Results outline for Meeting on Wednesday, 1:30pm

Datasets: roads\_levies2\_census\_9118.dta (voting and census), housesales\_9521\_slim.dta (housing data)

Code files:   
Stata:

1. ohio\_road\_housing\_census\_merge.do
2. road\_spending\_reg\_discontinuity\_agg.do

R:

1. housing\_data\_setup.R ohio\_road\_housing\_census\_merge.R
2. road\_spending\_reg\_discontinuity\_agg.R
3. covariates\_balance\_test.R
4. road\_spending\_reg\_discontinuity\_agg\_w\_covariates.R
5. Data Merges
   1. merged *housesales\_9521\_slim.dta* and *roads\_levies2\_census\_9118.dta* (both provided by Dr. Brasington) to create t-2, t-1, …., t+10 combined datasets containing **transaction-level** housing, voting and census data with renewals only and duration < 1000 –
      1. Datasets containing matches and non-matches (\_merge == 1, 2, 3)  
         housing\_roads\_census\_t\_minus\_1.dta, housing\_roads\_census\_t\_minus\_2.dta, housing\_roads\_census\_t\_plus\_1.dta ,……, housing\_roads\_census\_t\_plus\_10.dta
      2. Datasets containing matches only (\_merge == 3)  
         housing\_roads\_census\_t\_minus\_1\_matches.dta, housing\_roads\_census\_t\_minus\_2\_matches.dta, housing\_roads\_census\_t\_plus\_1\_matches.dta ,……, housing\_roads\_census\_t\_plus\_10\_matches.dta
   2. **Aggregated** transaction-level datasets to create county subdivision + year level dataset. Each observation represents median sale amount or median sale amount per sq feet for each county for each year and the associated county-level characteristics at the time of voting, along with voting result  
      Note: if duplicate was found (i.e. two road polls were conducted by a county in a year), then kept observation that was closest to the cutoff (see creation of roads\_and\_census.dta)
      1. Datasets containing matches only for median sale amount  
         housing\_agg\_roads\_census\_t\_minus\_1.dta, housing\_agg\_roads\_census\_t\_minus\_2.dta, housing\_agg\_roads\_census\_t\_plus\_1.dta, …. housing\_agg\_roads\_census\_t\_plus\_10.dta
      2. Datasets containing matches only for median sale amount per square feet  
         housing\_agg\_roads\_census\_per\_t\_minus\_1.dta, housing\_agg\_roads\_census\_per\_t\_minus\_2.dta, housing\_agg\_roads\_census\_per\_t\_plus\_1.dta, …. housing\_agg\_roads\_census\_per\_t\_plus\_10.dta  
           
         Note: median sale amount and median sale amount per square feet variables were in different datasets due to differences in aggregation caused by missing square feet values

So, analysis was performed on **housing\_agg\_roads\_census\_<t type>\_<year>.dta** and **housing\_agg\_roads\_census\_per\_<t type>\_<year>.dta** datasets.

They contain –

1. outcome variable: median sale amount and median sale amount per sq feet
2. Running variable: percent of votes where 50 is the cutoff
3. Treatment variable: whether county passed or failed a particular poll for road spending
4. Covariates: county-specific characteristics
5. Results  
   1. Density/Manipulation Tests & McCrary test: to check if running variable was manipulated i.e. discontinuity in distribution of running variable at cutoff; if some people were aware of exact outcome beforehand and purposefully voted just above
   2. RD plots: check whether discontinuity appears in graph b/w outcome and running variable
   3. Regressions without covariates:
   4. Covariate balance test:
   5. Regressions with covariates: