Large Tech Office Openings and the Onset of Gentrification

Benjamin Freyd*

July 21, 2022

1 Introduction

Gentrification has recently re-emerged as a major question in several US cities. Former low-income neighborhoods, especially in California and New York City, have seen a fresh inflow of wealthy residents bringing along substantial neighborhood changes. This phenomenon has a priori ambiguous welfare effects. On the one hand, it tends to increase property values, benefiting incumbent homeowners; it also improves local economic dynamism, with the potential creation of more businesses, especially in the service sector (Glaeser, Kim, and Luca, 2018). Both of these are seen as positive impacts of gentrification. On the other hand however, gentrification can out-price lower-income renters, thereby reinforcing geographical income segregation (Couture, Gaubert, et al., 2019; Berkes and Gaetani, 2018). Households who have to move out may not only lose their place, but they may have to move somewhere with worse economic opportunities (Ganong and Shoag, 2017). Overall, the evidence about the effects of gentrification remains mixed, suggesting a lot of heterogeneity effects (Meltzer, 2016; Meltzer and Ghorbani, 2017).

In the economic literature, gentrification is mostly seen as an endogenous, snow-balling effect: wealthy residents attract more of their kind through the development of specific neighborhood amenities. This endogenous amenity development is at the core of modern urban models (see e.g. Diamond, 2016). While many factors have been evoked (Hwang and Lin, 2016)little is known about what initially triggers a wave of gentrification. Policymakers could highly benefit from knowing whether their city is about to go through such a wave, in order to enact protective policies

*Department of Economics, University of California, Los Angeles. I am deeply grateful to Stuart Gabriel for his support and guidance on this research piece. This paper also benefited from insightful conversations with Barney Hartman-Glaser, Michela Giorcelli and seminar participants at UCLA. I would like to thank the UCLA Ziman Center for Real Estate's Rosalinde and Arthur Gilbert Foundation for generous funding. All errors are my own.

for renters or ensure sufficient new housing supply, both of which may take years to reach. So far, the presence of artistic and creative businesses has been associated with the future gentrification of a neighborhood (Behrens, Boualam, et al., 2018; Schuetz, 2014). Supermarket and coffee shop openings have also been associated with mild increases in house prices Glaeser, Kim, and Luca, 2018; D. G. Pope and J. C. Pope, 2015. But a clear causal link is hard to establish between the appearance of these businesses and a slowly ensuing gentrification process. In a lot of cases, either a fine enough spatial or temporal scale is missing and that is a threat to identification.

In this work, I study the impact of major technology office* openings on house prices in their neighborhood. These events have the potential of bringing a substantial amount of well-paid workers around the office location, thereby abruptly changing a neighborhood. Numerous anecdotal accounts of such effects have been reported (see e.g. Dave and Vincent, 2017; Rosenberg, 2018). Major industrial openings have also been shown to have sizeable effects on local labor and housing markets at a coarser scale (Greenstone, Hornbeck, and Moretti, 2010). I gather data on a number of major office openings, including their exact address and the year of opening. I look at the impact of these events on house prices by using transaction-level data on properties, controlling for housing quality through multiple property characteristics, and using fine-grained location fixed effects. To control for the endogeneity of firm location choices, I design treatment and control groups based on (i) their geographical proximity and (ii) their similarity in house prices prior to the opening event.

My results suggest a strong causal link between the office opening event and a rise in local house prices. In the main specification, I find that house prices rise 11% within 1 km of the new office, relative to matched areas located between 1 and 3 km away, within 2 years of the office opening. This difference persists at around 8% five years after. These effects are substiantially larger those found in the case of supermarket (D. G. Pope and J. C. Pope, 2015) and coffee shop (Glaeser, Kim, and Luca, 2018) openings. The results are qualitatively robust to a number of changes in the design of the treatment and control groups. I explore two important potential mechanisms behind these findings, to understand their size and persistence. I assess the role of agglomeration forces, whereby new technology offices attract similar companies, creating a long-lasting snow-balling effect (Davis and Dingel, 2019; Gaubert, 2018; Behrens, Duranton, and Robert-Nicoud, 2014;

*e.g. a Google office dedicated to software engineering.

Ellison, Glaeser, and Kerr, 2010; Rosenthal and Strange, 2001). I also assess the role of the development of consumption amenities around the office location, that can then attract a broader wealthy population (Couture and Handbury, 2017).

The paper is organized as follows. Section 2 presents the data. Section 3 describes the design of treatment and control groups and the main model specification for the empirical analysis. Section 4 presents the results and dicusses them. Section 5 explores the robustness of the results.

2 Data

2.1 Major Office Openings

I obtain data from various online and press sources about major openings of large technology offices. I obtain their exact addresses as well as their opening or opening-announcement year. In cases where I have announcement date information, I confirm that the office has opened and assume that it opened a year after the announcement. Table 1 describes the sample I use in this paper. The sample is made of 9 major openings spread out across the US and the last two decades. In the main specification, I regroup Amazon's and Apple's openings in Los Angeles due to strong their geographical and timing proximity.

2.2 House Prices

House prices are obtained from Zillow's ZTRAX database[†], a transaction level data set covering the entire US since the mid 1990's. ZTRAX provides sale amounts as well as an extensive set of property characteristics that I use to control for housing quality. I compute the sale price per built square foot and use it as the main outcome in this analysis. I further use information on the exact location (latitude/longitude) of the property, the number of bedrooms and bathrooms, the type (single family or condominium), I restrict the same to sales of single-family homes and condominiums with a minimum amount of $100,000 and a maximum of $3 million (2019 dollars). I also focus on square footage ranging between 300 and 6000. I restrict the sample to properties

[†]Zillow Group, 2020

built afte 1850. Since my analysis looks at highly urban areas, these restrictions only remove outliers. I only keep property transactions that happen within 5 kilometer of one of the office I have in my sample. At this stage, the sample contains 527,778 transactions. It will be further reduced through the design of the identification strategy, explained in section 3.1. I experiment further restrictions, removing pre-20th century building and very large homes: these do not affect the results significantly.

2.3 Demographics

I obtain demographic data from the Census LEHD. These data contain yearly counts of individuals down the the Census block level, split by broad demographic groups: age, race, industry of work. These counts are performed at residence and workplace levels, allowing a distinction between industrial and residential compositions. This is the only public data of the Census offering such high geographic granularity along with a yearly frequency, which are both needed for my analysis. These data are used to investigate the underlying mechanisms behing the changes in house prices observed following the opening of a major technology office. The industry of work information allows us to measure the composition of the workforce at a fine level and detect possible broader changes initiated by the opening of the office. This, in turn, can inform us about the development of the local consumption amenity sector, or the presence of agglomeration forces attracting firms that benefit from the presence of the new office.

2.4 Final Dataset

I generate a novel dataset by merging office opening information with the aforementioned property price and demographic data. To do so, I first obtain each office latitude/longitude coordinates from their exact address using Google Maps. I construct a 5 kilometer buffer around each office and include every property transaction that happens within that buffer (using their coordinates provided in ZTRAX) and within 5 years of the opening year (before and after). Using a geographical shapefile for 2010 Census block groups, I associate to every property transaction the code of the block group in which it happens. I can then merge the demographic information contained in the

LEHD data to obtain the final dataset.

3 Empirical Analysis

3.1 Treatment and Control Design

Treatment and control groups are primarily designed based on the geographical distance from the office opening. The desirability of a short commute[‡] the development of consumption amenities around large firms[§] and small-scale agglomeration economies[¶] are factors that suggest that the housing market response should be strongest around the opening location and wane over short distances. While commute time is a more direct measure of the how close two places are from a human perspective, my analysis is run at a scale small enough (a few kilometers) that commute times and geographical distances are interchangeable.

In the baseline specification, the treatment group is defined as transactions happening in Census block groups lying within a 1 kilometer radius of the opening event. Even at such small scales, there can be important differences in housing markets. Two close neighborhoods may have different dynamics that need to be taken into account. I use a pre-trend matching algorithm to select most control areas most comparable with the treatment area. The procedure first isolates potential control block groups witin a certain radius of the treatment area. It then calculates a 3-year pre-opening trend in log property price per square foot for the treatment area and all potential controls. It finally picks the bottom 10% (in the main specification) controls in terms of distances in trends with the treatment. More details on this procedure are given in appendix 7.

Figure 1 shows the size of treatment and control groups for each office opening event.

[‡]See e.g. Frenkel, Bendit, and Kaplan, 2013

[§]See e.g. Behrens, Boualam, et al., 2018

[¶]See e.g. Buzard et al., 2017

3.2 Baseline Comparisons

By comparing areas with similar property price pre-trends, our approach already accounts for most relevant differences in treatment and control groups. For the sake of completeness, and because assignment to treatment is never fully random in such analyses, I report a comparison of relevant characteristics of treatment and control groups as of one year prior to the opening event. I divide these characteristics into housing quality ones, measured at property level, and demographic ones, measured at block-group level. The offices in my sample open in highly different housing markets: some are in very dense urban locations, while some are in the surburbs. The treatment and control groups are also not well-balanced in terms of sample sizes. Doing a single, overall comparison of property characteristics would thus bear the risk of comparing urban areas with suburban ones. To avoid this, I report property-level comparisons within each office area. The results are shown in figure 2. For most characteristics and offices, treatment and control groups are not statistically distinguishable. While some differences appear, they should not directly be seen as a threat to the analysis. Having slightly different housing stocks does not preclude treatment and control groups from being counterfactually on the same price trend.

Demographic characteristics being obtained from the Census LEHD at block group level, I have a fairly small sample of them. I thus report baseline comparisons in these characteristics for the overall sample, without distinguishing by office area. At this fine geographic scale, LEHD data reports only a few demographic characteristics. I report the proportion of high-skill workers, as defined by those who work at companies with a two-digit NAICS code falling in the top quartile in terms of college-educated workforce share (education data is not directly available for most years). I also report comparisons in terms of the percentage of residents who are white, under 30 years old and whose monthly wage exceed \$3,333[‖]. As shown in table 2, treatment and control block groups do not significantly differ along any of these dimensions.

[‖] Wages above this value are the highest wage bracket available in LEHD data.

3.3 Event Study Model

I harness a difference-in-differences strategy to evaluate the impact of a large office opening on the local housing market. The main specification is an event study model, where I estimate differences in treatment and control house prices for each year around the opening event. This specification uses the following structure:

$$y_{iot} = \sum_{s=-3}^{5} \beta_s \Big[T_i \times \mathbf{1}(\tau_{ot} = s)\Big] + \sum_{s=-3}^{5} \gamma_s \, \mathbf{1}(\tau_{ot} = s) + \alpha_{c(i)} + \lambda_t + X'_{iot}\delta + \epsilon_{iot}$$

where:

* y_{iot} is the log sale price per square foot of property i, located around office o and sold in year t
* $T_i = 1$ if i is treated, 0 otherwise
* τ_{ot} is the difference, in years, between t and the opening of office o
* γ_s are a set of leads and lags fixed effects
* $\alpha_{c(i)}$ are Census tract fixed effects, λ_t year fixed effects
* X_{iot} are property-level controls: number of bedrooms, number of bathrooms, categorical variable for the decade the property was built, property type (SF home or condo)

The coefficients of interest in this specification are the β_s. For each year s relative to the office opening year, β_s measures the average difference between house prices in the treatment and control groups. These estimates are obtained controlling for census tract fixed effects[**], absolute time fixed effects (controlling for general housing market trends) and fixed effects for time relative to opening (capturing any other time varying effect that would systematically happen around an office opening event and have a common impact on treatment and control groups). β_s also account for differences in the housing stock pertaining to the controls on property characteristics listed above. For $s \leq 0$, we expect β_s to be systematically close to and not significantly different from 0, which suggests that the treatment and control groups were on the same house price trend before the office opening

[**]Census tract fixed effects also indirectly control for which office opening event the property relates to, because tracts are more granular. We thus don't need additional office or city fixed effect.

event. For $s > 0$, β_s measures, in each year after the opening, the possible difference in house prices between treatment and control groups induced by the office opening event.

3.4 Endogeneity Discussion

Technology firms decide on office locations based on a number of factors. Among the most important of them are the corporate tax regime, the availability of office space and the availability of skilled workers[††]. We believe that the fine spatial granularity of this analysis controls for these potential confounders. Corporate taxes are usually a state-level issue, and do not change across neighborhoods of the same city; they thus do not affect my analysis. The local availability of skilled workers is most likely the same within a radius of a few kilometers because such a distance can easily be covered by walking or biking; there is no reason to see a sharp difference in commute times for workers if the office instead opened in a control area.

A more general endogeneity concern is the possibility that other events caused both the companies to move in, as well as house prices to go up relative to nearby neighborhoods. Any such event that would affect both treatment and control areas would be controlled for by the leads/lags terms in the event-study specification. Regarding events that would affect treatment and control groups differentially, it seems unlikely that they happened systematically, given the very diverse timings and locations of openings in the sample. Moreover, the fact that house price increases are very focal to the office opening location (see results in section 4) leaves little room for other events to have caused them. To test that in a more thorough way, we run a set of geographical placebo tests, by randomizing the location of new offices. The results are presented in section 5.3.

[††]For impact of taxes on firm location choice, see e.g. Giroud and Rauh, 2019. Regarding the availability of skilled labor, see e.g. Audretsch, Lehmann, and Warning, 2004. Regarding the availability of land and other common factors, see e.g. Buczkowska and Lapparent, 2014.

4 Results

4.1 Event Study Results

Figure 3 shows the results of the main specification (I report also detailed results in table 3). It shows the estimates for the main coefficients of interest β_s around the opening year. Pre-opening trends in house prices are almost idential in the treatment and control groups, suggesting good comparability. The differences clearly appear following the opening of an office. Within two years after opening, property prices in the treatment group gain about 10% relative to those in the control group. This short term effect slightly subsides over time, to reach around 7% five years after opening. This slight dip could reflect some sprawl, with the aura of the major office opening eventually reaching further outwards. But a large difference persists after several years. That suggests the ignition of endogenous mechanisms following the opening event and the fact these effects are hyper-local[‡‡]. The potential role of different mechanisms is discussed in section 4.2.

We find much stronger property price effects that the existing literature working at a similar geographic scale. D. G. Pope and J. C. Pope (2015) find that Walmart and Target supermarket openings are associated with a 1–2% increase in prices within 0.5–1 mile (0.8–1.6 kilometers) over a 2 year period, while I find increases in the order of 10%. Glaeser, Kim, and Luca (2018) show that coffee shops openings are associated with a 0.5% increase in property prices within the same ZIP code over a 1-year period, while I find increases in the order of 7%. There are clearly major differences between the types of establishments that the literature has studied and those in my analysis. Coffee shop openings are most likely the consequence of an ongoing gentrification process and do not play a large part in that process individually. Supermarkets are more likely to play an important role, but not directly through the workforce they attract, being mostly composed of low-skill workers. Tech establishments tend to attract a large and highly skilled (hence highly paid) workforce, giving them a greater potential impact on the neighborhood they set foot in.

[‡‡] A result that is in line with e.g. Behrens, Boualam, et al., 2018

4.2 Discussion

This section considers the role of two distinct mechanisms in generating a large and persistent house price difference between treatment and control groups. We succesively discuss the roles of agglomeration forces (the new firm attracting other firms because of production spillovers) and consumption amenities (the new firm and its workers pushing the development of certain local services).

At small geographic scales (e.g. a few kilometers here) the microeconomic reasons for agglomeration economies are dominated by labor market pooling and knowledge spillovers (Rosenthal and Strange, 2001). In other words, firms benefit from close proximity with other firms either because they use the same type of labor, which can reduce job search costs, or because they use similar knowledge and technology for their production, which can benefit from more frequent human interactions. Both these factors suggest that small-scale agglomeration forces happen between firms within the same industry. Major offices have an anecdotal potential to be the trigger of an agglomeration phenomenon. Since they belong to well-established companies, these offices are fairly autonomous: they may not require a local ecosystem of technology firms to function properly, which allows them to set foot in locations that do not yet experience tech agglomerations. On the other hand, they often serve as incubators or venture capital providers for smaller promising firms. That gives smaller firms a reason to come and locate near these big establishements to benefit from the opportunities they offer.

Agglomeration Forces. A way to assess the important of agglomeration forces in our case is thus to see whether the tech industry locally booms following the opening of a major office. To proxy for this boom, I use Census LEHD data and look at the relative evolution of the workforce between treatment and control areas in three key sectors: finance & insurance, management, and professional/scientific services[§§]. Professional/scientific services notably include a number of activities associated to the technology sector, and is therefore an industry that we expect to boom

[§§]Finance is defined by individuals working in the sector with NAICS code 52. Management by NAICS code 55. Professional/Scientific Services by NAICS code 54.

as a result of agglomeration forces around the major opening. For each block group b in my sample, I compute the total number of people working in each of these three sectors and residing in b. I aggregate these counts to entire treatment areas and normalize their share within the entire workforce to 0 at the year of office opening. I plot this index for each sector over time in figure 4. The pattern is clear: treatment areas experience a substantial gain in residents related to the tech industry after the opening, while other major high-skill sectors such as finance and management do not show a comparable boom. This points towards an agglomeration mechanism. While the opening office can account for part of this rise, it most likely does not account for the entirety of it, suggesting the entry of additional similar firms.

Consumption Amenities. I perform a similar analysis as for agglomeration forces above. This time, I focus on industries that provide services to local residents: arts and entertainment, and accomodation and food. Moreover, workers in these sectors are often low-income and likely not resident in the affluent or gentrifying areas they work in. Therefore, I used *workplace-level* data (WAC data from the LEHD) instead of the residence-level data used for the analysis of agglomeration mechanisms. An inflow of highly-paid tech workers following the opening of a new office could boost the demand for such services, whose development could further attract new wealthier residents — not necessarily working in the tech industry. The graph presented on figure 5 does not support that mechanism. The share of residents working in these service industries exhibits an overall flat, although erratic trend over the entire period, and the opening of a tech office does not alter that trend much. It is to be noted that the changes observed on the graph are also much smaller than those reflecting potential agglomeration forces on figure 4. The maximum observed change relative to the opening year is -0.15 percentage points, in the arts and entertainment workforce three years after the opening. Overall, this analysis suggests that the development of local consumption amenities is a much weaker potential explanation than agglomeration for the rise in house prices following the opening of a new large tech office.

5 Robustness

5.1 Sensitivity of Results to Design Parameters

I first assess the robustness of my results to the choice of parameters in the empirical design. Three parameters govern the definition of control and treatment groups: the treatment radius, the control eligibility radius, and the share of control blocks retained after pre-trend matching. To make the sensitivity analysis of results to these parameters more concise, I collapse the event-study model into a simple pre/post difference-in-differences. The specification used here is:

$$y_{iot} = \beta \left[T_i \, \mathbf{1}(\tau_{ot} \geq 0) \right] + \gamma \, T_i + \lambda_t \alpha_{c(i)} + X'_{iot}\delta + \epsilon_{iot}$$

where all elements are defined as in the main specification. The main coefficient of interest is now β. I report estimates from this specification in figure 6. Top panels show estimates for the coefficient γ of the model. γ reflects average pre-opening differences between treatment and control groups. Small, unsignificant estimates are obtained in nearly all parameters configurations, showing that pre-trends are parallel and thus validating the identifying assumption of the difference-in-differences approach. Bottom panels show estimate for the coefficient β of the model, reflecting post-opening differences between treatment and control. The estimates are large and significant across the board, showing that my results are not qualitatively sensitive to the empirical design parameters. The results are generally more significant than in the full event-study specification (see table 3) because of a gain in power.

5.2 Time Placebo Tests

In this section I assess the relevance of opening events for triggering house price spikes. To do so, I run placebo tests by artificially shifting office opening years forward and backward. For this experiment to be directly comparable to my main result, I re-run the entire procedure on the artificially modified data, including the pre-trend matching phase. Figure 7 shows the result of these placebo tests in bottom panels, along with the main results in the top panel for reference. There are no significant changes in house prices around forward and backward placebo years. This reinforces the idea that office opening event are the true reason why we see house prices spike in the main result.

5.3 Geographical Placebo Tests

I further run geographical placebo test. Instead of using the actual office location, I redefine the center of the treatment area by a randomly drawn point within a 1 to 2-kilometer donut around the actual opening location. In other words, the office opening is now artificially moved to an outer ring around its actual location. I run 200 such random draws for each opening event and run the same regression analysis using these artificial locations. In the results, only 19% of the draws lead to significant estimates of the effect of office openings on house prices at the 5% confidence level. While a proportion around 5% would have been an ideal result to fully reject unobserved geographical factors, I take this result as a strong suggestion that these undesired effects are limited.

6 Conclusion

This article provides evidence that openings of large technology firms' offices have a positive and significant impact on local house prices, a key ingredient in the process of gentrification. These effects are strong and long-lasting relative to the findings of previous research, which have focused on the development of consumption amenities as a source of gentrification. Moreover, my results show that consumption amenities do not develop differentially close to the new office location, relative to similar locations further away. The growth of such businesses is overall only weakly associated with increases in house prices. Instead, my results point to an agglomeration story. The workforce in high-technology and related sectors grows much faster around new tech offices than further out, suggesting that the new office is attracting more related businesses in the area. While my data cannot provide direct evidence on that, these other businesses likely higher high-skill, wealthy workers that will increase the demand for housing in the area, hence its price. This finding is in line with theories of Marshallian agglomeration, like knowledge spillovers, labor market pooling and input-output sharing. Overall, my work thus suggest that technology firm have a key role in *igniting* a wave of gentrification and that this mainly goes through an agglomeration channel.

7 Appendix: Pre-trend Matching Algorithm

Consider an individual office opening event. The treatment area is defined as all block groups that intersect a buffer of radius d_T around the office location (defined by its latitude/longitude coordinates). $d_T = 1$ km in the main specification. I include all property transactions in a block group into the treatment, even when only part of that block group intersects with the treatment buffer. The reason to do so is twofold. First, block groups are very homogenous areas defined by the Census, and I do not want to create artificial boundary effects. Second, I use demographic data given at block-group level for balancing tests. To start the selection process of control block groups, I first consider as "eligible controls" the block groups that fall within a radius d_C of the treatment area (excluding, of course, the treatment area itself). $d_C = 2$ km in the main specification. Again, an entire block group is considered as soon as part of it intersects the d_C-radius buffer.

For each eligible control block group b, I compute the pre-opening trend in its average log property price per square foot $P_{C,b,t}$. I do so too for the treatment area *as a whole*. I get:

$$\forall t \leq \text{opening year}: \quad \left| \begin{array}{ll} \Delta P_{T,t} = P_{T,t} - P_{T,t-1} & \text{for the treatment} \\ \Delta P_{C,b,t} = P_{C,b,t} - P_{C,b,t-1} & \text{for each control block } b \end{array} \right.$$

Each eligible control block b's trend is then compared to the treatment trend through a mean squared distance (MSD):

$$\text{MSD}(b) = \sum_{t \leq \text{opening}} \left(\Delta P_{T,t} - \Delta P_{C,b,t} \right)^2$$

I select as a final control group the bottom λ_C share of eligible control blocks in terms of their MSD with the treatment. In the baseline specification, $\lambda_C = 10\%$. In the Google Venice example, this leads to the selection of the block groups colored in gray. This procedure is applied to every opening event separately.Figure 8 shows the final treatment and control areas for each opening event in my sample.

References

Audretsch, David B, Erik E Lehmann, and Susanne Warning (2004). *University Spillovers and New Firm Location*. en. Tech. rep. 0204. Papers on Entrepreneurship, Growth and Public Policy.

Behrens, Kristian, Brahim Boualam, et al. (2018). *Gentrification and pioneer businesses*. Tech. rep.

Behrens, Kristian, Gilles Duranton, and Frédéric Robert-Nicoud (2014). "Productive Cities: Sorting, Selection, and Agglomeration". In: *J. Polit. Econ.* 122.3, pp. 507–553.

Berkes, Enrico and Ruben Gaetani (2018). *Income Segregation and Rise of the Knowledge Economy*. Tech. rep. 213.

Buczkowska, Sabina and Matthieu de Lapparent (2014). "Location choices of newly created establishments: Spatial patterns at the aggregate level". In: *Reg. Sci. Urban Econ.* 48, pp. 68–81.

Buzard, Kristy et al. (2017). "Localized Knowledge Spillovers: Evidence from the Spatial Clustering of R&D Labs and Patent Citations".

Couture, Victor, Cecile Gaubert, et al. (2019). "Income Growth and the Distributional Effects of Urban Spatial Sorting".

Couture, Victor and Jessie Handbury (2017). "Urban Revival in America, 2000 to 2010".

Dave, Paresh and Roger Vincent (2017). "Snapchat has changed Venice, and the neighborhood isn't changing back". In: *Los Angeles Times*.

Davis, Donald R and Jonathan I Dingel (2019). "A Spatial Knowledge Economy". In: *Am. Econ. Rev.* 109.1, pp. 153–170.

Diamond, Rebecca (2016). "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000". In: *Am. Econ. Rev.* 106.3, pp. 479–524.

Ellison, Glenn, Edward L Glaeser, and William R Kerr (2010). "What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns". In: *Am. Econ. Rev.* 100.3, pp. 1195–1213.

Frenkel, Amnon, Edward Bendit, and Sigal Kaplan (2013). "Residential location choice of knowledge-workers: The role of amenities, workplace and lifestyle". In: *Cities* 35, pp. 33–41.

Ganong, Peter and Daniel Shoag (2017). "Why has regional income convergence in the U.S. declined?" In: *J. Urban Econ.* 102, pp. 76–90.

Gaubert, Cecile (2018). "Firm Sorting and Agglomeration". In: *Am. Econ. Rev.* 108.11, pp. 3117–3153.

Giroud, Xavier and Joshua Rauh (2019). "State Taxation and the Reallocation of Business Activity: Evidence from Establishment-Level Data". In: *J. Polit. Econ.* 127.3, pp. 1262–1316.

Glaeser, Edward L, Hyunjin Kim, and Michael Luca (2018). "Measuring Gentrification: Using Yelp Data to Quantify Neighborhood Change".

Greenstone, Michael, Richard Hornbeck, and Enrico Moretti (2010). "Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings". In: *J. Polit. Econ.* 118.3, pp. 536–598.

Hwang, Jackelyn and Jeffrey Lin (2016). "What Have We Learned About the Causes of Recent Gentrification?" In: *Cityscape* 18.3, pp. 9–26.

Meltzer, Rachel (2016). "Gentrification and Small Business: Threat or Opportunity?" In: *Cityscape*.

Meltzer, Rachel and Pooya Ghorbani (2017). "Does gentrification increase employment opportunities in low-income neighborhoods?" In: *Reg. Sci. Urban Econ.* 66, pp. 52–73.

Pope, Devin G and Jaren C Pope (2015). "When Walmart comes to town: Always low housing prices? Always?" In: *J. Urban Econ.* 87, pp. 1–13.

Rosenberg, Mike (2018). "Will Amazon's HQ2 sink Seattle's housing market?" In: *Seattle Times*.

Rosenthal, Stuart S and William C Strange (2001). "The Determinants of Agglomeration". In: *J. Urban Econ.* 50.2, pp. 191–229.

Schuetz, Jenny (2014). "Do art galleries stimulate redevelopment?" In: *J. Urban Econ.* 83, pp. 59–72.

Zillow Group (2020). *ZTRAX*. URL: https://www.zillow.com/research/ztrax/.

Figures

Figure 1: Sample composition: number of property transactions happening in the treatment and control groups, 3 years prior to 5 years after the opening, for each office in the sample.

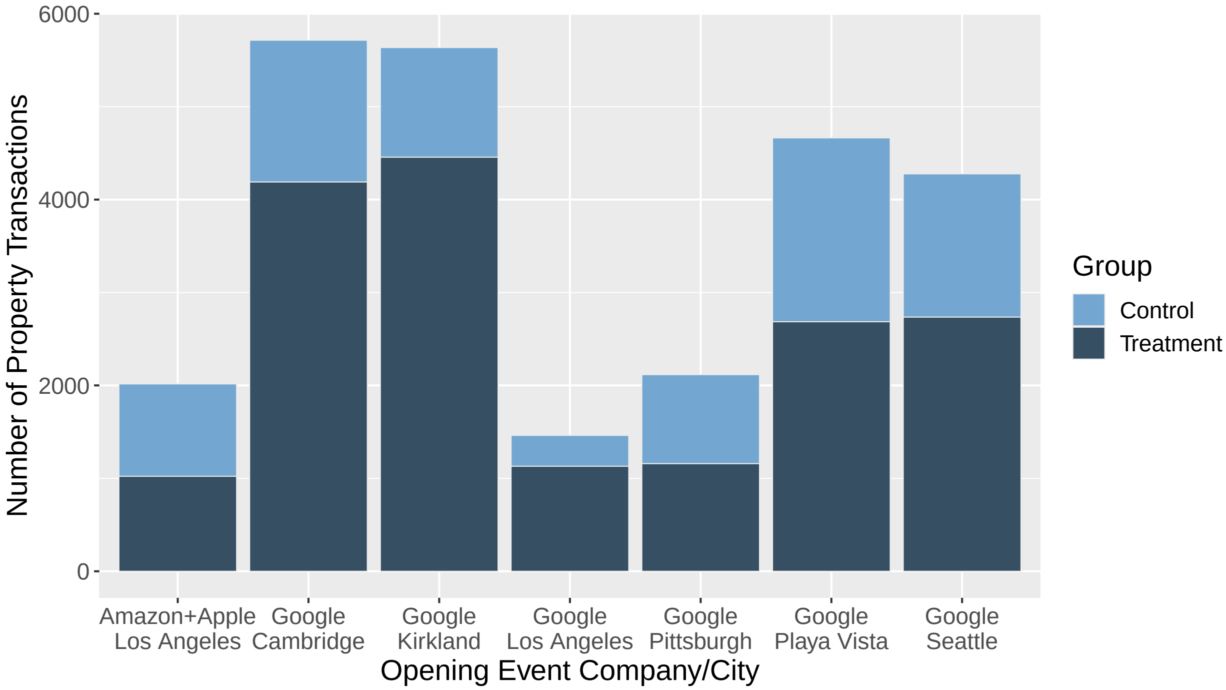


Figure 2: Baseline comparison of selected property characteristics between treatment and control groups, within each office area. Based on data one year prior to the office opening event. Dots represent mean estimates. Bars represent 95% confidence intervals.

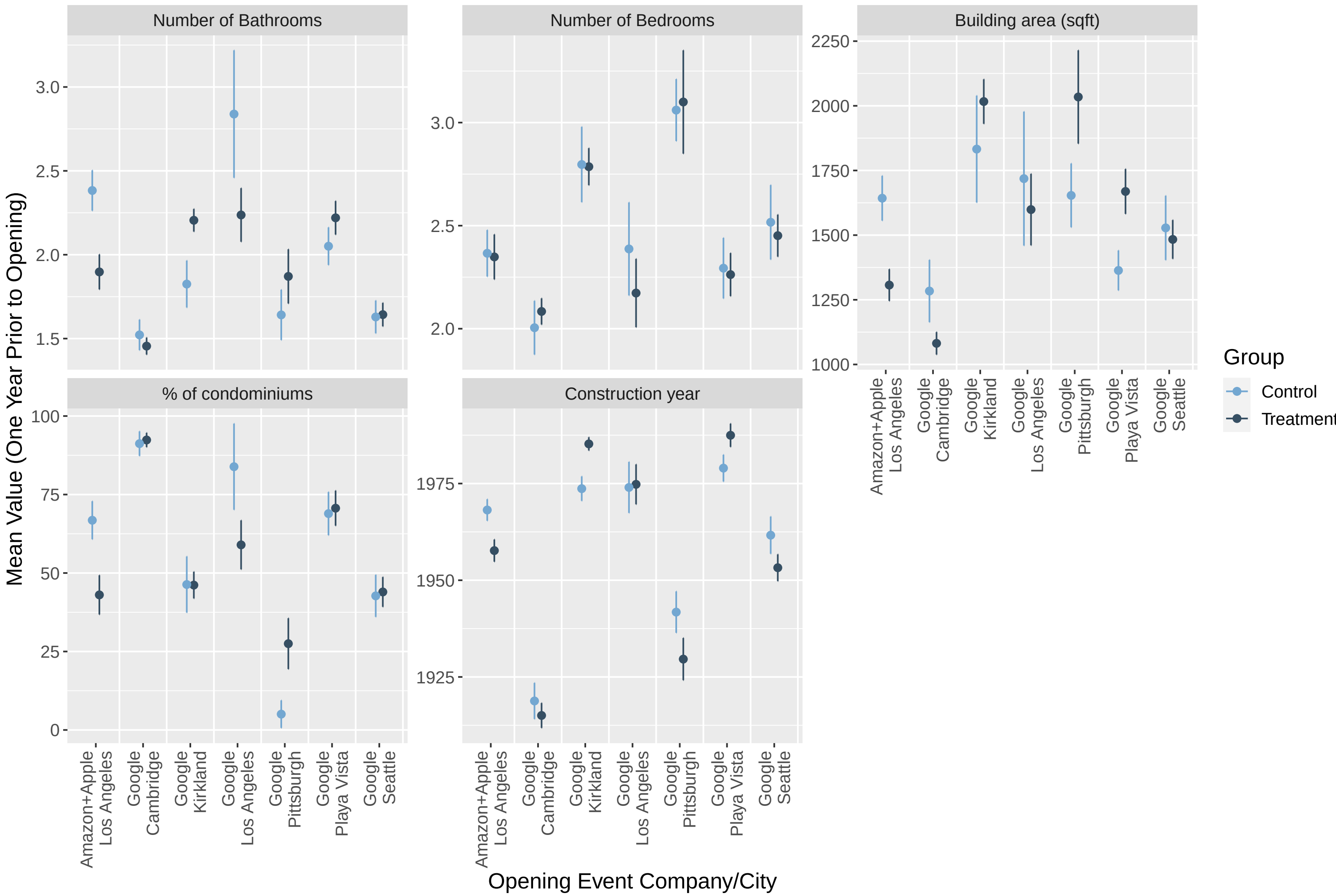


Figure 3: Event study results: estimates for β_s, $s \in [-3, 5]$. $N = 25{,}844$. $R^2 = .67$. Treatment group is defined as properties sold within 1 km of the office locations. Control group is defined as properties within a 2 km buffer of treatment area, belonging to block groups in the top 20% of pretrend matching. Property controls include number of beds and baths, an indicator for the decade of construction, a condo/single-family dummy. Error bars denote 95% confidence intervals. Robust standard errors clustered at block-group level.

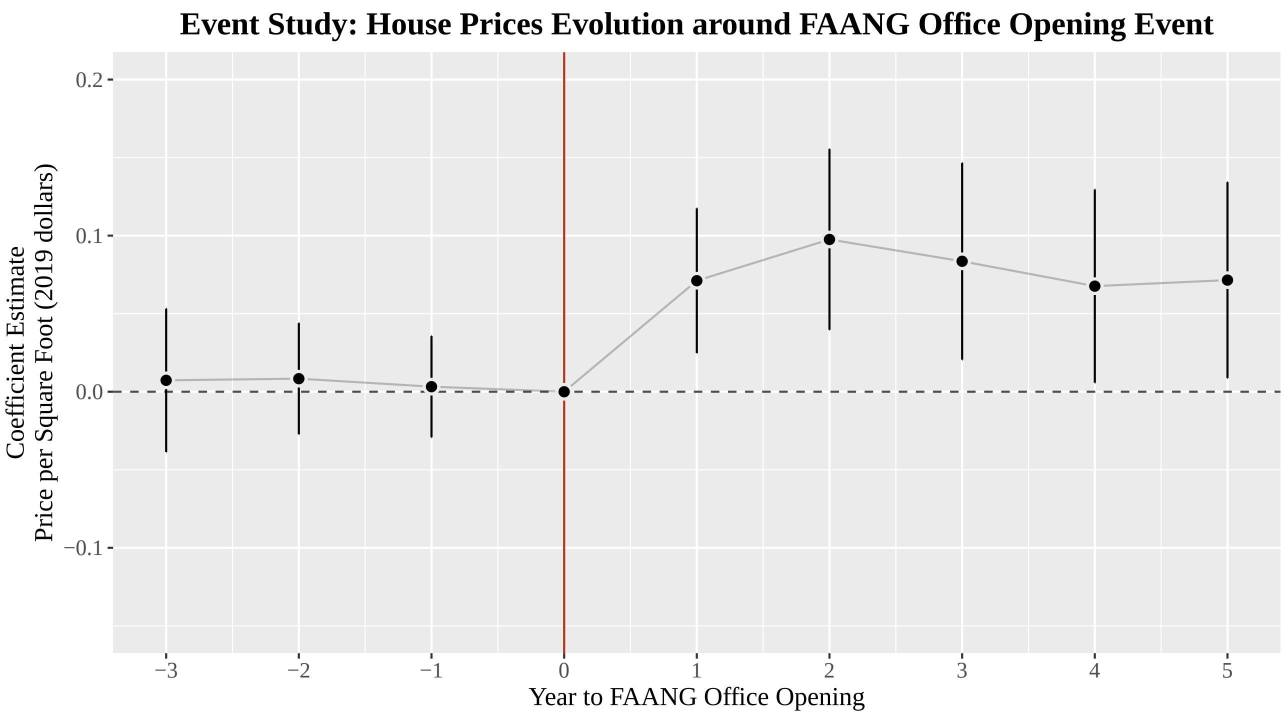


Figure 4: Evolution of the workforce residing in treatment areas in three sectors (finance, management and professional/scientific services). This plots the percentage shares of specific sectors' workforce within the entire workforce. Shares are normalized to their opening-year level, so the graph shows percentage points changes relative to that year.

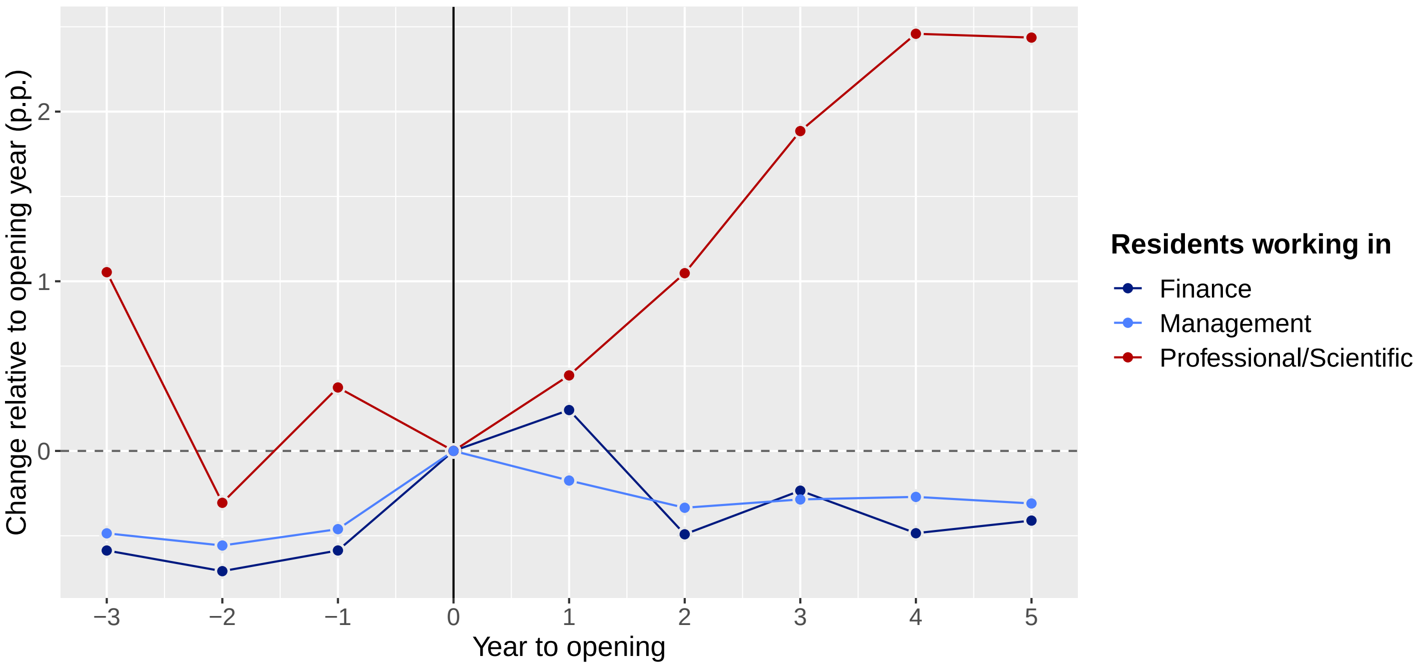


Figure 5: Evolution of the workforce working in treatment areas in two sectors (arts and entertainment, and accomodation and food). This plots the percentage shares of specific sectors' workforce within the entire workforce. Shares are normalized to their opening-year level, so the graph shows percentage points changes relative to that year.

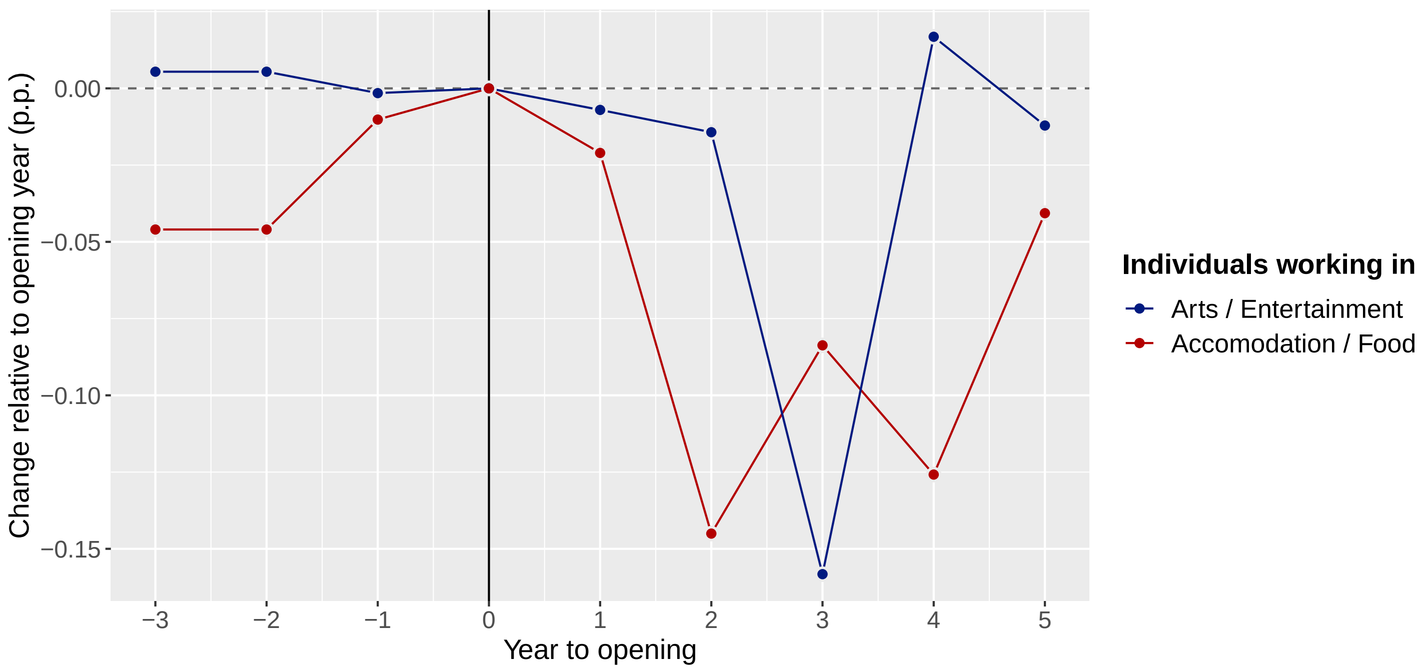


Figure 6: Robustness tests on 27 empirical design configurations. Varying treatment area radius (0.5, 1 and 1.5 km), control buffer size (1, 2 and 3km) and percentage of control block-groups kept in pretrend matching (10, 20 and 30%). The top panels show the coefficient on the treatment dummy, which measures the average log price difference between treatment and control, prior to office opening. The bottom panels show the coefficient on the interaction between treatment dummy and post dummy, measuring the additional difference in prices following the opening.

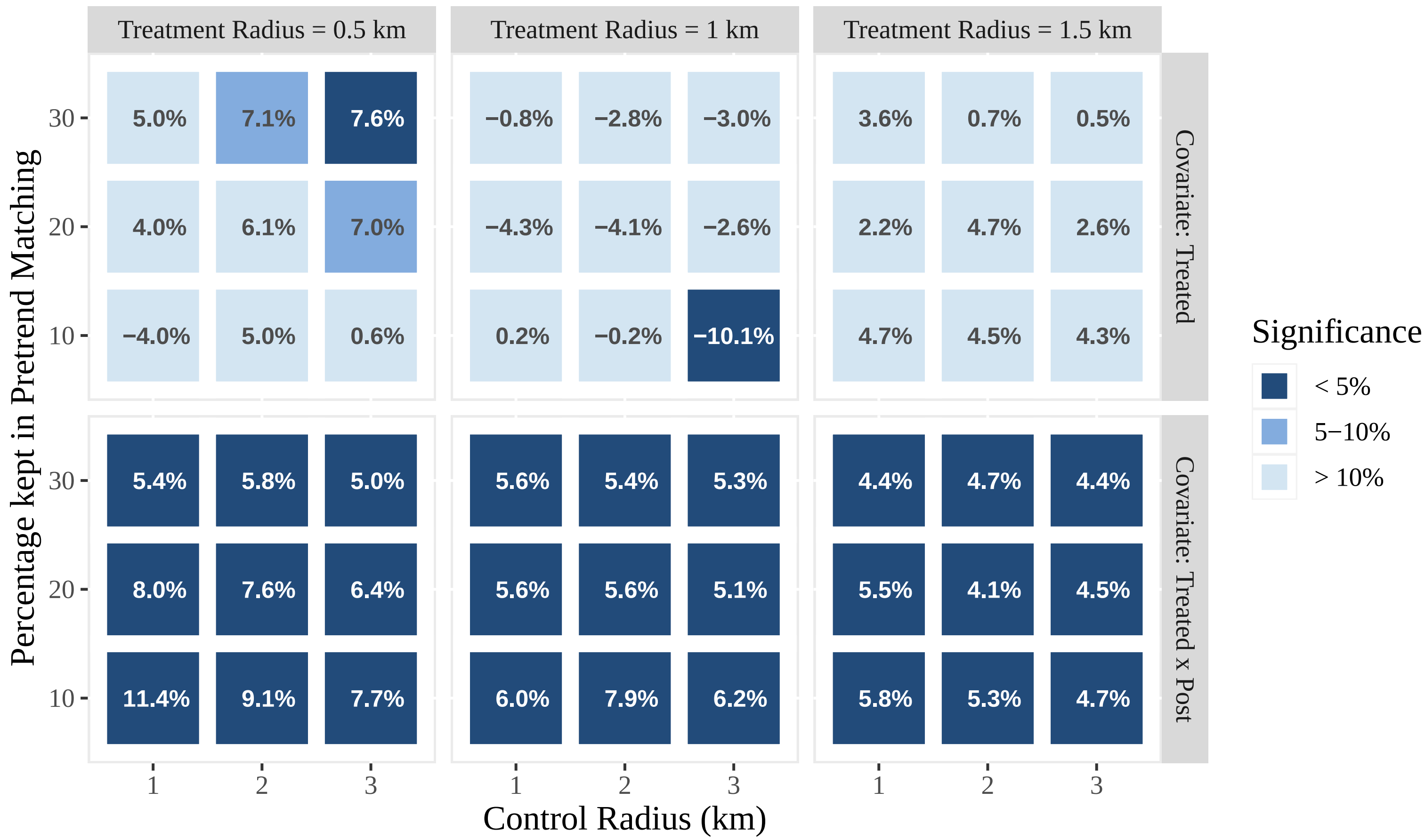


Figure 7: Placebo tests. Entire experiment, including pre-trend matching, is re-run using placebo years for office openings. The top panel shows the main results for reference. The bottom left panel shifts opening years 3 years forward. The bottom right panel shifts them 3 years backward.

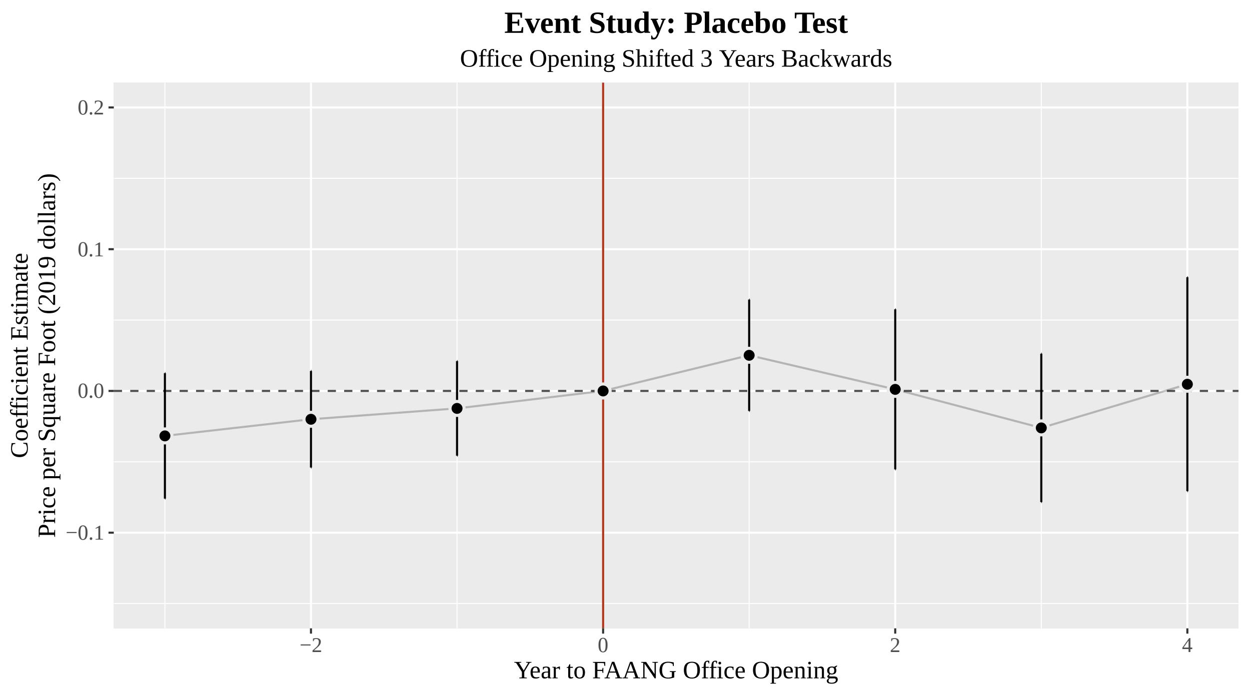
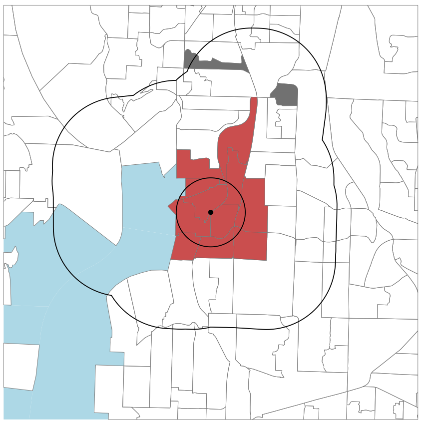


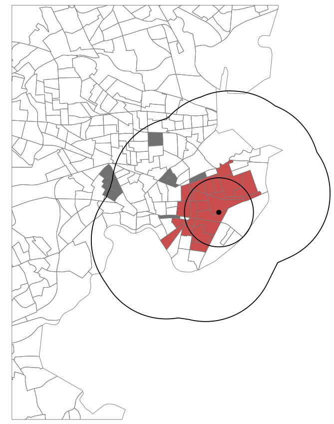
Figure 8: Treatment and control areas selected in the main specification. The dots represent new offices. The smaller circles represent treatment radius (1 km here). The red block groups are those that are considered treated. The larger black boundaries represent control eligibility areas (2 km around the treatment area here). The gray block groups are controls selected through pre-trend macthing procedure (top 10% here).

(a) Google – Kirkland

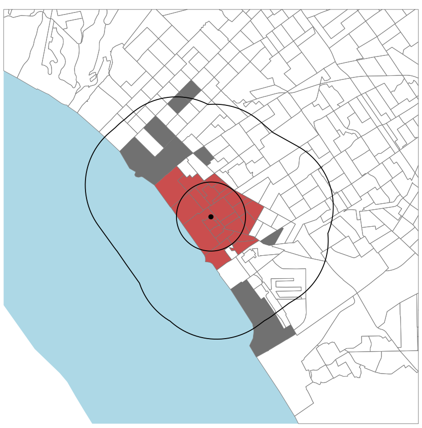


(b) Amazon and Apple – Los Angeles

(c) Google – Cambridge

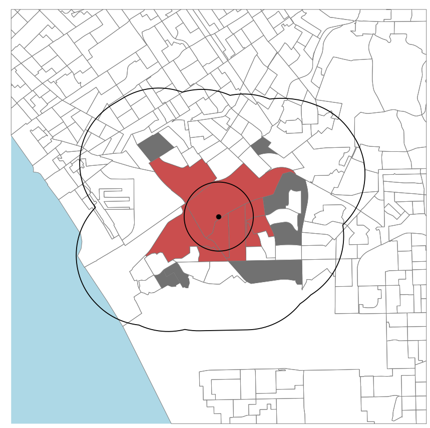


(d) Google – Los Angeles

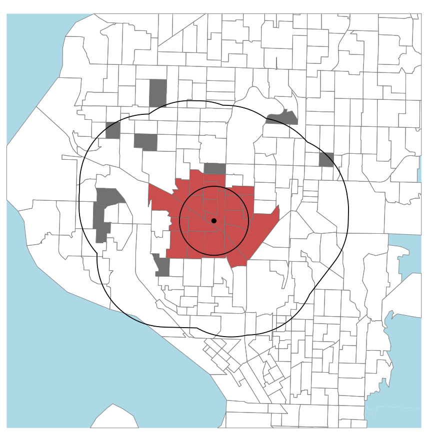


(e) Google – Pittsburgh

(f) Google – Playa Vista



(g) Google – Seattle



Tables

Table 1: Sample of tech office openings.

Company	City	State	Opening Year
Google	Kirkland	Washington	2004
Google	Seattle	Washington	2006
Google	Cambridge	Massachussetts	2008
Google	Pittsburgh	Pennsylvania	2010
Google	Playa Vista	California	2011
Google	Los Angeles	California	2011
Amazon	Los Angeles	California	2017
Apple	Los Angeles	California	(ann.) 2018

Table 2: Baseline comparison of selected demographic characteristics between treatment and control groups. The treatment and control columns represent proportion estimates in each group. Difference = treatment - control. The p-value is from a Student t-test on the difference, with H_0: difference = 0. Characteristics measured at block group level. Based on data one year prior to the office opening event. High-skill industries are defined as 2-digit NAICS codes in the top 25% in terms of share of college-educated workers.

Demographic Characteristic	Treatment	Control	Difference	P-value	N. Obs.
High Skill (%)	37.37	35.1	2.27	0.204	181
White (%)	76.49	76.06	0.43	0.887	109
Under 30 y.o. (%)	27.76	26.51	1.25	0.503	181
Monthly Wage > $3,333 (%)	47.54	45.81	1.72	0.367	181

Table 3: Estimates for the coefficients of interest β_s in the event-study specification. The baseline specification is Model 2, with a treatment radius of 1 km. I also report results under two alternative treatment radii in Model 1 and 3. Robust standard errors clustered at the block group level.

	Model 1	Model 2	Model 3
Treatment Buffer:	0.5 km	1 km	1.5 km
Treat × (YtO = −3)	−0.06	0.01	0.01
	(0.06)	(0.02)	(0.02)
Treat × (YtO = −2)	−0.02	0.01	0.00
	(0.03)	(0.02)	(0.02)
Treat × (YtO = −1)	−0.03	0.00	−0.01
	(0.03)	(0.02)	(0.01)
Treat × (YtO = 1)	0.06	0.07^{***}	0.06^{***}
	(0.05)	(0.02)	(0.02)
Treat × (YtO = 2)	0.08^{*}	0.10^{***}	0.07^{***}
	(0.04)	(0.03)	(0.03)
Treat × (YtO = 3)	0.04	0.08^{***}	0.04
	(0.04)	(0.03)	(0.03)
Treat × (YtO = 4)	0.08^{**}	0.07^{**}	0.04
	(0.03)	(0.03)	(0.03)
Treat × (YtO = 5)	0.03	0.07^{**}	0.05^{*}
	(0.05)	(0.03)	(0.03)
Property Controls	✓	✓	✓
Census Tract F.E.	✓	✓	✓
Adj. R^2	0.63	0.67	0.67
Num. obs.	14,934	25,844	39,266

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$