Union Membership by Congressional District, Based on the Current Population Study

Compiled by

Janet Box-Steffensmeier,
Department of Political Science
The Ohio State University
Columbus, OH 43210-373
jboxstef+@osu.edu

Laura W. Arnold
Department of Political Science
Southern Illinois University - Carbondale
Carbondale, IL 62901-4501
lwarnold@siu.edu

Christopher J. W. Zorn
Department of Political Science
Emory University
Atlanta, GA 30322
czorn@emory.edu

This paper describes the creation of a dataset containing the percentage of union membership for each of the 435 Congressional districts of the 103rd Congress (as defined after redistricting following the 1990 Census). These data are drawn from the area union membership files derived from the Current Population Study compiled by Professors Barry T. Hirsch and David A. Macpherson of The Department of Economics, Florida State University, and mapped to Congressional districts via the Consolidated Federal Funds Report (CFFR) Geographic Reference File for fiscal year 1992, as compiled by the United States Department of Commerce Bureau of the Census. For more information on this data, see Barry Hirsch and David MacPherson, "Union Membership and Coverage Files from the Current Population Surveys," 46 Industrial and Labor Relations Review (3): 574-578 (1993).

The Area Union Membership File is a disaggregation of the Current Population Study's (CPS) data on union membership throughout the United States. Individuals are counted as union members if they respond "yes" to the question: "On this job, is ... a member of a labor union or of an employee association similar to a union?". The CPS study includes all wage and salary workers in a given geographical area; self-employed individuals and military personnel are excluded.

The Area Union Membership File consists of union membership data for 247 Metropolitan Statistical Areas (MSAs) and 50 non-urban state areas, for a total of 297 areas. The latter include all individuals not residing in one of the 247 MSAs, so that all wage and salary workers, regardless of place of residence, are covered under one of the 297 areas. For each area, the number of union members (as defined above) is divided by the total number of wage and salary workers to obtain a percentage of total union membership in the area.

The Consolidated Federal Funds Report Geographic Reference File (hereafter CFFR) contains information on 45,976 distinct geographic areas, comprising all of the 50 United States plus its territories. Included are codes for Census Bureau designations, as well as state, county, and city governmental units. Also indicated are the number of Congressional districts in which a particular geographic area lies, and the identities of those districts. This file thus allows us to match data at the MSA level to its corresponding 103rd Congressional district location.

The process of doing so is as follows. First, data on union membership percentages were merged with CFFR data, so that every one of the 45,976 subunits in the CFFR file was assigned a union membership percentage based on its location in one of the 297 MSA/state areas. Thus, for example, the several local governmental units which make up the Omaha, NE MSA (Papillion, Ralston, etc.) are all assigned a membership percentage based on that of the greater Omaha area. This is an adequate representation of union membership in most instances, since the CPS data are based on a Census-based probability sample of the entire United States.

Once the membership percentages are assigned to the CFFR data, these data are then aggregated to the Congressional district level. This is done first by making one copy of each observation for each Congressional district in which it exists, so that every district contains a CFFR observation for all CFFR areas in their boundaries. For example, if the suburb of Papillion, NE is located in both the First and Second Congressional districts of Nebraska, that CFFR location is "copied" once. One of the observations is then given a "1" for its Congressional district number, while the other is given a "2". Next, the data are aggregated over the two districts.

In order to preserve the unique characteristics of a given Congressional district, a process of accounting for a CFFR unit's contribution to the total membership of the district is required. This is accomplished in the aggregation process. Union membership in a given area is determined by weighting a CFFR area's assigned union membership percentage by its population, then summing over the weighted membership scores and dividing by the total (aggregated) population of the Congressional district. In this way, the aggregation process is able to accurately account for varying levels of population as well as membership in generating membership levels for each Congressional district.

One significant problem with this approach is that areas which exist in more than one Congressional District are counted once for each district they are in. That is, there is no way to adequately determine what percentage of Papillion, NE is in which of the two districts, and whether union membership varies between those respective parts. Because of this, both the First and the Second districts in Nebraska would include a weighted measure of union membership in Papillion based upon its entire population, rather than just on that part which exists in each district. Of course, to the extent that this distinction represents an improvement over state-based measures of union membership, these problems are relatively minor. These problems are also mitigated by noting that 37,927 of the 45,976 CFFR units (82.5%) exist in only one Congressional district, and only 24 CFFR areas are spread across 5 or more districts. Also, because there is some overlap across types (e.g. the CFFR file contains data on both counties and the cities in those counties), in many cases these larger units can be broken into smaller subunits and parcelled out to their respective districts.

The data include information on public sector labor union membership (fourth column) and total union membership (fifth column), expressed as a proportion of the total workers in the district. Also included for identifying purposes are two-letter postal codes for the state in which the district lies (third column), two digit Bureau of the Census

Governmental Unit codes also indicating the state (first column), and a number indicating the number of the district within the state (second column). Taken together, this information allows for the unique identification of each Congressional district in the U.S. This information is current for district outlines for the 103rd Congress, i.e. incorporating the redistricting following the 1990 census.