PUBPOL 6090 Lecture 2 Notes

Thelonious Goerz

2023-08-23

Pre-lecture notes

Robustness versus efficiency

Suppose a model where individuals are indexed by a unit i and a time t.

$$y_{it} = \beta X_{it} + \alpha_i + \epsilon_{it}$$

The α_i coefficient can be modeled using a fixed effect or a random effect approach. This models whatever is time invariant at the individual level.

In the fixed effects approach, we assign unit level fixed effects that have t observations within each individual and we get a coefficient for X that is net of within person average variation. So, the fixed effect adjusts for the individual average variation and the coefficients of beta then capture all variation net of an individual's within-person variation. In this case, the α_i are dependent on the Xs where as the random effects approach assumes that α is independent of the Xs.

The random effects estimator is the *difference* between the weighted average of the *within* i observations and the between i *estimates* $RE = \bar{x_i} - \beta_x$. Random effects assume that $Cov(\alpha_i, x_i t) = 0$ Random effects allows for the estimation of *time constant* effects, while fixed effects does not.

The fixed effects estimator may be robust but less efficient in some cases.

Note: In hierarchical linear modeling, what are called fixed effects are actually random effects. The fixed and random effects *estimators* are distinct from the fixed and random effects concepts.

If the true state of the world is that the expectation of α_i conditional on X_i is equal to zeroˆ{Zero Conditional Mean assumption} then:

1. Using a fixed effect estimator will
    1. Produce consistently estimated $\hat{\beta}$
    2. But the estimator will not be efficient
2. Using a random effects estimator will
    1. Produce consistent $\hat{\beta}$
    2. Estimate very efficiently

    If the true state of the world is that the expectation of α_i and X_i are not equal then:

    1. Using a fixed effect estimator will be consistent but not efficient
    2. Using a random effects estimator will not be consistent

OLS (robust) versus flexible generalized least squares

Suppose

$$y_i = \beta X_i + \epsilon_i$$

If the $Var(\epsilon_i) = yZ_i + u_i$ then

1. OLS with Robust standard errors will be beta hat consistent, provide correct inference, but not be efficient
2. FGLS will be consistent, correct inference, and very efficient

If the $Var(\epsilon_i) = else$

1. OLS will be beta hat consistent, correct inference, and not efficient
2. FGLS will be consistent, incorrect inference, not efficient

Feasible Generalized Least Squares (FGLS) is based on our ability to know and model the heteroskedasticity of the errors.

Theoretical map of econometrics

Estimation We start with a population, that is unobserved. Then, we collect a sample of data, which are the Ys and Xs that we will model. Further, we choose a method (for instance, regression) and we choose a method of estimation (for instance, least squares or maximum likelihood). Finally, we obtain results, such as coefficients and standard errors, and interpret the results using statistical inference and relevant, credible assumptions about the data generating process and relationship between variables.

Results Results are usually in the context of the sampling distribution (frequentist). We obtain a sample of the data out of many potential samples that we could draw. For each sample, we could have estimated our coefficients which would have yielded *different* results in each sample. So, we think about the sampling distribution of $\hat{\beta}$s that we could have obtained. Different methods of estimation will have different properties for their sampling distributions.

- An OLS result will be, on average, unbiased estimate of the true population parameter β
- The variability of the sample, the sampling standard deviation is the broader concept, or the standard error is the efficiency of the estimator. A smaller standard error is more efficient.
- The BLUE result states that among alternatives, OLS is the *most* efficient estimator.
- The Plim as N gets bigger will be equal to the true β which is called consistency
- Importantly, the least squares coefficients are random variables because they have a distributions, the β_k has a distribution of alternative samples.
- Via the central limit theorem, the distribution of β is asymptotically normal $N(\beta_k, \sigma^2$. As the plim of the sample grows, the distribution of the sample means will be approximately normally distributed.

Questions Q1: The Freedman paper is not a "what's the effect of X on Y" paper. How do you describe its goals? What type of paper is it?

Freedman critiques what he believes to be an over-reliance on regression modeling for the purpose of drawing causal inferences in social science instead of careful reasoning, domain knowledge, and shoe leather – in reference to the work of Jon Snow, which he believes demonstrates what social science research should aspire to. Freedman takes offers a criitique of the paradigm of using model-based inference to attempt to control away and adjust specific aspects of estimation to identify a causal quantity. Rather, he advocates for researchers to pend time thinking through estimation and functional form assumptions as well as utilizing design-based inference approaches such as natural experiments to identify causal quantities. Overall, he argues that the convenience of regression as a research tool has lead researchers to begin to formulate hypotheses as

whether certain coefficients are significant or not rather than doing careful investigative work like that of Snow.

Q2: For getting at causal impacts, Freedman clearly dislikes using regression analysis or other statistical techniques. What does he propose as alternatives?

Freedman highlights natural experiments and investigative work, kike that of Jon Snow, as alternatives to using regression-based methods. From the outset of the paper, he does not entirely discredit regression, explaining that he believes that in certain contexts it can be used effectively, but advocates for a more careful assessment of assumptions about specification and the data generating process.

Q3: What's the research question (narrowly defined) of the DiNardo and Pishkie paper? What's the broader research question of the DiNardo and Pishkie paper?

DiNardo and Pischke examine whether the wage differentials associated with computer use reflect productivity differences due to the introduction of computers in the workplace. More broadly, they attempt to offer an explanation about whether these returns explain large wage differentials observed in the last 15 years or whether this is the result of unobserved heterogeneity.