

# Online Appendix: Cracking Down, Pricing Up: Housing Supply in the Wake of Mass Deportation

Troup Howard\*   Mengqi Wang<sup>†</sup>   Dayin Zhang<sup>‡</sup>

October 2024

---

\*University of Utah. Email: [troup.howard@eccles.utah.edu](mailto:troup.howard@eccles.utah.edu).

<sup>†</sup>Amherst College. Email: [mewang@amherst.edu](mailto:mewang@amherst.edu).

<sup>‡</sup>University of Wisconsin-Madison. Email: [dayin.zhang@wisc.edu](mailto:dayin.zhang@wisc.edu).

# A Sample Construction: CoreLogic Microdata

All quantity results for new construction are based on 4.22 million records of new construction transactions, representing nearly the entire universe of residential constructions between 2005 and 2012. Price results for new construction are based on a subset of 2.07 million records that contain fully populated attribute data, allowing us to adequately adjust for observable quality. Resale house price results are based on 10.91 million records of resales with a full set of house characteristics. This section details each step of the data construction process.

## A.1 County-Level New Construction Quantity Data

We assemble data on all new residential construction completed in local markets from CoreLogic deeds data. We start with all transaction records flagged as new construction and with a built year between 2005 and 2012. Notably, this includes records with a sales year through 2022. We also permit the recorded sales year to be one year before the built year, to include preconstruction sales. We remove nonresidential properties by restricting to property types of single-family homes, condos, and duplexes. We also remove approximately 100k records that despite being flagged as new construction have a prior transaction associated with the same property PIN. Geographically, we keep only properties which we can match to a county launch date for Secure Communities. Practically, this means excluding US Territories and nine small counties with irregular governance structures. Finally, we drop properties with abnormal square footage.<sup>1</sup> In total, this yields 4.22 million records. This is the core dataset for all quantity results.

## A.2 New Construction Price Data

For tests of new construction prices, we impose five further filters. Figure S1 demonstrates the overview of the data construction process. First, we focus on single family constructions (3.39 million). Second, to ensure an accurate signal about market value, we restrict attention to arms-length transactions using general deeds, which removes 83k observations. We also remove 37k observations with a second transaction occurring within 60 days of the focal transaction, to avoid inadvertently including partial-interest transactions. Third, we remove 142k transactions with sales price greater than \$5M or less than \$100,000, as well as 1k properties with abnormal numbers of bedrooms or bathrooms. Fourth, we remove 114k transactions which sell more than 10 years after their built year for two reasons: (i) at

---

<sup>1</sup> We drop properties with square footage above 99,999 or below 0.

this age, it’s no longer clear that such a transaction accurately reflects “new construction” pricing, and (ii) as such delays are unusual, we are concerned about large property-specific unobservables. Our fifth filter concerns the availability of hedonic attributes. We use these attributes to provide both direct and indirect tests of endogenous shifts in building characteristics. As usual in real estate microdata, there is only partial coverage of structural characteristics. We focus on standard hedonic attributes in the real estate literature: square footage, age, bedrooms, bathrooms, and census tract. Property attributes are contained in CoreLogic tax assessor files; each observation in the deeds data contains a unique PIN that permits straightforward linking to the assessor data. Because it is quite common for assessors to receive building information at a meaningful lag, we allow a look-ahead period of five years. That is: for a property transacting in year  $t$ , we take the first set of fully populated hedonic controls in any year between  $t$  and  $t + 5$ . After this, for records missing attribute data, we match to MLS records and augment them with listing characteristics. This allows us to recover attributes for additional new construction observations.

In total, within the subset of 3.01 million observations that pass our first four price filters, we have a full set of hedonic attributes for 2.07 million records (69%). To hold the sample constant between regressions that do and do not use quality controls, we use the set of 2.07 million observations as our core dataset for price regressions.

### **A.3 Resale House Price Data**

For tests of resale house prices, we collect 57.85 million residential resale transactions from 2005 to 2012, and apply the following five filters to get the core price regression data. Figure S2 demonstrates the overview of the data construction process. First, we focus on single family constructions (49.73 million), to be comparable with new construction price results. Second, to ensure an accurate signal about market value, we restrict attention to arms-length transactions using general deeds, which removes 21.65 million observations. We also remove 3.27 million observations with a second transaction occurring 60 days before or after the focal transaction, to avoid inadvertently including partial-interest transactions. Third, we remove “flip” sales, where property owners aggressively change house characteristics over a short period to increase home value. These rapid changes can lead to less accurate house characteristic information in the assessment records. We designate a transaction as likely to be a flip sale if the seller purchased the property less than two years prior to the sale, resulting in the removal of 1.69 million observations. Fourth, we keep transactions with sales price greater than \$5M or less than \$100,000. We also remove properties with abnormal characteristics: square footage equal to or larger than 99,999; or number of bedrooms greater

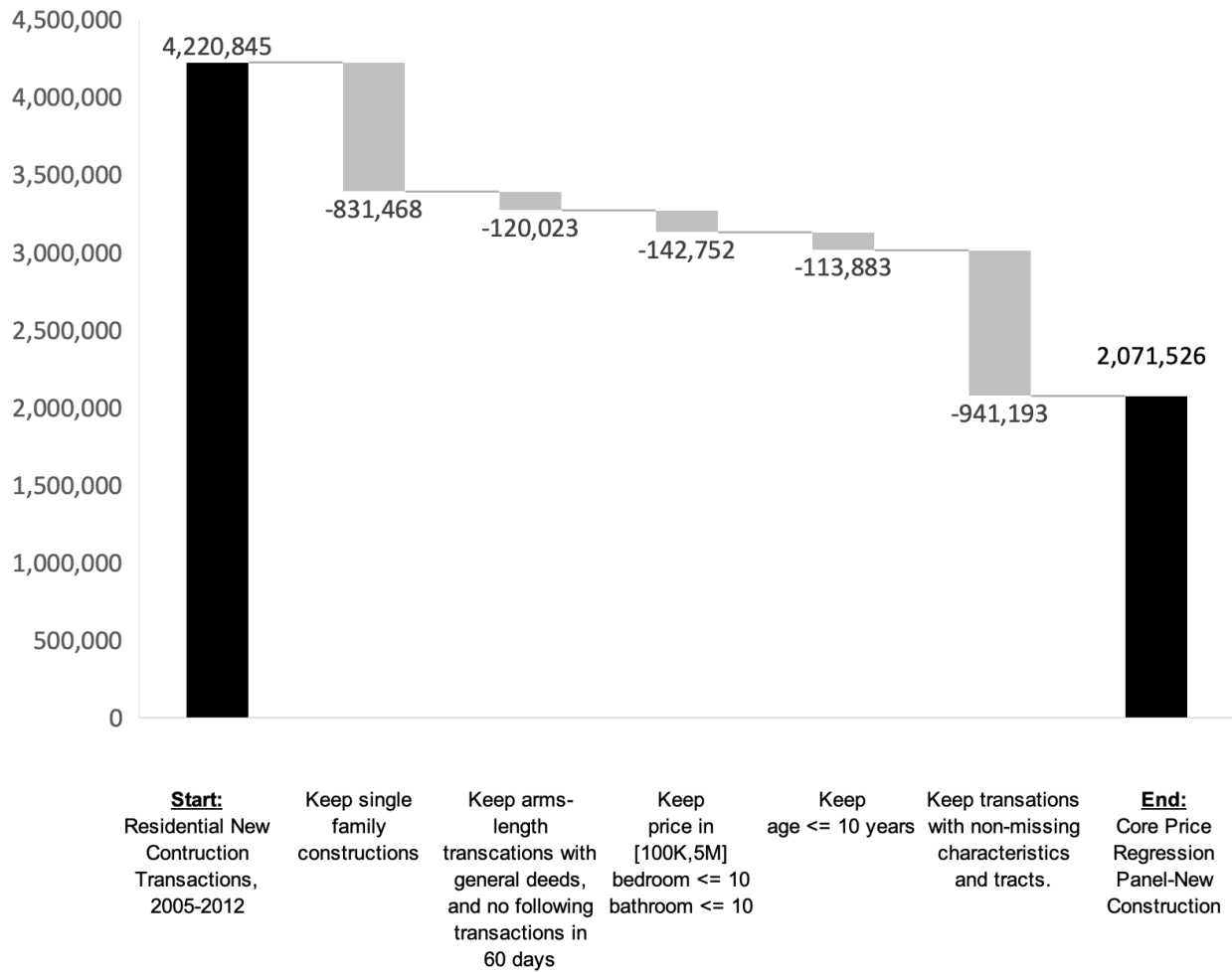
than 10; or number of bathrooms greater than 10; or more than 200 years old.

Finally, we adopt the same approach to filtering on the availability of hedonic attributes as described above in [A.2](#). In total, we have a full set of hedonic attributes for 10.91 million transactions which form our core dataset for existing-stock price regressions.

## References

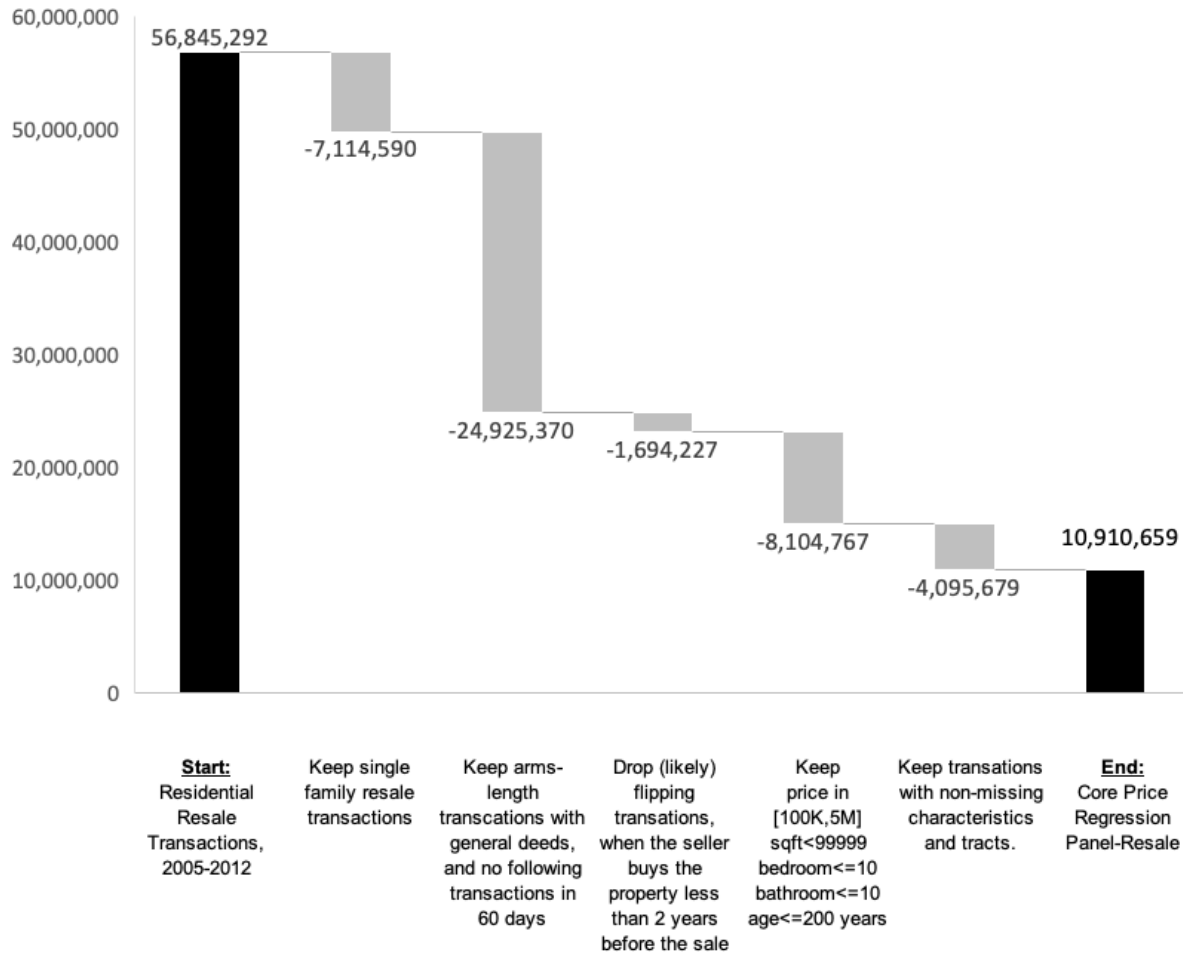
Gardner, J. (2022). Two-stage differences in differences. *arXiv preprint arXiv:2207.05943*.

Figure S1: Overview of Price Panel Construction (New Constructions)



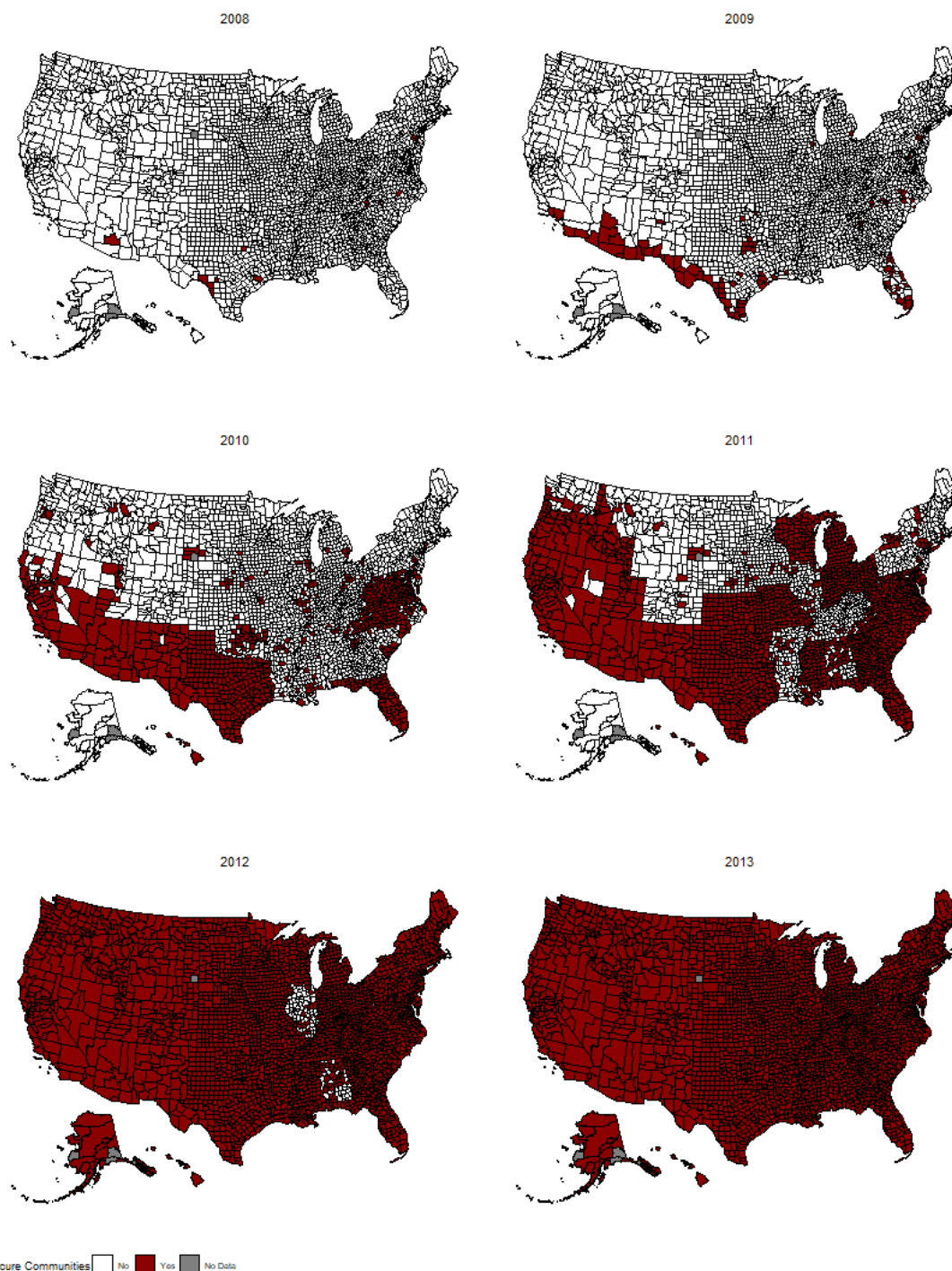
*Note:* This figure provides an overview of each step in the price data of new constructions. Additional detail on each step is available in Section [A.2](#).

Figure S2: Overview of Price Panel Construction (Resales)



*Note:* This figure provides an overview of each step in the price data of house resales. Additional detail on each step is available in Section A.3.

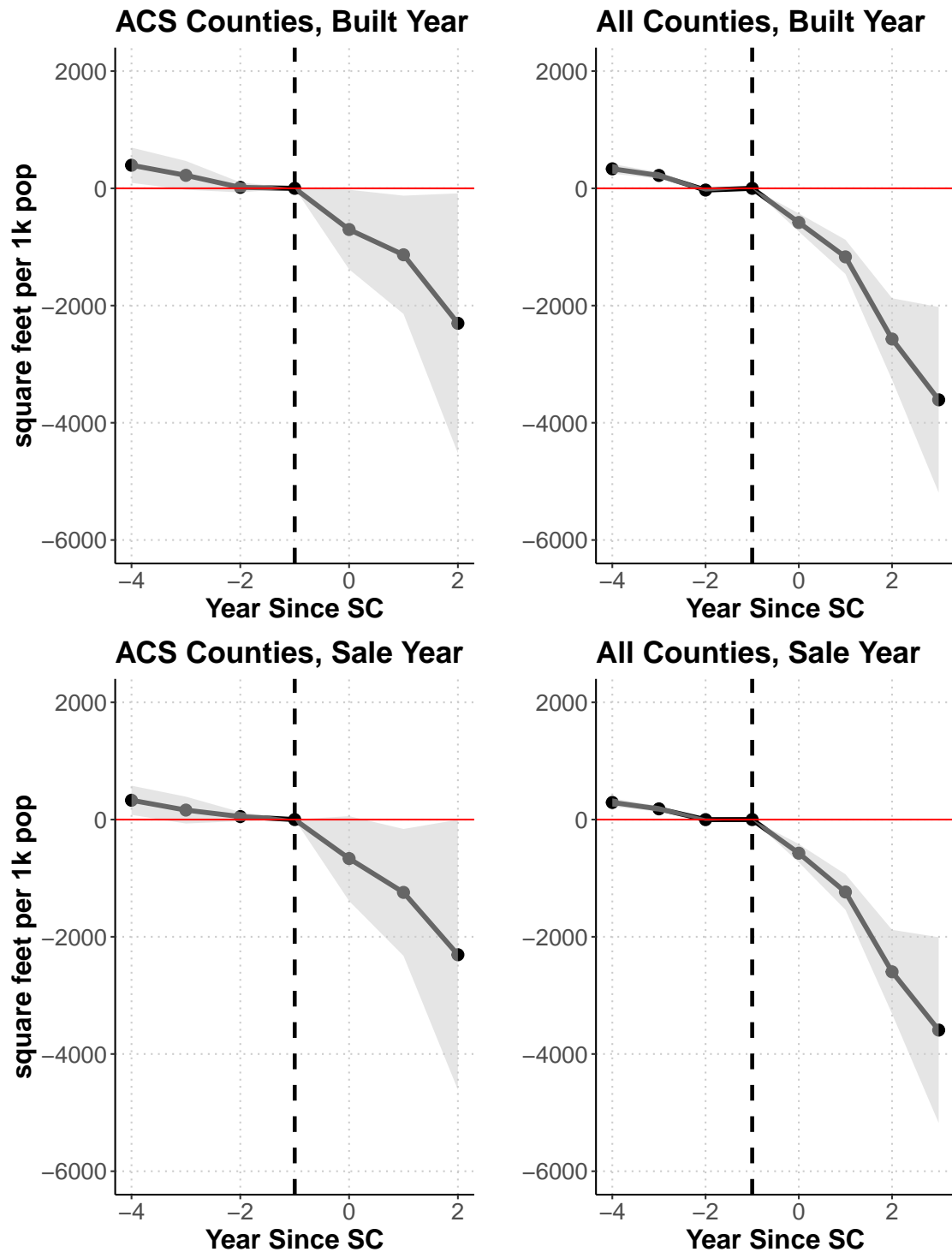
Figure S3: Staggered Rollout of Secure Communities (by exact date)



*Note:* Each panel of this figure shows the counties that implement SC within each year. This map reflects treatment based on exact date of implementation: a county is coded as treated in year  $t$  if the launch date falls at any point within year  $t$ . All regressions in this paper assign annual treatment status only to counties which have been treated for at least half a year; a corresponding map of this empirical treatment indicator is shown in Figure 2.

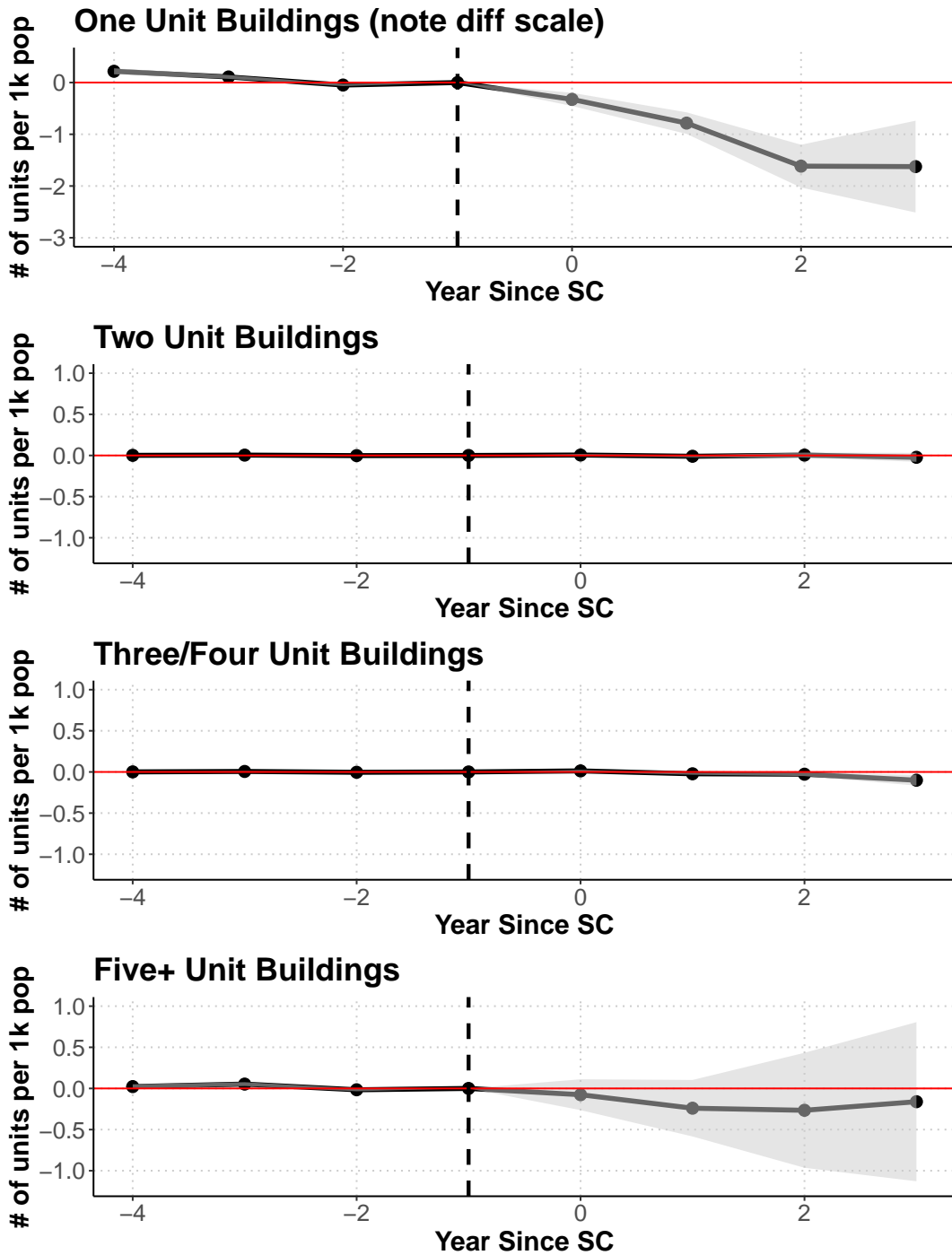


Figure S4: New Construction, Without Square Footage Imputation



*Note:* This figure plots the impact of SC on residential construction activity measured by observed new construction (completed new construction), with the approach of [Gardner \(2022\)](#) and specification (2). This figure repeats the analysis of Figure 8 but does not impute square footage for any observation missing that information. The outcome variable is total square footage normalized by 2005 county-wide population. 95% confidence intervals are plotted around the point estimations, and the standard errors are clustered at the county level. The unit of one period is a year.

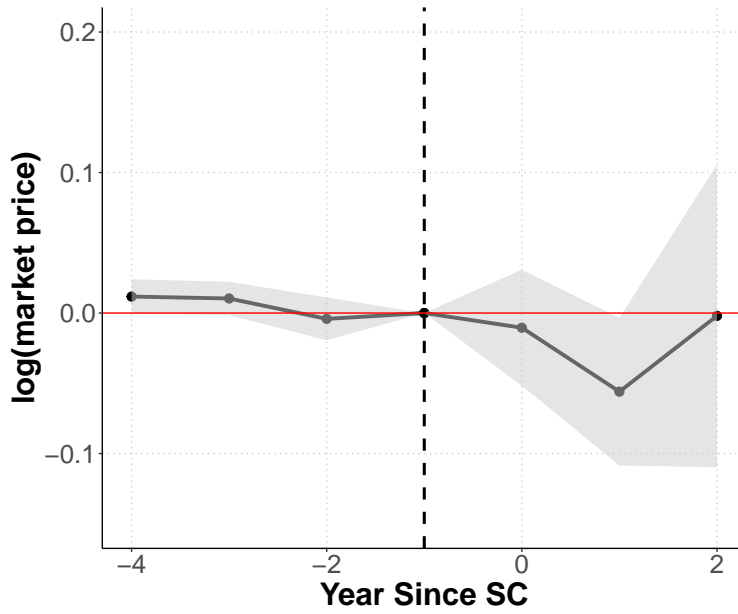
Figure S5: Permitted Units by Building Class



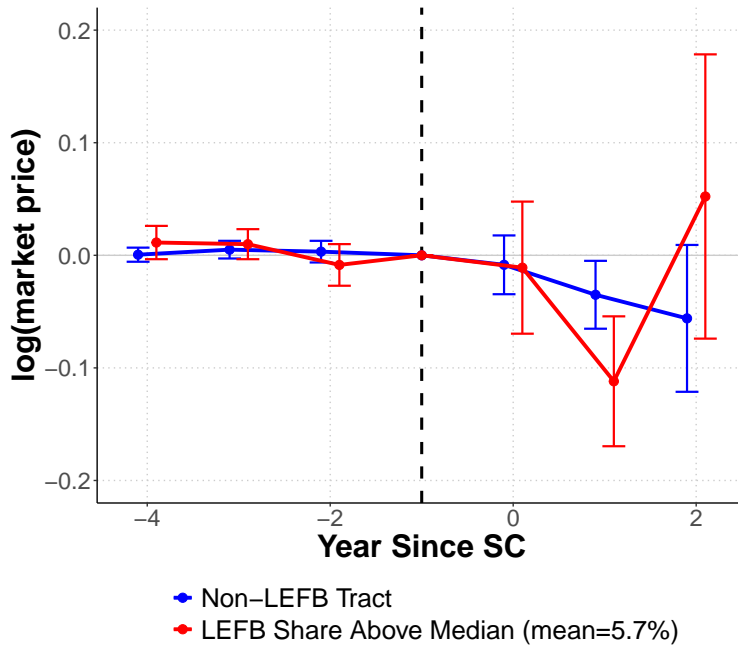
*Note:* This figure plots the impact of SC on residential permits (intended construction) by building size, with the approach of [Gardner \(2022\)](#) and specification (2). The four panels examine the impact on permitting buildings of one-unit, two-unit, three/four-unit, and five/more-unit. Estimations in this figure are based on the national subsample. We use total permitting buildings per 1,000 residents as the outcome variables. 95% confidence intervals are plotted around the point estimations, and the standard errors are clustered at the county level. The unit of one period is a year.

Figure S6: Resale Price Response, Single-Family Homes, ACS Counties

(a) Average Price Response

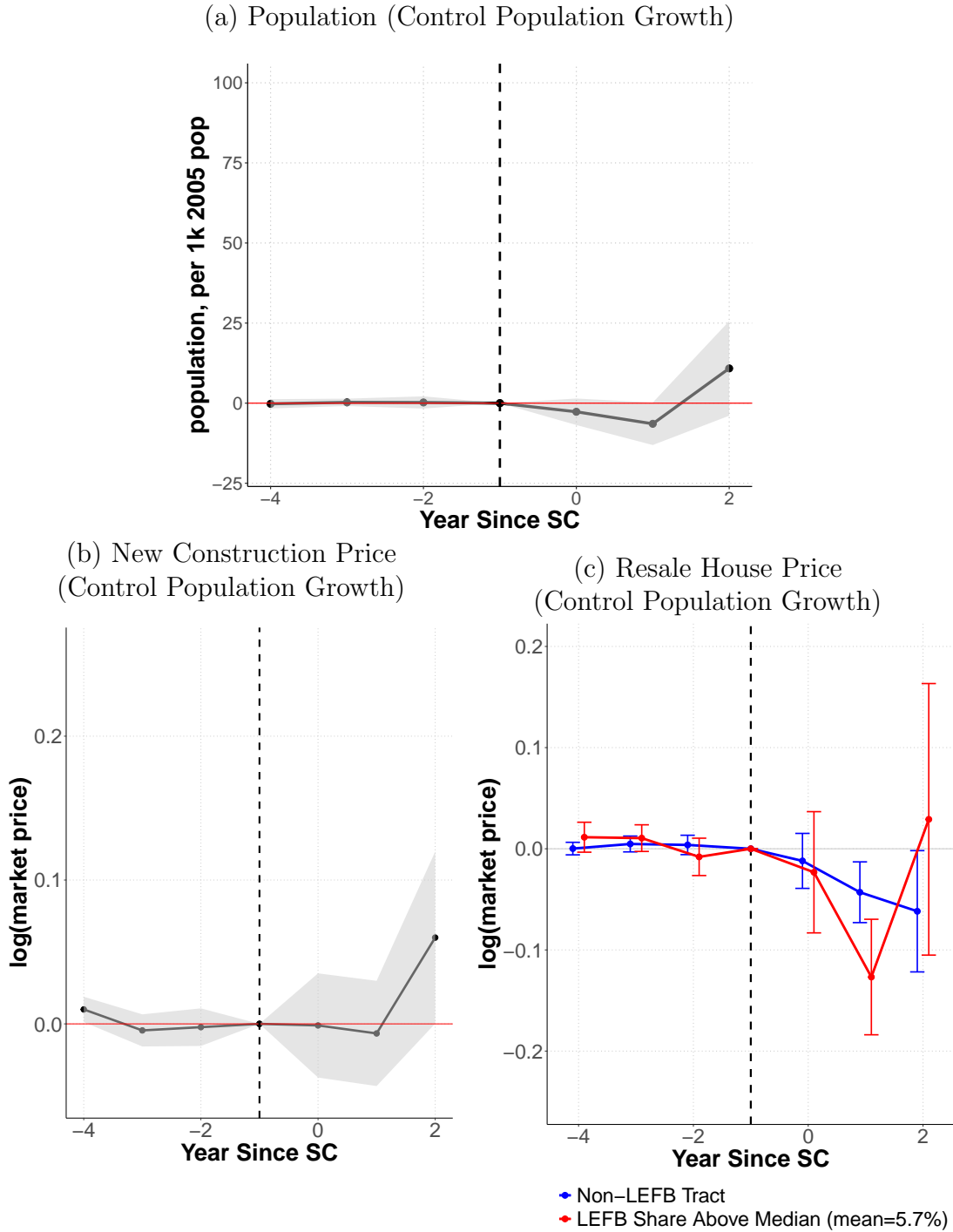


(b) By Tract LEFB Share



*Note:* This figure plots the impact of SC on the resale prices of the existing single-family home stock, with the approach of Gardner (2022) and specification (2). The top panel shows the average effect. The bottom panel shows the price responses for tracts with 0% LEFB share (blue), and for tracts with more than 1% LEFB share (red). The outcome variable is the natural log of recorded market price. Results are based on the subset of counties separately identifiable in the ACS microdata, and results on Figure 10 are the national sample, and. 95% confidence intervals are plotted around the point estimations, and the standard errors are clustered at the county level. The unit of one period is a year.

Figure S7: House Price Responses, Control Population Growth, ACS Counties



*Note:* This figure repeats several estimations with the inclusion of growth-tercile-by-year fixed effects. Results are based on the subset of counties separately identifiable in the ACS microdata. Panel (a) estimates the SC effect on population (paralleling the top of Figure A4). Panel (b) shows the SC effect on new construction prices (paralleling the top left of Figure 9). Panel (c) shows the house resale price responses for tracts with 0% LEFB share (blue), and for tracts with above median share (red) (paralleling the bottom of Figure S6). 95% confidence intervals are plotted around the point estimations, and the standard errors are clustered at the county level. The unit of one period is a year.