

PS207 Homework #3: Missing Data

There is a Stata dataset on my website called ps207hw3. This is survey data from the 1994 Federal election in Germany. The variables in this dataset are:

- *FavOut*, coded 0 if an individual was unhappy with the election outcome, 1 if happy
- *LiveEast*, coded 0 for people living in West Germany, 1 for people living in East Germany
- *income* in Deutschmarks, divided by 10,000
- *age* in years
- *union*, coded 0 for someone from a non-union household, 1 for someone from a union household
- *female*, coded 0 for men and 1 for women);
- *SocClass*, a 5 point scale with higher values indicating higher self-reported social class
- *educ*, an 8 point scale with higher values indication more formal education
- *rejectEI*, coded 0 for people that believe economic individualism is a good thing, 1 for those who disagree
- *exppfin* coded -1 for those who expect their personal finances to get worse in the coming year, 0 for those who think it will stay the same, and 1 for those who expect things to get better

Download this dataset and do the following three problems. Give me a couple of paragraphs writeup on each problem, as well as the log file.

Problem 1

Go to the section of my website for publications and research, and take a look at my paper with Roberto Weber entitled “Is there a relationship between election outcomes and perceptions of personal economic well-being? A test using post-election economic expectations.” Replicate Model 1 in Table 2 using the homework dataset. Don’t forget to create and include the appropriate interaction terms. (Note you don’t need much of a writeup for this problem, as you’ve done this before).

Problem 2

Using *Amelia*, fill in the missing values in the data with multiple imputation, making any transformations you need to the data so that the assumptions of this MI program are met (consult the *Amelia* manual for guidance on this. Create 5 imputed datasets.

Where was most of the missing data in the original dataset? Do the imputations produced by *Amelia* look plausible?

Problem 3

Use *Clarify* to replicate Model 1 in Table 2 again, this time estimating the model across all 5 imputed datasets and combining the results. Are there any differences between these results and the results in Problem 1?