

How to Assess Power Law Behavior Using Stochastic Process Methods: A Note of Caution

Matthias Fatke¹

LHS Stuttgart, Kronenstraße 20, 70173 Stuttgart, Germany. Email: matthias.fatke@uni-konstanz.de

Abstract

Power law distributions have a certain appeal to researchers, not least because they often insinuate a general empirical law. Hence, searching for them in data generated by social and political processes has become popular. In the political science literature, however, power law behavior has rarely been assessed rigorously. Relying mainly on qualitative appraisals of log–log plots, merely a necessary, but no sufficient, condition of power law behavior is tested. This letter therefore seconds the note of caution expressed recently. Moreover, it showcases the use of a principled statistical framework to test power law behavior in a quantitative manner. Applying this method to a seminal case in political science, the results of the analysis invite an empirical as well as theoretical refinement of the claim that changes in public budgets follow a power law. In a more general sense, the letter wishes to contribute to a more thorough practice of stochastic process methods in political science.

Keywords: power law, heavy-tailed distributions, Pareto, stochastic process methods, punctuated equilibrium theory

In a seminal contribution, Breunig and Jones (2011) showcase the use of stochastic process methods in political science by applying them to budgetary data. Claiming that budget changes follow a power law distribution, they conclude to have found a general empirical law (cf. Jones *et al.* 2009). Yet, their evidence is largely based on the analysis of frequency distributions using log–log plots: if size and frequencies of budget changes scale linearly on log scales, it is, according to their argumentation, evidence of a power law distribution.¹ But as Clauset, Shalizi, and Newman (2009, 675) point out, “being roughly straight on a log–log plot is a necessary but not sufficient condition for power-law behavior.” Clauset, Shalizi, and Newman (2009), in fact, demonstrate that several nonnormal, heavy-tailed distributions can appear as a straight line on log–log scales. Thus, “this method and other variations on the same theme generate significant systematic errors under relatively common conditions, [. . .] and as a consequence the results they give cannot be trusted” (Clauset, Shalizi, and Newman 2009, 665).

In light of these challenges, this letter would like to revisit the claim of a power law distribution in budgetary data. To that end, I apply the principled statistical framework suggested by Clauset, Shalizi, and Newman (2009), involving the following steps: first, we need to estimate the exponent via maximum likelihood, which is, however, conditional on knowing the lower tail. Having obtained estimates of the lower tail and, after that, of the exponent, we can then estimate the goodness of fit between the data and the presumed power law. Second, we can compare the power law fit to fits of other heavy-tailed distributions such as exponential, log-normal, and Weibull via likelihood ratio tests. The *powerLaw* package by Gillespie (2015) offers an implementation of this

Author’s note: I would like to thank Paul C. Bauer, Lukas Rudolph, Tobias Schwarzbözl, my reviewers, and the editors of *Political Analysis* for their input and helpful comments. In particular, I want to thank Christian Breunig and Bryan D. Jones for sharing their data, and Simon Heuberger for his support. Replication material for this study including R code as well as their original data is available at the Harvard Dataverse (Fatke, Breunig, and Jones 2019).

Jones *et al.* (2009), as well as subsequent analyses of power law behavior, also use the l-kurtosis to assess how leptokurtic a distribution is. While this is a convenient way to identify heavy-tailed distributions in general, it does not clearly indicate which function best describes the heavy-tailed distribution.

Political Analysis (2020)
vol. 28:134–138
DOI: 10.1017/pan.2019.33

Published
30 August 2019

Corresponding author
Matthias Fatke

Edited by
Jeff Gill

© The Author(s) 2019. Published
by Cambridge University Press
on behalf of the Society for
Political Methodology.

Table 1. Statistics and parameter estimates (standard deviations in brackets) for budget changes (Breunig and Jones 2011).

	n	Median	σ	x_{\max}	\hat{x}_{\min}	$\hat{\alpha}$	n_{tail}	ρ
Outlays	204	1.062	0.251	2.853	1.118 (0.051)	6.116 (1.394)	60 (48)	0.246
Domestic	204	1.069	0.184	2.426	1.098 (0.065)	7.761 (1.572)	69 (64)	0.622
Defense	204	1.088	0.337	3.408	1.211 (0.094)	4.817 (1.144)	48 (64)	0.422

procedure for R. The R code for this analysis is available at the Harvard Dataverse (Fatke, Breunig, and Jones 2019).

The letter is structured according to these steps and applies them to the budgetary data by Breunig and Jones (2011).² As the analysis does not reveal unambiguous support for a power law distribution in public budgets, the findings invite an empirical as well as theoretical refinement. Since, in a wider sense, the contribution intends to showcase the use of a more thorough, yet straight-forward approach to stochastic process methods in political science, the concluding section offers some more general remarks.

1 Parameter Estimation

This section details the estimation of the power law parameters and its goodness-of-fit measures using the budgetary data of Breunig and Jones (2011). Results for overall, domestic, and defense outlays are shown in Table 1. (Note that the scaling differs from the original values, since absolute values of budgetary growth rates are used and linearly transformed to have a minimum of 1. Further below, I also report the results for separate analyses of positive and negative values.) Power law functions are defined as $f(x) = x^{-\alpha}$. Thus, we need to estimate the exponent α (also called scaling parameter). However, power laws typically appear only above a threshold x_{\min} (also called lower bound). In other words, the estimate of α is conditional on the value of x_{\min} , which has to be estimated first. This should be done with care because a high threshold implies discarding valuable data points. Clauset, Shalizi, and Newman (2009) propose to set \hat{x}_{\min} so that the remaining data is as similar as possible to a power law distribution. In practice, they advocate using the Kolmogorov–Smirnov (K–S) statistic to minimize the distance between the cumulative distribution functions (CDFs) of the data and the fitted model. To estimate the uncertainty of \hat{x}_{\min} a nonparametric bootstrap procedure can be applied that repeats the K–S statistic for randomly sampled values of the original data. In the transformed budgetary data \hat{x}_{\min} is estimated to be 1.12 leaving 60 observations as n_{tail} . Clauset, Shalizi, and Newman (2009, 669) suggest that at least 50 observations are necessary for reliable parameter estimation. With the lower bound set, we can calculate the maximum likelihood estimator for the scaling parameter $\hat{\alpha} = 1 + n[\sum_{i=1}^n \ln \frac{x_i}{x_{\min}}]^{-1}$, which is asymptotically normal and consistent. The standard error can be computed analytically.³

Having estimated the most likely $\hat{\alpha}$ and \hat{x}_{\min} for our distribution, we can assess whether it indeed follows a power law. Alas, the challenge to conclusively identify power law behavior in empirical data is not trivial. Since a power law can be fitted to any empirical distribution, it is, in fact, only possible to test whether an empirical distribution of x is consistent with the hypothesis that it is drawn from a distribution of the form $f(x) = x^{-\alpha}$ rather than from

- 2 The data they collected has been made available at the Harvard Dataverse for the purpose of replication (Fatke, Breunig, and Jones 2019).
- 3 It should be noted that the principled statistical framework does not explicitly address the issue of measurement error. However, in contrast to the inspection of log–log plots, the bootstrap procedure does report the accuracy of the estimates by randomly sampling values of the original data (values in Table 1 in brackets). Moreover, if the goal is to assess whether data resembles power law behavior, measurement error should not work in principle in favor of detecting false positives. Nonetheless, stochastic process methods should be more considerate in that regard. If, for instance, information on the accuracy of a measure is available, it would be possible to use multiple separate imputations so that for any given observation, uncertainty in the measure is given by the differences across these imputations.

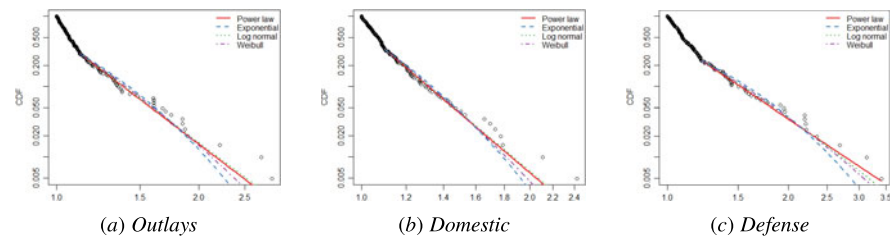


Figure 1. CDFs of budget changes (Breunig and Jones 2011) and corresponding fits.

(heavy-tailed) distributions of other forms. In contrast to qualitative appraisals using CDFs or log–log plots, which can only show a necessary but no sufficient condition, Clauset, Shalizi, and Newman (2009) suggest a quantitative assessment using, again, the K–S statistic. To that end, multiple (typically between 1,000 and 10,000) synthetic data sets with the same parameters as estimated before are generated from the power law function. For each, we calculate the K–S statistic to assess its fit to the power law distribution, and compare it to the K–S statistic of the original distribution. The p value corresponds to the fraction of synthetic data sets whose K–S statistic is larger (i.e., goodness of fit is poorer) than for the empirical data. Thus, p values close to one suggest that differences between the empirical distribution and the power law model can be attributed to statistical fluctuations. If $p \leq 0.1$, a power law is ruled out. In the case of budgetary changes, p values range above this cutoff between 0.25 and 0.62. Yet, this does not allow concluding that budgetary changes do follow a power law since other heavy-tailed distributions might describe the data equally well. The next section, therefore, presents likelihood ratio tests to compare the power law fit to fits of exponential, log-normal, and Weibull distributions.

2 Comparison to Other Distributions

Whereas the previous section assessed how well the data fits the power law function, comparing goodness of fits between several heavy-tailed functions can be insightful, too. Ideally, there are competitive theoretical expectations implying different distributions. But even if this is not the case, a comparative assessment is more conservative. It should be noted though that without alternative theoretical expectations the analysis is more explorative in nature. To assess, which function provides the best fit, we first fit several nonnormal and heavy-tailed functions to the data using the same $\hat{\chi}_{\min}$ as found above for the power law distribution. Figure 1 illustrates CDFs for the three cases of budgetary changes and corresponding fits of exponential, log-normal, Weibull, and power law distributions. Log-normal and Weibull in particular deviate only marginally and seem to describe the empirical data equally well.

Having fitted several distributions to the data, we can perform likelihood ratio tests to assess under which of two competing distributions the empirical distribution is more likely. Following Vuong’s method, we test H_1 that one of the two distributions is closer to the empirical distribution by inspecting whether the ratio of the two estimated log-likelihoods is significantly different from zero. If that is the case, the sign of the ratio indicates whether the alternative distribution better describes the data than the power law distribution, or not. Table 2 displays log-likelihood ratios and p values of two-sided tests of the log-likelihood ratio being significantly different from zero. As none of the p values indicates significant differences, we cannot reject H_0 that both distributions are equally far from the empirical distribution.

3 Further Analyses

The maximum likelihood estimator for the exponent α , being equivalent to the Hill estimator, is potentially biased in small samples. An alternative procedure is to regress the logarithm of the values on the logarithm of their rank minus 1/2 (Gabaix and Ibragimov 2011). These estimates,

Table 2. Log-likelihood ratios and p values (in brackets) for budget change data (Breunig and Jones 2011) compared to power law fits.

	Outlays		Domestic		Defense	
Exponential	1.420	(0.156)	1.239	(0.215)	0.554	(0.579)
Log-normal	0.072	(0.943)	0.069	(0.945)	-0.081	(0.935)
Weibull	0.990	(0.322)	0.958	(0.338)	0.009	(0.993)

reported in Table A1 and Figure A1 in the supplementary material, are by and larger similar, but somewhat lower (4.603, 6.157, 3.820, respectively) for all three types of budget changes. Using the bias adjusted exponent estimates for the assessment of power law behavior, yields the same results. Again, p values are too large to rule out power law behavior, while log-likelihood ratios do not indicate better fits of the power law distribution than of other heavy-tailed distributions.⁴

When dealing with data on changes in public budgets, we should also consider the possibility that increases exhibit different behavior than decreases. For that reason, Jones *et al.* (2009) conduct their analysis separately for positive and negative values. Specifically, the “contagion of urgency” (Jones *et al.* 2009, 871) leading to power law distributed budget changes might apply when it comes to extending budgets. Cutting, in contrast, involves more negotiations and concessions, and thus makes extreme punctuation less likely. To that end, I repeat the analyses separately for positive and negative values. The results in Tables A2 and A3 as well as in Figure A2 in the supplementary material do not differ meaningfully from the previous results of the overall distribution. None of the p values let us rule out power law behavior. At the same time, most of them are not as high as to attribute differences between the empirical distribution and the power law model only to statistical fluctuations (except in the case of cuts to domestic outlays). Comparing the fits to other distributions, we again cannot confidently favor one model over another. Given the low number of observations in the separate analyses, however, these estimates should be treated with care.

In addition to formal likelihood ratios tests presented so far, it is also possible to infer graphically the distinctiveness of the empirical distribution. Figure A3 in the supplementary material includes multiple histograms of random data sets generated from various heavy-tailed distributions with the same parameters as estimated before. After reverting to the initial scaling and joining with the empirical values for $x < x_{\min}$, they are plotted along with the original one in Breunig and Jones (2011, 107). The procedure can be reproduced using the provided R code. Evidently, it is challenging to pick the one histogram, which is based on empirical data, as all tails look rather similar and hardly any distribution seems to stand out clearly. While it is generally difficult to discern heavy-tailed distributions, this can be seen as further, and rather intuitive evidence that budget changes might just as well be described by another heavy-tailed function.

4 Concluding Remarks

Power law distributions appear rather commonly in nature. Not only does the heavy tail of a power law allow for a higher frequency of extreme events than in Gaussian distributions, but its scale-free characteristic, moreover, imbues “them with a vague and mistakenly mystical sense of universality” (Stumpf and Porter 2012, 666). It is thus understandable that they are sought after in data generated by social or political processes. However, the excitement should not come at the expense of a thorough assessment. This letter intends to draw the attention to what Clauset, Shalizi, and Newman (2009) have proved: log–log plots are not an appropriate tool to establish

⁴ This result mirrors the conclusion by Ibragimov, Ibragimov, and Kattuman (2013) who report that the results using the adjusted estimates are similar to those using Hill’s estimation procedure. Yet, it seems promising to implement robust versions of the Hill estimator in the principled statistical framework (e.g., Beran and Schell 2012).

empirical support, because the scaling property on log scales is a necessary, but no sufficient condition. Instead, researchers should follow the principled statistical framework suggested by Clauset, Shalizi, and Newman (2009) and showcased in this letter with budgetary data by Breunig and Jones (2011).

Several findings of supposed power laws have not fared well against this scrutiny (Clauset, Shalizi, and Newman 2009; Stumpf and Porter 2012). A more thorough assessment will inevitably lead to more conservative estimates, and thus fewer startling results. On the one hand, a general empirical law of public budgets might be another case in point. For empirical analyses it is not necessarily inconvenient if a distribution does not follow a power law since it may become easier to calculate moments and conduct estimations. That does not imply it would not make sense to think about power laws in political data at all. There are many political phenomena with heavy tails covering several orders of magnitude. And finding evidence of power law behavior would indeed have implications for the data generating process as well as for applicable statistical methods. On the other hand, the result of applying the principled statistical framework did not rule out a power law distribution of budget changes either. As with many other political processes, data on budget changes, particularly regarding the tail of the distribution, are probably simply too sparse to confidently conclude power law behavior. This poses a general challenge because real systems are finite, whereas models of power law behavior assume infinite systems (Stumpf and Porter 2012, 665).

This letter therefore seconds the note of caution expressed by Stumpf and Porter (2012). In addition to the application of the principled statistical framework, investigating power law behavior should comprise a theoretical aspect as well. Genuine insights can be gained when we are not only able to describe the data accurately, but also to formulate a model whose mechanism predicts a power law distribution. So far, there are no alternative models implying Weibull or log-normal distributions. But like distributions of income, explanations can evolve when more detailed evidence of the underlying functions becomes available. In that sense, the letter can be seen as invitation to revise and refine, empirically as well as theoretically, what has come to be known as punctuated equilibrium theory (PET) in policy analysis.

Supplementary Material

For supplementary material accompanying this paper, please visit <https://doi.org/10.1017/pan.2019.33>.

References

- Beran, J., and D. Schell. 2012. "On Robust Tail Index Estimation." *Computational Statistics & Data Analysis* 56(11):3430–3443.
- Breunig, C., and B. D. Jones. 2011. "Stochastic Process Methods with an Application to Budgetary Data." *Political Analysis* 19(1):103–117.
- Clauset, A., C. R. Shalizi, and M. E. J. Newman. 2009. "Power-law Distributions in Empirical Data." *SIAM Review* 51(4):661–703.
- Fatke, M., C. Breunig, and B. D. Jones. 2019. "Replication Data for: How to Assess Power Law Behavior Using Stochastic Process Methods: A Note of Caution." <https://doi.org/10.7910/DVN/8DWAZW>, Harvard Dataverse, V1, UNF:6:fwel8mNrf05YOWL2U+0v/Q== [fileUNF].
- Gabaix, X., and R. Ibragimov. 2011. "Rank-1/2: A Simple Way to Improve the OLS Estimation of Tail Exponents." *Journal of Business & Economic Statistics* 29(1):24–39.
- Gillespie, C. S. 2015. "Fitting Heavy Tailed Distributions: The PowerLaw Package." *Journal of Statistical Software* 64(2):1–16.
- Ibragimov, M., R. Ibragimov, and P. Kattuman. 2013. "Emerging Markets and Heavy Tails." *Journal of Banking & Finance* 37(7):2546–2559.
- Jones, B. D., F. R. Baumgartner, C. Breunig, C. Wlezien, S. Soroka, M. Foucault, and A. Franc et al. 2009. "A General Empirical Law of Public Budgets: A Comparative Analysis." *American Journal of Political Science* 53(4):855–873.
- Stumpf, M. P. H., and M. A. Porter. 2012. "Critical truths about power laws." *Science* 335(6069):665–666.