CODEBOOK: DECOMPOSING THE SOURCE OF THE GENDER GAP IN LEGISLATIVE COMMITTEE SERVICE

September 24, 2021

Julia Payson – New York University
Andrew B. Hall – Stanford University
Alexander Fouirnaies – Harris School at University of Chicago

This codebook describes the dataset `women.dta`. The dataset contains legislator-, district, and legislature-level information.

Each row in the dataset is uniquely identified by a state legislator in a given term. The dataset contains 41 variables. Each variable is briefly described below.

`state`	String variable. String variable that reports the state. The variable has 89,848 non-missing observations and 50 distinct values.
`chamber`	String variable. String variable that reports the legislator serves in the upper chamber (==senate) or lower chamber (==house) in the state legislature in a particular term. The variable has 89,848 non-missing observations and 2 distinct values.
`year`	Numeric variable. A discrete variable that reports the year in which a legislator was elected to office. The variable has 89,848 non-missing observations and 29 distinct values.
`StateChamberYear`	Numeric variable. Discrete variable where each value uniquely identifies a legislative chamber in a given state in a particular year. The variable has 89,848 non-missing observations and 1,338 distinct values.
`DistId`	Numeric variable. Discrete variable where each value uniquely identifies a state legislative district. The variable has 89,848 non-missing observations and 26,299 distinct values.
`CandId`	Numeric variable. Discrete variable that uniquely identifies an individual state legislator. This is also the id in the Klarner dataset. The variable has 89,848 non-missing observations and 25,896 distinct values.
`CandName`	String variable. String variable reporting the name of the legislator. The variable has 89,848 non-missing observations and 25,669 distinct values.

`FirstName`	String variable. String variable reporting the first name of the legislator. The variable has 89,848 non-missing observations and 2,149 distinct values.
`female`	Numeric variable. A dummy variable indicating whether the legislator is a women (==1) or not (==0). The variable has 89,642 non-missing observations and 2 distinct values.
`cmt_women`	Numeric variable. Dummy variable that indicates if a legislator serves on the Education Committee or Health Committee in their legislature. The variable has 89,848 non-missing observations and 2 distinct values.
`cmt_men`	Numeric variable. Dummy variable that indicates if a legislator serves on the Commerce Committee or Finance Committee in their legislature. The variable has 89,848 non-missing observations and 2 distinct values.
`log_women_bills`	Numeric variable. Continuous variable that reports the log of health and education bills sponsored by a legislator in a particular term. The variable has 14,899 non-missing observations and 129 distinct values.
`log_sponsored`	Numeric variable. Continous variable that measures the log of bills+1 sponsored by a legislator. The variable has 15,592 non-missing observations and 247 distinct values.
`log_first_women`	Numeric variable. Continuous variable that reports the log of campaign contributions from health and education sectors in the legislator's first electoral campaign. The variable has 37,598 non-missing observations and 6,684 distinct values.
`party`	String variable. String variable reporting the party of the legislator. The variable has 89,848 non-missing observations and 3 distinct values.

`NumberOfSeats`	Numeric variable. Discrete variable that reports the number of legislators elected in the state legislative district. The variable has 89,848 non-missing observations and 14 distinct values.
`cmt_women`	Numeric variable. Dummy variable that indicates if a legislator serves on the Education Committee or Health Committee in their legislature. The variable has 89,848 non-missing observations and 2 distinct values.
`cmt_men`	Numeric variable. Dummy variable that indicates if a legislator serves on the Commerce Committee or Finance Committee in their legislature. The variable has 89,848 non-missing observations and 2 distinct values.
`cmt_energy`	Numeric variable. Dummy variable indicating that the legislator served on the Energy Committee. The variable has 89,848 non-missing observations and 2 distinct values.
`cmt_trans`	Numeric variable. Dummy variable indicating that the legislator served on the Transport Committee. The variable has 89,848 non-missing observations and 2 distinct values.
`cmt_health`	Numeric variable. Dummy variable indicating that the legislator served on the Health Committee. The variable has 89,848 non-missing observations and 2 distinct values.
`cmt_fin`	Numeric variable. Dummy variable indicating that the legislator served on the Finance Committee. The variable has 89,848 non-missing observations and 2 distinct values.
`cmt_ag`	Numeric variable. Dummy variable indicating that the legislator served on the Agriculture Committee. The variable has 89,848 non-missing observations and 2 distinct values.

cmt_educ	Numeric variable. Dummy variable indicating that the legislator served on the Education Committee. The variable has 89,848 non-missing observations and 2 distinct values.
cmt_commerce	Numeric variable. Dummy variable indicating that the legislator served on the Commerce Committee. The variable has 89,848 non-missing observations and 2 distinct values.
cmt_labor	Numeric variable. Dummy variable indicating that the legislator served on the Labor Committee. The variable has 89,848 non-missing observations and 2 distinct values.
cmt_welfare	Numeric variable. Dummy variable indicating that the legislator served on the Welfare Committee The variable has 89,848 non-missing observations and 2 distinct values.
cmt_waysandmeans	Numeric variable. Dummy variable indicating that the legislator served on the Ways and Means Committee The variable has 89,848 non-missing observations and 2 distinct values.
cmt_social	Numeric variable. Dummy variable indicating that the legislator served on the Committee on Social Affairs The variable has 89,848 non-missing observations and 2 distinct values.
cmt_rules	Numeric variable. Dummy variable indicating that the legislator served on the Rules Committee The variable has 89,848 non-missing observations and 2 distinct values.
cmt_judiciary	Numeric variable. Dummy variable indicating that the legislator served on the Judiciary Committee The variable has 89,848 non-missing observations and 2 distinct values.
cmt_ethics	Numeric variable. Dummy variable indicating that the legislator served on the Ethics Committee The variable has 89,848 non-missing observations and 2 distinct values.

`cmt_approp`	Numeric variable. Dummy variable indicating that the legislator served on the Appropriations Committee The variable has 89,848 non-missing observations and 2 distinct values.
`first_total`	Numeric variable. Continous variable measuring the total campaign contributions donated to legislators in their first election. The variable has 37,598 non-missing observations and 12,542 distinct values.
`first_educ`	Numeric variable. Continous variable measuring the total campaign contributions donated by the education sector to legislators in their first election. The variable has 37,598 non-missing observations and 5,449 distinct values.
`log_first_women`	Numeric variable. Continuous variable that reports the log of campaign contributions from health and education sectors in the legislator's first electoral campaign. The variable has 37,598 non-missing observations and 6,684 distinct values.
`log_first_energy`	Numeric variable. Continuous variable measuring the log of campaign contributions from the energy sector to legislators in their first electoral campaign. The variable has 37,598 non-missing observations and 2,564 distinct values.
`log_first_trans`	Numeric variable. Continuous variable measuring the log of campaign contributions from the transport sector to legislators in their first electoral campaign. The variable has 37,598 non-missing observations and 1,727 distinct values.
`log_first_health`	Numeric variable. Continuous variable measuring the log of campaign contributions from the health care sector to legislators in their first electoral campaign. The variable has 37,598 non-missing observations and 3,101 distinct values.

`log_first_fin` Numeric variable. Continuous variable measuring the log of campaign contributions from the financial sector to legislators in their first electoral campaign. The variable has 37,598 non-missing observations and 3,834 distinct values.

`log_first_ag` Numeric variable. Continuous variable measuring the log of campaign contributions from the agriculture sector to legislators in their first electoral campaign. The variable has 37,598 non-missing observations and 1,863 distinct values.

`log_first_educ` Numeric variable. Continuous variable measuring the log of campaign contributions from the education sector to legislators in their first electoral campaign. The variable has 37,598 non-missing observations and 5,445 distinct values.

`log_first_commerce`
Numeric variable. Continuous variable measuring the log of campaign contributions from the commerce related businesses to legislators in their first electoral campaign. The variable has 37,598 non-missing observations and 4,044 distinct values.

`log_first_labor` Numeric variable. Continuous variable measuring the log of campaign contributions from the labor unions to legislators in their first electoral campaign. The variable has 37,598 non-missing observations and 3,883 distinct values.