sciendo

Lukasz Wordliczek[1]

Between incrementalism and punctuated equilibrium: the case of budget in Poland, 1995-2018

ABSTRACT: *Incrementalism and punctuated equilibrium theory (PET) have secured their standing in public policy research when studying change in budgetary data. On the other hand, however, new empirical evidence is constantly developed to confront it with theoretical assumptions. In line with the above, the aim of the paper is threefold. First, it is examined if budgetary outlays in Poland follow either incrementalism or PET's core premises. Second, the paper aims at facilitating discussion on identifying punctuations. It is claimed that any cut-off point should be data-driven, category-responsive, and generalizable across different types of outliers. And third, it is investigated which of the budget categories have the most punctuations. Methodologically, the study is based on descriptive and distributional statistics provided to tackle the above two issues comprehensively. Consequently, the paper aims at filling the gap in theory-driven literature on Polish budget shifts and their empirical rigorous explanations. Thus, it is claimed that the Polish case study contributes to the debate on the verification of empirical research on public policy agendas and public policy change.*

The data will obviously not determine directly the outcome of debate between various schools of thought; it does, however, influence the conflict by defining what battlefield positions must be.

(Sims, 1980, p. 30)

INTRODUCTION

The question of stability and changes in budgets is critical for at least two reasons. First, it serves policymakers in their assessment of the structure of spending in a given political entity (nation-state, region, municipality, etc.). It may be reasonably argued that evidence-based decisions are being made through the budgetary process (C. Breunig & Koski, 2006; M. M. Jordan, 2003). After all, deliberation on spending and revenues calls for consideration of the political environment and its prominent component—budget structure—due to its intrinsic and decisive character. Second, budgetary fluctuations research is informative for scientific reasons. The relevant body of research is decades old and debate is prolific in terms of theoretical argument, methodological sophistication, and empirical verification (C. Breunig & Jones, 2011; Davis, Dempster, & Wildavsky, 1974; Dempster & Wildavsky, 1979; Jones et al., 2009; Padgett, 1980; Wildavsky, 1964). Yet, notwithstanding the achievements of budgetary public policy studies, there are still issues that call for consideration, such as the detection of abrupt changes in data at hand.

One of the most perplexing issues in public policy scholarship is still a shortage of studies going beyond a pool of countries covered in the Comparative Agendas Project (hereafter CAP; https://www.comparativeagendas.net/). As of the writing of this article, the CAP list includes twelve countries: Australia, Belgium, Brazil, Denmark, France, Germany, Hungary, Italy, the Netherlands, Spain, the United Kingdom, and the United States. Some data and research is also available for the state of Pennsylvania and the European Union. As the list shows, there is only one—albeit notable—example of a Central/Eastern European country. This scarcity makes any comparative approaches well limited and empirically constrained (for literature review, please refer to the "Theory and terminology" section). Thus, the current paper serves as a relatively modest attempt toward filling up the void. Also, hopefully, it will

---

1 Institute of American Studies and Polish Diaspora, Jagiellonian University, Krakow, Poland

encourage scholars and researchers to develop the CAP framework in other Central/Eastern European countries. As was the case with any of the entities already covered by the CAP, this would be beneficial to the development of case studies as well as any comparative research designs.

There are two main aims of the following paper. First, it is investigated if budget outlays in Poland follow assumptions based on incrementalism or punctuated equilibrium theory (PET). There is a good deal of conjecture that it should follow the pattern present in other countries. This issue, however, to the author's best knowledge is severely underestimated.[2] And, second, if the pendulum of argument would move toward PET, it will be researched which of the budget categories experienced punctuations. Obviously, this leads to the very accurate question: When and why prospective punctuations have occurred. There is a myriad of possible explanations that contribute to the structure of budgetary decisions. Yet, some of them may be operationalized by independent variables. First, it may be argued that abrupt budget changes are related to the electoral cycle: in presidential and parliamentary election years, policymakers may be inclined toward spending more money to gain the electoral advantage over contenders through breaking constituents in and pleasing them (Davis et al., 1974; Kamlet & Mowery, 1987; Mueller, 2003). Second, a switch in the parliamentary majority may contribute to major budget relocation of resources: if only people voted for other political forces, this implies that they were charged with making a difference to what was done before; if one looks for a continuum, s/he would probably vote for incumbents and continuity (Bozeman, 1977; Cox, Hager, & Lowery, 1993). Also, the macroeconomic factor may be considered: in times of pressure resulting from the budget deficit, there may be a "window of opportunity" open for making necessary budget adjustments (Davis et al., 1974; Noguchi, 1980; Wanat, 1974). Yet, these three issues, however critical for the explanation, call for a separate explanatory study. The starting point is to investigate data structure and, specifically, to determine which of the data points available may be labeled as punctuations.[3] This would hopefully allow for a more rigorous understanding of policy processes. All in all, the paper contributes to theory-driven literature on budget shifts and their empirical explanations. Specifically, the assumption on punctuations/incrementalism mechanisms is tested through some descriptive and distributional statistics. This allows for data-driven consideration of cross-function and over-time changes in budget outlays. Importantly, it is argued to be rather a departure point for future research, not a final conclusion.

Consequently, the paper is organized as follows. First, theoretical considerations on budget changes are scrutinized and terminology is explained. These serve as a basis for the next section that explains in detail data and methodology. Here, the focus is on identifying punctuations in budget categories. Part three presents and discusses the results whereas the closing section critically approaches conclusions and shows some possible alleys for future research.

THEORY AND TERMINOLOGY

There are two main strains of theory-oriented research in public policy on budget fluctuations: incrementalism and punctuated equilibrium theory. For obvious reasons, there is no need to repeat or review both approaches in much detail, but, however, for the sake of clarity of the following argument, some basic considerations are evoked.

Chronologically, incrementalism should be discussed first. Its main assumptions are based on foundational works published in the 1950s (Lindblom, 1959; Simon, 1955), and some empirical evidence is based on observations of relatively parsimonious budget shifts (Davis et al., 1974; Wildavsky, 1964). This "stability logic" contributes to a specific methodological assumption. Since there are mostly minor adjustments in outlays, the data points are heavily clustered around zero point ("no change") and gradually fade out toward distribution's tails. This makes incremental budgeting to follow a normal (Gaussian) distribution (C. Breunig & Koski, 2006, p. 3). Interestingly enough, there is some evidence that incremental budgetary considerations were also relevant for communist countries (Bunce & Echols, 1978), notwithstanding their peculiarities and distinctiveness.

Based on the above findings, PET's departure point also acknowledges that the budget structure is governed by equilibrium. This, however, does *not* rule out that the budget is more "dynamic" in terms of its changes. Indeed, occasionally robust and abrupt punctuations take place. This structure, on the other hand, is followed not by a bell-shaped distribution but a leptokurtic ("fat tailed")

---

2 For a notable departure from this scarcity, please refer to (Bunce & Echols, 1978). As publication date shows, however, the research is relevant for historical context only.

3 Hereafter, the terms "punctuations" and "outliers" will be used interchangeably.

Tab. 1: Thresholds defining incremental and punctuated budget changes

Author(s)	Range cut-off points
Dezhbakhsh et al., 2003	40 and 45 percentile point below/above zero
Sebők & Berki, 2017	40% below/above zero
Flink, 2017; Flink & Robinson, 2020; Robinson, Caver, Meier, & O'Toole, 2007; Robinson, Flink, & King, 2014	{+35.5%; -33%} ±5% margin
M. M. Jordan, 2003	{+35%; -25%}
Bailey & O'Connor, 1975; Wildavsky, 1964	30%
Gist, 1974	20%
Jones et al., 1998	{+20%; -15%}
Fenno, 1966	{10%; -10%} and {20%; -20%}
Kemp, 1982; Wanat, 1974	10%
Baumgartner & Epp, 2013; Kanter, 1972	5%

Note: Entries refer to bands and thresholds based on percent change in the budget. Please refer to the sources for details.

one. The main reason behind such a phenomenon is the policy system and its limitations; "institutional friction" is just one of the most prevalent. Furthermore, if the policymaking process is structured in terms of information system (as PET assumes), this also contributes to two contradictory features manifested in one system: equilibrium and punctuations (Baumgartner, Green-Pedersen, & Jones, 2006; Baumgartner & Jones, 2002; True, Jones, & Baumgartner, 1999, 2007). Empirical verification made the PET's "Founding Fathers" to describe the pattern as a "general empirical law of public budgets" (Jones et al., 2009).

Notwithstanding theoretical grounding, the research has at least one disturbing aspect: identifying punctuations. How to come up with abrupt changes? When do we deal with equilibrium-like departures? When a change may be called a substantial one? Is every change a punctuation or, to be more specific, when does a change become a punctuation? To come to terms with questions like these, the following analysis treats punctuations as outliers—or anomalies—in their statistical meaning (Schubert, Wojdanowski, Zimek, & Kriegel, 2012, p. 1). For the sake of clarity, a dictionary definition may be of some relevance here. Merriam-Webster Dictionary defines an outlier as "a statistical observation that is markedly different in value from the others of the sample" whereas anomaly is "something different, abnormal, peculiar, or not easily classified. . . .deviation from the common rule" ('Merriam-Webster Online Dictionary', 2019). Consequently, for the current purposes, anomaly/outlier/punctuation is an observation in time series data that substantially differ from the rest of data points (Grubbs, 1969, p. 1). The obvious question is then: What does it mean to be *substantially* different? The majority of approaches employ user-defined thresholds to label certain observations an anomaly; see Table 1 for some detailed coverage.

This plentiful and illustrative survey underlines the critical importance of setting the right cut-off point. It seems justifiable to introduce the measure that would satisfy the following criteria *collectively*: (1) to be user-independent, i.e., not being arbitrarily determined but rather data-determined, (2) to be determined by each budget category instead of setting any universal measure across the entire dataset in order to follow different extreme values across categories (M. M. Jordan, 2003, p. 352; Munir, Siddiqui, Dengel, & Ahmed, 2019, p. 1997), and (3) to be able to differentiate well across various kinds of outliers. To put it in other words, any threshold should be, respectively, data-dependent, category-responsive, and of generalizability potential. Further analysis aims at introducing such an anomaly detector.

Previous research on anomaly detection falls into several categories. Since the topic has already been covered in the literature, there is no need to repeat others' work (Agyemang, Barker, & Alhajj, 2006; Chandola, Banerjee, & Kumar, 2009; Goldstein & Uchida, 2016; Hodge & Austin, 2004; Khan & Madden, 2014; Xu, Liu, & Yao, 2019). At the same time, however, for the sake of the clarity of the argument some basic review seems to be grounded. There are several possible classifications offered. One of the most elementary covers the following three categories:

(1) *Distribution-based* models that fit a specified probability distribution and then, based on discordancy tests, outliers are identified (Barnett & Lewis, 1994). The category accommodates statistical techniques to fit a model to data thus parametric and

nonparametric variants are available. The former assumes that distribution (usually Gaussian) is known a priori and belongs to one of the first approaches used in the area of outlier detection. An example of a classic formula of 3 and its extensions belongs here (Grubbs, 1969; Laurikkala, Juhola, & Kentala, 2000; Shewhart, 1923; Tukey, 1977). Another simple technique is using histograms. To put it succinctly, in its basic form, a histogram of a variable of interest is plotted, and then it is examined if an observation falls in any of the bins of the histogram. If it does, the case is normal, otherwise it is abnormal (Denning, 1986; Endler, 1998; Helman & Bhangoo, 1997; Yamanishi, Takeuchi, Williams, & Milne, 2004). On the other hand, nonparametric approaches are more flexible since they do not need to determine a specific data distribution in advance, but it is inferred from data at hand. This makes them scalable to other applications (Laptev, Amizadeh, & Flint, 2015) at the cost of computational complexity. There are at least two limitations of distributional techniques. First, in its parametric version, they preconceive that the data belongs to a particular distribution that is often not true; multidimensional data is a case in point. And second, applying any of the discordancy tests is often not a straightforward task since there are no generic heuristics to choose one particular statistics out of over one hundred available tests (Barnett & Lewis, 1994).

(2) *Distance-based* approach offers various outlier measures calculated with observations' distances to its kth nearest neighbor in the dataset. One of the main advantages is a possibility not only to identify anomalies but also to rank them in terms of their outlyingness through outputting an anomaly score (Knorr & Ng, 1998; Knorr, Ng, & Zamar, 2001; Ramaswamy, Rastogi, & Shim, 2000).

(3) *Density-based* models aim at estimating the global density for each data instance through counting the number of neighbors in a hypersphere of a given radius. Obviously, the observation that is in a neighborhood with low density is declared to be an outlier while an instance that lies in high-density area is declared to be normal (Amer & Goldstein, 2012; M. Breunig, Kriegel, Ng, & Sander, 2000; Papadimitriou, Kitagawa, Gibbons, & Faloutsos, 2002). Here, it is also possible to assign a degree of being an outlier but contrary to distance-based models, a probability measure instead of a score is implemented. The major limitation is related to datasets that contain regions with varying densities.

Notwithstanding their drawbacks, one shall acknowledge that researchers have developed some remedies to advance the three methodological designs (for details please consult literature cited).

The above review of incrementalism, PET, and anomaly detection serves as a departure point for further analysis. It is aimed at complementing to existing research on budgetary issues in the economy (de Crombrugghe & Lipton, 1994). Here, it is argued that public policy studies may also contribute to the debate on analyzing budget structure.

DATA AND METHOD

For the sake of robust verification of the above assumptions, the author collected data on budget outlays in Poland between 1995 and 2018. Data were derived from official records provided by the Central Statistical Office (https://stat.gov.pl/en/). Dataset consists of time series with 25 budget functions for 1995–2000 and 24 functions for 2001–2018. The reason for two subsets instead of one stems from a change in the classification of budget categories since 2001. This poses some analytical challenges; please refer to the below discussion for details. All the items effective from 2001 are specified in details in a decision issued by the Ministry of Finance; as of writing the article, the most current version was released in 2014 (Ordinance of the Minister of Finance, 2014). Furthermore, in the 2017 fiscal year, a new budget function was added—Family—but due to few observations, this category was dropped from the analysis. Overall, this makes data inconsistent in terms of the structure. To approach the issue, a crosswalk was prepared based on the Polish Classification of Activities manual; please refer to Table 2 for details. As may be clearly seen, some of the categories vary in terms of their names whereas, more importantly, some were merged into broader categories to make them as consistent throughout the time window as possible.

The above approach combines two major types of budget outlays in Poland: mandatory and discretionary. The former constitutes the majority of the budget and they are automatically obligated by virtue of enacted laws. National defense and public debt financing are two examples of mandatory appropriations in Poland. On the other hand, discretionary appropriations are set on a yearly basis as specified in statutory provisions; salaries and wages of public sector employees is one example of discretionary appropriations. They

Tab. 2: Budget major functions crosswalk

Budget major functions, 1995–2000	Budget major functions, 2001–2018
Industry	Mining and quarrying
Construction	+ Manufacturing
Agriculture	Agriculture and hunting
Forestry	Forestry
Transportation	Transport and telecommunication
Communication	
Trade: domestic	Trade
Trade: foreign	
Miscellaneous material services	Services
Municipal services	Communal services and environmental protection
Housing economy and intangible municipal services	Dwelling economy (housing)
Science	Science
Education	Education + educational care
Higher education	Higher education
Culture and art	Culture and national heritage
Health care	Health care
Social welfare*	Social assistance and other social policy issues
Physical education and sport	Physical education and sport
Tourism and recreation	Tourism
State administration	Public administration
Administration of justice and public prosecutor's office	Administration of justice
Public safety	Public safety and fire care
Finance	Public debt servicing
Social security	Compulsory social security
National defense	National defense

Until 2003
Source: Statistical Yearbook of the Republic of Poland, 1995–2018 editions, categories according to the Polish Classification of Activities (PKD 2007).

also serve as a vehicle for attempts toward making the budget balanced. All in all, mandatory and discretionary appropriations reflect the policy priorities of a given government and parliamentary majority in a given fiscal year.[4]

Budget outlays delivered by the Central Statistical Office are originally available in nominal values (millions of zloty) in a given year. To make the data to serve well the research objectives, two transformations were introduced (Bunce & Echols, 1978; Dezhbakhsh, Tohamy, & Aranson, 2003; Jones, Baumgartner, & True, 1998; Jones, True, & Baumgartner, 1997). First, amounts were adjusted to inflation for the values to be more comparable across 24 years. To accomplish the goal, nominal values were recalculated with the inflation rate set at 100 in 1995 as a reference point. And second, based on the above real values, yearly percent changes for each category were used for substantial investigation. This would make our units of analysis to be in line with theoretical

---

4 The Fiscal year in Poland overlaps the calendar year. This provision is regulated by Article 109(4) of the Public Finance Act of 2009 (Public Finance Act, 2009). Consequently, budget preliminary studies begin in February of the preceding year, and the budget act is passed by the Parliament and signed by the President in January at the latest of a given budget year. The Parliament assesses execution of the budget till mid-July—at the latest—of the next calendar year.

assumptions on budget fluctuations since it is argued that current budget allocations are based on previous values as points of departure for any adjustments. Also, such transformation allows for controlling for nonstationarity of time series, i.e., changing over time its distributional characteristics. The issue is critical since time series exhibiting a trend may inflate R^2 thus making classical regression invalid (Dezhbakhsh et al., 2003, p. 536). Also, some statistical tests, for example, augmented Dickey–Fuller test, may help to come to terms with the nonstationarity issue.

Consequently, for the current purposes 22 budget functional categories were identified and operationalized in percent yearly changes across 1996–2018 (thus, there are no values for the first year in the real values dataset, i.e., 1995).

The above structure of budget functions classification is, however, not without its price. The suggested merger of the most similar categories from two time periods (i.e. 1995–2000 and 2001–2018) results in at least one observation that seems to be not in line with other data points. Specifically, the Dwelling economy (Housing) budget function results in a value of 4 840% change between 2000 and 2001. This seems to be due to error since other values are within the range between 77.9% and 431% of yearly changes. Also, substantial investigation of the policy behind the Dwelling economy category between 2000 and 2001 did not reveal any empirical grounding for such a significant increase in outlays. Consequently, the observation was dropped from further analysis.

All in all, a complete dataset consists of 22 columns and 23 rows minus one data point, totaling 505 observations. For further analysis, data was transformed to just one dependent variable, i.e., yearly percent changes in budget outlays across functions between 1996 and 2018. Dataset is available from the author upon request.

The resulting dataset was used to determine the validity of the incrementalism/PET hypothesis in relation to Polish budgetary policy. To accomplish the goal, several descriptive and distributional statistics were accommodated. As it was already mentioned, one of the critical points in the budget analysis is to determine data distribution. If it would result in a normal one, there is a good point to acknowledge that we deal with incremental changes since most of the observations will be clustered around 0 and there will be no outliers. On the other hand, punctuations assumed by PET tend to be manifested in leptokurtic distribution. To come to terms with these assumptions, (1) statistical normality tests were calculated, and (2) a density plot of budget yearly percent changes in outlays across time and categories was used.

The above descriptive and distributional approach allowed for calculating frequency distributions of budget yearly changes. This was aimed at statistical investigation of punctuations in data. Please refer to the next section for details.


RESULTS AND DISCUSSION

In line with the above assumptions, the empirical part of the research was divided into two parts. First, several descriptive statistics and tests were used to control for the normal distribution of the dependent variable. Specifically, two versions of the Kolmogorov–Smirnov test (K-S) were run: with defined parameters for normal distribution with a zero mean and variance = 1, and with estimated parameters. Since the power of the K-S test is questioned, also other normality tests were run, including Shapiro–Wilk, Lilliefors, and Chen–Shapiro. Detailed results, not reported here due to space limitations, show that with any of the above tests, the normality assumption may be rejected at a 0.05 significance level. As using any normalcy test has its own limitations, further investigation was performed with the use of a density plot of budget yearly percent changes in outlays across time and categories (see Figure 1).[5]

Visual investigation of the distribution of percentage shifts for all budget functions clearly confirms that the Polish budget structure in terms of its dynamics is in line with other case studies. Specifically, the budget leptokurtic distribution indicates that the majority of changes are incremental since they are clustered around zero.[6] But there is also a higher probability of some extreme values (i.e., punctuations) than a normal distribution would assume (Jones, Sulkin, & Larsen, 2003, p. 164). This observation is especially true for positive values and it is consistent with core PET theoretical assumptions and empirical findings on underreacting

---

5 Another option was to calculate the variable's kurtosis value (Baumgartner & Epp, 2013; C. Breunig, 2006; Jones et al., 2003). This, however, is not a robust approximation since one of its limitations is sensitivity to extreme values (Jones et al., 2003, p. 158; Robinson et al., 2007, p. 149). On the other hand, kurtosis allows for the initial assessment of outliers in data: the higher the kurtosis, the more extreme outliers are in a given distribution. Kurtosis has a value of three for the normal distribution, and distributions with values greater than 3 are called leptokurtic. They tend to have slender peaks and heavy tails producing more outliers. For values less than 3 (platykurtic distribution), it is the other way around. The calculated kurtosis value of 56.27 confirms the visual analysis of Figure 1 and its leptokurtic distribution.

6 Obviously, a simple **run-sequence plot** repeats information from the histogram in Figure 1. Therefore, it is not reported here.

Fig. 1: Annual percent changes across budget functions, 1996–2018
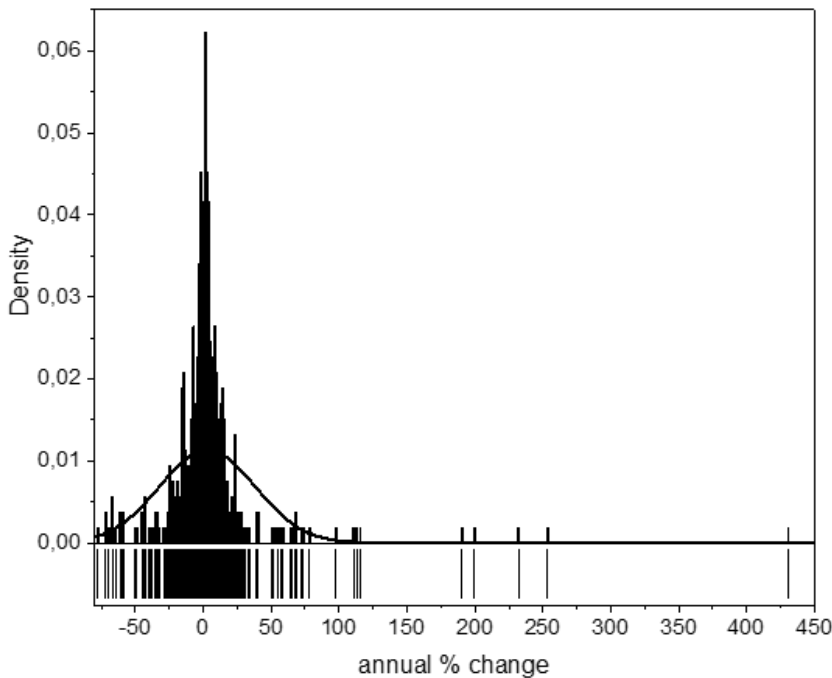
and overreacting in policy processes. Interestingly enough, punctuations are skewed toward positive values. This implies that abrupt changes are easier to force through when increasing budget items are at stake. On the other hand, negative changes are relatively modest in terms of year-to-year reductions of outlays. This suggests that cutting funding is basically a more cumbersome process when compared to budget increases.

The above findings contribute to a general thesis on punctuated equilibrium that describes the dynamics of Polish budget outlays between 1996 and 2018. This, however, does not contribute to identifying punctuations themselves. Eyeballing over Figure 1 may easily suggest that the most right-side observation is an obvious candidate for punctuation. But when we move to the left, the question arises: Are other data points—these more toward the "base" of zero—also outliers? How far shall we move to stop considering observations as outliers? Where are inliers–outliers limits set? Here, the issue of finding the right threshold for defining incremental and punctuated changes shows its potential to be explored.

Since we already know that data normal distribution assumption does not hold true, some convenient techniques—such as Grubb's test, Dixon's Q test, interquartile range boxplots (Tukey, 1977), or Chauvenet's criterion—are not viable options to formally test whether observations are outliers. To come to terms with distributional obstacles, a more robust rule-based approach seems to be feasible. One of the possible options is the Median Absolute Deviation (MAD) test. Notwithstanding its relative simplicity, yet it is more resilient to outliers than any of other standard techniques based on the mean and the standard deviation (Bartolucci, 2016; Hampel, 1971; Leys, Ley, Klein, Bernard, & Licata, 2013). Formally speaking, the MAD belongs to nonparametric techniques what is also a relevant argument here. Furthermore, the MAD is argued to work well notwithstanding the sample size (Leys et al., 2013, p. 2).

Specifically, the MAD test computes the median of the absolute deviations from the median of original input data:

$$\text{MAD} = \text{median}(|x_i - M|) \tag{1}$$

where x_i is an original observation, and M is the median of the original data set. In the case of our pooled data, its MAD = 7.25. Now, the question of the rejection criterion of the value emerges. Following the literature (Bartolucci, 2016; Leys et al., 2013; Miller,

1991), the value of 3 is proposed as a conservative cut-off point for identifying outliers based on MAD. The relevant test statistic stems from the following formula:

$$\frac{x_i - M}{MAD} > 3 \tag{2}$$

To put it in other words, our decision boundary is set as $M - 3 \times MAD < x_i < M + 3 \times MAD$ (Leys et al., 2013, p. 3). This yields 1.16 ± (3 × 7.25) = {-20.59; 22.91} rejection limits marking nearly 20% of all observations as outliers. Considering the above approach to be rather generous in terms of its ability to sort out abnormal data points, it seems relevant to look at them more closely.

There is a huge variation in the percentage of outliers across budget functions: for one category ("Trade") the number is as high as more than 50%, whereas six categories are marked with no outliers at all. At first sight, this may seem to be troublesome since it suggests no pattern in data. Closer investigation, however, allows for acknowledging that the more outliers in a budget function, the more likely it is a discretionary part of the budget. To put it in other words: most mandatory budget categories are associated with relatively few outliers in budget outlays. The possible rationale for such an equilibrium-like mechanism is that it is relatively difficult to substantially change such items since they are most often planned in advance. On the other hand, discretionary funding is based on ad hoc decisions that allow for more variation in terms of spending limits. What is also important is the fact that mandatory spending is gradually making most of the total budget followed by its relative lack of flexibility. This observation is consistent with other countries, for example, the United States (A. A. Jordan, Taylor, Meese, Nielsen, & Schlesinger, 2009, p. 197).

The above analysis was based on the distribution estimated through pooling the annual change in budget for all functions and all years together. Through putting all observations into one basket, however, some patterns may be masked. Thus, the next analytical step was to run the MAD test that uses the annual variation of budget changes across categories and—separately—across time. Consequently, the statistics varies accordingly to, respectively, fluctuations in budget functions and in time. This allows for control for outliers, i.e., substantial percentage changes in budget outlay with a focus on functions and time *separately*.

According to findings in Table 3, the cross-category approach does not deliver substantially different results, either in terms of absolute or relative values. It allows, however, for a more balanced identification of outliers in data: it reduces high numbers present in the first approach and finds more outliers in other categories sparsely labeled with outliers in pooled data. Thus, the budget category-centered procedure may be treated as a more stable one. Over-time distribution in Figure 2 confirms the finding on a more balanced structure since the number of outliers in any given year varies between one and seven.

The above three specifications—pooled data, cross-category, and over-time distribution of changes—were based on the median measure. Yet, in spite of the fact that is it some conceptual improvement, it still may be unclear where the possible outliers are exactly located in terms of their position in the dataset. To put it in other words, there is a need to merge the above specifications into one to get the fullest possible picture of outlyingness. Thus, the research followed with some extensions.

Specifically, data was analyzed through a standardized distribution in order to set statistically based bands for identifying outliers in a dataset normalized around its central moment. But since we already know that our data is not normally distributed, any classical mean-, covariance-, and standard deviation-based analysis is not the viable option since such features may easily affect the outlier detection performance. For this single reason, (Dezhbakhsh et al., 2003) design toward accounting for class- and time-variation was used, albeit with some major corrections. To put it succinctly, it was modified to accommodate a more robust statistic of central tendency: the median instead of the mean (and others based on it such as the standard deviation) (Hampel, 1971; Rousseeuw & Hubert, 2018; Zimek & Filzmoser, 2018, p. 11). Specifically, there are two rearrangements of the (Dezhbakhsh et al., 2003) approach: (1) percent changes were used instead of real values, and (2) a modified quartile deviation (QD) was used as a robust measure of scale instead of the standard deviation. The rationale for using percent values was already discussed earlier, whereas the applicability of the QD stems from the fact that this statistics robustly addresses dispersion of data in heavily centered distributions. And since we already know that majority of our observations tend to lie densely around the central moment of the dataset, the quartile deviation metric seems to be a viable option. The details on the modification of the QD are discussed later.

Finally, the rates of percent budget changes were standardized. Here, also, one modification was done, i.e., the distribution was standardized according to an outlier-resilient left side of the equation (2) above, i.e., values were subtracted from the median, and next they were divided by the MAD of the pooled data. This strategy allows for accounting for differentiating incremental and abrupt budget changes through the implementation of a statistical-based band in a combined cross-category and over-time distribution. To put it in other words, such a procedure provides a measure not only for the variation across 22 budget functions but also across 23 fiscal years.

Tab. 3: Budget categories and number of outliers based on the Median Absolute Deviation test

Budget category	No. of cases	Pooled data		Cross-function	
		No. of outliers	% of outliers	No. of outliers	% of outliers
Trade	23	12	52.17%	6	26.01%
Physical education and sport	23	10	43.48%	4	17.39%
Mining and quarrying + manufacturing	23	9	39.13%	4	17.39%
Transport and telecommunication	23	9	39.13%	3	13.04%
Education + educational care	23	8	34.78%	5	21.74%
Communal services and environmental protection	23	8	34.78%	6	26.01%
Forestry	23	7	30.43%	3	13.04%
Tourism	23	7	30.43%	7	30.43%
Dwelling economy (housing)*	22	5	22.73%	4	18.18%
Agriculture and hunting	23	5	21.74%	5	21.74%
Culture and national heritage	23	5	21.74%	4	17.39%
Health care	23	4	17.39%	7	30.43%
Social assistance and other social policy issues	23	3	13.04%	9	39.13%
Services	23	2	8.69%	2	8.69%
National defense	23	1	4.35%	3	13.04%
Public debt servicing	23	1	4.35%	1	4.35%
Science	23	0	0%	3	13.04%
Public administration	23	0	0%	2	8.69%
Public safety and fire protection	23	0	0%	4	17.39%
Administration of justice	23	0	0%	2	8.69%
Higher education	23	0	0%	7	30.43%
Compulsory social security	23	0	0%	0	0%
Total	505	96	19%	91	18%

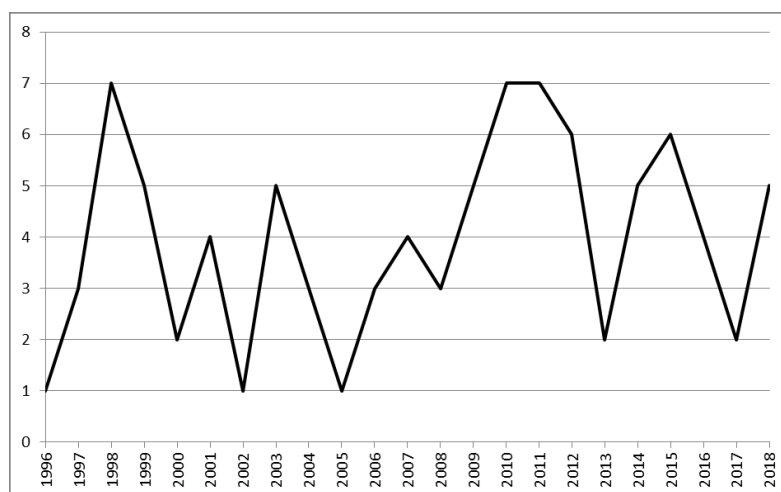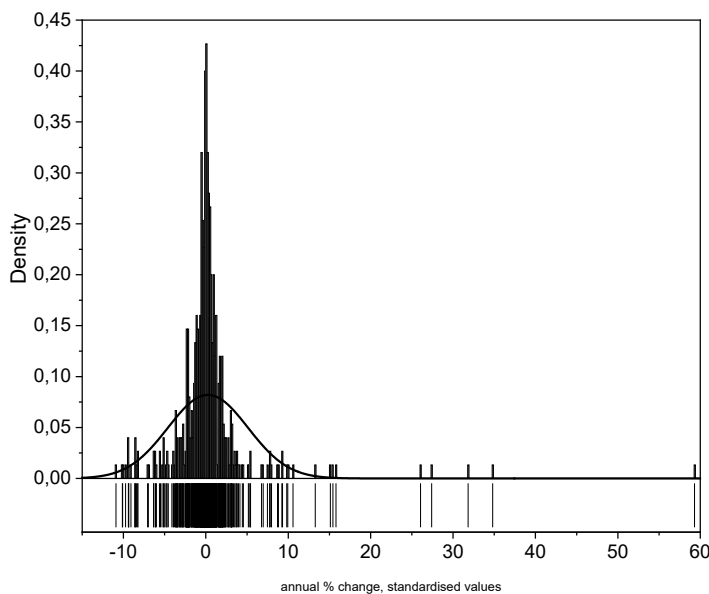*Note: Data was trimmed for the (*) category. See body text for details.*

Fig. 2: Fiscal years and number of outliers based on the Median Absolute Deviation test

Fig. 3: Annual percent changes across budget functions, 1996–2018, standardized values

As is usually the case, the substantial analysis starts with some basic statistical description of data. Figure 3 shows a histogram of the distribution of the budget percentage changes.

By eyeballing of Figure 3, it is evident that our standardized data is not normally distributed.[7] Furthermore, normality departure is twofold. First, distribution is heavily centered around its 0% point. To be more specific, data kurtosis of 54.3 is well over the value of 3 that is customarily assigned to normal distributions, and the skewness measure of 5.5 indicates asymmetry to the right. And second, some points are far in the tails of the distribution—again much further off than the density of normally distributed data should be. Both points are against the classic incrementalism assumptions.

The last analytical puzzle was to set critical values of the standardized distribution. The use of the quartile deviation follows with a 25% margin on both sides of the median. This, however, seems to be too liberal since it labels too many observations as outliers. To tackle the issue, a theory-driven approach was introduced. As we already know from Table 1, several possible options are available to serve as cut-offs. Here, the common measure was accommodated, i.e., two 90 percentile bands designating observations on both sides of the median of the distribution were set leaving 10% of data labeled as outliers/punctuations in each tail. Importantly, and that is the last departure from (Dezhbakhsh et al., 2003), bands were clustered not symmetrically around the central moment level of the distribution. It is argued that margins should follow the data distribution and since it is skewed to the right—consequently bound applied to the positive values is further from the central moment than that for negative values. For clarity, data is visualized in Figure 4.

Interestingly enough, the above approach does not immune from extreme values: the best illustrative example is a data point for Mining and quarrying/Manufacturing in 2001 (with its value close to 60). On the other hand, other extreme values are behind 90 percentile bands. The reason for such inconsistency stems from data scarcity: when too few observations were available, the band tends to embrace variable extreme values. Consequently, for some cases, there was no difference in setting a 90, 95, or 99 percentile cut-off since all of them embraced the most outer data points. This argument follows with a call for more data-ample research design; please refer to the concluding section below for some possible extensions.

Table 4 summarizes the distribution of outliers across categories and fiscal years.

Some of the categories are especially prone to contain outlying observations: mining and quarrying/manufacturing, trade, and physical education/sport are the most "contaminated" whereas science, higher education, health care, public administration, national

---

7 This observation is confirmed by formal tests of data normality. Details are available from the author.

Fig. 4: Bands for cross-category and over-time standardized distribution
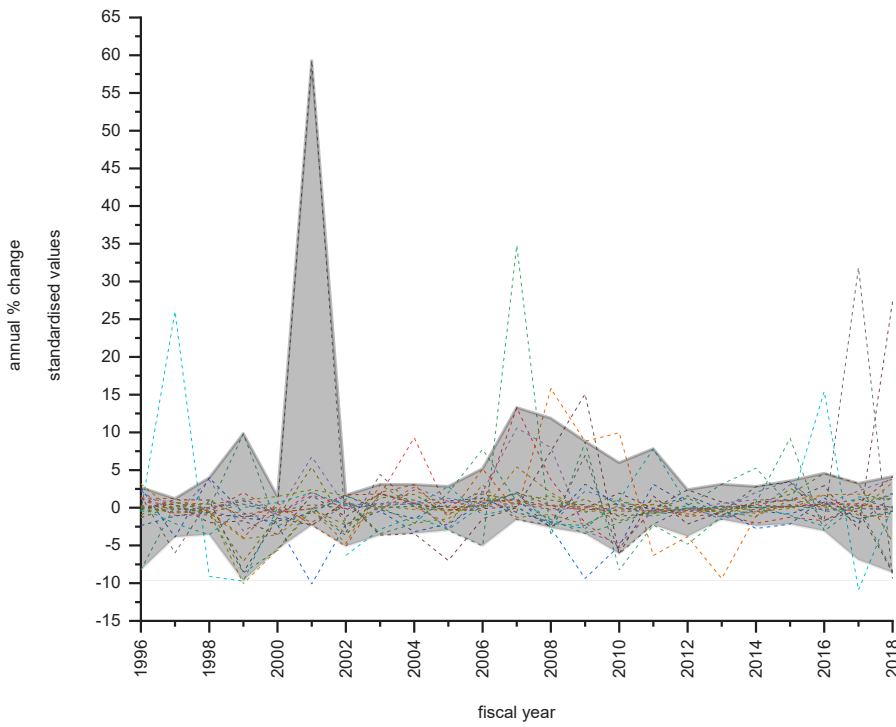
Note: Shaded area indicates bands for 90 percentile point below and above zero.

Tab. 4: Distribution of outliers across categories and fiscal years

	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	TOTAL
Mining and quarrying + Manufacturing		*				*		*				*			*				*			*	*	8
Agriculture and hunting												*										*		2
Forestry	*	*				*							*	*										5
Trade										*	*	*	*		*	*	*			*	*			9
Transport and telecommunication																								0
Tourism	*							*																2
Dwelling economy (housing)			*	*			*								*			*						5
Services																	*					*		2
Science																								0
Public administration																								0
National defense																				*				1
Compulsory social security					*																			1
Public safety and fire protection																								0

Continued **Tab. 4:** Distribution of outliers across categories and fiscal years

	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	TOTAL
Administration of justice							*																	1
Public debt servicing																			*	*				2
Education + Educational care	*				*						*			*				*	*					6
Higher education									*															1
Health care			*																					1
Social assistance and other social policy issues		*								*											*	*		4
Communal services and environmental protection						*		*		*				*								*		5
Culture and national heritage																					*	*		2
Physical education and sport							*		*		*		*	*	*	*	*	*						9
Total	3	2	2	2	2	3	3	3	2	3	3	3	3	4	3	3	3	3	3	3	3	3	4	

defense, compulsory social security, public safety and fire protection, and administration of justice seem to be more "stable," i.e., have no or just one outlier. A different story is for the temporal dimension: here outliers are more equally distributed across all fiscal years.

CONCLUSIONS

Acknowledging punctuations/incrementalism assumptions through descriptive and distributional statistics is important, yet it is just the first step in further analysis. There are at least two points that seem to be critical here: one on methodological extensions and one on theoretical assumptions. Let us take them in turns.

First, a methodologically sound research agenda is feasible. As it was already discussed, the focus here was on examining changes in funding for specific programs operationalized in major budget functions but not minor functions. This lack of comprehensiveness made funding of specific agencies or departments out of scope here. Future research might expand the scope to cover budget beneficiaries in order to check if "institutions matter." Preliminary research has already identified the relevant data sources, but due to their wide range, a separate and more rigorous approach would be mandatory. Please refer to Table 5 for details.

Using major and minor budget functions would result in a substantial improvement in terms of data availability: 688 minor categories across 23 years means 15,824 data points, assuming consistency in typology. For illustrative purposes, details on the 757 budget major code are shown in Table 6.

Notwithstanding possible methodological developments, future research agenda would also invoke theoretical considerations since domain knowledge and expertise are crucial for final decisions in outlier analysis (Zimek & Filzmoser, 2018, p. 8). Thus, let us turn to the second closing point: theory.

Last but not least, on theoretical grounding, there is still an open space for theory-driven investigation of the causes and timing of punctuations, i.e., going further beyond just incrementalism/punctuations identification (Dempster & Wildavsky, 1979, p. 378; Sebők & Berki, 2017). This would directly address the question of mechanisms behind outliers through referring to one of the classic

Tab. 5: Budget major categories: a comprehensive coverage

Budget major code	Number of minor categories	Budget category
010	42	Agriculture and hunting
020	9	Forestry
050	15	Fishing and fisheries
100	12	Mining and quarrying
150	21	Industrial processing
400	10	Production and distribution of electrical energy, gas, and water
500	11	Trade
550	9	Hotel and restaurant services
600	31	Transport and communication
630	9	Tourism services
700	19	Dwelling economy (housing economy)
710	25	Services
720	7	Information technology
730	14	Science
750	48	Public administration
751	18	Offices of supreme bodies of the central government, control and protection of law and judiciary
752	24	National defense
753	17	Compulsory social security
754	26	Public safety and fire protection
755	16	Administration of justice
756	37	Revenue from legal persons, natural persons, and other units without legal personality, and expenses related to its collection
757	5	Public debt service
758	31	Other transfers and settlements
801	35	Education
803	13	Higher education
851	34	Health care
852	27	Welfare (social assistance)
853	23	Other social policy issues
854	23	Educational care (educational social services)
900	29	Communal economy and environmental protection
921	28	Culture and protection of national heritage
925	10	Botanical and zoological gardens, nature sites, and nature reserves
926	10	Physical education and sports

Note: As of 2014.
Source: Ordinance of the Minister of Finance, 2014, pp. 5–7.

Tab. 6: Major and minor codes: public debt service

Major code	Minor code	Budget function: public debt service
757	75701	Servicing of foreign debt, debtors, and other foreign transactions
	75702	Servicing of securities, credits, and loans of local government units
	75703	Servicing of national Treasury securities and other financial instruments on the domestic market
	75704	Settlements under sureties and guarantees provided by the State Treasury or local government units
	75705	Servicing of domestic entities' credits of other public finance sector units

Source: (Ordinance of the Minister of Finance, 2014, pp. 21–22)

definitions of outliers: "an observation which deviates so much from other observations as to arouse suspicions that it was generated by a different mechanism" (Hawkins, 1980, p. 1). Indeed, such a research would inform us about forces shaping policymaking processes—which is one of the key areas in public policy scholarship.

In line with the above alleys for future research, the current piece shall be considered a modest but necessary starting point. Through inference based on statistical assumptions, it is possible to consider cross-function and over-time changes in budget outlays with at least some level of objectivity, i.e., indifference to a researcher's a priori argument but instead toward "letting the data speak for themselves" (Gould, 1981).

REFERENCES

Agyemang, M., Barker, K., & Alhajj, R. (2006). A comprehensive survey of numeric and symbolic outlier mining techniques. *Intelligent Data Analysis*, *10*(6), 521–538. https://doi.org/10.3233/IDA-2006-10604

Amer, M., & Goldstein, M. (2012). Nearest-Neighbor and Clustering based Anomaly Detection Algorithms for RapidMiner. In *3rd RapidMiner Community Meeting and Conferernce*. https://doi.org/10.5455/ijavms.141

Bailey, J. J., & O'Connor, R. J. (1975). Operationalizing Incrementalism: Measuring the Muddles. *Public Administration Review*, *35*(1), 60–66. https://doi.org/10.2307/975202

Barnett, V., & Lewis, T. (1994). *Outliers in Statistical Data* (3rd Ed.). Wiley.

Bartolucci, A. A. (2016). Methodologies in Outlier Analysis. In A. A. Bartolucci, K. P. Singh, & S. Bae (Eds.), *Introduction to Statistical Analysis of Laboratory Data* (pp. 79–111). John Wiley & Sons.

Baumgartner, F. R., & Epp, D. A. (2013). *Explaining Punctuations* (Paper presented at the annual meetings of the Comparative Agendas Project, Antwerp, Belgium, June 27–29, 2013). Retrieved from https://unc.live/3afbnxB

Baumgartner, F. R., Green-Pedersen, C., & Jones, B. D. (2006). Comparative studies of policy agendas. *Journal of European Public Policy*, *13*(7), 959–974.

Baumgartner, F. R., & Jones, B. D. (Eds.). (2002). *Policy dynamics*. Chicago: University of Chicago Press.

Bozeman, B. (1977). The Effect of Economic and Partisan Change On Federal Appropriations. *Western Political Quarterly*, *30*(1), 112–124. https://doi.org/10.1177/106591297703000111

Breunig, C. (2006). The more things change, the more things stay the same: a comparative analysis of budget punctuations. *Journal of European Public Policy*, *13*(7), 1069–1085. https://doi.org/10.1080/13501760600924167

Breunig, C., & Jones, B. D. (2011). Stochastic Process Methods with an Application to Budgetary Data. *Political Analysis*, *19*(1), 103–17.

Breunig, C., & Koski, C. (2006). Punctuated Equilibria and Budgets in the American States. *Policy Studies Journal*, *34*(3), 363–379. https://doi.org/10.1111/j.1541-0072.2006.00177.x

Breunig, M., Kriegel, H.-P., Ng, R., & Sander, J. (2000). LOF: Identifying Density-Based Local Outliers. In *ACM Sigmod Record* (Vol. 29, pp. 93–104). Dallas, TX. https://doi.org/10.1145/342009.335388

Bunce, V., & Echols, J. M. (1978). Power and Policy in Communist Systems: The Problem of 'Incrementalism'. *The Journal of Politics*, *40*(4), 911–932. https://doi.org/10.2307/2129902

Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly Detection: A Survey. *ACM Computing Surveys*, *41*(3), 1–58. https://doi.org/10.1145/1541880.1541882

Cox, J., Hager, G., & Lowery, D. (1993). Regime Change in Presidential and Congressional Budgeting: Role Discontinuity or Role Evolution? *American Journal of Political Science*, *37*(1), 88–118. https://doi.org/10.2307/2111525

Davis, O. A., Dempster, M. A. H., & Wildavsky, A. (1974). Towards A Predictive Theory of Government Expenditure: US Domestic

Appropriations. *British Journal of Political Science*, 4(4), 419–452. https://doi.org/DOI: 10.1017/S0007123400009650

de Crombrugghe, A., & Lipton, D. (1994). The Government Budget and the Economic Transformation of Poland. In O. Blanchard, K. Froot, & J. Sachs (Eds.), *Transition in Eastern Europe* (Vol. 2, pp. 111–136). Chicago, Illinois: University of Chicago Press.

Dempster, M., & Wildavsky, A. (1979). On Change: Or, There is No Magic Size for An Increment. *Political Studies*, 27(3), 371–389. https://doi.org/10.1111/j.1467-9248.1979.tb01210.x

Denning, D. (1986). An Intrusion-Detection Model. In *IEEE Transactions on Software Engineering* (Vol. 13, pp. 118–133). IEEE. https://doi.org/10.1109/SP.1986.10010

Dezhbakhsh, H., Tohamy, S. M., & Aranson, P. H. (2003). A new approach for testing budgetary incrementalism. *Journal of Politics*, 65(2), 532–558. https://doi.org/10.1111/1468-2508.t01-3-00014

Endler, D. (1998). Intrusion detection. Applying machine learning to Solaris audit data. In *Proceedings 14th Annual Computer Security Applications Conference* (pp. 268–279). Phoenix, AZ. https://doi.org/10.1109/CSAC.1998.738647

Fenno, R. F. (1966). *The Power of the Purse*. Boston: Little Brown.

Flink, C. M. (2017). Rethinking Punctuated Equilibrium Theory: A Public Administration Approach to Budgetary Changes. *Policy Studies Journal*, 45(1), 101–120. https://doi.org/10.1111/psj.12114

Flink, C. M., & Robinson, S. E. (2020). Corrective policy reactions: Positive and negative budgetary punctuations. *Journal of Public Policy*, 40(1), 96–115. https://doi.org/10.1017/S0143814X18000259

Gist, J. R. (1974). Mandatory Expenditures and the Defense Sector: Theory of Budgetary Incrementalism. In R. B. Ripley (Ed.), *A Sage Professional Paper* (pp. 5–39). London: Sage.

Goldstein, M., & Uchida, S. (2016). A comparative evaluation of unsupervised anomaly detection algorithms for multivariate data. *PLoS ONE*, 11(4), 1–31. https://doi.org/10.1371/journal.pone.0152173

Gould, P. (1981). Letting the Data Speak for Themselves. *Annals of the Association of American Geographers*, 71(2), 166–176.

Grubbs, F. E. (1969). Procedures for Detecting Outlying Observations in Samples. *Technometrics*, 11(1), 1–21. https://doi.org/10.1080/00401706.1969.10490657

Hampel, F. R. (1971). A General Qualitative Definition of Robustness. *The Annals of Mathematical Statistics*, 42(6), 1887–1896.

Hawkins, D. M. (1980). *Identification of outliers*. London: Chapman and Hall. https://doi.org/10.1007/978-94-015-3994-4

Helman, P., & Bhangoo, J. (1997). A Statistically Based System for Prioritizing Information Exploration under Uncertainty. In *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics* (Vol. 27, pp. 449–466). IEEE Press. https://doi.org/10.1109/3468.594912

Hodge, V., & Austin, J. (2004). A Survey of Outlier Detection Methodologies. *Artificial Intelligence Review*, 22(2), 85–126. https://doi.org/10.1023/B:AIRE.0000045502.10941.a9

Jones, B. D., Baumgartner, F. R., Breunig, C., Wlezien, C., Soroka, S., Foucault, M., ... John, P. (2009). A General Empirical Law of Public Budgets: A Comparative Analysis. *American Journal of Political Science*, 53(4), 855–73.

Jones, B. D., Baumgartner, F. R., & True, J. L. (1998). Policy Punctuations: US Budget Authority, 1947-1995. *Journal of Poliitics*, 60(1), 1–33.

Jones, B. D., Sulkin, T., & Larsen, H. A. (2003). Policy Punctuations in American Political Institutions. *American Political Science Review*, 97(1), 151–169.

Jones, B. D., True, J. L., & Baumgartner, F. R. (1997). Does Incrementalism Stem from Political Consensus or from Institutional Gridlock? *American Journal of Political Science*, 41(4), 1319–1339.

Jordan, A. A., Taylor, W. J., Meese, M. J., Nielsen, S. C., & Schlesinger, J. (2009). *American National Security* (6th ed.). Baltimore: Johns Hopkins University Press.

Jordan, M. M. (2003). Punctuations and Agendas: A New Look at Local Government Budget Expenditures. *Journal of Policy Analysis and Management*, 22(3), 345–360. https://doi.org/10.1002/pam.10136

Kamlet, M. S., & Mowery, D. C. (1987). Influences on Executive and Congressional Budgetary Priorities, 1955–1981. *American Political Science Review*, 81(1), 155–178. https://doi.org/DOI: 10.2307/1960783

Kanter, A. (1972). Congress and the Defense Budget: 1960–1970. *American Political Science Review*, 66(1), 129–143.

Kemp, K. (1982). Instability in budgeting for federal regulatory agencies. *Social Science Quarterly*, 63(4), 643–660.

Khan, S. S., & Madden, M. G. (2014). One-class classification: Taxonomy of study and review of techniques. *The Knowledge Engineering Review*, 29(3), 345–374. https://doi.org/10.1017/S026988891300043X

Knorr, E. M., & Ng, R. T. (1998). Algorithms for Mining Distance-Based Outliers in Large Datasets. In *Proceedings of the 24rd International Conference on Very Large Data Bases* (pp. 392–403). San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Knorr, E. M., Ng, R. T., & Zamar, R. (2001). Robust Space Transformations for Distance-based Operations. In *Proceedings of the 7th International Conference on Knowledge Discovery and Data Mining (KDD01)* (pp. 126–135). https://doi.org/10.1145/502512.502532

Laptev, N., Amizadeh, S., & Flint, I. (2015). Generic and Scalable Framework for Automated Time-Series Anomaly Detection. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1939–1947). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/2783258.2788611

Laurikkala, J., Juhola, M., & Kentala, E. (2000). Informal identification of outliers in medical data. In N. Lavrač, S. Miksch, & B. Kavšek (Eds.), *Fifth International Workshop on Intelligent Data Analysis in Medicine and Pharmacology IDAMAP-2000 Berlin, 22 August. Organized as a workshop of the 14th European Conference on Artificial Intelligence ECAI-2000.* https://doi.org/10.1142/s0217979205027834

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, *49*(4), 764–766. https://doi.org/10.1016/j.jesp.2013.03.013

Lindblom, C. E. (1959). The Science of 'Muddling Through'. *Public Administration Review*, *19*(2), 79–88. https://doi.org/10.2307/973677

Merriam-Webster Online Dictionary. (2019). Retrieved 14 November 2019, from https://www.merriam-webster.com/dictionary

Miller, J. (1991). Short Report: Reaction Time Analysis with Outlier Exclusion: Bias Varies with Sample Size. *The Quarterly Journal of Experimental Psychology Section A*, *43*(4), 907–912. https://doi.org/10.1080/14640749108400962

Mueller, D. C. (2003). *Public Choice III* (3rd ed.). Cambridge: Cambridge University Press. https://doi.org/DOI: 10.1017/CBO9780511813771

Munir, M., Siddiqui, S. A., Dengel, A., & Ahmed, S. (2019). DeepAnT: A Deep Learning Approach for Unsupervised Anomaly Detection in Time Series. *IEEE Access*, *7*, 1991–2005. https://doi.org/10.1109/ACCESS.2018.2886457

Noguchi, Y. (1980). A Dynamic Model of Incremental Budgeting. *Hitotsubashi Journal of Economics*, *20*, 11–25.

Ordinance of the Minister of Finance (2014).

Padgett, J. F. (1980). Bounded Rationality in Budgetary Research. *American Political Science Review*, *74*(2), 354–372. https://doi.org/DOI: 10.2307/1960632

Papadimitriou, S., Kitagawa, H., Gibbons, P. B., & Faloutsos, C. (2002). *LOCI: Fast Outlier Detection Using the Local Correlation Integral* (No. CMU-CS-02-188). Pittsburgh, PA. https://doi.org/10.1184/R1/6607028.v1

Public Finance Act (2009).

Ramaswamy, S., Rastogi, R., & Shim, K. (2000). Efficient Algorithms for Mining Outliers from Large Data Sets. In *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data* (pp. 427–438). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/342009.335437

Robinson, S. E., Caver, F., Meier, K. J., & O'Toole, L. J. (2007). Explaining policy punctuations: Bureaucratization and budget change. *American Journal of Political Science*, *51*(1), 140–150. https://doi.org/10.1111/j.1540-5907.2007.00242.x

Robinson, S. E., Flink, C. M., & King, C. M. (2014). Organizational History and Budgetary Punctuation. *Journal of Public Administration Research and Theory*, *24*(2), 459–471. https://doi.org/10.1093/jopart/mut035

Rousseeuw, P. J., & Hubert, M. (2018). Anomaly detection by robust statistics. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *8*(2), 1–14. https://doi.org/10.1002/widm.1236

Schubert, E., Wojdanowski, R., Zimek, A., & Kriegel, H. P. (2012). On evaluation of outlier rankings and outlier scores. In *Proceedings of the 12th SIAM International Conference on Data Mining, SDM 2012* (pp. 1047–1058). https://doi.org/10.1137/1.9781611972825.90

Sebők, M., & Berki, T. (2017). Incrementalism and punctuated equilibrium in Hungarian budgeting (1991-2013). *Journal of Public Budgeting, Accounting and Financial Management*, *29*(2), 151–181. https://doi.org/10.1108/jpbafm-29-02-2017-b001

Shewhart, W. A. (1923). *Economic Control of Quality of Manufactured Product*. London: D.van Nostrand Company.

Simon, H. A. (1955). A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics*, *69*(1), 99–118. https://doi.org/10.2307/1884852

Sims, C. A. (1980). Macroeconomics and Reality. *Econometrica*, *48*(1), 1–48.

True, J. L., Jones, B. D., & Baumgartner, F. R. (1999). Punctuated-equilibrium theory: explaining stability and change in American policymaking. In P. A. Sabatier (Ed.), *Theories of the policy process*. Boulder, CO: Westview Press.

True, J. L., Jones, B. D., & Baumgartner, F. R. (2007). Punctuated-Equilibrium Theory: Explaining Stability and Change in Public Policymaking. In P. A. Sabatier (Ed.), *Theories of the Policy Process* (pp. 155–88). Boulder, CO: Westview Press.

Tukey, J. (1977). *Exploratory Data Analysis*. Reading, MA: Addison-Wesley.

Wanat, J. (1974). Bases of Budgetary Incrementalism. *American Political Science Review*, *68*(3), 1221–1228.

Wildavsky, A. (1964). *The politics of the budgetary process*. Boston: Little, Brown.

Xu, X., Liu, H., & Yao, M. (2019). Recent Progress of Anomaly Detection. *Complexity*, 1–11. https://doi.org/10.1155/2019/2686378

Yamanishi, K., Takeuchi, J., Williams, G., & Milne, P. (2004). On-Line Unsupervised Outlier Detection Using Finite Mixtures with Discounting Learning Algorithms. *Data Mining and Knowledge Discovery*, *8*(3), 275–300. https://doi.org//10.1023/B:DAMI.0000023676.72185.7c

Zimek, A., & Filzmoser, P. (2018). There and back again: Outlier detection between statistical reasoning and data mining algorithms. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *8*(6). https://doi.org/10.1002/widm.1280