```
*******************************************************************************
****************************** Section 1: Regressions ***************************
*******************************************************************************

***************************** VAR1: PROP APPRECIATION ***************************

* check without intervention (base trend)
xtreg ia_total_appr ow, fe vce(cluster PIN_n)

* Preliminary step: ensure PINs are numbers, not strings
encode PIN_num, gen(PIN_n)
* Step 1: check for duplicates of same PIN/time combination, and set data as ///
panel data
duplicates list PIN_n time
xtset PIN_n time
* Step 2: run an xtreg. include controls for time fixed effects and zoning ///
(i. makes it categorical))
xtreg ia_total_appr ow##post_int i.time, fe vce(cluster PIN_n)
xtreg ia_total_appr ow##post_int i.time i.z_num, fe vce(cluster PIN_n)
* controlling for just zoning reduces the part of the overall observed trend ///
that might be attributal to zoning. Adding an ineraction between zoning and ///
the treatment tests whether the treatment had a different effect in different ///
zoning types -- only report on interaction coeff. if part of my main question ///
this is not what i want, want to know if intervention had diff effect in diff ///
zoning types, right?
xtreg ia_total_appr ow##post_int i.time i.z_num ow##i.z_num, ///
```

```
fe vce(cluster PIN_n)
* check for multicollinearity
vif
coefplot

*** EXCLUDING RESIDENTIAL ***
* Preliminary step 1: create variable for only non-res properties
gen appr_comm = ia_total_appr if z_num==2
* Preliminary step 2: ensure PINs are numbers, not strings
encode PIN_num, gen(PIN_n)
xtreg appr_comm ow##post_int i.time, fe vce(cluster PIN_n)



****************************** VAR2: LAND ACTIVATION ****************************
* Preliminary step 1: create variable for only once-vacant properties
gen vac_new = new_total1 if vacant_base_yr==1
drop if vacant_base_yr==0
* Preliminary step 2: ensure PINs are numbers, not strings
encode PIN_num, gen(PIN_n)
* Step 1: check for duplicates of same PIN/time combination, and set data as ///
panel data
duplicates list PIN_n time
xtset PIN_n time
* Step 2: run an xtreg. include controls for time effects
xtreg new_total ow##post_int i.time, fe vce(cluster PIN_n)
xtreg new_bldg ow##post_int i.time, fe vce(cluster PIN_n)
```

xtreg new_land ow##post_int i.time, fe vce(cluster PIN_n)

******** CLEANER TABLES ***********

asdoc xtreg new_total ow##post_int i.time, fe vce(cluster PIN_n)

asdoc xtreg new_bldg ow##post_int i.time, fe vce(cluster PIN_n)

asdoc xtreg new_land ow##post_int i.time, fe vce(cluster PIN_n)

* (less good)

eststo model1: xtreg new_bldg ow##post_int i.time, fe vce(cluster PIN_n)

eststo model2: xtreg new_land ow##post_int i.time, fe vce(cluster PIN_n)

esttab using mytable.rtf, replace

*** LAND ACTIVATION OF VACANT PROPERTIES ONLY ***

****

********** Section 2: Charts **********

****

* NOTE: Some changes to chart titles and appearances made within the graph ///

editor; not recorded here

********** VAR1: PROP APPRECIATION **********

```
        ****************************** BOX CHART ******************************
* TITLE: Property Value Appreciation across time for the Treatment and  ///
Control Groups
* Land Activation levels across time for the Treatment and Control Groups

* Step 1: represent the total appr variable as a percentage
gen ia_prctg = ia_total_appr*100
* Step 2: Create the box chart
graph box ia_appr_prctg, over(ow) over(time) nooutsides asyvars                    ///
    title("title")                    ///
    legend(rows(1) position(12))                    ///
          ytitle("Appreciation")                    ///
   xtitle("Time")

*** simple box chart (starting point) ***
        graph box ia_total_appr, over(ow) over(time) nooutsides



        ***************************** LINE BTWN MEANS ************************

* TITLE: Mean Property Value Appreciation in each Time Period for the ///
Treatment and Control Groups
* Step 1: represent the total appr variable as a percentage
gen ia_prctg = ia_total_appr*100
* Step 2: Generate mean variables for each time period, for control and trtment
collapse (mean) ia_prctg, by(ow time)
```

```
reshape wide ia_prctg, i(time) j(ow)
* Step 3: create line chart connecting means for the trtmnt and control
line ia_prctg0 ia_prctg1 time, ///
   lpattern(solid dash) ///
   lcolor(blue navy) ///
   lwidth(medthick medthick) ///
   legend(label(1 "control") label(2 "treatment")) ///
   title("title") ///
   ytitle("mean") ///
   xtitle("time")

* TITLE: Median Property Value Appreciation in each Time Period for the ///
Treatment and Control Groups
* Step 1: Make the total appr variable percentage
gen ia_prctg = ia_total_appr*100
* Step 2: Generate median variables for each time period, for control and trtment
collapse (median) ia_prctg, by(ow time)
reshape wide ia_prctg, i(time) j(ow)
* Step 3: create line chart connecting medians for the trtmnt and control
line ia_prctg0 ia_prctg1 time, ///
   lpattern(solid dash) ///
   lcolor(blue navy) ///
   lwidth(medthick medthick) ///
   legend(label(1 "control") label(2 "treatment")) ///
   title("title") ///
   ytitle("mean") ///
```

```
   xtitle("time")


*** EXCLUDING RESIDENTIAL ***
* TITLE: Mean Property Value Appreciation for Only Commercial and Industrial ///
Properties
* Step 1: Generate a total appr variable that excludes residential ///
properties, and represent as a percentage
gen noR_total = ia_total_appr if zc_num!=1
gen noR_totalprctg = noR_total*100
* Step 2: Generate mean variables for each time period, for control and trtment
collapse (mean) noR_totalprctg, by(ow time)
reshape wide noR_totalprctg, i(time) j(ow)
* Step 3: create line chart connecting means for the trtmnt and control
line noR_totalprctg0 noR_totalprctg1 time, ///
   lpattern(solid dash) ///
   lcolor(blue navy) ///
   lwidth(medthick medthick) ///
   legend(label(1 "Two-way") label(2 "One-way")) ///
   title("title") ///
   ytitle("Mean Appreciation Value") ///
   xtitle("Time")


* TITLE: Median Property Value Appreciation for Only Commercial and ///
Industrial Properties
* Step 1: Generate a total appr variable that excludes residential ///
properties, and represent as a percentage
```

```
gen noR_total = ia_total_appr if zc_num!=1
gen noR_totalprctg = noR_total*100
* Step 2: Generate median variables for each time period, for control and trtment
collapse (median) noR_totalprctg, by(ow time)
reshape wide noR_totalprctg, i(time) j(ow)
* Step 3: create line chart connecting medians for the trtmnt and control
line noR_totalprctg0 noR_totalprctg1 time, ///
   lpattern(solid dash) ///
   lcolor(blue navy) ///
   lwidth(medthick medthick) ///
   legend(label(1 "Two-way") label(2 "One-way")) ///
   title("title") ///
   ytitle("Median Appreciation Value") ///
   xtitle("Time")

                  ************************** MARGINS PLOT ************************
*** POTENTIALLY MORE POWERFUL THAN RAW MEANS... but harder to interpret
xtreg ia_total_appr ow##i.time, fe vce(cluster PIN_n)
margins ow, at(time=(1 2 3 4 5))
marginsplot, title("Predicted Property Appreciation Over Time") ///
   legend(rows(1) position(12))                              ///
         legend(label(1 "Two-way") label(2 "One-way")) ///
   title("title")

* include intervention? (DIDN'T WORK) ///
xtreg ia_total_appr ow##post_int i.time, fe vce(cluster PIN_n) ///
```

```
margins ow, at(time=(1 2 3 4 5)) ///
marginsplot, title("Predicted Property Appreciation Over Time")

**************************** VAR2: LAND ACTIVATION ****************************
        **************************** BAR CHART ********************************
* TITLE: Land Activation Levels across Time for the Treatment and ///
Control Groups
* Step 1: For each time period and group (trtmnt or control), create variable ///
= count of properties that either: (a) went from land value=0 to positive ///
land value (developable land creation) or (b) went from bldg value=0 to ///
positive bldg value (new construction)
bysort time: egen count_new = total(new_total >= 1) if ow==1
bysort time: egen count_newC = total(new_total >= 1) if ow==0
replace count_new = count_newC if count_new==.
drop count_new0
* Step 2: Create Bar chart
graph bar (mean) count_new, over(ow) over(time) asyvars     ///
    title("title")     ///
          legend(rows(1) position(12))

        ************************* LINE BTWN COUNTS ****************************
* TITLE: Land Activation Levels across Time for the Treatment and ///
Control Groups
* Step 1: Generate mean variables for each time period, for control and trtment
collapse (mean) count_new, by(ow time)
reshape wide count_new, i(time) j(ow)
```

```
* Step 2: Create Line Graph
line count_new0 count_new1 time, ///
   lpattern(solid dash) ///
   lcolor(blue red) ///
   lwidth(medthick medthick) ///
   legend(label(1 "control") label(2 "treatment")) ///
   title("Title") ///
   ytitle("Number of Activated Properties") ///
   xtitle("Time")

*** LAND ACTIVATION OF VACANT PROPERTIES ONLY ***
* TITLE: Land Activation of Vacant Properties for the Treatment and ///
Control Groups
* Step 1: create variable of only properties vacant in the base year
gen vacant = new_total if vacant_base==1
* Step 2: Generate mean variables for each time period, for control and trtment
bysort time: egen count_new = total(vacant_base >= 1) if ow==1
bysort time: egen count_new0 = total(vacant_base >= 1) if ow==0
replace count_new = count_new0 if count_new==.
drop count_new0
* Step 3: Create Line Graph
line count_new0 count_new1 time, ///
   lpattern(solid dash) ///
   lcolor(blue red) ///
   lwidth(medthick medthick) ///
   legend(label(1 "control") label(2 "treatment")) ///
```

```
   title("Title") ///
   ytitle("Number of Activated Properties") ///
   xtitle("Time")
```

********** Notes **********

* Controls: factor variables (i.___) and interaction (ow*i.___)

********** SCRAP **********

* mean_var1: the mean of variable1 by time.

* sd_var1: the standard deviation of variable1 by time.

* upper_var1: the upper bound of the standard deviation (mean + SD).

* lower_var1: the lower bound of the standard deviation (mean - SD).

```
twoway (scatter total_appr time_jitter if ow==0, msymbol(circle) lcolor(red)) (scatter total_appr time_jitter if ow==1, msymbol(circle) lcolor(blue)) (rcap lower_appr_TW upper_appr_TW time)          (rcap lower_appr_OW upper_appr_OW time), ysc(range(-0.3 0.3))
```

```
twoway (scatter total_appr time_jitter if ow==0, msymbol(circle) lcolor(red)) (scatter total_appr time_jitter if ow==1, msymbol(circle) lcolor(blue)), ysc(range(-0.3 0.3))
        title("Property Appreciation with Standard Deviation Bars")
        xlabel(, grid) ylabel(, grid)
        xtitle("Time") ytitle("Property Appreciation")
```

```
* Scatter plot w std. dev.
* Calculate the means and SDs across each time period
bysort time: egen mean_appr_OW = mean(total_appr) if ow==1
bysort time: egen sd_appr_OW = sd(total_appr) if ow==1

gen upper_appr_OW = mean_appr_OW + sd_appr_OW
gen lower_appr_OW = mean_appr_OW - sd_appr_OW


bysort time: egen mean_appr_TW = mean(total_appr) if ow==0
bysort time: egen sd_appr_TW = sd(total_appr) if ow==0

gen upper_appr_TW = mean_appr_TW + sd_appr_TW
gen lower_appr_TW = mean_appr_TW - sd_appr_TW

* zoom in without outliers
* remove outliers: A common rule is to define outliers as values that fall outside the range of 1.5 times the IQR above the upper quartile or below the lower quartile.
* Step 1: Calculate the 25th (Q1) and 75th (Q3) percentiles for each time period
bysort time: egen p25 = pctile(total_appr), p(25)
bysort time: egen p75 = pctile(total_appr), p(75)

* Step 2: Calculate the IQR (Interquartile Range) for each time period
bysort time: gen IQR = p75 - p25
```

```
* Step 3: Define the lower and upper bounds for outliers
bysort time: gen lower_bound = p25 - 1.5 * IQR
bysort time: gen upper_bound = p75 + 1.5 * IQR

* Step 4: Drop observations that fall outside the IQR bounds
bysort time: drop if total_appr < lower_bound | total_appr > upper_bound

* Step 5: Drop the temporary variables created for bounds
drop p25 p75 IQR lower_bound upper_bound
* prop_analysis_1.txt
```