To: Frank Baumgartner and Bryan Jones
Re: Are All Budgets Punctuated?
Date: June 6, 2006

This memo provides
(1) an overview of the newly updated data sets
(2) descriptive statistics for each data set, and
(3) graphical analysis of punctuations

(1) Overview

- The collected data encompasses 12 individual budget outputs from seven countries
- The accompanying data file (AllBudgetsFinal.dat) contains the entire data.
- The data in organized in the following order:

Table 1. Data sets and their source

Data set number	Data Source
1	US Central Government Budget Authority
2	US State Government Outlays
3	Texas School District Expenditures
4	Danish National Government Expenditures
5	Danish Local Government Expenditures
6	French Government Overall Spending Levels
7	French Government Programmatic Spending
8	Canadian Government Expenditures
9	Belgian Government Expenditures
10	US Historical Outlays
11	German National Government Expenditures (Sub-Functions)
12	UK Government Spending

- The analyzed data are the annual changes for each budget category (*diffpct*). The change measure is based on annual percent changes, i.e. $(count_t - count_{t-1})/ count_{t-1}$. I followed Jones et al (2003) and cut off all budget changes that are greater than tenfold.

(2) Descriptive Statistics

- As Table 2 indicates, the amount of observations ranges from 156 (FR overall spending) to 15174 (Danish local level spending).
- All the data sets have a central tendency (particular the median) that is positive and "close" to zero. The central tendency indicates the "on average" growth of budgets over time.
- The measures of dispersion are usually small too. The exceptions are US BA (1), FR programmatic spending (7), and BEL expenditures (9).
- The same data sets (1, 7, and 9) also contain very extreme maxima (and many other high values [see Table 3]), when compared to the other data sets.

Table 2. Basic Descriptive Statistics

Data Set	mean	median	sd	IQR	min	max	unique	n	miss
1	0.113	0.032	0.605	0.164	-1.000	9.721	3077	3109	301
2	0.056	0.043	0.239	0.114	-1.000	9.761	8586	8998	501

3	0.025	0.021	0.120	0.089	-0.793	3.828	12469	12502	1051
4	0.029	0.018	0.183	0.086	-0.707	2.762	340	832	26
5	0.021	0.007	0.144	0.078	-0.886	3.996	862	15174	1086
6	0.038	0.021	0.160	0.084	-0.495	0.986	181	180	3
7	0.099	0.022	0.522	0.156	-0.806	6.630	1051	1050	299
8	0.000	0.002	0.163	0.099	-0.871	0.663	156	156	0
9	0.113	0.015	0.850	0.145	-1.000	7.926	229	229	38
10	0.109	0.026	0.555	0.156	-0.815	7.477	545	560	3
11	0.066	0.008	0.380	0.145	-0.748	6.324	983	982	32
12	0.017	0.015	0.109	0.084	-0.399	0.480	175	266	0

Table 3. Basic Descriptive Statistics –Five biggest changes

Data Set	Five Biggest Changes				
1	6.825	6.967	8.135	9.290	9.721
2	3.702	4.237	4.295	5.854	9.761
3	2.272	2.303	2.784	2.786	3.828
4	1.145	1.170	1.201	1.970	2.762
5	3.203	3.241	3.774	3.781	3.996
6	0.406	0.704	0.733	0.774	0.986
7	3.792	4.778	5.807	5.970	6.630
8	0.446	0.470	0.497	0.577	0.663
9	2.129	3.393	4.233	7.318	7.926
10	2.868	3.161	4.524	5.381	7.477
11	2.239	2.749	3.068	3.364	6.324
12	0.409	0.418	0.429	0.466	0.480

- Table 4 provides a statistical assessment of the degree of punctuation.
- The K-S and S-W tests reject that the 12 data sets are Normally distributed.
- A Normal distribution has a kurtosis measure of 0 and an L-kurtosis score of ca. 0.123.
- The kurtosis statistics indicate that budgets are leptokurtic, but variation among the sets exists. The US BA and the FR programmatic spending are the most leptokurtic, budgets in the UK, GER (major categories), and CAN are the least leptokurtic.
- The difference between the L-kurtosis and kurtosis score might be due to the impact of a few positive outliers

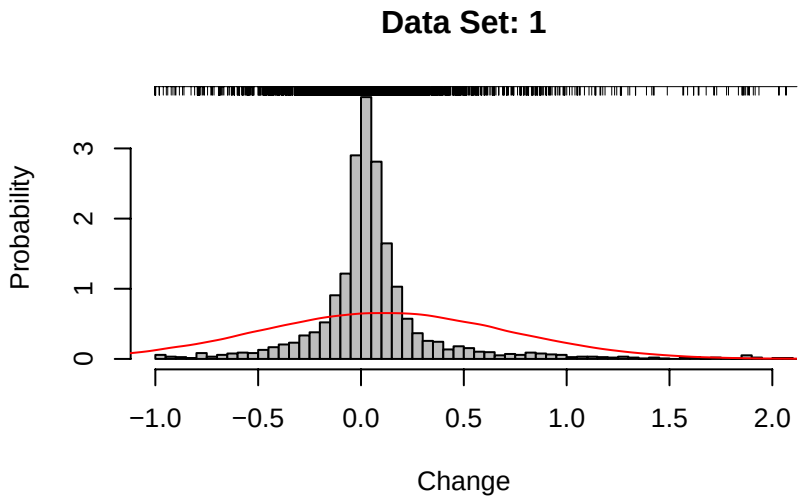
Table 4. Statistical Assessment of Kurtosis

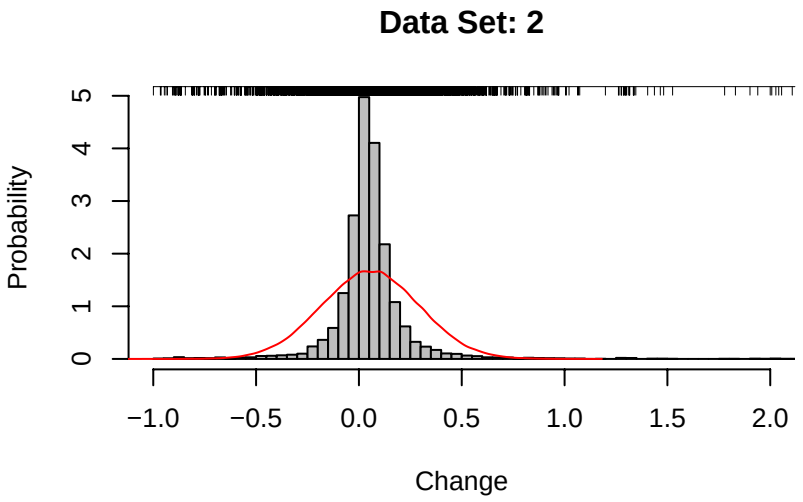
Data Set	kurtosis	L-kurtosis	KS-stat	KS-p	SW-stat	SW-p
1	77.488	0.512	0.322	0.000	0.455	0.000
2	380.030	0.403	0.388	0.000	0.488	0.000
3	175.191	0.293	0.411	0.000	0.555	0.000
4	80.618	0.421	0.389	0.000	0.563	0.000
5	194.946	0.363	0.408	0.000	0.606	0.000
6	11.589	0.424	0.396	0.000	0.741	0.000
7	67.170	0.505	0.346	0.000	0.432	0.000
8	7.749	0.379	0.372	0.000	0.833	0.000
9	54.346	0.611	0.322	0.000	0.359	0.000
10	78.452	0.509	0.346	0.000	0.424	0.000

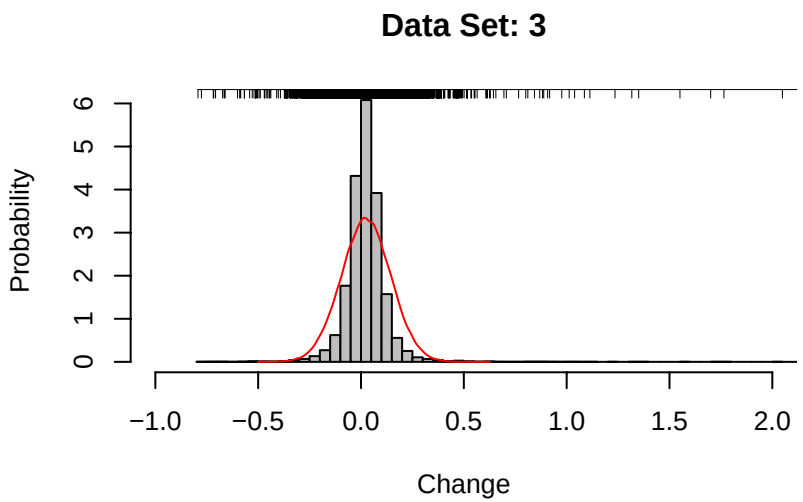
11	89.747	0.456	0.362	0.000	0.510	0.000
12	4.422	0.319	0.404	0.000	0.893	0.000

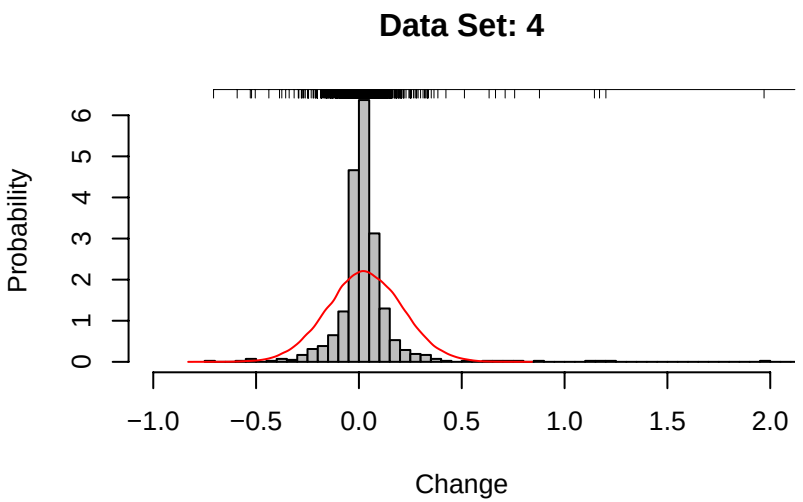
(3) Graphical Analysis

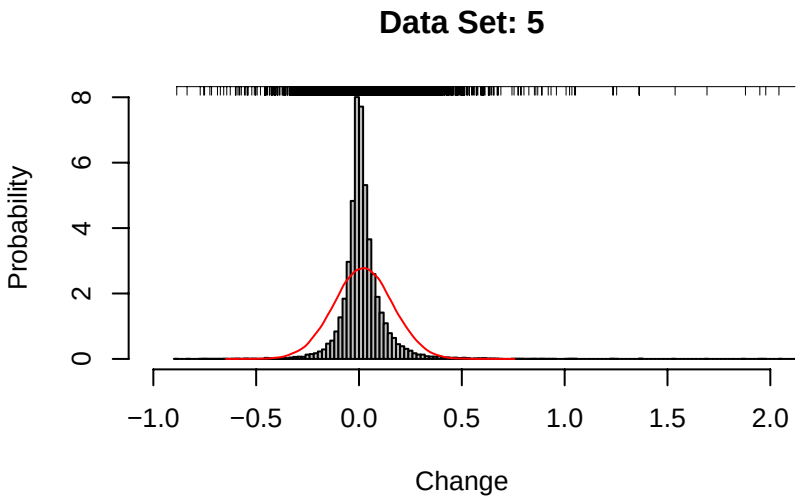
- Figure 1 displays the histograms of the 12 sets. Bin size is determined by Sheather-Jones and depends on distribution of the data as well as observation numbers. The bins are not based on cumulative frequencies.
- The rug on top of each histogram shows that French overall spending (6), Canadian (8), and UK spending (12) does not produce more than two-fold changes.
- None of the budgets looks Normally distributed but the statistical variation in leptokurtosis is graphically confirmed.
- The flat Normal distribution (red line) for the US BA, FR programmatic data, and BEL budget are due to the high dispersion of the data.

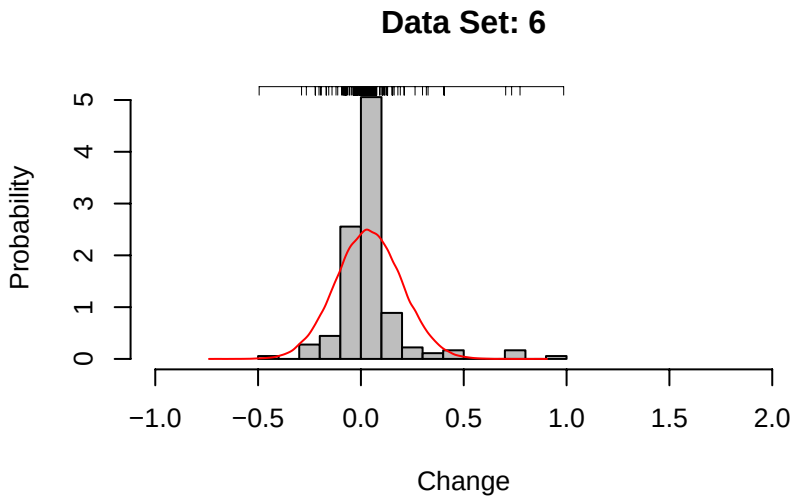

Data Set: 1
Probability
0
1
2
3
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
Change

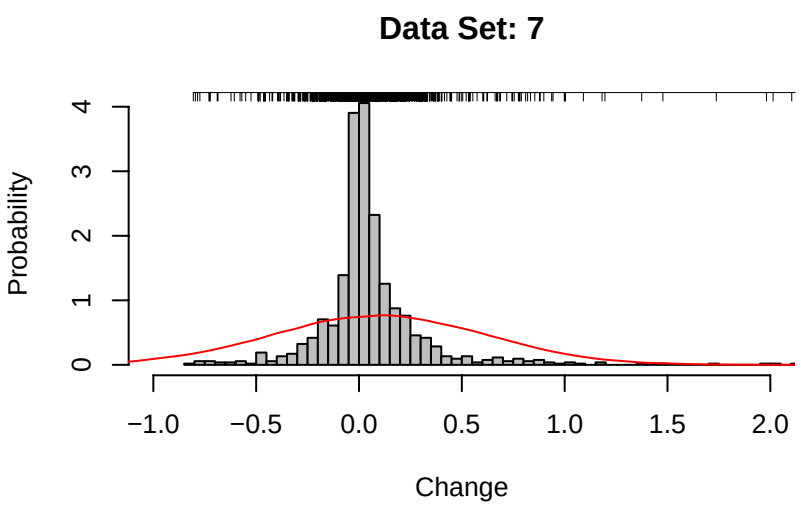

Data Set: 2
Probability
0
1
2
3
4
5
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
Change

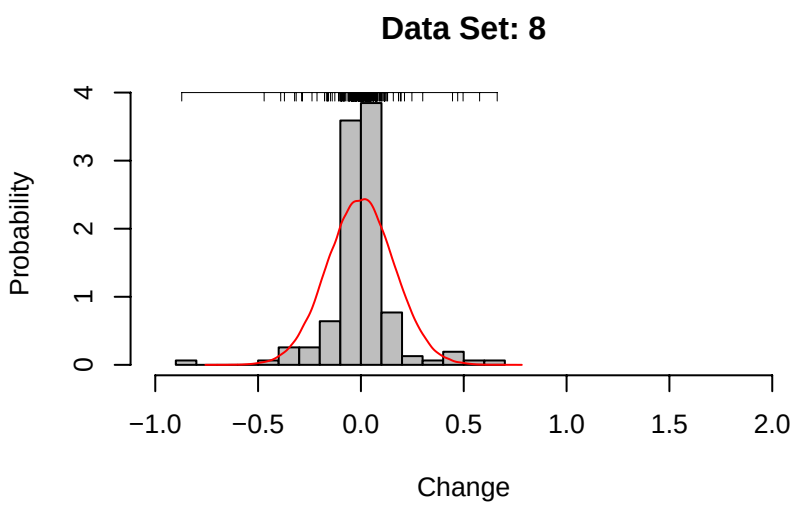

Data Set: 3
Probability
0
1
2
3
4
5
6
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
Change

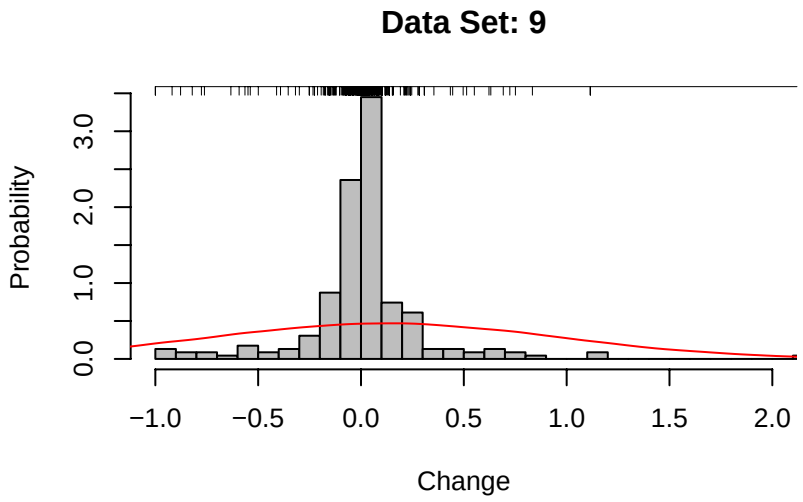

Data Set: 4
Probability
0
1
2
3
4
5
6
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
Change

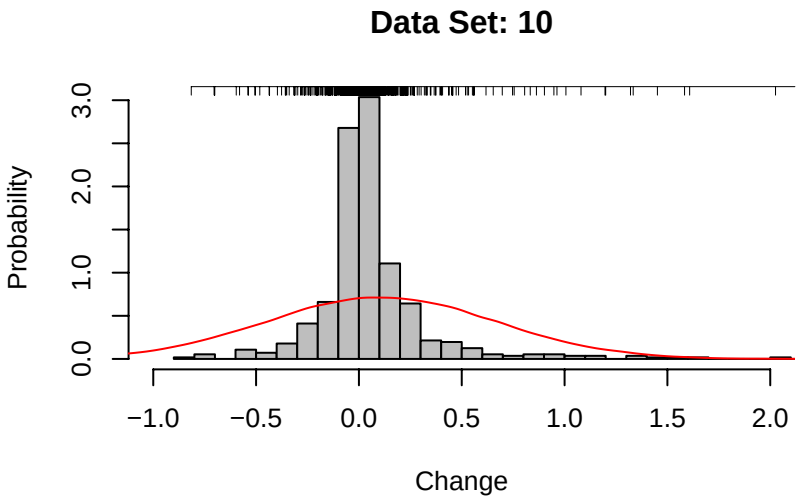

Data Set: 5
Probability
0
2
4
6
8
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
Change

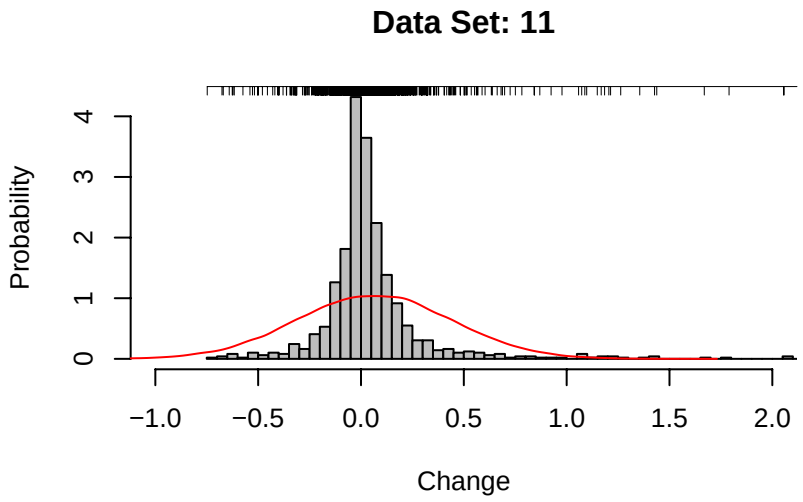

Data Set: 6
Probability
0
1
2
3
4
5
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
Change

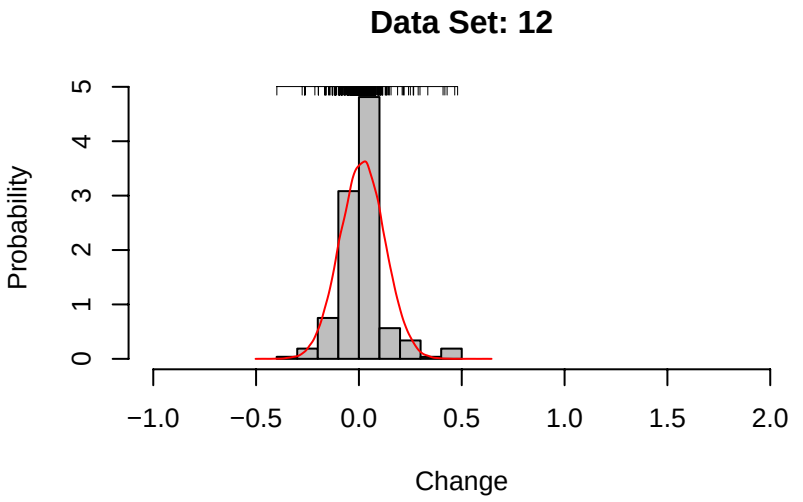

Data Set: 7
Probability
0
1
2
3
4
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
Change


Data Set: 8
Probability
0
1
2
3
4
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
Change


Data Set: 9
Probability
0.0
1.0
2.0
3.0
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
Change


Data Set: 10
Probability
0.0
1.0
2.0
3.0
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
Change


Data Set: 11
Probability
0
1
2
3
4
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
Change


Data Set: 12
Probability
0
1
2
3
4
5
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
Change