al. 2003; Stimson 2004; Wlezien 1996). But major changes also often occur in interelection periods, at least in the United States (Jones and Baumgartner 2005a, 84). Since all elections by no means cause major policy changes, we speculate that those elections that do are characterized by a similar sense of urgency to action. This is an additional strong reason for employing a distributional approach to studying policy change, since elections alone cannot be responsible for all major policy punctuations.  

---

6Padgett (1980) further derived budget decisional models that are leptokurtic and in some cases Paretoian.

7If decision makers are able to adjust proportionately, output distributions will be Gaussian even if decision makers are updating from past information. See Jones and Baumgartner (2005b).

8In Barabasi’s model, the tails of a distribution of response time represent delayed action whereas the peak indicates the urgency associated with short processing times. In the study of shifting policy priorities, the peaks of a budget distribution indicate the lack of urgency, while the tails, indicating big shifts in budgetary allocations, point to urgency.
Lots of physical systems, such as earthquakes and avalanches, have frequency distributions characterized by power laws (Schroeder 1991; Sornette 2006a), but of course they do not have conscious priority-setting mechanisms. As Per Bak’s sandpile experiments have shown, physical systems with friction are capable of generating power functions, even when inputs (grains of sand) are incrementally added. His sandpiles generated either very small landslides or very large ones, but few moderate-sized slides (Bak 1997). Bak’s systems can be characterized as error accumulation models in that the sandpile has “underadjusted” to the accumulation of pressures with small landslides, and then must adjust in one fell swoop. So there can be clear similarities in distributional attributes between physical and social systems.

The Empirical Analysis of Budget Distributions

A number of studies have shown that budget change distributions are highly leptokurtic, with strong central peaks and extended tails, and clearly not Gaussian (Baumgartner, Foucault, and Françoise 2006; Breunig 2006; Breunig and Koski 2006; John and Margetts 2003; Jones and Baumgartner 2005a; Jordan 2003; Mortensen 2005; Robinson 2004; Soroka, Wlezien, and McLean 2006; True, Jones, and Baumgartner 2007). Yet systematic comparisons across different political systems are lacking, and the particular probability distribution functions have only occasionally been studied.

To remedy this, we have assembled datasets on public budgets from seven national and two subnational governmental units. For two long series for France and the United States, we can analyze year-to-year inflation-adjusted percentage changes in Total or, in the United States, Domestic and Defense spending; for the other datasets, where the series are considerably shorter, we have to pool across budget categories (and across the subunits for subnational governments), again using annual percentage changes. This is necessary in the latter case to ensure that the distributions are not dominated by one or two really large budget categories; it is desirable in the former to enable comparison. Table 1 briefly describes these datasets.

Government expenditure data are notoriously unreliable at any but the most aggregate level, because categories are added and subtracted for accounting purposes but are not generally adjusted backwards to ensure comparability with earlier data. “Off the shelf” budget datasets should generally not be used for analysis across categories. Moreover, national governments do not usually maintain separate capital budgets, so budget decisions and the outlays generated by those decisions can occur in different fiscal years. As a consequence, it has been necessary for us to make certain that all series are internally comparable.

### Table 1  Dataset Descriptions

<table>
<thead>
<tr>
<th>Dataset Type</th>
<th>Date</th>
<th>Units Pooled</th>
</tr>
</thead>
<tbody>
<tr>
<td>National Governments (long series)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>United States Outlays</td>
<td>1800-2005</td>
<td>Years</td>
</tr>
<tr>
<td>U.S. Domestic Outlays</td>
<td>1800-2004</td>
<td>Years</td>
</tr>
<tr>
<td>U.S. Defense Outlays</td>
<td>1800-2004</td>
<td>Years</td>
</tr>
<tr>
<td>France Outlays</td>
<td>1820-2002</td>
<td>Years</td>
</tr>
<tr>
<td>National Governments (pooled)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>United States Budget Authority</td>
<td>1947-2005</td>
<td>Years, 60 OMB programmatic subfunctions</td>
</tr>
<tr>
<td>France Outlays</td>
<td>1868-2004</td>
<td>Years, 7 ministries</td>
</tr>
<tr>
<td>Germany Outlays</td>
<td>1962-2000</td>
<td>Years, 26 functions</td>
</tr>
<tr>
<td>Great Britain Outlays</td>
<td>1981-1999</td>
<td>Years, 14 functions</td>
</tr>
<tr>
<td>Belgium Outlays</td>
<td>1991-2000</td>
<td>Years, 27 functions</td>
</tr>
<tr>
<td>Denmark Outlays</td>
<td>1971-2003</td>
<td>Years, 26 functions</td>
</tr>
<tr>
<td>Canada Outlays</td>
<td>1990-2004</td>
<td>Years, 12 functions</td>
</tr>
<tr>
<td>Subnational Governments (pooled)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U.S. State Operating Outlays</td>
<td>1984-2002</td>
<td>Years, 10 functions, 50 states</td>
</tr>
<tr>
<td>Danish Local Operating Outlays</td>
<td>1991-2005</td>
<td>Years, 4 functions, 265 municipalities</td>
</tr>
</tbody>
</table>

*Note: Full descriptions available from the authors.*
This accounts for the fairly short time periods covered by some of the datasets.

The somewhat shorter time series on disaggregated budget data is more than offset by the advantage these data offer: a direct assessment of changing priorities of government. “Off the shelf” budget data are not acceptable exactly because of this: the creation of new categories and the failure to update older series will cause the investigator to mistake accounting adjustments for shifts in priorities (Soroka, Wlezien, and McLean 2006).

Figure 1 displays one of the long time series we analyzed. At least until the recent financial crisis, the volatility of budget series declined over time. This reduced budgetary volatility can be clearly seen in Figure 1, which shows inflation-adjusted expenditures for the U.S. national government from 1800 to 2004.

To test for the Paretoan, we take the logarithm of both sides of the expression $y = ax^b$, where $x$ represents the strength of the signal (e.g., the percentage change in the budget) and $y$ represents the cumulative frequencies associated with each value (e.g., how many observations have that level of change, or higher). This yields $\log(y) = \log(a) + b \log(x)$, which will plot as a straight line if the distribution is Paretoan. Figure 2 depicts frequency distributions and log-log plots for our two longest budget series, calculated as percentage changes. The exponents for the left side of the frequency distribution (the negative tail) are reversed in sign, so that both sides may be plotted on one graph. In the cases of both France and the United States, inflation-adjusted outlays follow a power function. Looking separately at domestic and defense outlays in the United States (in the second panels) makes little difference—again, the figures reveal signature power function frequency distributions.

The exponents relating to these figures (alongside exponents for subsequent figures) are listed in Table 2. The table also shows the L-kurtosis (L-K) measure, a standardized measure of kurtosis that adjusts for overresponsiveness of kurtosis to extremes, for each series. As expected, both the U.S. and French data reveal a good degree of leptokurtosis. Regarding the exponent estimates, those for the United States and France, both series center on $-0.9$ (with France slightly lower than the United States) for the right tail of the distribution, but are higher for the left tail. That is, viewed as absolute values, the right tail exponent is larger than the left tail exponent. This indicates fatter tails on the positive side; that is, increases in expenditures tend to occur in bursts while decreases occur somewhat

---

9 Figure 1 also illustrates the “war ratchet” of Peacock and Weisman (1967): when war occurs, both defense and domestic expenditures go up, and domestic expenditures tend to stay at the higher level.

10 Researchers need to use considerable caution when using log-log plots to detect power function PDFs. Coefficients of determination are generally very high, because of accumulation in bins, and statistical tests are not reliable (see Clauset, Shalizi, and Newman 2007 and Sornette 2006). It is imperative to compare the log-log plot with a semi-log plot (which estimates an exponential distribution) and to use care when studying the tails of the distribution. Researchers usually use cumulative distribution functions (CDFs) because they are better behaved and bins may be cumulated without changing the underlying distribution—the cumulative power function distribution is also a power function.
FIGURE 2  Frequency Distribution and Log-Log Plots for the Long Budget Series: Total Outlays in the United States and France

(a) U.S. Total Outlays, Frequency Distribution

(b) U.S. Total Outlays, Log-Log Plot

(c) U.S. Domestic Outlays, Log-Log Plot

(d) U.S. Defense Outlays, Log-Log Plot

(e) French Total Expenditures, Frequency Distribution

(f) French Total Expenditures, Log-Log Plot
TABLE 2  Exponent Estimates for Power Functions of Tails of Distributions

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Positive Tail</th>
<th>Negative Tail</th>
<th>R²</th>
<th>L-K</th>
</tr>
</thead>
<tbody>
<tr>
<td>National Governments (long series)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>United States</td>
<td>−0.911</td>
<td>1.396</td>
<td>.949</td>
<td>0.509</td>
</tr>
<tr>
<td>U.S. Domestic</td>
<td>−1.094</td>
<td>1.400</td>
<td>.933</td>
<td></td>
</tr>
<tr>
<td>U.S. Defense</td>
<td>−0.976</td>
<td>1.602</td>
<td>.963</td>
<td></td>
</tr>
<tr>
<td>France</td>
<td>−0.885</td>
<td>1.091</td>
<td>.962</td>
<td>0.424</td>
</tr>
<tr>
<td>National Governments (pooled)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>United States</td>
<td>−1.024</td>
<td>1.789</td>
<td>.916</td>
<td>0.512</td>
</tr>
<tr>
<td>France</td>
<td>−1.019</td>
<td>1.353</td>
<td>.924</td>
<td>0.505</td>
</tr>
<tr>
<td>Germany</td>
<td>−1.387</td>
<td>1.629</td>
<td>.960</td>
<td>0.456</td>
</tr>
<tr>
<td>Great Britain</td>
<td>−1.490</td>
<td>1.797</td>
<td>.970</td>
<td>0.319</td>
</tr>
<tr>
<td>Belgium</td>
<td>−1.543</td>
<td>1.293</td>
<td>.992</td>
<td>0.611</td>
</tr>
<tr>
<td>Denmark</td>
<td>−1.565</td>
<td>2.179</td>
<td>.984</td>
<td>0.421</td>
</tr>
<tr>
<td>Canada</td>
<td>−1.245</td>
<td>1.549</td>
<td>.915</td>
<td>0.379</td>
</tr>
<tr>
<td>Subnational Governments (pooled)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U.S. State</td>
<td>−1.926</td>
<td>2.007</td>
<td>.910</td>
<td>0.403</td>
</tr>
<tr>
<td>Danish Local</td>
<td>−1.810</td>
<td>2.000</td>
<td>.965</td>
<td>0.363</td>
</tr>
</tbody>
</table>

more gradually. As we see, this is a general characteristic of all budget distributions.

Figure 3 shows both frequency distributions and log-log plots for U.S. Budget Authority aggregated over Office of Management and Budget programmatic subfunctions, and Figure 4 shows German and French programmatic expenditures over several functions. (Again, the corresponding exponents are listed in Table 2.)

The distributions of all three series follow power functions, and in all three cases growth punctuations are more probable than cutback punctuations. Indeed, the negative tail for the United States is not discernibly distinct from an exponential fit. This can be seen in the log-linear plot in Figure 3c. Log-linear plots estimate simple exponential distributions, which are much less fat-tailed than power functions. Apparently modern governments find it more difficult to cut back programs significantly than to expand them dramatically.

Figure 5 depicts log-log plots for the other national governments for which we have reliable data. All show power function frequency distributions, and most (with the exception of Canada) show a tendency to have more difficulties in cutting programs in a very large fashion than in increasing them greatly. None of the national governments, however, show the strong difference between the tails that is evident in the U.S. plot.

We conclude that national governments shift priorities according to a power function and in a manner generally consistent with punctuated equilibrium. This holds for both the United States, with its presidential system, as well as for parliamentary democracies. Moreover, the governments we have examined generally experienced more resistance, or friction, in cutting programs than in expanding them. In particular, the shoulders of the negative tail for the national distributions are considerably stronger than those for the positive side, suggesting more resistance—to the point of approaching the exponential distribution for the United States.

Of course, national governments can borrow money to fund operating expenses, and this may allow for a more mellow approach to cutting programs than may be evident for subnational governments for whom borrowing is not possible. Moreover, in harsh economic times, program cuts by national governments can contribute to declines in economic demand, and this adds a policy justification for what may be less extreme cuts in national budgets. Is the distribution of budgetary change notably different, then, for subnational governments?

U.S. state and Danish local government data are depicted in Figures 6 and 7. In both cases, we do indeed detect considerable differences from the national governments. Both may be classified as power functions and the U.S. state governments unambiguously so. For Danish local governments, the situation is not quite as clear, although the tails of both sides of the distribution approximate a power function and the distribution is closer to a power function than a pure exponential.

Also, in comparison to the national governments the distributions for these subnational governments are remarkably symmetrical and have somewhat less extreme tails (the exponent estimates indicate flatter tails). This is clear in Table 2; it is also clear in Table 3, which presents pooled results for the (a) long-term national, (b) other national, and (c) subnational governments. Subnational governments appear about as likely to cut budgets as to raise them; one gets a sense of an ongoing reprioritization that is far more moderate than happens at the national level. It is likely that this is a function of mandates

---

11 Both increases and decreases of expenditures occur relative to a long-run positive mean due to increasing economies.

12 We may rule out this being an artifact of using percentages (proportions), as the left tail of these distributions terminates before reaching 100%. Moreover, an examination of first differences for these series indicates no censored data issues.

13 One possibility is the stretched exponential, which has fat tails, but they are not as pronounced as in the case of power functions. See LaHerre and Sornette (1998).
imposed on these governments by their superior governmental units. States in the United States can borrow only for capital needs, and hence cannot borrow to tide over required cuts. Generally, local governments meet restrictions on the money they raise locally and the grants they receive from higher levels of government. The result is a far more balanced fiscal system than what occurs at the national level.

Looking across the preceding results, we have established the ubiquity of the Paretian probability distribution in describing budget changes in a variety of governmental settings, a finding that goes substantially beyond the vaguer description of budget changes as leptokurtic.\textsuperscript{14} We now turn to exploring how to account for what may be a general law of public budgets.

\textsuperscript{14}Peter Erdi (2008) and his colleagues have recently reanalyzed these data and suggests that the Generalized Extreme Value Distribution may fit a whole budget change distribution more efficiently than two power functions, but it requires an additional parameter to estimate. One advantage is that this distribution does not include a singularity at the mode. Our findings here stand in any case: the power function is robust in describing budget change distributions. Unfortunately the GEV distribution will not allow the direct comparisons across countries that we pursue shortly, and Erdi notes that the choice of functions may depend on the problem at hand when an extra parameter must be estimated.
Explaining the Generalization

Several different mechanisms are associated with power function probability distributions in the natural world, but many other mechanisms, such as the pure incrementalist model of decision making, are ruled out (Jones and Baumgartner 2005). In the remainder of the article, we demonstrate the plausibility of a very important set of mechanisms that generate power functions in the natural world: stick-slip friction dynamics.

In the natural world, there are various specifications for friction, but all of them involve the interaction of two forces: a retarding force and a force directed at overcoming the retarding force. Even such seemingly simple systems have defied top-down models capable of point predictions. Earthquakes, for example, are complex nonlinear systems whose full dynamics have not yet been specified, and whose underlying stress-strain dynamics are either inaccessible to measurement or fundamentally not measurable (Rundle et al. 1996). Moreover, there seem to be complex connections between spontaneous occurrences caused by plate tectonics and triggering events, such as previous earthquakes and other causes, including man-caused events such as the pressure from constructed reservoirs. This mixture of exogenous and endogenous causes also characterizes many social systems (Sornette 2006b). In the face of such complexity, geophysicists have proceeded by observation and simulation, as well as by trial-and-error modeling.

While the specific dynamics of earthquakes are not yet known, the general processes include what is known as stick-slip dynamics. The earth’s tectonic plates are held
in place by a retarding force, the “friction” of the plates, while the dynamic processes generated by activities in the earth’s core push on these plates. When the forces acting on the plates are strong enough, the plates release, and, rather than slide incrementally in adjustment, slip violently, resulting in the earthquake. The problem is that there is no system for predicting exactly when and where the very modest earthquakes become violent ones. The Gutenberg-Richter Law, a power function, is largely explained by the stick-slip nature of earthquake dynamics, but it is a stochastic law and does not offer point predictions.

Political systems, like many social systems, are characterized by considerable friction. Standard operating procedures in organizations, cultural norms, and facets of human cognitive architectures provide stability of behavior in a complex world. In politics, ideology and group identifications provide stable guides to behavior in complex circumstances. In politics, however, a second source of friction exists: institutional rules that constrain policy action. In the United States, policies can be enacted only when both houses of Congress and the president reach agreement on a measure. Federalism can produce friction in a similar way, at least when governments share jurisdiction in particular policy domains. And in parliamentary democracies, especially ones with proportional electoral systems, action may be constrained by multi-party governing coalitions. Institutional rules “congeal” preferences (Riker 1980), making it difficult for new policies to enter the political arena.

In the past, scholars characterized these systems using comparative statics, a method of analysis that
**FIGURE 6**  Frequency Distributions and Transformed Plots for U.S. State Outlays

(a) U.S. States: Frequency Distribution

(b) U.S. States: Log-Log Plot

**FIGURE 7**  Frequency Distribution, Log-Linear, Probit of Log-Linear, and Log-Log Plots for Danish Local Government

(a) Frequency Distribution

(b) Log-Linear Plot

(c) Probit of Log-Linear

(d) Log-Log Plot
concentrates on equilibrium processes based on the preferences of decision makers (Krehbiel 1998; Shepsle and Weingast 1987). Change was admitted primarily through the replacement of governing parties through elections, which established a new preference-based equilibrium to which the policymaking system quickly adjusted. But the comparative statics approach ignores the ongoing information-processing needs of an adaptive system, and political systems are clearly adaptive systems. They dynamically respond to incoming information, not just the preferences of those making decisions.

Punctuated equilibrium has provided an alternate analytical frame to the preference-based analyses of comparative statics (Baumgartner and Jones 1993; True, Jones, and Baumgartner 2007). The stability imposed by the two kinds of friction, cognitive/organizational friction and institutional friction, does not cause universal gridlock, with a system awaiting elections to point to change. But it is a retarding force that interferes with the smooth adjustment of a political system to changing information signals from the policymaking environment. This force resembles the friction that occurs in the physical world, in that change occurs but only when the informational signal from the external world either is extraordinarily strong or when the signals accumulate to overcome the friction. These latter processes are described as error accumulation models (Larkey 1978), in that the deviation between the external world and the system response gets increasingly out of kilter until the system can no longer ignore the deviation.

Systems characterized by friction remain stable until the signals from outside exceed a threshold, and then they lurch forward; they will continue moving only if the external signal continues at greater than threshold strength. Otherwise they resume “equilibrium.” At present we cannot specify mathematically the specific interacting forces that lead to such political stick-slip dynamics, but trying to specify precisely these processes is premature until we can mathematically describe the global processes and can verbally describe the likely internal generating mechanisms.

Certainly we understand the sticking part of the process better than the slipping part, even if we don’t understand very well how cognitive and institutional friction interact. But we understand the dynamics of slip at only the vaguest level. It is likely that political systems overcome friction when a sense of urgency about the external world drives decision makers to reprioritize their preferences. Urgency causes collective attention to focus on a very limited number of issues out of the panoply that are candidates for government action; these issues are rewarded by disproportionate attention, often leading to large changes in budget allocations. Such situations generally lead to considerable contagion.

Whatever the internal processes, the maladaptability of the system must often reach a high threshold before the accumulated errors have an impact. While error correction models are well known in social science (Engle and Granger 1987), there is no formal specification of error accumulation. But they are a special case of nonlinear error correction models—they lead to much larger deviations before the “error” is corrected than the more typical model (Escribano and Miar 2002). They generate leptokurtic outputs (such as the power function distribution) from linear inputs, but of course other systems could also be consistent with such outputs.

Power function frequency distributions characterize many market-based transactions, and the exponents for these transactions are similar for different kinds of markets and transactions (Gabalx et al. 2003; Mandelbrot and Hudson 2004). Market transactions differ from political interchanges in one very important sense: in modern markets, there are limited formal decision costs in choosing to pursue a transaction. In politics, collective decision-rules limit the freedom of choice of any set of actors. Markets may be governed by cognitive friction that is overwhelmed occasionally by the sense of urgency (to buy or to sell), but the institutional friction so important in

<table>
<thead>
<tr>
<th></th>
<th>Budget Increases (Positive Tail)</th>
<th>Budget Cuts (Negative Tail)</th>
<th>Average L-Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>National Governments</td>
<td>$P(b &gt; x) = x^{-\xi}$</td>
<td>$P(b &lt; x) = x^{-\xi}$</td>
<td>.467</td>
</tr>
<tr>
<td>(long series)</td>
<td>$[.89, .91]$</td>
<td>$[1.09, 1.40]$</td>
<td></td>
</tr>
<tr>
<td>National Governments</td>
<td>0.9</td>
<td>1.2</td>
<td></td>
</tr>
<tr>
<td>(pooled)</td>
<td>$[1.02, 1.57]$</td>
<td>$[1.29, 2.18]$</td>
<td>.458</td>
</tr>
<tr>
<td>Subnational Governments</td>
<td>1.3</td>
<td>1.67</td>
<td></td>
</tr>
<tr>
<td>(pooled)</td>
<td>$[1.76, 1.93]$</td>
<td>$[2.01, 2.58]$</td>
<td>.353</td>
</tr>
</tbody>
</table>
Comparative Institutional Friction

Past work argued that institutional friction increases as institutions add costs to the translation of inputs into outputs (Jones, Sulkin, and Larsen 2003). The distribution for U.S. Budget Authority consequently displays a greater degree of leptokurtosis than does one for congressional hearings, which in turn is more punctuated than one for New York Times stories. Indeed, an analysis of 15 different U.S. series suggested a strong relationship between leptokurtosis and a rough ranking of the “costs” in different political venues. Similarly, these costs are associated with declining representation of the public’s policy preferences—as institutional costs increase, the priorities of the public are decreasingly represented in the policy actions of American policymaking institutions (Jones, Larsen-Price, and Wilkerson 2009). A recent study of Belgium, Denmark, and the United States reports similar associations (Baumgartner et al. 2009).

Our data allow for a more rigorous quantitative comparison for budgetary outputs. Recall that we suggest that the nonnormality of policy change distributions will be a function of both cognitive/organizational friction and institutional friction. Cognitive/organizational differences across countries are of course difficult to measure, though it seems reasonable to assume that cognitive costs do not vary systematically across countries. Organizational costs will certainly vary, but it is not clear how to measure them—the efficiency of a bureaucracy is hard to get at.

Institutional differences are simpler to capture, however, and we provide an initial analysis here of the link between political institutions and the punctuatedness of budgetary series—or, more broadly stated, an analysis of the way in which political institutions affect the way in which information is processed. This first cut relies in large part on a veto players approach (Tsebelis 2002): institutional friction is expected to increase with the number of veto players involved in the policymaking process.

The classic literature on the performance of democratic institutions (e.g., Lijphart 1977, 1999; Sartori 1997) points to an approach for measuring in a reliable and straightforward fashion aspects of institutions that may systematically be related to friction. Friction can be expected to increase where mechanisms inhibit reaction to input stimuli, and where institutions are designed to mitigate “overreaction.” In particular, we expect friction to decrease under conditions of (1) parliamentary government, (2) single-party governments, (3) unicameralism, and (4) a unitary state. More precisely, ceteris paribus:

(1) friction should be greater in presidential systems than in parliamentary systems, since policy change in the former is dependent on approval from several different bodies;
(2) friction should be greater in coalition or minority governments, since coalitions require more internal bargaining;
(3) friction should increase with bicameralism, since policy is vetted by two rather than one legislative chamber; and
(4) friction should be greater in federal systems, where multiple governments are involved in making policy decisions.

Each of these institutional criteria has been very well studied. There are several measures available of each; Table 4 shows some of the standard measures across each of the countries for which we also have reliable budgetary data.

This is by no means an exhaustive list. But we regard these as four of the institutional features that most clearly affect the potential for adding friction to the policymaking process. Each of these institutions, in general, reduces the government’s discretion to act. In so doing, each makes it more difficult for the government to adjust policy in response to inputs of various types, including the electoral demands of voters. A pent-up demand for policy change is a natural result. As institutional friction increases, therefore, policy will demonstrate more leptokurtosis. This can be directly tested, and we do so below.

While we argue that institutional friction begets punctuated outcomes, in one important study more responsiveness to public opinion was found in presidential systems. Soroka and Wlezien (2005) argue that policymaking in presidential systems should be more responsive to public opinion than in parliamentary systems, in effect because such responsiveness is one thing on which parties in a system of divided government can agree. It may thus be that parliamentary governments exhibit more discretion but less responsiveness to the public, and yet more responsiveness to other changing circumstances facing
Table 4  Political Institutions and Friction

<table>
<thead>
<tr>
<th></th>
<th>Executive Dominance</th>
<th>Single-Party Governments</th>
<th>Bicameralism</th>
<th>Decentralization</th>
<th>Friction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belgium</td>
<td>1.95</td>
<td>28.8</td>
<td>3</td>
<td>2.8</td>
<td>19</td>
</tr>
<tr>
<td>Canada</td>
<td>4.17</td>
<td>95.2</td>
<td>3</td>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td>Denmark</td>
<td>2.09</td>
<td>23.9</td>
<td>5</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>France</td>
<td>5.52</td>
<td>63.5</td>
<td>3</td>
<td>4.7</td>
<td>10</td>
</tr>
<tr>
<td>Germany</td>
<td>5.52</td>
<td>46.2</td>
<td>2</td>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>5.52</td>
<td>93.3</td>
<td>3.5</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>United States</td>
<td>1</td>
<td>80.1</td>
<td>2</td>
<td>1</td>
<td>21</td>
</tr>
</tbody>
</table>

In any case, we treat this as an open question here.

Table 4 shows our measures of these variables in the countries for which we have budget data. These institutional measures are drawn from Lijphart's *Patterns of Democracy*, and rely on data gathered for the 1971–96 period. The first column of Table 4 shows the percentage of time during which the country was governed by a single-party (that is, noncoalition) majority government. The index of executive dominance is based on the average duration of cabinets, except for the United States, where the duration of cabinets is fixed. Executive dominance over the legislature in the United States is of course quite low, so the country is assigned a value of 1. The federalism measure is impressionistic—a 1 to 5 scale, where Lijphart assigns each country to one of five categories. The Belgian value represents an average of the pre- and post-1993 decentralization. Bicameralism is a 1 to 4 scale, where 4 is two “symmetrical and incongruent chambers,” and 1 is unicameralism. (Full details of each measure are available in Lijphart 1999.)

There are other measures of these criteria, to be sure, and as with any broad comparative institutional measure there are contestable values in each. We wish to stake no claim to the specific extent of federalism in one country versus another, however. For this preliminary work, we wish only to establish a general ranking of the countries for which we have budgetary data, on each criterion, and then on all criteria combined. Ordinal rankings (1 through 7) are included in parentheses in Table 4. The final column then shows the sum of these ordinal rankings, where higher values suggest greater degrees of friction. The United Kingdom, with 93.3% single-party majority governments, strong executive dominance, a centralized state (for the period for which we have spending data), and only weak bicameralism, receives a total score of 6. We accordingly expect much less friction in the United Kingdom than in Germany, or particularly the United States—decentralized and bicameral, with little to no executive dominance over the legislature.

Is there a systematic relationship between this rough approximation of institution constraints on information processing and the distributional statistics for budgetary series across the seven countries? Figure 8 plots the L-kurtosis scores for each country alongside the institutional friction scores from Table 4. The correlation in the figure is striking (Spearman’s Rho = .75). It suggests that leptokurtosis is systematically (and strongly) related to the arrangement of political institutions. Where the number of decision-making bodies is greater—where there are greater impediments to reactive policymaking—so too is there a greater degree of punctuatedness in budgetary policy.16

---

15Lijphart shows that the scale is highly correlated with central governments’ tax shares, a more nuanced, though still incomplete, measure of decentralization.

16The correlation between L-kurtosis and the parameter estimate for the positive (budget increase) slope is .17 for all countries, but dropping Belgium, which is plagued with a short series and hence is dominated by its cross-section observations, the correlation is .71. Moreover, the scatterplot reveals that France and the United States share a close affinity on both the kurtosis measure and the exponent measure, suggesting that our calculation based on Lijphart’s measure may underestimate institutional friction in France.
There is a convincing connection between the kurtosis of national budget distributions and our parameter estimates of the exponents for the power functions that characterize those distributions in our data. The parameter estimates for the power exponents seem to reflect in large measure institutional friction. While the power function is the general law of budgets, differences in exponents capture differences in institutional arrangements.

**Conclusions**

Firm empirical generalizations provide a critical role in science, one that has been somewhat unappreciated in political science. In this article, we have established a robust empirical generalization, that government budget changes in Western democracies follow power function probability distributions. This implies that they display periods of quiescence interrupted by bursts of frenetic activity. Because budgets are reflections of priorities, and budget change distributions reflect changing priorities, the dynamics of budget changes could well indicate the occasional occurrence of bursts of urgency about the external world.

This is the general budgetary law, but within that law there exists systematic variation across political systems. This variation is associated with institutional friction. Limitations on the ease of action due to institutional rules and facets of human decision making damp down easy adjustment to changing inputs, requiring larger mobilizations to overcome them.

Exponents for national governments are variable, but a strong tendency exists for bursts of spending increases to dominate budget changes on the positive tail, while cuts are subject to less severe bursts. Orgies of spending are not fully offset by equally exuberant cutting.

Exponents for subnational governments are both very similar (for the two disparate situations we studied) and quite symmetrical. Subnational government budgets are less punctuated—less subject to bursts of budgetary activity—than national government budgets. While orgies of spending and cutting both occur, they are more muted than in the case of national governments. While national governments exhibit considerable country-to-country variation, they tend to display more dramatic dynamics than the subnational governments.

We suggest that these patterns may be explained by reference to a particular form of friction—the stick-slip dynamics of earthquakes and other natural phenomena. Friction is also a characteristic of political systems; it holds in place the status quo through both formal means (such as supermajority requirements in the United States, and the need to construct coalition governments in many parliamentary democracies) and informal means (such as the cognitive screen of political ideology). But stability will
not allow a system to respond proportionally to changing external circumstances. Demands outside the political system build up, in a type of “error accumulation” process; when these errors exceed a threshold, friction is overcome.

We studied variability in the power function exponents, which indicate the extent of radical budget behavior, across nations. We found that institutional friction, as we assessed it, corresponds in general to the extent of punctuations, assessed by the power function exponents.

A combination of internal reprioritization and organizational friction seems best able to explain the patterns we have observed: strong budgetary conservatism represented by the peaks of the distribution of budget changes; weak shoulders, indicating the inability to respond to incoming information in a moderate, proportionate way; and fat tails, representing frenetic bursts of activity. The contagion of urgency overcomes the friction of order and leads to the dynamics of public budgeting.

Generally speaking, a distribution should approach Normal as a government increases its cognitive/organizational capacity and reduces institutional impediments to reactive policymaking. More dramatic power function distributions should result from governments with poor cognitive/organization capacity, and many impediments to reactive policymaking. It’s hard to assess cognitive/organizational capacity, but institutional impediments—veto players, in large part—are easier to think about.

Indeed, the effect of institutional design on friction in budgetary policy is evident in the results above. We expect friction to be greater in presidential systems than in parliamentary systems, for instance, since policy change in the former is dependent on approval from several different bodies. We also expect friction to increase with coalition government, low party discipline, and federalism, at least where federalism is structured in a way that requires the agreement of multiple governments for single policy decisions. Our analyses bear out these expectations.

Appendix

Budget Data Source Descriptions

All of the series we studied were corrected for category consistency, or the issue was not relevant to the dataset (as in the case of fully aggregated data).

U.K. budgetary data consist of data for 14 major functions, consistently defined from fiscal years 1980 to 1999. Data are from Stuart Soroka and Christopher Wlezien, Total Expenditure on Government Services in the United Kingdom, 1980–2000, U.K. Data Archive (SN 4980). Details are available at http://www.data-archive.ac.uk/. Fiscal years in the United Kingdom run from April of one year to March of the following year.

Canadian budgetary data are for the 12 major functions for Federal General Government Expenditure, consistently defined from fiscal years 1989 to 2002. Data are available from CANSIM (Matrix 3950002). Details are available at http://cansim2.statcan.ca/. Note that the dataset used here excludes a few very minor expenditure categories as well as some unspecified intergovernmental transfers (mainly to provincial governments) which cannot be allocated by function. Fiscal years in Canada run from April of one year to March of the following year.

Belgian budget data are for 27 categories of spending over the period of 1991 to 2000, and originate from the Belgian Political Agenda-setting Project. The project (2001–04) was funded by the “Federale Diensten voor Wetenschappelijke, Technische en Culturele Aangelegenheden” (DWTC). It was conducted by Stefaan Walgrave (coordinator, UA), Lieven De Winter, André Frognier, Frédéric Varone and Benoit Rihoux (UCL), Patrick Stouthuysen (VUB), and Marc Swynge-douw (KUL). Details are available at http://www.ua.ac.be/main.aspx?c=m2p.

Danish local spending data consist of inflation-adjusted local spending figures using four consistently defined categories of spending from 1991 to 2005 pooled across 271 Danish municipalities. The data originally come from Tables “BUD32” and “BUD32X,” available online from Statistics Denmark (http://www.statistikbanken.dk). See Mortensen (2005) and www.agendasetting.dk for further documentation.

The dataset on Danish national spending consists of inflation-adjusted public spending figures using 26 consistently defined categories of spending from 1971 to 2003, using data originally made available by Statistics Denmark, Section of Public Finances (www.dst.dk). Further documentation is available at www.agendasetting.dk.

The sources for national-level French budgetary data are the INSEE (Institut National de la Statistique et des Etudes Economiques) Statistical Handbook (annual). The historical data (1868 through 1939) are gathered through a retrospective series published in the 1951 French Statistical Handbook. All other data have been computed from the annual INSEE Statistical Handbooks. For data after the Second World War, we have used the Statistical Handbook 1947–87 published by the INSEE. From 1988 onwards, we have used the annual publication of INSEE called Tableaux de l’Economie Française, which provides a complete presentation of public spending adopted by the
Parliament through the Finance Law. Total expenditure is made up of separate series for Defense and Civilian public spending. Each statistical series is originally produced and delivered by the Direction of National Public Accounts (a division of the Ministry of Finance). Data are expressed in current francs and were then adjusted into constant francs using the Consumer Price Index (CPI) as supplied in the INSEE publications.

U.S. Budget Authority Data are derived from the Office of Management and Budget Sources, which adjusts categories for consistency after 1976. The Policy Agendas Project (www.policyagendas.org) applied consistent adjustments back to 1947. Data are adjusted for inflation using GDP deflators, with 2005 as the base year.

U.S. Government Outlays are from Historical Statistics of the United States, compiled by the U.S. Census Bureau, updated from the Office of Management and Budget website, Historical Statistics, Table 1.1. The Consumer Price Index was used to adjust for inflation due to the absence of GDP deflators for the early part of the series, with June 1984 = 100.

References


