Final project - Data Analysis: Data and Programming for Public Policy II

Cristian Bancayan, Sol Rivas Lopes & Claudia Felipe

2024-12-04

Set-up

```
Directory changed to: C:\Users\clfel\Documents\GitHub\Python
II\python_final_proj
```

Data cleaning and merging

Country-Aggregated Education Outcomes over Time

In this section, we create visualizations to compare the median values of key outcome variables over time between rural and urban areas for all Latin American countries with conditional cash transfer (CCT) programs, excluding Colombia and Argentina. This approach allows us to observe trends and differences across the region, providing insights into the potential impact of CCT programs. By focusing on median values, we minimize the influence of outliers and better capture central tendencies in the data.

```
alt.LayerChart(...)
```

```
alt.LayerChart(...)
```

```
alt.LayerChart(...)
```

Education outcomes by country and region type

In this section, we analyze the mean values of key outcome variables across Latin American countries with conditional cash transfer (CCT) programs. We calculate the mean for each variable, distinguishing between rural and urban areas, and grouping by the presence or absence of CCT programs. This analysis provides insights into the average impact of CCT programs at the country level.

```
alt.Chart(...)
```

```
alt.Chart(...)
```

```
alt.Chart(...)
```

Differential growth in years of education and enrollment pre- and post-CCT, per country

This section produces graphs showing the differential increse in education outcomes by country, disaggregated by region type.

```
alt.Chart(...)
```

```
alt.Chart(...)
```

```
alt.Chart(...)
```

T-test

```
Difference of Differences Results:
     Country           Variable  Rural Increment  Urban Increment  \
0     Brazil      years_edu_all         1.761709         1.908782
1     Brazil   enrollment6_12yo         9.443256         3.912245
2     Brazil  enrollment13_17yo        16.868877         8.069285
3      Chile      years_edu_all         2.026280         1.610190
4      Chile   enrollment6_12yo         4.831182         0.888070
5      Chile  enrollment13_17yo        22.189369         5.801917
6     Mexico      years_edu_all         1.505164         1.150179
7     Mexico   enrollment6_12yo         5.017859         1.750852
8     Mexico  enrollment13_17yo        19.877480         8.191788
9   Paraguay      years_edu_all         1.479305         1.570265
10  Paraguay   enrollment6_12yo         4.705052         2.149144
```

11	Paraguay	enrollment13_17yo	15.229327	6.619711
12	Peru	years_edu_all	0.887155	0.731799
13	Peru	enrollment6_12yo	3.103534	0.407910
14	Peru	enrollment13_17yo	12.192505	3.155340

	Difference of Differences	t-stat	p-value
0	0.147072	-0.603600	5.498954e-01
1	-5.531011	11.893812	1.052610e-11
2	-8.799592	7.011017	1.108841e-07
3	-0.416090	1.206698	2.487465e-01
4	-3.943112	15.954646	1.573648e-07
5	-16.387452	11.843053	7.910643e-07
6	-0.354985	1.456123	1.594543e-01
7	-3.267007	8.285836	6.281733e-08
8	-11.685692	5.756274	1.675278e-05
9	0.090960	-0.457260	6.502325e-01
10	-2.555909	5.723570	5.436598e-06
11	-8.609616	5.536294	1.171162e-05
12	-0.155355	1.514825	1.418201e-01
13	-2.695624	9.606900	1.139337e-10
14	-9.037165	6.439323	2.587994e-06

Education and quality of dwellings

```
alt.Chart(...)
```

Correlation calculation

	Country	Correlation
0	Brazil	-0.918320
1	Chile	-0.953974
2	Mexico	-0.974280
3	Peru	-0.356715
4	Paraguay	-0.469723

Quality of Dwellings post-CCT, graph

```
alt.LayerChart(...)
```

```
alt.LayerChart(...)

alt.LayerChart(...)

alt.LayerChart(...)

alt.LayerChart(...)
```

Regression Analysis

In this section, we perform a correlation analysis to explore the relationships between key variables and the implementation of conditional cash transfer (CCT) programs. We separately analyze rural and urban areas, focusing on variables related to education outcomes, infrastructure, and living conditions.

```
Rural data shape: (68, 12)
Urban data shape: (68, 12)

Correlations with CCT Active (Rural):
cct_active                    1.000000
year                          0.777479
enrollment3_5yo_rural         0.742998
enrollment6_12yo_rural        0.732965
enrollment13_17yo_rural       0.624852
hygienic_restrooms_rural      0.597938
water_rural                   0.544289
years_edu_all_rural           0.531501
sewerage_rural                0.512202
electricity_rural             0.499567
dwellings_low_quality_rural  -0.029018
Name: cct_active, dtype: float64

Correlations with CCT Active (Urban):
cct_active                    1.000000
year                          0.777479
enrollment3_5yo_urban         0.760834
enrollment6_12yo_urban        0.612219
years_edu_all_urban           0.578741
enrollment13_17yo_urban       0.433408
electricity_urban             0.431030
hygienic_restrooms_urban      0.418414
water_urban                   0.381919
```

```
sewerage_urban                  0.183597
dwellings_low_quality_urban     0.085650
Name: cct_active, dtype: float64
```

In this section, we conduct fixed effects regressions to examine the relationship between the implementation of conditional cash transfer (CCT) programs and key educational outcomes in rural and urban areas. The regressions are run separately for rural and urban datasets, allowing us to identify differences in the impact of CCT programs across these contexts. By using a fixed effects approach, we account for unobserved heterogeneity within countries over time, providing robust estimates of the effects of the CCT programs.

```
--- Fixed Effects Regressions for Rural Data ---

Fixed Effects Results for Years_edu_all (Rural):
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:     years_edu_all_rural   R-squared:
0.9098
Estimator:                    PanelOLS   R-squared (Between):
0.4816
No. Observations:                   68   R-squared (Within):
0.9098
Date:                 Sat, Dec 07 2024   R-squared (Overall):
0.6970
Time:                         16:15:03   Log-likelihood
-5.7676
Cov. Estimator:             Unadjusted
                                         F-statistic:
                                         117.05
Entities:                            5   P-value
0.0000
Avg Obs:                        13.600   Distribution:
F(5,58)
Min Obs:                        6.0000
Max Obs:                        22.000   F-statistic (robust):
117.05
                                         P-value
                                         0.0000
Time periods:                       31   Distribution:
F(5,58)
Avg Obs:                        2.1935
Min Obs:                        1.0000
```

```
Max Obs:                         4.0000

                                  Parameter Estimates
==========================================================================================
                          Parameter  Std. Err.     T-stat    P-value    Lower
                          CI    Upper CI
------------------------------------------------------------------------------------------
const                        2.9797     0.2734     10.899     0.0000
2.4325      3.5270
cct_active                   0.4756     0.1436     3.3112     0.0016
0.1881      0.7631
electricity_rural           -0.0031     0.0063    -0.4976     0.6207
-0.0157      0.0094
sewerage_rural               0.0134     0.0135     0.9890     0.3268
-0.0137      0.0405
hygienic_restrooms_rural     0.0227     0.0044     5.1333     0.0000
0.0138      0.0315
water_rural                  0.0118     0.0056     2.1030     0.0398
0.0006      0.0230
==========================================================================================

F-test for Poolability: 45.565
P-value: 0.0000
Distribution: F(4,58)

Included effects: Entity


Fixed Effects Results for Enrollment3_5yo (Rural):
                           PanelOLS Estimation Summary
================================================================================
Dep. Variable:     enrollment3_5yo_rural   R-squared:
0.9608
Estimator:                     PanelOLS   R-squared (Between):
-1.3688
No. Observations:                    68   R-squared (Within):
0.9608
Date:                  Sat, Dec 07 2024   R-squared (Overall):
0.4857
Time:                          16:15:03   Log-likelihood
-194.32
Cov. Estimator:              Unadjusted
```

```
                                        F-statistic:
                                        284.60
Entities:                            5  P-value
0.0000
Avg Obs:                        13.600  Distribution:
F(5,58)
Min Obs:                        6.0000
Max Obs:                        22.000  F-statistic (robust):
284.60
                                        P-value
                                        0.0000
Time periods:                       31  Distribution:
F(5,58)
Avg Obs:                        2.1935
Min Obs:                        1.0000
Max Obs:                        4.0000
```

Parameter Estimates

	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI
const	4.1462	4.3753	0.9476	0.3472	-4.6119	12.904
cct_active	3.1404	2.2986	1.3662	0.1771	-1.4607	7.7416
electricity_rural	0.0766	0.1004	0.7628	0.4487	-0.1244	0.2776
sewerage_rural	-0.2997	0.2168	-1.3821	0.1722	-0.7337	0.1343
hygienic_restrooms_rural	0.5900	0.0708	8.3384	0.0000	0.4484	0.7316
water_rural	0.4265	0.0898	4.7511	0.0000	0.2468	0.6062

```
F-test for Poolability: 96.378
P-value: 0.0000
Distribution: F(4,58)

Included effects: Entity
```

Fixed Effects Results for Enrollment6_12yo (Rural):

PanelOLS Estimation Summary

Dep. Variable:	enrollment6_12yo_rural	R-squared:	0.7146
Estimator:	PanelOLS	R-squared (Between):	-4.7992
No. Observations:	68	R-squared (Within):	0.7146
Date:	Sat, Dec 07 2024	R-squared (Overall):	-0.1047
Time:	16:15:03	Log-likelihood	-136.78
Cov. Estimator:	Unadjusted		
		F-statistic:	29.039
Entities:	5	P-value	0.0000
Avg Obs:	13.600	Distribution:	F(5,58)
Min Obs:	6.0000		
Max Obs:	22.000	F-statistic (robust):	29.039
		P-value	0.0000
Time periods:	31	Distribution:	F(5,58)
Avg Obs:	2.1935		
Min Obs:	1.0000		
Max Obs:	4.0000		

Parameter Estimates

	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI
const	85.192	1.8773	45.380	0.0000	81.434	88.949
cct_active	2.5326	0.9863	2.5679	0.0128	0.5584	4.5069
electricity_rural	0.1247	0.0431	2.8946	0.0053	0.0385	0.2109

```
sewerage_rural              -0.2903     0.0930    -3.1205     0.0028
-0.4765     -0.1041
hygienic_restrooms_rural     0.0450     0.0304     1.4838     0.1433
-0.0157      0.1058
water_rural                  0.0092     0.0385     0.2381     0.8126
-0.0679      0.0863
================================================================================================

F-test for Poolability: 4.6958
P-value: 0.0024
Distribution: F(4,58)

Included effects: Entity


Fixed Effects Results for Enrollment13_17yo (Rural):
                            PanelOLS Estimation Summary
================================================================================
Dep. Variable:     enrollment13_17yo_rural   R-squared:
0.8802
Estimator:                         PanelOLS   R-squared (Between):
-2.4370
No. Observations:                        68   R-squared (Within):
0.8802
Date:                     Sat, Dec 07 2024   R-squared (Overall):
0.1882
Time:                             16:15:03   Log-likelihood
-173.98
Cov. Estimator:                 Unadjusted
                                           F-statistic:
                                           85.245
Entities:                                5   P-value
0.0000
Avg Obs:                            13.600   Distribution:
F(5,58)
Min Obs:                            6.0000
Max Obs:                            22.000   F-statistic (robust):
85.245
                                           P-value
                                           0.0000
Time periods:                           31   Distribution:
F(5,58)
Avg Obs:                            2.1935
```

```
Min Obs:                          1.0000
Max Obs:                          4.0000

                                Parameter Estimates
==============================================================================================
                          Parameter  Std. Err.     T-stat    P-value    Lower
                          CI    Upper CI
----------------------------------------------------------------------------------------------
const                        50.407     3.2442     15.538     0.0000
43.913      56.901
cct_active                   4.7416     1.7044     2.7820     0.0073
1.3299      8.1532
electricity_rural            0.2643     0.0745     3.5503     0.0008
0.1153      0.4134
sewerage_rural              -0.1224     0.1608    -0.7613     0.4495
-0.4442      0.1994
hygienic_restrooms_rural     0.0841     0.0525     1.6037     0.1142
-0.0209      0.1892
water_rural                  0.0997     0.0666     1.4981     0.1395
-0.0335      0.2330
==============================================================================================

F-test for Poolability: 40.135
P-value: 0.0000
Distribution: F(4,58)

Included effects: Entity


Fixed Effects Results for Dwellings_low_quality (Rural):
                            PanelOLS Estimation Summary
================================================================================
Dep. Variable:     dwellings_low_quality_rural   R-squared:
0.5125
Estimator:                            PanelOLS   R-squared (Between):
-0.2084
No. Observations:                           68   R-squared (Within):
0.5125
Date:                         Sat, Dec 07 2024   R-squared (Overall):
-0.0070
Time:                                 16:15:03   Log-likelihood
-200.72
Cov. Estimator:                     Unadjusted
```

```
                                          F-statistic:
                                          12.195
Entities:                              5  P-value
0.0000
Avg Obs:                          13.600  Distribution:
F(5,58)
Min Obs:                          6.0000
Max Obs:                          22.000  F-statistic (robust):
12.195
                                          P-value
                                          0.0000
Time periods:                         31  Distribution:
F(5,58)
Avg Obs:                          2.1935
Min Obs:                          1.0000
Max Obs:                          4.0000

                                Parameter Estimates
==============================================================================================
                          Parameter  Std. Err.     T-stat    P-value    Lower
                          CI    Upper CI
----------------------------------------------------------------------------------------------
const                       24.074     4.8071     5.0080     0.0000
14.452      33.697
cct_active                 -1.8142     2.5255    -0.7183     0.4754
-6.8695      3.2412
electricity_rural           0.0901     0.1103     0.8170     0.4173
-0.1307      0.3110
sewerage_rural             -0.9414     0.2382    -3.9517     0.0002
-1.4183     -0.4645
hygienic_restrooms_rural   -0.2661     0.0777    -3.4236     0.0011
-0.4218     -0.1105
water_rural                 0.2259     0.0986     2.2909     0.0256
0.0285      0.4234
==============================================================================================

F-test for Poolability: 40.249
P-value: 0.0000
Distribution: F(4,58)

Included effects: Entity
```

```
--- Fixed Effects Regressions for Urban Data ---

Fixed Effects Results for Years_edu_all (Urban):
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:     years_edu_all_urban   R-squared:
0.9014
Estimator:                    PanelOLS   R-squared (Between):
0.1383
No. Observations:                   68   R-squared (Within):
0.9014
Date:                 Sat, Dec 07 2024   R-squared (Overall):
0.6426
Time:                         16:15:03   Log-likelihood
-3.7586
Cov. Estimator:             Unadjusted
                                         F-statistic:
                                         106.00
Entities:                            5   P-value
0.0000
Avg Obs:                        13.600   Distribution:
F(5,58)
Min Obs:                        6.0000
Max Obs:                        22.000   F-statistic (robust):
106.00
                                         P-value
                                         0.0000
Time periods:                       31   Distribution:
F(5,58)
Avg Obs:                        2.1935
Min Obs:                        1.0000
Max Obs:                        4.0000

                                Parameter Estimates
================================================================================
                          Parameter  Std. Err.     T-stat    P-value    Lower
                          CI    Upper CI
--------------------------------------------------------------------------------
const                       1.0687     2.9380     0.3638     0.7174
-4.8123      6.9498
cct_active                  0.4663     0.1252     3.7254     0.0004
0.2158      0.7169
```

```
electricity_urban           -0.0321     0.0374    -0.8574     0.3947
-0.1070      0.0428
sewerage_urban               0.0113     0.0164     0.6866     0.4951
-0.0215      0.0440
hygienic_restrooms_urban     0.0935     0.0150     6.2416     0.0000
0.0635      0.1235
water_urban                  0.0049     0.0197     0.2493     0.8040
-0.0345      0.0443
================================================================================

F-test for Poolability: 80.829
P-value: 0.0000
Distribution: F(4,58)

Included effects: Entity


Fixed Effects Results for Enrollment3_5yo (Urban):
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:     enrollment3_5yo_urban   R-squared:
0.9198
Estimator:                     PanelOLS   R-squared (Between):
-35.779
No. Observations:                    68   R-squared (Within):
0.9198
Date:                  Sat, Dec 07 2024   R-squared (Overall):
-1.3519
Time:                          16:15:03   Log-likelihood
-202.13
Cov. Estimator:              Unadjusted
                                          F-statistic:
                                          132.98
Entities:                             5   P-value
0.0000
Avg Obs:                         13.600   Distribution:
F(5,58)
Min Obs:                         6.0000
Max Obs:                         22.000   F-statistic (robust):
132.98
                                          P-value
                                          0.0000
```

```
Time periods:                          31   Distribution:
F(5,58)
Avg Obs:                           2.1935
Min Obs:                           1.0000
Max Obs:                           4.0000

                                Parameter Estimates
==============================================================================================
                          Parameter  Std. Err.     T-stat    P-value    Lower
                          CI    Upper CI
----------------------------------------------------------------------------------------------
const                       -110.85     54.327    -2.0403     0.0459
-219.59     -2.0981
cct_active                   9.8471     2.3146     4.2544     0.0001
5.2139      14.480
electricity_urban            0.2504     0.6919     0.3618     0.7188
-1.1347      1.6354
sewerage_urban              -1.0924     0.3030    -3.6056     0.0006
-1.6989     -0.4859
hygienic_restrooms_urban     2.6814     0.2771     9.6777     0.0000
2.1268      3.2360
water_urban                 -0.0525     0.3642    -0.1440     0.8860
-0.7815      0.6766
==============================================================================================

F-test for Poolability: 52.119
P-value: 0.0000
Distribution: F(4,58)

Included effects: Entity


Fixed Effects Results for Enrollment6_12yo (Urban):
                           PanelOLS Estimation Summary
================================================================================
Dep. Variable:     enrollment6_12yo_urban   R-squared:
0.6722
Estimator:                       PanelOLS   R-squared (Between):
-0.3724
No. Observations:                      68   R-squared (Within):
0.6722
Date:                   Sat, Dec 07 2024   R-squared (Overall):
0.5563
```

```
Time:                           16:15:03   Log-likelihood
-78.762
Cov. Estimator:               Unadjusted
                                           F-statistic:
                                           23.785
Entities:                            5     P-value
0.0000
Avg Obs:                          13.600   Distribution:
F(5,58)
Min Obs:                          6.0000
Max Obs:                          22.000   F-statistic (robust):
23.785
                                           P-value
                                           0.0000
Time periods:                        31     Distribution:
F(5,58)
Avg Obs:                          2.1935
Min Obs:                          1.0000
Max Obs:                          4.0000
```

Parameter Estimates

	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI
const	97.130	8.8527	10.972	0.0000	79.409	114.85
cct_active	0.7179	0.3772	1.9034	0.0620	-0.0371	1.4729
electricity_urban	-0.1628	0.1127	-1.4442	0.1541	-0.3885	0.0629
sewerage_urban	0.0146	0.0494	0.2965	0.7679	-0.0842	0.1135
hygienic_restrooms_urban	0.1150	0.0451	2.5481	0.0135	0.0247	0.2054
water_urban	0.0612	0.0593	1.0309	0.3069	-0.0576	0.1800

```
F-test for Poolability: 6.4380
P-value: 0.0002
Distribution: F(4,58)
```

```
Included effects: Entity


Fixed Effects Results for Enrollment13_17yo (Urban):
                            PanelOLS Estimation Summary
================================================================================
Dep. Variable:     enrollment13_17yo_urban   R-squared:
0.8383
Estimator:                         PanelOLS   R-squared (Between):
0.0215
No. Observations:                        68   R-squared (Within):
0.8383
Date:                     Sat, Dec 07 2024   R-squared (Overall):
0.4975
Time:                             16:15:03   Log-likelihood
-122.00
Cov. Estimator:                 Unadjusted
                                           F-statistic:
                                           60.149
Entities:                                5   P-value
0.0000
Avg Obs:                            13.600   Distribution:
F(5,58)
Min Obs:                            6.0000
Max Obs:                            22.000   F-statistic (robust):
60.149
                                           P-value
                                           0.0000
Time periods:                           31   Distribution:
F(5,58)
Avg Obs:                            2.1935
Min Obs:                            1.0000
Max Obs:                            4.0000

                                Parameter Estimates
================================================================================
                          Parameter  Std. Err.     T-stat    P-value    Lower
                          CI    Upper CI
--------------------------------------------------------------------------------
const                        61.645     16.719     3.6872     0.0005
28.179      95.112
cct_active                   1.5865     0.7123     2.2273     0.0298
0.1607      3.0123
```

```
electricity_urban           -0.2879     0.2129    -1.3521     0.1816
-0.7141      0.1383
sewerage_urban               0.1404     0.0932     1.5062     0.1374
-0.0462      0.3271
hygienic_restrooms_urban     0.2251     0.0853     2.6394     0.0106
0.0544      0.3957
water_urban                  0.2936     0.1121     2.6198     0.0112
0.0693      0.5180
================================================================================

F-test for Poolability: 14.957
P-value: 0.0000
Distribution: F(4,58)

Included effects: Entity


Fixed Effects Results for Dwellings_low_quality (Urban):
                            PanelOLS Estimation Summary
================================================================================
Dep. Variable:     dwellings_low_quality_urban   R-squared:
0.5238
Estimator:                            PanelOLS   R-squared (Between):
-0.5048
No. Observations:                           68   R-squared (Within):
0.5238
Date:                         Sat, Dec 07 2024   R-squared (Overall):
-0.5051
Time:                                 16:15:03   Log-likelihood
-137.97
Cov. Estimator:                     Unadjusted
                                               F-statistic:
                                               12.758
Entities:                                    5   P-value
0.0000
Avg Obs:                                13.600   Distribution:
F(5,58)
Min Obs:                                6.0000
Max Obs:                                22.000   F-statistic (robust):
12.758
                                               P-value
                                               0.0000
```

```
Time periods:                                31   Distribution:
F(5,58)
Avg Obs:                                 2.1935
Min Obs:                                 1.0000
Max Obs:                                 4.0000

                                Parameter Estimates
==============================================================================================
                           Parameter  Std. Err.     T-stat    P-value    Lower
                           CI    Upper CI
----------------------------------------------------------------------------------------------
const                        58.854     21.145     2.7833     0.0073
16.528      101.18
cct_active                  -0.4887     0.9009    -0.5425     0.5896
-2.2920      1.3146
electricity_urban           -0.9025     0.2693    -3.3513     0.0014
-1.4416     -0.3635
sewerage_urban              -0.1440     0.1179    -1.2213     0.2269
-0.3801      0.0920
hygienic_restrooms_urban    -0.4017     0.1078    -3.7254     0.0004
-0.6176     -0.1859
water_urban                  0.9095     0.1418     6.4161     0.0000
0.6258      1.1933
==============================================================================================

F-test for Poolability: 62.276
P-value: 0.0000
Distribution: F(4,58)

Included effects: Entity
```

Dif in Dif

```
--- Difference-in-Differences Analysis for Rural Data ---

DiD Results for years_edu_all (Rural):
                            OLS Regression Results
==============================================================================
```

```
Dep. Variable:     years_edu_all_rural   R-squared:
0.424
Model:                             OLS   Adj. R-squared:
0.407
Method:                  Least Squares   F-statistic:
23.96
Date:                 Sat, 07 Dec 2024   Prob (F-statistic):
1.60e-08
Time:                         16:15:04   Log-Likelihood:
-95.836
No. Observations:                   68   AIC:
197.7
Df Residuals:                       65   BIC:
204.3
Df Model:                            2
Covariance Type:             nonrobust
==============================================================================================
                     coef    std err          t      P>|t|      [0.025
                     0.975]
----------------------------------------------------------------------------------------------
Intercept           3.6051      0.239     15.098      0.000       3.128
4.082
cct_active          0.7369      0.348      2.117      0.038       0.042
1.432
post                0.6146      0.154      4.002      0.000       0.308
0.921
cct_active:post     0.6146      0.154      4.002      0.000       0.308
0.921
==============================================================================
Omnibus:                        5.892   Durbin-Watson:
0.252
Prob(Omnibus):                  0.053   Jarque-Bera (JB):
5.948
Skew:                           0.717   Prob(JB):
0.0511
Kurtosis:                       2.787   Cond. No.
2.11e+16
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is
correctly specified.
[2] The smallest eigenvalue is 3.58e-31. This might indicate that there are
```

strong multicollinearity problems or that the design matrix is singular.

```
DiD Results for enrollment6_12yo (Rural):
                            OLS Regression Results
==============================================================================
Dep. Variable:     enrollment6_12yo_rural   R-squared:
0.591
Model:                                OLS   Adj. R-squared:
0.578
Method:                     Least Squares   F-statistic:
46.88
Date:                    Sat, 07 Dec 2024   Prob (F-statistic):
2.48e-13
Time:                            16:15:04   Log-Likelihood:
-155.92
No. Observations:                      68   AIC:
317.8
Df Residuals:                          65   BIC:
324.5
Df Model:                               2
Covariance Type:                nonrobust
==============================================================================
                      coef    std err          t      P>|t|      [0.025
                      0.975]
------------------------------------------------------------------------------
Intercept          91.7915      0.578    158.872      0.000      90.638
92.945
cct_active          4.7520      0.842      5.642      0.000       3.070
6.434
post                1.0814      0.372      2.910      0.005       0.339
1.824
cct_active:post     1.0814      0.372      2.910      0.005       0.339
1.824
==============================================================================
Omnibus:                       52.189   Durbin-Watson:
0.333
Prob(Omnibus):                  0.000   Jarque-Bera (JB):
249.012
Skew:                          -2.218   Prob(JB):
8.47e-55
Kurtosis:                      11.259   Cond. No.
2.11e+16
```

```
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is
correctly specified.
[2] The smallest eigenvalue is 3.58e-31. This might indicate that there are
strong multicollinearity problems or that the design matrix is singular.


DiD Results for enrollment13_17yo (Rural):
                            OLS Regression Results
==============================================================================
Dep. Variable:     enrollment13_17yo_rural   R-squared:
0.428
Model:                                 OLS   Adj. R-squared:
0.411
Method:                      Least Squares   F-statistic:
24.36
Date:                     Sat, 07 Dec 2024   Prob (F-statistic):
1.28e-08
Time:                             16:15:04   Log-Likelihood:
-235.86
No. Observations:                       68   AIC:
477.7
Df Residuals:                           65   BIC:
484.4
Df Model:                                2
Covariance Type:                 nonrobust
==============================================================================
                     coef    std err          t      P>|t|      [0.025
                     0.975]
------------------------------------------------------------------------------
Intercept          71.6476      1.872     38.274      0.000      67.909
75.386
cct_active         11.1457      2.729      4.084      0.000       5.696
16.596
post                2.5004      1.204      2.077      0.042       0.096
4.905
cct_active:post     2.5004      1.204      2.077      0.042       0.096
4.905
==============================================================================
Omnibus:                        7.695   Durbin-Watson:
0.427
```

```
Prob(Omnibus):                  0.021   Jarque-Bera (JB):
8.001
Skew:                          -0.838   Prob(JB):
0.0183
Kurtosis:                       2.876   Cond. No.
2.11e+16
================================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is
correctly specified.
[2] The smallest eigenvalue is 3.58e-31. This might indicate that there are
strong multicollinearity problems or that the design matrix is singular.


DiD Results for dwellings_low_quality (Rural):
                            OLS Regression Results
==============================================================================================
Dep. Variable:     dwellings_low_quality_rural   R-squared:
0.003
Model:                                     OLS   Adj. R-squared:
-0.027
Method:                          Least Squares   F-statistic:
0.1076
Date:                         Sat, 07 Dec 2024   Prob (F-statistic):
0.898
Time:                                 16:15:04   Log-Likelihood:
-280.47
No. Observations:                           68   AIC:
566.9
Df Residuals:                               65   BIC:
573.6
Df Model:                                    2
Covariance Type:                     nonrobust
==============================================================================
                     coef    std err          t      P>|t|      [0.025
                     0.975]
------------------------------------------------------------------------------
Intercept          22.3031      3.607      6.182      0.000      15.098
29.508
cct_active         -2.2491      5.259     -0.428      0.670     -12.752
8.253
```

```
post                0.9289      2.320      0.400      0.690      -3.705
5.562
cct_active:post     0.9289      2.320      0.400      0.690      -3.705
5.562
==============================================================================
Omnibus:                       24.928   Durbin-Watson:
0.503
Prob(Omnibus):                  0.000   Jarque-Bera (JB):
35.031
Skew:                           1.588   Prob(JB):
2.47e-08
Kurtosis:                       4.511   Cond. No.
2.11e+16
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is
correctly specified.
[2] The smallest eigenvalue is 3.58e-31. This might indicate that there are
strong multicollinearity problems or that the design matrix is singular.



--- Difference-in-Differences Analysis for Urban Data ---

DiD Results for years_edu_all (Urban):
                            OLS Regression Results
==============================================================================
Dep. Variable:     years_edu_all_urban   R-squared:
0.464
Model:                             OLS   Adj. R-squared:
0.447
Method:                  Least Squares   F-statistic:
28.08
Date:                 Sat, 07 Dec 2024   Prob (F-statistic):
1.62e-09
Time:                         16:15:04   Log-Likelihood:
-89.282
No. Observations:                   68   AIC:
184.6
Df Residuals:                       65   BIC:
191.2
Df Model:                            2
```

```
Covariance Type:             nonrobust
==============================================================================
                     coef    std err          t      P>|t|      [0.025
                     0.975]
------------------------------------------------------------------------------
Intercept           6.4123      0.217     29.572      0.000       5.979
6.845
cct_active          0.8623      0.316      2.728      0.008       0.231
1.494
post                0.5505      0.139      3.947      0.000       0.272
0.829
cct_active:post     0.5505      0.139      3.947      0.000       0.272
0.829
==============================================================================
Omnibus:                        4.782   Durbin-Watson:
0.231
Prob(Omnibus):                  0.092   Jarque-Bera (JB):
2.189
Skew:                          -0.065   Prob(JB):
0.335
Kurtosis:                       2.131   Cond. No.
2.11e+16
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is
correctly specified.
[2] The smallest eigenvalue is 3.58e-31. This might indicate that there are
strong multicollinearity problems or that the design matrix is singular.


DiD Results for enrollment6_12yo (Urban):
                            OLS Regression Results
==============================================================================
Dep. Variable:     enrollment6_12yo_urban   R-squared:
0.421
Model:                                OLS   Adj. R-squared:
0.403
Method:                     Least Squares   F-statistic:
23.59
Date:                    Sat, 07 Dec 2024   Prob (F-statistic):
1.98e-08
```

```
Time:                            16:15:04   Log-Likelihood:
-108.38
No. Observations:                      68   AIC:
222.8
Df Residuals:                          65   BIC:
229.4
Df Model:                               2
Covariance Type:                nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025
                 0.975]
------------------------------------------------------------------------------
Intercept          96.6459      0.287    336.579      0.000      96.072
97.219
cct_active          1.6024      0.419      3.828      0.000       0.766
2.438
post                0.4184      0.185      2.265      0.027       0.050
0.787
cct_active:post     0.4184      0.185      2.265      0.027       0.050
0.787
==============================================================================
Omnibus:                       20.488   Durbin-Watson:
0.340
Prob(Omnibus):                  0.000   Jarque-Bera (JB):
38.159
Skew:                          -1.015   Prob(JB):
5.17e-09
Kurtosis:                       6.057   Cond. No.
2.11e+16
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is
correctly specified.
[2] The smallest eigenvalue is 3.58e-31. This might indicate that there are
strong multicollinearity problems or that the design matrix is singular.


DiD Results for enrollment13_17yo (Urban):
                            OLS Regression Results
==============================================================================
Dep. Variable:     enrollment13_17yo_urban   R-squared:
0.217
```

```
Model:                                OLS   Adj. R-squared:
0.193
Method:                     Least Squares   F-statistic:
8.993
Date:                    Sat, 07 Dec 2024   Prob (F-statistic):
0.000356
Time:                            16:15:04   Log-Likelihood:
-199.92
No. Observations:                      68   AIC:
405.8
Df Residuals:                          65   BIC:
412.5
Df Model:                               2
Covariance Type:                nonrobust
================================================================================
                     coef    std err          t      P>|t|      [0.025
                     0.975]
--------------------------------------------------------------------------------
Intercept          87.4490      1.103     79.248      0.000      85.245
89.653
cct_active          3.5861      1.609      2.229      0.029       0.374
6.799
post                1.0991      0.710      1.549      0.126      -0.318
2.516
cct_active:post     1.0991      0.710      1.549      0.126      -0.318
2.516
==============================================================================
Omnibus:                       15.691   Durbin-Watson:
0.315
Prob(Omnibus):                  0.000   Jarque-Bera (JB):
18.055
Skew:                          -1.219   Prob(JB):
0.000120
Kurtosis:                       3.657   Cond. No.
2.11e+16
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is
correctly specified.
[2] The smallest eigenvalue is 3.58e-31. This might indicate that there are
strong multicollinearity problems or that the design matrix is singular.
```

```
DiD Results for dwellings_low_quality (Urban):
                            OLS Regression Results
==============================================================================
Dep. Variable:     dwellings_low_quality_urban   R-squared:
0.021
Model:                                     OLS   Adj. R-squared:
-0.009
Method:                          Least Squares   F-statistic:
0.7122
Date:                         Sat, 07 Dec 2024   Prob (F-statistic):
0.494
Time:                                 16:15:04   Log-Likelihood:
-218.67
No. Observations:                           68   AIC:
443.3
Df Residuals:                               65   BIC:
450.0
Df Model:                                    2
Covariance Type:                     nonrobust
==============================================================================
                     coef    std err          t      P>|t|      [0.025
                     0.975]
------------------------------------------------------------------------------
Intercept           7.3933      1.454      5.085      0.000       4.490
10.297
cct_active         -0.0474      2.119     -0.022      0.982      -4.280
4.185
post                0.9051      0.935      0.968      0.337      -0.962
2.772
cct_active:post     0.9051      0.935      0.968      0.337      -0.962
2.772
==============================================================================
Omnibus:                        6.561   Durbin-Watson:
0.542
Prob(Omnibus):                  0.038   Jarque-Bera (JB):
6.571
Skew:                           0.760   Prob(JB):
0.0374
Kurtosis:                       2.914   Cond. No.
2.11e+16
==============================================================================
```

```
Notes:
[1] Standard Errors assume that the covariance matrix of the errors is
correctly specified.
[2] The smallest eigenvalue is 3.58e-31. This might indicate that there are
strong multicollinearity problems or that the design matrix is singular.
```

Final DiD Results:

	Outcome	Interaction_Coeff	p-value
0	years_edu_all_rural	0.614602	0.000163
1	enrollment6_12yo_rural	1.081422	0.004939
2	enrollment13_17yo_rural	2.500356	0.041772
3	dwellings_low_quality_rural	0.928885	0.690192
0	years_edu_all_urban	0.550484	0.000197
1	enrollment6_12yo_urban	0.418350	0.026825
2	enrollment13_17yo_urban	1.099080	0.126308
3	dwellings_low_quality_urban	0.905113	0.336606