April 2, 2021

Replication Package for

"Local Effects of Large New Apartment Buildings in Low-Income Areas"

Brian Asquith, Evan Mast, and Davin Reed

Preliminaries

This document describes the steps necessary to replicate the results in the paper "Local Effects of Large New Apartment Buildings in Low-Income Areas." All figures and tables in the paper use proprietary data that are not publicly available. In Section I, we describe this data and offer instructions for how other researchers can obtain it. In Section II, we describe the code that processes the RCA data on new buildings into the form used in the paper. In Section III, we describe the code that processes the raw Zillow data into the form used in the paper. In Section IV, we do the same for the Infutor data. In Section V, we describe the code that produces the results in the paper. The vast majority of programs are in Stata (v16), and one appendix exercise is in R and one in Matlab.

I. Data Documentation and Instructions for Access

The main data sets in this paper are provided by Zillow (rent information), Infutor Data Solutions (migration information), and Real Capital Analytics (new building information). All are proprietary and cannot be shared without permission from the provider. Both the Infutor and Real Capital Analytics data are easily available for purchase and interested researchers can contact sales representatives of those companies for pricing and other information. The email address for Infutor sales is sales@infutor.com , and RCA inquiries can be submitted at https://www.rcanalytics.com/request-info/ . The Zillow data is not available for purchase. We obtained access through a research partnership. Interested researcher can contact Zillow researchers or representatives and inquire about possible partnerships. Inquiries can be submitted at https://www.zillow.com/research/contact-us/ .

II. Creating Sample of Large New Buildings

This section describes how we process the building data from RCA and neighborhood data from the Census and American Community Survey. The code first joins the neighborhood data across years to create a longitudinal file. It then cleans the RCA data on new buildings and joins it to the neighborhood data.

1) Building_processing_directory.do calls the programs that execute these steps.
2) Building_processing_macros.do contains the underlying programs.

III. Creating Rent Analysis File

This section describes how we combine the rent data from Zillow with the new buildings data from RCA to create the analysis files for the rent results. The code first pairs rental listings with nearby buildings, measures the distance between, and appends the new building characteristics. It then uses that master data set to construct the sample for each specific analysis (e.g. near-far comparison, near-near comparison, et cetera).

3) Rent_processing_directory.do calls the programs that execute these steps.
4) Rent_processing_macros.do contains the underlying programs.

IV. Creating Migration Analysis File

This section describes how we combine the migration data from Infutor with the new buildings data from RCA data to create the analysis files for the migration results. The code first uses string matching techniques to identify the RCA new buildings in the Infutor data. It then pulls all moves that are close to (or directly to) one of the new buildings and flags the moves that are directly to a new building. Characteristics of new buildings are appended to this master file. The code then uses that master data set to construct the sample for each specific analysis (e.g. near-far comparison, near-near comparison, et cetera).

1) Migration_processing_directory.do calls the programs that execute these steps.
2) Migration_processing_macros.do contains the underlying programs.

V. Main Empirical Results

The code in this section replicates the results in Tables 1-5 as well as Figures 1-4.

1) Figure 1: figure_1.do
    a. This file uses the migration data samples created in Section IV to make the plot of in-migration to the area near new buildings shown in Figure 1 of the paper.
2) Figure 2: figure_2.do
    a. This file uses the rent data samples created in Section III to make the event study plot in Figure 2 of the paper.
3) Figure 3 simply plots estimates from other tables and figures. Each estimate can be created using the corresponding do-file. The coefplot package was used to generate the figure.
4) Figure 4: file.do
    a. This file uses the migration data samples created in Section IV to make the event study plot in Figure 4 of the paper.
5) Table 1: table_1.do
    a. This file uses the new building data created in Section II to make the summary statistics shown in Table 1 of the paper.
6) Table 2: table_2.do
    a. This file uses the new building data created in Section II to make the summary statistics shown in Table 2 of the paper.
7) Table 3: table_3.do
    a. This file uses the rent data samples created in Section III to generate the estimates shown in Table 3 of the paper.
8) Table 4: table_4.do
    a. This file uses the rent samples created in Section III to generate the estimates shown in Table 4 of the paper.
9) Table 5: table_5.do
    a. This file uses the migration samples created in Section IV to generate the estimates shown in Table 5 of the paper.

VI. Appendix Empirical Results

The code in this section replicates the results in the appendix

1) Appendix Figure 1: appendix_figure_1.do
    a. This figure plots the Zillow listings in Chicago in 2018
2) Appendix Figure 2: appendix_figure_2.do
    a. This figure plots selection into the Infutor data
3) Appendix Figure 3: appendix_figure_3.do
    a. This figure compares move rates in the Infutor data and Census data
4) Appendix Figure 4: appendix_figure_4.do
    a. This figure shows net low-income migration around new buildings
5) Appendix Figure 5: appendix_figure_5.do
    a. This figure plots the near-far sample in two cities
6) Appendix Figure 6: appendix_figure_6.do
    a. This figure plots the near-far and near-near event studies with a 400m treatment radius
7) Appendix Figure 7: appendix_figure_7.R
    a. This figure plots the spatial competition exercise described in Section 3.2.
8) Appendix Figure 8: appendix_figure_8.M
    a. This figure computes the empirical derivative figure.
9) Appendix Figure 9: appendix_figure_9.do
    a. This figure plots the near-far and near-near event studies without the pioneer restriction
10) Appendix Figure 10: appendix_figure_10.do
    a. This figure plots the long-run near-far effect for buildings constructed before the baseline sample
11) Appendix Figure 11: appendix_figure_11.do
    a. This figure plots long-run near-far pre-trends for buildings completed in 2019
12) Appendix Figure 12: appendix_figure_12.M
    a. This figure uses the empirical derivative methodology to compute long-run pre-trends near the sites of 2019 buildings
13) Appendix Figure 13: appendix_figure_13.do
    a. This figure shows near-far rent event studies for buildings in neighborhoods of all incomes
14) Appendix Figure 14: appendix_figure_14.do
    a. This figure shows near-near rent event studies for buildings in neighborhoods of all incomes
15) Appendix Figure 15: appendix_figure_15.do
    a. This figure shows near-far and near-near rent event studies for buildings in only high-income neighborhoods
16) Appendix Figure 16: appendix_figure_16.do
    a. This figure shows near-near migration event studies with a 400m treatment radius.
17) Appendix Table 1: appendix_table_1.do
    a. This table shows mean rent in each CBSA x year

18) Appendix Table 2: appendix_table_2.do
    a. This table shows mean one-bedroom rent in each CBSA x year
19) Appendix Table 3: appendix_table_3.do
    a. This table compares the neighborhoods around the treatment and control buildings of the near-near specification.
20) Appendix Table 4: appendix_table_4.do
    a. This table shows diff-in-diff results with unit characteristics as the dependent variable
21) Appendix Table 5: appendix_table_5.do
    a. This table shows rent diff-in-diff results with a 400m treatment radius
22) Appendix Table 6: appendix_table_6.do
    a. This table shows the “far-far” diff-in-diff for rent outcomes
23) Appendix Table 7: appendix_table_7.do
    a. This table shows robustness of the near-far rent results
24) Appendix Table 8: appendix_table_8.do
    a. This table shows robustness of the near-near rent results
25) Appendix Table 9: appendix_table_9.do
    a. This table shows rent results with a 400m treatment radius and buildings in neighborhoods at all income levels
26) Appendix Table 10: appendix_table_10.do
    a. This table shows robustness of the near-near migration results for the dependent variable of log(origin income)
27) Appendix Table 11: appendix_table_11.do
    a. This table shows robustness of the near-near migration results for the dependent variable of 1(low-income arrival)
28) Appendix Table 12: appendix_table_12.do
    a. This table shows the far-far migration results for the dependent variable of log(origin income)