

Data Appendix
for
Dynamics of Union Organizations:
A Look at Gross Flows in the LORS Files
by
Thomas J. Holmes and Michael Walrath
June 2007

We begin with two LORS data sets, the 2000 file (current at the end of 1999) and the 2007 file (current at the beginning of 2007). We start with the raw files from LORS and perform a number of checks and make necessary corrections.

First the 2000 file. The raw LORS file has 45,736 observations. We delete observations that are inactive; this step leaves us with 31,502 active observations. We then correct known mistakes regarding discrepancies between `aff_name` and `aff_code` variables. (Unions have an affiliate name, referring to the national organization it is affiliated with, such as United Auto Workers. Each affiliate name has a three digit affiliate code assigned to it). There are 227 observations with some discrepancy between `aff_name` and `aff_code` that we correct.¹ There are six affiliate names that represent intermediate groups (called Trades or Councils). We delete all observations with these affiliate names, deleting 408 observations.

The data set contains a TLO (Type Labor Organization) variable that indicates if the organization is national (TLO=1), intermediate (TLO=2) or local (TLO=3). We only want to keep local organizations (or national organizations that have no locals). To do this, we use the TLO variable. First we make sure this variable is correct. Comparing TLO in 2000 to TLO in 2007, we find three organizations that have TLO change from 2 to 3.

¹ There are fifteen local organizations involved with Bakery & Tobacco Workers that have an `aff_name` that does not match the `aff_code`. To correct this, we take `aff_codes` and `aff_names` from the most recent LORS report. There are seven observations affiliated with the Service Workers that are assigned the `aff_code` for the Service Employees International Union; we correct these `aff_codes`. There are five observations with the `aff_code` for Unite, but the incorrect `aff_name`; we correct the `aff_name`. The Aluminum and Brick Glass Workers are affiliated with the Steelworkers, so we fix the `aff_name` and `aff_code` for these 157 observations. The Mailhandlers, LIUNA should be included directly with LIUNA (Laborers) so we fix the `aff_name` and `aff_code` for these 38 observations.

Checking the union name, it seems these organizations are locals so we change the TLO for the 2000 file. We check to see if any TLO 3's are actually intermediate groups by checking the `d_name` variable which describes the type of organization. This check reveals that the Air Line Pilots Association has organizations with TLO 3 that are actually intermediate organizations that parallel actual locals; we delete 100 of these such organizations. We then delete all observations with a TLO of 2; this deletes 1,670 observations. We then delete national organizations (TLO 1) that have locals affiliated with them; this step delete 124 observations. Since we are interested in locals in the US, we delete 184 observations with charter state (state local organized in) not in the 50 states (i.e. PR, VI). The 2000 file contains 29,016 observations (local union organizations).

Now that we have the observations of interest, we want to check for data entry error, specifically for the members variable. In order to look for possible errors in the membership variable, we look at the ratio between receipts and members. Specifically, we look at all observations such that receipts divided by members is less than 10 and members is greater than 1000. (It is unlikely an organization would be spending less than \$10 per member.) There are 41 observations that meet this criteria. We looked at each of the 41 observations, comparing the 2000 file membership to membership reported in early 2001. In 16 cases the 2000 file mis-reported the membership and we replaced it with the membership number in the 2001 file. We also checked for any observations with more than 500,000 members, we found two such observations. Both were incorrect entries and we corrected them by using the 2001 file.² These corrections reduce total membership (from 29,016 observations) from 14,905,134 to 12,894,216. Another problem with the members variable could be the number of observations with either a 0 or a missing value. Of the 29,016 observations, 3,676 observations have either a 0 or have no information reported for membership. We take LORS files collected roughly once a year between 2000-2006 and check if any of these files (file closest to 2000) have non-zero membership information. There are 848 observations

² In these cases, the 6 digit `f_number` (the unique identifier for each organization) was entered in the membership column. These two mistakes were not caught in the receipts/membership check because both of these organizations reported 0 for receipts. We further checked if there were any observations with more than 50,000 members (suspiciously large) with 0 receipts, there are none.

that have no membership information in the 2000 file, but do have membership information in another file. With these corrections, membership in the 2000 file increases from 12,894,216 to 13,215,340.

As mentioned in the paper, there is a variable that keeps track of when an organization exits and merges with another organization. This variable is called an xref. If organization A is merged in with organization B, organization A reports as an xref organization B's f_number (unique identifier for an organization). We take all xref's from the 2000 file and check whether the absorbing organizations still exist in the 2007 file. The 2000 xref's that refer to organizations still alive in 2007 we denote as "live" in 2007. In the 2000 file there are 2,076 observations with an xref; 1,868 of those observations were merged into an organization still existing in 2007. We also want to note if the absorbing organization existed before 2000 (January 1, 2000) or if it was created after 2000. We use the files from 1960, 1971, 1980, 1990, 2000 and 2007 to put an age on each organization. There were 22 organizations, representing 8,240 members, that were xref'd to an organization that had no age. Since f_numbers are issued chronologically we were able to determine if these organization with no age were pre-2000 or post-2000. We used the f_number 541595 as the last pre-2000 f_number and 541596 as the first post-2000 f_number.³ There are 1,983 organizations xref'd to an organization existing before 2000; there are 93 organizations xref'd to an organization created after 2000.

The 2007 file is created in a very similar fashion. We start with the raw report received in early 2007. It contains 47,670 observations, we delete all inactive organizations. We are left with 26,690 active observations. In the 2007 file there are a number of mergers between affiliates that were not all recorded at the local level. There are 3,262 observations for which we correct the aff_name or aff_code.⁴

³Technically there were a couple f_numbers before 541596 with early 2000 establishment dates. The xref's without an age were far away enough from this cut-off value that we can be confident about whether they are pre or post-2000.

⁴There are 1,106 locals affiliated with either Locomotive Engineers or Maintenance of the Way that should be affiliated with the Teamsters due to a merger. There are 1,155 locals affiliated with PACE that should be merged in with the Steel Workers. 656 observations from HERE and UNITE should be assigned to the newly formed UNITE HERE. 38 observations from Electrical IUW and Flight Attendants should be

With affiliations corrected we then begin deleting observations to create a set with locals only. We delete specific observations we know to be intermediate organizations (including the six affiliate names that are Trades or Councils (as we did above), the Change to Win Coalition (a new federation), Air Line Pilots Association); this deletes 475 observations. We then delete all observations with a TLO of 2; this deletes 1,437 observations. We then delete national organizations (TLO 1) that have locals affiliated with them; this step delete 115 observations. Since we are interested in locals in the US, we delete 153 observations with charter state (state local organized in) not in the 50 states (i.e. PR, VI). The 2007 file contains 24,510 observations (local union organizations).

We then checked for mistakes in the membership numbers. We again checked if the ratio of receipts to members was less than 10, while membership was greater than 1000. This check produced 37 observations. We checked them each by hand, 7 of them seem to be data entry mistakes. We take the membership figure from the report closest to 2007 and use that number for membership in 2007. By performing these corrections, membership falls in from 12,807,845 to 12,674,577. We performed one additional check; we looked at all observations with ratio of receipts to members less than 100 with membership greater than 1,000 and compared to previous years membership. This yielded one additional observation. This reduces membership to 12,673,047. There are 2,583 observations that report zero or no membership. We find non-zero membership numbers for 23 of these observations by checking reports filed before 2007. These corrections add 4,101 members, bringing the total membership for the 2007 file to 12,677,148.

Having created a 2000 file and a 2007 file, we then merge the two together. This combined set has an observation for every organization in either the 2000 or 2007 file. If an organization appeared in both files there is one observation with information for both years. (22,923 observations in both 2000 and 2007, 6,093 observations in 2000 but not 2007, 1,587 observations in 2007 but not in 2000. A total of 30,603 observations in the combined set.)

assigned to the CWA. We also fixed 38 locals with the Teamsters who had the incorrect affiliate name (they had 'Teamsters' instead of the official, 'Teamsters AFL-CIO'). The Mailhandlers LIUNA, with 36 observations, should be affiliated with LIUNA directly. 233 observations affiliated with the Transportation Communication Union should be affiliated with the Machinists.

We have quantitative variables (such as membership, assets and receipts) and descriptive variables (union name, location, affiliate name, affiliate number, xref) for each year.

Since for some specific unions we are comparing membership from 2000 to 2007, we want to be careful about locals that changed affiliations from 2000 to 2007. We are specifically concerned about this for the 10 large "minimal problem" unions, plus the Steelworkers. When looking at change in membership from 2000 to 2007, we want to include the membership of locals that became affiliated with these 11 unions some time between 2000 and 2007 (if we did not include these locals in the 2000 membership count, growth would be overstated and destruction would be understated). For each of these 11 unions we look at locals that exist in 2000 and 2007, are affiliated with one of the 11 unions in 2007, but are not affiliated with those unions in 2000. We created the variables `aff_name2000b` and `aff_code2000b` that would match `aff_code2007` and `aff_name2007`. We also took into account mergers. For example, the Maintenance of the Way Employees merged with the Teamsters between 2000 and 2007. There are a number of Maintenance of the Way locals that existed in 2000 but exited some time before 2007. We assign these locals `aff_code2000b` and `aff_name2000b` corresponding with the Teamsters.

We also create a (less detailed) data set going back to 1960. We have raw reports from 1960, 1971, 1980, 1990, 2000 and 2007. The 1960-1990 files include only active union organizations founded within the 50 states. For the 2000 file we need to delete inactive observations; for the 2007 file we need to delete observations that are inactive or founded outside the 50 states. We then merge in these 6 different files, using the unique file number (`f_number`) for each organization. We denote whether or not each organization appears in each of the six files. We assign each observation an age. The age of an observation is the year of the first report in which that observation appears. There are 89,827 different organizations in this set (includes national, intermediate, and local groups). Due to changing report requirements over this time frame we want to delete affiliates that are government or postal affiliated. Since the early files do not have a variable indicating this, we delete all unions with affiliate names containing the following terms: Postal, Post Off, Letter, Federal Emp, Government, Treasury, Air Traf, Weather, Civilian Tech. This deletes 15,182 observations, leaving 74,645 observations. The 2000 and 2007 file include a variable indicating if the

organization is government, postal or other. We can use this variable to check for how well we did at clearing out government and postal organizations. Of the 74,645 observations, 315 of them are either government or postal according to the 2000 or 2007 reports. These 315 observations have 73 different affiliate names, implying that we are not missing any government or postal unions.

In more recent years the reports from LORS include more detail about types of members (active, retired, etc.). We want to look at this detailed level of membership and then possibly take certain unions out of the sample when analyzing the data. We specifically look at the 2007 LORS report. In LORS we can see membership reported by the national organization of a union ($TLO = 1$) and the sum of all the locals for that union. The report from the national will include Canadian members. (Since we only have reports from US locals, the sum over locals is automatically US only.) Canada reports membership for large unions, so we can take the $TLO=1$ report and subtract off Canadian membership for the unions of interest.⁵ In Table 2 we compare the membership totals reported by the national versus the sum of the local membership reports. We are also interested in the type of member, specifically how many retirees the unions are counting. We use the detailed information on membership.⁶ We obtain information on retirees both from the national report and the sum of the local reports. To calculate shares we take the number of retirees and divide by the number of members (either reported by national, or sum of locals). This information is reported in Table 2.

We use the joint 2000, 2007 file to look at changes in membership. For each observation that is in both 2000 and 2007 we noted if the observation grew or shrank over this period and recorded the change in membership; we observed both creation and destruction from

⁵The source of the data for Canada is Human Resources and Social Development Canada (2006). No information is given for the Communication Workers (CWA). We estimated this by using the web and finding any CWA locals in Canada. We found only one, TNG (a newspaper guild) with 9,000 members.

⁶ There are two ways to obtain this information. One way is to go to the DOL website and submit a query for membership information, specifying which unions and time period. Also available on the DOL website is text files with the detailed membership information for all unions. Unfortunately, at this time these two sources are not identical. We use information obtained from a query specifically for specific unions of interest.

continuing locals. Membership added from a local that exists in 2007 but not 2000 is growth due to creation. Membership lost due to a union existing in 2000 but not 2007 is destruction due to exit. When dealing with creation and exit we need to take into account locals that were merged into other locals; we use the xref variable to do this. We distinguish between locals that merged into other locals that still exist in 2007 versus locals merged into locals that do not exist in 2007. We also distinguished between locals merged into locals that exist before 2000 and locals created after 2000. When looking at changes in membership for particular unions (Steelworkers, Auto Workers, Service Employees) we first use the affiliation assigned to a local in 2007. If the local did not exist in 2007 we used the aff_codeb (discussed above) assigned to the local in 2000 (using the aff_codeb takes care of locals that exit before 2007 but were part of an affiliate that merged with another affiliate before 2007).

We use the data set going back to 1960 to calculate conditional survival rates. We already have an age variable (indicating the year of data set in which an f_number first appeared) and a variable indicating if an f_number was in a file for a particular year. We create a variable indicating whether an observation appears in some starting year and some future year (1960 and 1971, 1960 and 1980, etc., then 1971 and 1980, 1971 and 1990, etc.,). For each year, conditional on age, we calculate the fraction of organizations that exist in a certain year that also exist in a future year. These survival probabilities are seen in Table 5. For Table 6 we use survival rates calculated by using all organizations existing in a certain year and calculating the probability of surviving to the next period (we do not condition on age). Using the period-long survival rate, we calculate the annual rate. One minus this annual survival rate gives the estimated exit rate reported in Table 6. The birth cohort is the number of organizations with a certain age, beginning and end of period counts come from whether or not an organization was in a certain report. Estimated Level of Entry takes into account that in between report years organizations are both dying and being born. Between 1960 and 1970 we seen that 10,550 organizations are born, but there could be many organizations that were created and then died between 1960 and 1970 (that we never see). To account for this, we solve: $10,550 = N + N(1-d) + N(1-d)^2 + \dots + N(1-d)^{10}$ where d is the annual destruction rate and N is the number of new organizations (we assume

this to be constant). We want to look at ratio of new members organized to the private labor force. We use the estimate for number of new organizations and multiply by 380⁷ to obtain number of new union members being organized. We use data from the Economic Report of the President for size of non-union private labor force. Using our estimate for number of new private union members organized and the level of non-union private employment we obtain Estimated New Organization Rate.

We then look at current (2007) membership and see how members are distributed across locals of different ages. We start with the file with observations from 1960 through 2007. We are only interested in locals existing in 2007 and that have a variable indicating they represent private workers (the 2007 file contains this variable, we delete observations that represent public or postal). We then use the age variable to determine the percent of members that belong to organizations of different ages. For specific unions we want to see how different affiliates that have been merged in account for current membership. We keep all observations from the 2007 file for the unions of interest. For 1960, 1971, 1980, 1990, 2000 using the f_number we determine with whom the organization was affiliated with. If the local did not exist in that year we labeled it as "future entry." For each decade we attributed current membership to the affiliate it belonged to at that time.

⁷This is our estimate for average size of organization. We use private union membership compiled by Hirsch for 1971, 1980, 1990, 2000, 2007 and use our counts for number of union organizations in each of these years to obtain average membership. This average membership is 311, 342, 329, 393, 396 for 1971, 1980, 1990, 2000, 2007 respectively. We chose to use 380.