To stress test the causal relationship in our Predictions -I introduce a more sophisticated causal inference technique by including lagged dependent variables. **This technique is known as the Arellano-Bond Estimator**

- I am trying to estimate the within effect of our model using the Arellano-Bond estimator.

This estimation technique differs from the dynamic panel or cross-sectional model because it deals with a single dependent variable.

- The reason to include this to control for unobserved heterogeneity factors

Timeline

Description automatically generated

- Let’s take a simpler model, without lag as the predictor as the first step: In our data, I use the lagged grad rate as the predictor

Though I saw some impact on the Grad rate on the enrollment, with the above first step, I see some inconsistent estimations.

To deal with this endogeneity problem, I apply the **First Difference**. This will eliminate the effects of unobserved variables

On doing the first difference, I will cancel out the time invariant factors.

Ie. the difference of enrollment between 2017 and 2016 – regressed with Graduation rate difference between 2016 and 2015. Here, I am good to do with the estimators controlling the effects of endogeneity issues/control the unobserved heterogeneity

* IMPORTANT and FINAL STEP
* Now, when I have lagged dependent variable in the predictors list, and do the First-difference. And get rid of the unobserved effect using the first differencing

I then included a difference of enrollment between 2016 and 2015(TotalEnrollmentit-2) as the dependent variable

**So doing the First Difference will eliminate the effects of unobserved variables**

Graphical user interface

Description automatically generated

**Inferences:** Ultimately, the sequence of tests has allowed me to make a stronger claim to the causal impact prior period graduation rates have on enrollments. This was achieved by using lag variables and first differences to control for many of the endogeneity inherent in the process