Budget Authority Data Codebook

Policy Agendas Project
Updated 2018

An Overview of the Data Set

The Policy Agendas Project Budget Authority data set is based on the *Budget of the United States Government* and was originally compiled by James True. It measures budget authority and provides an inflation adjusted overview of the budgetary decision-making process. (Additional adjustments are described below). The data is organized by functions and subfunctions of the budget, representing the long-term purposes these budgets were meant to serve. The data set was designed for analysts interested in comprehensive budget data that is reliable and consistent over a long period, hence our measures begin in FY 1947 (July 1, 1946 – June 30, 1947) and continue through the most recent fiscal year.

This data codebook lists a series of notes important to researchers using these data, provides information about each of the variables in the dataset, and summarizes the general procedures we use to update the dataset. For a full description of functions and subfunctions, detailed source information, and a glossary of related terminology, please see the comprehensive budget codebook compiled by James True (available on request, please email poliyagendas@gmail.com). We also provide a 'crosswalk' of budget functional categories with Policy Agendas Project topic codes.

Research Notes

1. Note that on July 30, 2013, we converted the data set to utilize a different OMB deflator (2005) and inflation-adjusted variables have been replaced for the entire dataset. The legacy dataset is available for replication purposes on request.

2. Analysts who use the data in the historical tables should take care with what we call *financial subfunctions* (as listed in the comprehensive codebook on Page 1 and described in the detail starting on Page 51) as well as with the various treatments of trust fund figures over the years.

3. All fiscal year periods contain data for twelve months. Figures have been eliminated for the three-month transition quarter before fiscal year 1977.

4. Analysts may also be interested in the *Historical Tables* of each FY *Budget,* which provide subfunctional outlay data beginning with FY 1962 and subfunctional budget authority data beginning with FY 1976.

5. Recent historical tables do not include offsetting receipts, and we make adjustments to ensure reliability with prior years that do.

6. Current year historical tables may include revised values for the previous years. We check these and make any updates as necessary.

7. Previous calculations of change scores resulted in large negative change scores – those that exceed -100%. Greater than -100% change scores can come from two sources: 1) categories that go from positive to negative and 2) categories that are negative and that more than double in

their negative values. Negative BA can mean many things – a rescission, a change in mandatory spending (CHIMP), or a "profit" from something. Researchers analyzing these data ought to consider the negative BA categories – 657 cases out of 8165 – as potentially distinct from other categories (see item 2 above). Most categories that experience negative BA are financial subfunctions (e.g. within 900; other notable financial subfunctions are 371 (Mortgage Credit), and occasionally 155 (International Financial Programs). Other categories such as 453 (Disaster Relief and Insurance), and 274 (Emergency Energy Preparedness) yield negative BA given that these are subfunctions that collect and payout money depending on, well, disasters. For specific subfunction analysis of these categories, researchers may wish to consider increases in the magnitude of negative numbers as positive values, even though mathematically these changes are large negative changes. Or at least consider them separately

For our purposes, we truncate all negative changes beyond -100% at -100%. There are 113 cases for which truncation are necessary. Researchers should carefully consider what kinds of questions they wish to answer using the data before using the change term we have constructed here.

Variable Names and Descriptions

ID
This column records the unique identifier for each observation. It has no substantive application.

Year
This column records the corresponding *fiscal year*.

Deflator
This column records the value of the deflator used to adjust for inflation over time for each year. We use the 2009 deflator values provided by OMB.

Dllr
This column records the *adjusted* dollar value (in millions) of each code.

chngdllr
This column records the simple year-over-year difference in dollars across codes.

pctChng
This column records the year-over-year percentage change in dollars across codes. Note that blank observations refer to year-over-year undefined changes (i.e. a reduction to zero), while zeros refer to year-over-year change of zero.

Amount
This column records the *un-adjusted* dollar value (in millions) of each code.

AlphaCode
This column records the alpha-numeric OMB code for functions and sub-functions.

MajorFunction

This column records the number OMB code for functions.

SubFunction

This column records the number OMB code for sub-functions, with the addition of a decimal that identifies particular categories of sub-functions. This is left blank for authority that only corresponds to a major budget function.

FunctionName

This column records the OMB name for the function or sub-function.

IsItMajorFunction

This column records if the authority in question corresponds to a major budget function (coded 1) or not (coded 0).

IsItSubFunction

This column records if the authority in question corresponds to a budget subfunction (coded 1) or not (coded 0).

IsItSubSubFunction

This column records if the authority in question corresponds to a category within a budget subfunction (coded 1) or not (coded 0).

Programmatic

This column records if the authority in question is programmatic (coded 1) or not (coded 0).

Discretionary

This column records if the authority in question is discretionary (coded 1) or not (coded 0).

Mandatory

This column records if the authority in question is mandatory (coded 1) or not (coded 0).

Congress

This column records the corresponding session of Congress.

Data Sources

Data are derived from four primary sources, all created by OMB

Office of Management and Budget Historical Tables:
Table 5.1—Budget Authority by Function and Subfunction: 1976–20XX
Table 10.1—Gross Domestic Product and Deflators Used in the Historical Tables: 1940–20XX
Table 13.1—Cash Income, Outgo, and Balances of the Social Security and Medicare Trust Funds: 1936– 20XX

Office of Management and Budget Analytical Perspectives
Table 26-1. Budget Authority and Outlays by Function, Category, and Program (the number of this table changes from year to year; the title largely stays the same)

Updating Procedures

1. Table 26.1 is checked against previous years Table 26.1 to see if spending categories have shifted from one subfunction to another.

2. Start with OMB Historical Table 5.1 is then rearranged such that rows correspond with order of PAP decimal coding system (not alphanumeric) of functions and subfunctions. Koski developd a SORTING HAT key in excel.

3. Check simple differences for any errors and OMB revisions in current year in previous year values.

4. Copy values into PAP Budget file

5. Make adjustments as needed. These include:
-Medicare & Social Security to account for how the programs have evolved over time (i.e. premiums). We calculate these separately from the historical budget table and enter them into the dataset.
-Fish and Wildlife to account for federal government change of location (i.e. we move authority from the Conservation and Land Management (302) to Recreational Resources (303, the historical category). Look at *Analytical Perspectives ("Budget Authority and Outlay")* - Afterward, change Total BA after making above adjustments.

6. Check OMB current deflator is the presently utilized deflator. If not, update for all observation-years and recalculate inflation-adjusted variables (i.e. Dllr, chngdll, pctChng).

7. Create change formulae that zeros out #DIV/0 issues and truncates values <-100.