Lab Exercise #7: Leveraging panels to increase sample size of ORG responses linked to supplement data

Learning Goals

- Understand how to leverage panel structure to increase sample size of linked ORG response data and supplement-specific measures
- Understand additional particularities of validating jobs across time
- Link augmented ORG responses paired with supplement-specific measures across two years
- Identify sources for universe information

Summary

The Outgoing Rotation Group (ORG)/Earner Study includes data on wages and union status as well as other more detailed job-related data. ORG data are collected monthly, but only for respondents with MISH values of 4 or 8; this is only 25% of the sample. In this exercise, we'll explore how to leverage the panel component of the CPS to combine labor union data collected as part of the ORG with data on volunteerism from the Volunteer supplement, even when the respondent's ORG data are fielded in a different month than the Volunteer supplement. We increase our sample size for exploring the relationship between volunteering and union status substantially by combining Volunteer supplement data for those in MIS 1-3 and 5-7 with their ORG data observed in future months.

Make an extract that includes the Volunteer supplement and the three subsequent months for 2007 and 2008 for a total of eight samples (four from 2007 and four from 2008). Be sure to include UNION and ELIGORG from the ORG/Earner Study and VLSTATUS from the Volunteer supplement. Additionally, include MISH, EMPSTAT, LABFORCE, EMPSAME, OCC1990 and IND1990 as well as variables for validating links (AGE, SEX, RACE).

Part 1: Augmenting Sample Size by Linking

1. How many observations are there per month/year combination? Fill in Table 1.

Table 1

Month	2007	2008
September		
October		
November		
December		

2. How many Volunteer supplement respondents are in the data?
   a. 2007: _____
   b. 2008: _____
3. Why is the number of Volunteer supplement respondents lower than the total number of respondents in September?

   Hint: See Supplement Documentation on IPUMS-CPS website.

   _____

   _____

4. Using ELIGORG, how many respondents are eligible for the ORG questions? Fill in Table 2.

Table 2

Month	2007	2008
September		
October		
November		
December		

5. What is the universe for the ORG questions?

   Hint: See ORG User Notes on IPUMS-CPS website.

   _____

   _____

For the remainder of part 1, use observations from 2007 only.

6. How many people are members of a union in the 2007 Volunteer supplement? How many people are members of a union and volunteer? How many people are not members of a union and volunteer? Fill in Table 3 below.

7. What are the MISH values of people in the Volunteer supplement with information on their union membership? About how many could we expect to retain in the sample if we linked forward to ORG responses for people in MISH 1-3 and 5-7 in the Volunteer supplement?

_____

Hint: for the next step, you will be linking within 2007 for everyone not in MIS 4 or 8 at the time of Volunteer supplement data collection. Your final dataset should be the September CPS where the missing ORG values now have the ORG values from future months.

To increase the sample size of Volunteer supplement respondents with a union membership response, link the Volunteer supplement data to the following three months to recover the UNION variable from each respondent's MISH 4 or 8 response in 2007.

8. How many of these links validate based on Sex? _____ Race? _____
   Age? _____ All three characteristics? _____

Table 3

	Question 6: *Volunteer Supplement Only*		*Questions 9-10:* *Volunteer Supplement Combined with ORG*		*Question 13:* *Volunteer Supplement Combined with ORG and validated employment*	
	In Volunteer Supplement	**Volunteers Only**	**In Volunteer Supplement**	**Volunteers Only**	**In Volunteer Supplement**	**Volunteers Only**
Union						
Not in Union						

9. Using the validated dataset linking Volunteer supplement and ORG data, how many respondents are included in the Volunteer supplement? Fill in Table 3.

10. Using the validated dataset linking Volunteer supplement and ORG data, how many people are members of a union and volunteer? How many people are not members of a union and volunteer? Fill in Table 3.

11. How much larger is the validated dataset linking Volunteer supplement and ORG data from subsequent months than the sample of only those members whose ORG data are collected in the same month as the Volunteer supplement data?

_____

_____

Part 2: Validating Job Transitions

This exercise raises another flavor of linking validation. In this exercise, we are examining the relationship between volunteer status of a person in September of 2007 and union membership of a person (either in the same or subsequent months, depending on MISH at the time of the Volunteer supplement). The linking validation from Part 1of this exercise does not verify that the job referenced in UNION in October, November, or December of 2007 is the same job referenced in the September Volunteer supplement. This is not a problem when we retrieve union membership and volunteerism in the same month. However, if we are linking union membership information from later survey months, another important validation is comparing a person's job at the time of reporting volunteer information to their job at the time of reporting union membership.

12. Compare OCC1990 and IND1990 in September to the appropriate linked ORG data in subsequent months. How many people match occupations between the two time periods? _____ Industry? _____ Both? _____

13. How many respondents in the Volunteer supplement have the same job when they respond to UNION? How many people are members of a union and volunteer? How many people are not members of a union and volunteer? Fill in the final columns of Table 3 from part 1 of this exercise.

Part 3: Linking Across Years

So far, we have only linked respondents within a single year in order to increase the number of people for whom we have both Volunteer supplement data and ORG data. Researchers may also be interested in studying changes in volunteer status across years. To do so, we can link people across two years.

First, create a linked ORG/Volunteer supplement sample for 2008.

14. About how many of our currently-linked and validated 2007 records are eligible to link to our 2008 linked ORG/Volunteer supplement sample?

_____

_____

Now, link the 2007 sample to the 2008 sample.

15. How many people link across these two years before validation? _____ How many links validate based on Sex? _____ Race? _____ Age? _____ All three characteristics? _____

16. For those whose links are validated, examine transitions in volunteering and union status. How many people change volunteer status across these two years? _____ How many people change union status across these two years? _____

Part 1: Augmenting Sample Size by Linking

1. How many observations are there per month/year combination? Fill in Table 1.

Table 1

Month	2007	2008
September	134,647	133,585
October	133,962	133,482
November	134,285	132,812
December	133,672	131,026

2. How many volunteer supplement respondents are in the data?
   a. 2007: 96,268
   b. 2008: 94,944
3. Why is the number of Volunteer supplement respondents lower than the total number of respondents in September?
   The universe for the Volunteer supplement is different than the universe for the Basic monthly. The volunteer supplement includes only people 15 and older whereas the Basic monthly includes data on all persons.
4. Using ELIGORG, how many respondents are eligible for the ORG questions? Fill in Table 2.

Table 2

Month	2007	2008
September	14,772	14,463
October	14,915	14,674
November	14,725	14,315
December	14,833	14,361

5. What is the universe for the ORG questions?
   The universe for the ORG questions is, in addition to being month-in-sample 4 or 8, includes only civilians age 15 and older who are currently employed as a wage or salaried worker (that is, not self-employed).
6. How many people are members of a union in the 2007 Volunteer supplement? How many people are members of a union and volunteer? How many people are not members of a union and volunteer? Fill in Table 3 below.

Table 3

	Question 6: Volunteer Supplement Only		Questions 9-10: Volunteer Supplement Combined with ORG		Question 13: Volunteer Supplement Combined with ORG and validated employment	
	In Volunteer Supplement	**Volunteers Only**	**In Volunteer Supplement**	**Volunteers Only**	**In Volunteer Supplement**	**Volunteers Only**
Union	1,591	494				
Not in Union	11,482	3,211				

7. What are the MISH values of people in the Volunteer supplement with information on their union membership? About how many could we expect to retain in the sample if we linked forward to ORG responses for people in MISH 1-3 and 5-7 in the Volunteer supplement?
   MISH 4 and MISH 8. Theoretically, we retain the entire sample if we link forward.
8. How many of these links validate based on Sex? 134,557 Race? 134,576 Age? 134,032
   All three characteristics? 133,952
9. Using the validated dataset linking Volunteer supplement and ORG data, how many respondents are included in the Volunteer supplement? Fill in Table 3.
10. Using the validated dataset linking Volunteer supplement and ORG data, how many people are members of a union and volunteer? How many people are not members of a union and volunteer? Fill in Table 3.
   How much larger is the validated dataset linking Volunteer supplement and ORG data from subsequent months than the sample of only those members whose ORG data are collected in the same month as the Volunteer supplement data?
   (43,642 + 6,112) - (11,482 + 1,591) = 36,681

Table 3

	Question 6: Volunteer Supplement Only		Questions 9-10: Volunteer Supplement Combined with ORG		Question 13: Volunteer Supplement Combined with ORG and validated employment	
	In Volunteer Supplement	**Volunteers Only**	**In Volunteer Supplement**	**Volunteers Only**	**In Volunteer Supplement**	**Volunteers Only**
Union	1,591	494	6,112	2,054		
Not in Union	11,482	3,211	43,642	12,929		

Part 2: Validating Job Transitions

12. Compare OCC1990 and IND1990 in September to the appropriate linked ORG data in subsequent months. How many people match occupations between the two time periods? 129,253 Industry? 129,972 Both? 128,702

13. How many respondents in the Volunteer supplement have the same job when they respond to UNION? How many people are members of a union and volunteer? How many people are not members of a union and volunteer? Fill in the final columns of Table 3 from part 1 of this exercise.

Table 3

	Question 6: Volunteer Supplement Only		Questions 9-10: Volunteer Supplement Combined with ORG		Question 13: Volunteer Supplement Combined with ORG and validated employment	
	In Volunteer Supplement	**Volunteers Only**	**In Volunteer Supplement**	**Volunteers Only**	**In Volunteer Supplement**	**Volunteers Only**
Union	1,591	494	6,112	2,054	5,767	1,963
Not in Union	11,482	3,211	43,642	12,929	39,599	11,858

Part 3: Linking Across Years

14. How many of our currently-linked and validated 2007 records are eligible to link to our 2008 linked ORG/Volunteer supplement sample
50% (MISH 1 - 4)

15. How many people link across these two years before validation? 48,884 How many links validate based on Sex? 48,351 Race? 48,884 Age? 45,502 All three characteristics? 45,417

16. For those whose links are validated, examine transitions in volunteering and union status. How many people change volunteer status across these two years? 3,153 + 2,705 = 5,858 How many people change union status across these two years? 570 + 563 = 1,133