Estimating the Effect of San Francisco's Impact Fees on Housing Production *

Salim Damerdji

September 2023

Abstract

Housing shortages across the country have raised rents and displaced tenants, putting housing policy on the agenda at the state and local level. Yet, policymakers lack visibility into how much housing would get built under alternative proposed reforms. This paper analyzes the causal effect of reducing impact fees on housing production and finds that even a 50% reduction in impact fees may fail to push the needle on housing production. Using San Francisco as a case study, I analyze adjacent parcels where one parcel is subject to a 50% higher impact fee than its neighboring parcel. Over a 2014-2023 timeframe, I find no difference in the probability of development of adjacent lots, no difference in the expected unit count conditioned on development occurring, and, by consequence, no difference in housing production. This result is corroborated by a Double/Debiased Machine Learning algorithm that uses machine learning to control for nearly one hundred parcel-specific factors and their non-linear interactions. Policymakers motivated to promote housing production should identify policy solutions that do more than reduce impact fees - like those under study - by 50%.

Keywords: Housing Supply, Spatial Econometrics, RDD, Double/Debiased ML

*My deepest thanks to Professor Frank Windmeijer for his thoughtful guidance and invaluable feedback on my dissertation, from which this paper hatched. And a warm thank you to Reza Amindarbari of San Francisco Planning for his generous help in explaining and providing data from the City. Many thanks to Professor Chris Elmendorf and Dr. Issi Romem for their kind verbal feedback and wise advice when I first set out to pursue this paper.

1 Introduction

In high-demand cities, new housing construction reduces rents.[18][16][11][1][12][8] To redress regional housing shortages,[9] policymakers nationwide are contemplating proposals to increase housing production.[25] Yet, there are as many policy levers to produce more housing as there are regulations that currently constrain housing production, of which there are many.[4][9] Policymakers motivated to tackle housing shortages, thus, face a prioritization question: given a slew of possible reforms, which reforms should be prioritized? A twin question is: which reforms match the scale of the housing crisis?

One policy option is to reduce impact fees. An impact fee is a financial exaction that local jurisdictions levy on new development to pay for the costs of public services associated with new development.[2] Though property taxes and other local taxes historically paid for infrastructure, cities starting in the 1920s gravitated towards using impact fees to pay for a wider and wider scope of infrastructure.[2] In California, the state with the most extensive use of impact fees,[2] impact fees can add tens of thousands of dollars to the cost of development for each unit of an apartment, though the figure ranges widely from city to city.[13] Typically, cities require impact fees to be paid before the apartment is built, which means that high-interest loans compound the costs of these fees.[21][19]

Reducing impact fees lowers the cost of development and should, in theory, spur housing production. The strength of this relationship, however, remains an open question. The existing literature primarily focuses on the relationship between impact fees and the affordability of housing, and most studies find that impact fees raise housing prices.[7] But the focus on home prices side-steps an essential question: do prices rise because of decreased supply or increased demand?[14] The latter possibility is raised by authors who hypothesize that impact fees increase demand for neighborhoods by paying for amenities that otherwise would not exist - or would only otherwise exist with increased property taxes.[2]

Existing research on the causal effect of impact fees on housing supply is sparse and mixed. Skidmore and Peddle studied a panel dataset from 1977 to 1992 in DuPage County, Illinois, and the authors found a 25% reduction in the rate of housing development for cities that had an impact fee.[24] Using a richer, nationwide dataset, Mayer and Somerville find no statistically significant relationship between impact fees and housing production.[14]

This paper contributes to this literature by proposing a novel identification strategy. Using San Francisco as a case study, I compare neighboring parcels of land where one parcel is subjected to a 50% higher impact fee than its neighboring parcel. To control for selection bias, I use a comprehensive, parcel-level dataset describing zoning, existing structures, recent improvements, and more. (See **Table 5** in the Appendix for the extended list.) If controlling for this large set of variables is insufficient to control for selection bias, I argue that any unobserved sources of selection bias vary continuously along the border between two neighboring parcels subject to different fees. To make this concrete, **Figure 1** provides a street view of this boundary; my assumption is that, if there are unobserved confounders, that their effect on housing production varies smoothly across the boundary. For example, if local opposition to housing both causes higher impact fees and reduces housing production, then my assumption is that a local community group would be similarly opposed to housing developments on neighboring parcels separated by the fee boundary.[1] If this assumption holds, then by framing the boundary between these parcels as a sharp, geospatial regression discontinuity, I can show that, after controlling for observed variables, any discontinuity in housing production is caused by the impact fees.

Figure 1: Example of two neighboring parcels on the boundary.

The impact fees in question regulate development in four neighborhoods in San Francisco: South of Market (SoMa), the Mission, the Showplace Square/Potrero Hill neighborhood, and the Central Waterfront. The City and County of San Francisco refers to these four neighborhoods as the Eastern Neighborhoods for shorthand. Starting in 2011, the City implemented policies to encourage housing construction in these neighborhoods. In the last decade, 32% of all units built in San Francisco have been built in these four neighborhoods.

Pursuant to the Mitigation Fee Act of 1987, San Francisco conducted a Nexus Study in 2008, which found that new development in these four neighborhoods would require additional expenditures to pay for library expansions, childcare facilities, transit infras-

[1] To be precise, the success of the identification strategy does not hang on whether the local opposition is equivalent on either side of the boundary: I merely need that the degree of opposition does not drop off suddenly at the boundary. With other words, identification requires that unobserved confounders vary continuously at the boundary.

tructure, and parks.[6] Rather than charge all development at the highest amount legally allowed by the Nexus Study, the City opted to charge higher impact fees - which the City refers to as Tier 2 fees - to parcels that were upzoned more as part of the area plan updates. Most parcels in these neighborhoods, however, are subjected to a lower impact fee, which the City refers to as the Tier 1 fee. Both fees are charged on net new square footage built for residential use. Annual fees increases are indexed to the Annual Infrastructure Construction Cost Inflation Estimate. As of 2023, the lower impact fee costs \$14.72 per square foot, and the higher impact fee, which is indexed to cost 50% more, is \$22.08 per square foot.[22] For a prototypical 100-unit apartment building, the higher impact fee raises the per-unit development cost by \$7,500.[2] That's about 0.8% of the total development costs and, after accounting for high interest loans used to pay the fee, the fee and the additional interest incurred may account for 1.2% of total development costs for a prototypical project.[19]

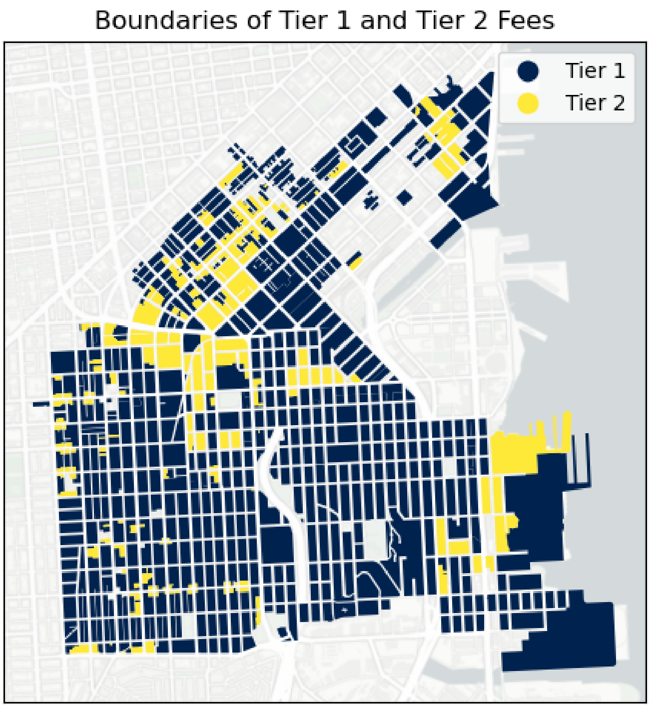

Figure 2: Parcels in yellow are subject to a 50% higher impact fee than parcels in dark blue.

This paper addresses two difficulties in studying parcel-level housing production in this case study. First, geospatial regression discontinuity designs (RDD) frequently involve compounding treatments at the border,[10] and that is the case with these impact fees under study. As explained earlier, San Francisco imposes higher fees on parcels that ben-

[2]These calculations are derived from a 2022 pro forma kindly provided by a developer in SF.

efitted more from upzoning during the 2011-2014 time period. As a result, estimates of the effect of the impact fee will be biased unless one controls for the effect of zoning on the parcel. Yet, if one only models the effect of zoning based on parcels near the fee boundary under study, the zoning and fee variables will be highly collinear, and so the fitted coefficients will be unstable, with large standard errors. I solve this challenge by relying on a unique source of data that provides helpful simplifications of San Francisco's zoning codes for 153,010 parcels citywide. By using parcels from across the city to estimate the contribution of zoning to housing production, I am able to decrease the variance of my estimates and, by consequence, adequately control for this compounding treatment.

The second challenge I face is in modelling housing production. Typical count data models like Poisson regression will not suffice for modelling housing development for two reasons: first, the factors that determine *whether* housing is built overlap with, but are not identical to the factors that determine *how much* housing will get built if anything is. For example, if a parcel has a dilapidated building on it, that plainly increases the probability of development occurring; but there's no obvious reason as to how an existing dilapidated building would alter the expected unit count of a potential project on that parcel. For standard count data models, there is no separate way to model zero counts distinct from positive counts, but modelling these two phenomena separately would reasonably improve one's ability to model housing production. The second limitation of conventional count data models is that there's overdispersion in the data: on most parcels, nothing is built; on a tiny fraction of parcels, most of the City's new housing is built. As a result, there are more undeveloped parcels and also more intensively-redeveloped parcels in the data than a Poisson or negative binomial distribution would predict. Both of these shortcomings are addressed by using a hurdle model that, firstly, uses logistic regression to model the probability of development occurring and, secondly, uses negative binomial regression to model how many homes would be built on a parcel, if it's developed.

Using a hurdle model also provides an additional interpretive benefit in that it disaggregates overall housing production into its two component factors: a parcel's probability of development $\mathbb{P}[D]$, and a parcel's expected unit count if developed, $\mathbb{E}[U|D]$. This disaggregation is useful for policymakers because, while pro-supply policymakers may be agnostic as to which lever to pull, some policymakers prefer to increase $\mathbb{E}[U|D]$, but not $\mathbb{P}[D]$, for concern of development displacing tenants, while other policymakers may prefer to increase $\mathbb{P}[D]$, not $\mathbb{E}[U|D]$, by way of an aesthetic preference for views where no one building sticks out.

Contrary to Skidmore and Peddle, my findings ultimately agree with the null findings from Mayer and Somerville. Using a geospatial RDD with a hurdle model, I find no difference in the probability of development of neighboring lots subjected to different fees, no difference in the expected unit count conditioned on development occurring, and, taking these results together, no difference in overall housing production. My findings represent a conservative estimate of the causal effect of the impact fees because, by construction, I do not allow for the possibility that the impact fees were necessary to make it politically feasible for the City to upzone the eastern side of the city.

To confirm that this result does not stem from parametric modelling assumptions, I perform a robustness check with Double/Debiased Machine Learning, which uses machine learning to control for nearly one hundred parcel-specific factors and their potentially complex non-linear interactions. This robustness check corroborates the same basic story that a 50% reduction in impact fees does not register as a statistically significant constraint removal.

My findings suggest that policymakers motivated to increase housing supply should prioritize other housing reforms over impact fee reductions. While there is much variation in impact fees from city to city, and there are surely some cities successfully leveraging impact fees in an exclusionary way, impact fee reductions are unlikely to prompt much, if any, housing production in a city like San Francisco.

2 Data

2.1 Sources

This project merges six data sources provided by the City and County of San Francisco with neighborhood-level home price data.

The treatment variable of interest comes from San Francisco's Neighborhood-Specific Impact Fee Areas dataset, a geospatial dataset of fees charged for residential development in some parts of the city but not others.

The outcome of interest - how much housing was built on a parcel - came from San Francisco's Department of Building Inspection's (SF DBI) dataset of permits, which indicates the date a permit was created, how many homes were built, and where they were built. Specifically, I filtered for relevant SF DBI permits based on the latest available list of completed housing developments from San Francisco's Planning Department.[3]

I joined this dataset with San Francisco's BlueSky dataset, which tracks roughly 150,000 San Francisco parcels from 2001 to 2016 and includes information on each parcel's his-

[3]This data was generously provided by SF Planning Department's Reza Amindarbari, a manager in the Data & Analytics Group, via private correspondence in July 2023.

torical status, residential status, the existing building envelope, and the potential buildable envelope given the parcel's zoning designation.[23] When the City prepared the BlueSky dataset, they removed duplicative parcel identifiers, including condos, and they removed parcels with no potential for residential capacity (such as parks). This dataset provides several variables that a priori are predictive of where housing will get built: parcels listed as a historical resource are harder to redevelop into housing due to California's Environmental Quality Act; parcels that have existing homes with tenants are harder to redevelop due to San Francisco's tenant protections that make demolitions costly; and the parcel's zoning determines how much revenue can be gained by redeveloping the parcel, as more square footage can be built with looser zoning. As a result, controlling for these variables improves the precision of my estimate.

This dataset is joined with data from the county tax assessor, which includes information on the age of the property, the construction type, the property's square footage, the basement area, lot area, lot shape, the ownership status, the prior sale date of the land, the assessed improvement value, the assessed land value, the number of bedrooms, baths, stories, and units, and more.

Because steep lots pose unique construction costs, I join this dataset with a topological map of San Francisco. Created in 2019, this dataset is post-treatment but causally unaffected by the treatment, while strongly predicting development. Thus, its inclusion increases the precision of my estimate.

Finally, I incorporate lag indicators for whether the landowner applied for a permit for each parcel to build, tear down, improve, or alter something on the parcel by joining the panel dataset with SF DBI's dataset of permits. A priori, one would think that permits to improve a parcel are negative indicators that the owner is interested in tearing down the property to rebuild. Conversely, it is reasonable that demolition permits are lead indicators for future development on the parcel.

The economic feasibility of building homes depends strongly on neighborhood rents, and so I join my dataset with Zillow's Home Value Index for All Homes (both single family and condos) for annual rent per neighborhood in San Francisco.[4]

In all, there are nearly a hundred covariates in addition to the treatment and outcome. Key explanatory variables are presented in **Table 5** in the Appendix.

[4]This is the only neighborhood-level data provided by Zillow.

2.2 Cleaning

All but two of the aforementioned datasets are geospatial. That is, for most of my sourced datasets, each observation is accompanied by a set of coordinates that indicates the observation's shape and location on a map, as well as a Coordinate Reference System (CRS) for identifying which map. For consistency and accuracy, I translate all geospatial datasets to the California Albers transformation of North American Datum of 1983, a coordinate reference system which minimizes errors in estimating the distance of points in California.

To dovetail the datasets together, I primarily relied on geospatial joins through the Python package geopandas. Two joins had to be handled differently because two sourced datasets lacked geospatial information: San Francisco's BlueSky dataset and the Zillow dataset on neighborhood-level home prices. For SF's BlueSky dataset, I joined this dataset using the block lot parcel identifier with San Francisco's geospatial dataset of all active and retired parcels, thus yielding a geospatial version of the BlueSky dataset. For Zillow's neighborhood-level home price dataset, performing a join was non-trivial since there was no equivalent geospatial dataset containing all and only the neighborhood names in Zillow's historical home price dataset. Instead, the best I could find was a 2011 Zillow geospatial dataset of neighborhoods via the WayBack Machine, which did not match up exactly with the neighborhoods now used by Zillow's neighborhood-level home price dataset in 2023. To reduce the number of parcels with missing values for local home values, I first performed the join based on how the tax assessor lists the neighborhood, a join which succeeds for most parcels in my panel dataset, and then used the geospatial information for a join where missing values remained.

3 Exploratory Data Analysis

Homebuilding is relatively rare. In the decade under study, there are 1,970 parcels where housing is built in a dataset of 153,010 parcels. Because 98.7% of all parcels see no development, the dataset contains an inflated number of zeros relative to a negative binomial or Poisson distribution. When housing is built, most housing projects add just one or two homes, but a few housing projects make up the bulk of what the city builds, as indicated by **Table 1**. The five largest projects alone account for more units built than the 1393 single-unit projects. To inform policymakers about the supply effects of potential policies, it is plainly important, therefore, to not just model how policies affect the rate of development, but also the size of the projects that are built.

Development trends vary dramatically across San Francisco. **Figure 3** shows where housing gets built. Each purple circle reflects a housing development, with the size indicating

Table 1: Number of Observations for Different Buckets of Residential Project Size

Units Built	1	2	3	4	5-10	11-199	200-399	400+
Count	1393	255	85	48	55	116	15	3

how many units of housing were built as part of the development: large circles reflect large apartment (or condo) complexes. As can be seen, most of the total units built are built on the east side of the city. The west side of the city, in comparison, sees development, but of a very different type: namely, accessory dwelling units. Much of this variation is downstream of zoning, which underscores why it is important to control for zoning capacity in estimating the effect of impact fees, as argued in Section 2.1.

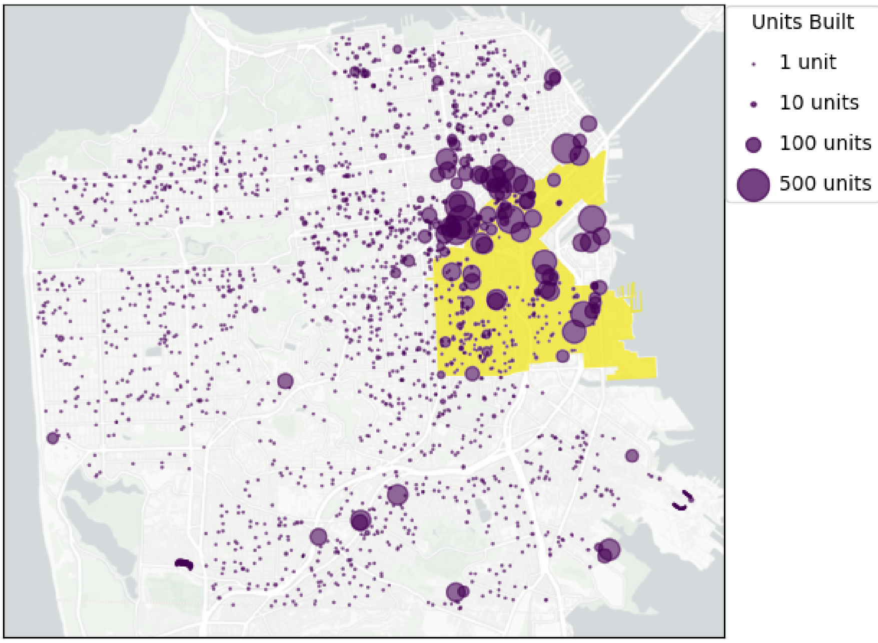

Figure 3: Spatial trends in homebuilding. The four neighborhoods under study - SoMa, Mission, Central Waterfront, and Protrero Hill - are shaded in yellow.

In **Figure 4**, there is evidence that properties with existing single family residences and government uses are less likely to be redeveloped into housing. Furthermore, it's striking that parcels with an existing use that's categorized as Miscellaneous/Mixed-Use are much more likely to be developed into housing, at roughly three times the rate of other parcels.

Per **Table 2**, most variables are weakly correlated with whether or not housing is built. The variables displayed are those whose pearson correlations have the largest absolute values, and so most variables have virtually no correlation at all with whether housing is developed. A priori, however, one should not expect strong linear, univariate relationships

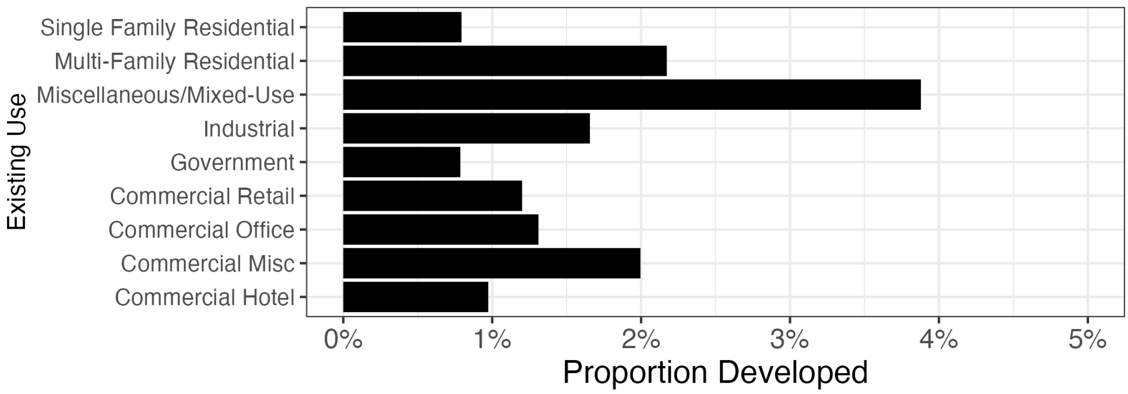

Figure 4: Development rates of parcels based on the existing use.

between these variables and development for the simple reason that development is mediated by financial metric analyses that involve the non-linear interaction of half a dozen variables, including rent, land values, zoning, and various neighborhood-level factors like impact fees.[5][17]

Variable	Correlation with Development
Year Property Built	-0.055
4+ Unit Zoning	0.045
Homeowner Exemption Value	-0.044
Vacant Lot	0.040
Form-Based Zoning	0.034

Table 2: Pearson Correlation Coefficients with Binary Indicator for Development

Given that housing is built, a slightly different set of factors are relevant to how much housing is built, as outlined in **Table 3**. For example, the top two variables are both functions of the square footage that can built on a parcel: buildable envelope is the square footage permitted under the City's zoning; and revenue is the product of the buildable envelope and neighborhood rent per square foot. This is trivially what we would expect to see.

Table 3: Pearson Correlation Coefficients with Count of Homes Built on Parcel, Conditioned on Development

Variable	Correlation with # of Homes Built
Revenue	0.612
Buildable Envelope	0.544
Existing Residential Use	-0.454
Office/Commercial Zoning	0.439
Lot Area	0.425

The disparity in which variables are correlated with development as opposed to units built, again, motivates the view that the decision to build housing is a separate decision from the choice of how much housing to build. For example, it's intuitive that the historical status of a building affects *whether* a parcel is redeveloped, but less so how much housing is built, conditioned on there being development. A parcel's historical status adds expensive CEQA red tape to build anything, but that red tape does not scale linearly as the unit count increases.

Turning our attention to the four eastern neighborhoods under study, it's helpful to look at how certain key variables vary near the border between two fees. A histogram breaks down the distribution of parcels with Office/Commercial Zoning in **Figure 5**. The histogram bars plotted to the left of zero are those where the lower fee is paid, and the bars to the right of zero reflect parcels where the higher fee is paid. Clearly, the two groups are very unalike as a whole: there's much more office zoning for the parcels subject to the higher fee (those to the left of 0) than for the parcels subject to the lower fee (those to the right of 0). However, along the border, parcels are zoned much more similarly. For this reason, I estimate the effect of the higher fee only by comparing those parcels just along the boundary.

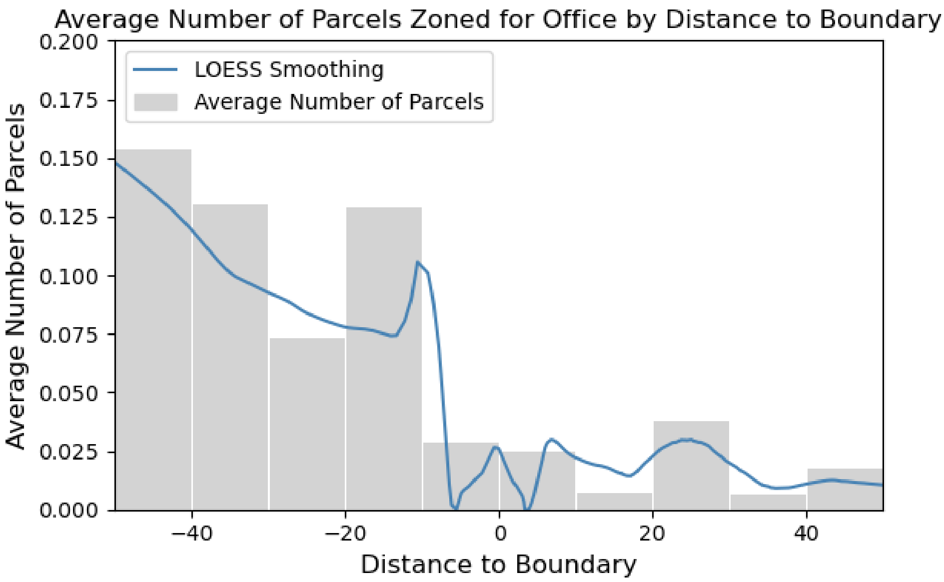

Figure 5: RDD Boundary versus Important Predictor, Office Zoning.

Using a 5 meter threshold for the distance between any two neighboring parcels, the parcels selected to estimate the local average treatment effect are shown in **Figure 6**. In all, 1,043 neighboring parcels are selected to estimate the effect of the fees on housing production. Of these parcels, 594 are subject to the lower fee, and 449 are subject to the higher fee.

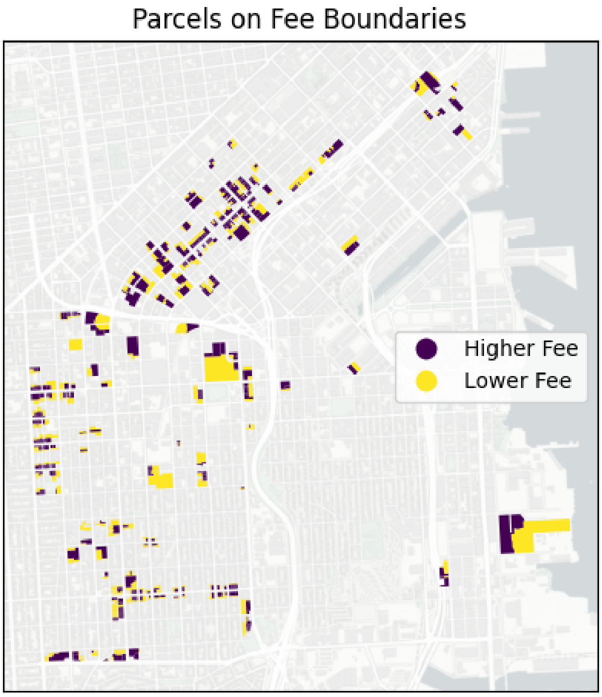

Figure 6: Parcels matched across the boundary.

Even among this subset of parcels, however, there is a compounding treatment in that parcels that were upzoned more are subject to the higher impact fees; the City's ordinance explicitly requires this. Evidence of this compounding treatment is provided in **Figure 7**. Thus, not all covariates vary continuously along the fee boundary. To avoid obtaining biased estimates, an adequate model must, at minimum, control for variables that are a function of zoning.

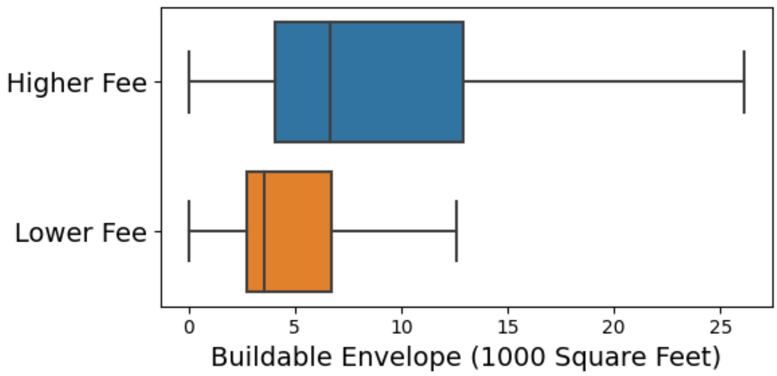

Figure 7: Comparing parcels within five meters of the fee boundary, parcels with the lower fee have stricter zoning than parcels with the higher fee.

4 Method

The fee boundaries can be viewed as discontinuities in a sharp regression discontinuity design. I will estimate the local average treatment effect of the higher fee F at a discontinuity. I will refer to the higher fee - the second tier fee - as the treatment variable F. The treatment F is determined by:

$$F = \sum_{c \in C} I[(c_{lat}^{min} <= x_{lat} <= c_{lat}^{max}) \cap (c_{lon}^{min} <= x_{lon} <= c_{lon}^{max})] \tag{1}$$

where C is a set of tuples $c := (c_{lat}^{min}, c_{lat}^{max}, c_{lon}^{min}, c_{lon}^{max})$ of latitude and longitude coordinates that demarcate the boundaries of each rectangle where a higher fee is charged. Because these boundaries are non-overlapping, $F \in \{0, 1\}$. This RDD has two forcing variables x_{lat} and x_{lon} that determine the treatment.

The average treatment effect of F is defined as:

$$\tau := \mathbb{E}[y_1 - y_0 \Big| I(\exists c \in C \text{ s.t. } x_{lat} \in \{c_{lat}^{min}, c_{lat}^{max}\}, x_{lon} \in \{c_{lon}^{min}, c_{lon}^{max}\})] \tag{2}$$

where y_1 and y_0 are the potential outcomes under treatment and control. In plain English, the average treatment effect is, then, the average difference in housing built if a parcel in the Eastern Neighborhoods is subjected to the higher fee instead of the lower fee. The local average treatment effect is just the average treatment effect evaluated along the border between the two fees under study.

Though, in principle, there are two forcing variables x_{lat} and x_{lon}, we can simplify the problem to one with a single forcing variables x^d which denotes the Euclidean distance from the nearest boundary. This simplification preserves the physical sense in which a parcel can be near a boundary.[10]

Now we can view the treatment as a binary treatment of applying the higher fee. The response variable is the number of homes built on that parcel between 2014 and 2023. I model the outcome Y, the number of net new homes built per parcel, with a hurdle model:

$$P(Y_i = y_i) = \begin{cases} p_i, & \text{if } y_i = 0, \\ \frac{(1-p_i)p(y_i;\mu_i)}{1-p(y_i=0;\mu_i)}, & \text{if } y_i > 0, \end{cases} \tag{3}$$

where p_i is the probability that parcel i is not redeveloped into net new housing, and

$p(y_i; \mu_i)$ represents the probability under a negative binomial distribution with mean μ_i and dispersion ϕ that y_i homes are built. A parcel's probability of development is estimated with logistic regression. Under the negative binomial distribution, a redeveloped parcel's expected net new unit count is μ_i, which is modelled as:

$$log(\mu_i) = \alpha + \tau F_i + \beta_1 \mathbb{I}(x_i^d < h) + \sum r_{ij}\beta_j \tag{4}$$

where $h > 0$ and the covariates r_{ij} are additional regressors that control for selection bias and the compounding treatment.

Observations in the Eastern Neighborhoods with $x_i^d > h$ are dropped from the analysis. Notably, the estimate for the LATE of F is fitted locally solely from parcels near the boundary, but estimates for the β_j are fitted globally using both parcels near the boundary and parcels outside of the Eastern Neighborhoods. This is done to reduce the variance in estimating many parameters, while retaining a low h of $h = 5$ meters, which reduces bias.[26]

I use forward stepwise feature selection to select the variables to include in the logistic regression model and the negative binomial model. At each iteration, the model drops a term to minimize the Bayesian information criterion (BIC), which provides a parsimonious model.[5] This algorithm is completed separately for each component of the hurdle model because, as argued in the exploratory data analysis section, the relevant variables for *whether* development occurs are different from the relevant factors pertaining to *how much* housing would get built, if any does get built. The estimated effect of F and other model summary statistics are presented in **Table 4**.

Per the R_M^2 metric of 0.65, the hurdle model provides an excellent fit to the data.[6] The confidence intervals for the local average treatment effect of the higher fee are calculated using robust standard errors that correct for overdispersion in the data. The results indicate that the higher fee has no statistically significant effect on the odds of development, nor on the expected number of homes built, if developed. The corresponding p-values of 0.77 and 0.16 are not close to reasonable cutoffs for statistical significance.

[5]The forward stepwise feature selection optimized for BIC and not Aikake's Information Criterion (AIC) because optimizing for BIC reduces the risks associated with post-selection inference, as BIC encourages simpler models less prone to overfitting the data. Additionally, due to the complexities of correcting for post-selection inference, one should only assign the canonical interpretation of p-values to the statistical significance of the treatment, as the treatment was guaranteed to be included in the resulting model.

[6]McFadden's R_M^2 metric is $1 - \ell(M)/\ell(M_0)$, where ℓ is the log-likelihood, M is the fitted model, and M_0 is the intercept-only model. This metric is standard use for hurdle models, for which other R-squared metrics like R_{KL}^2 cannot be computed. McFadden writes that as a general rule of thumb a R_M^2 between 0.2 and 0.4 reflects an excellent fit.[15]

Table 4: Hurdle Model Summary

36 count model coefficients and 10 logistic model coefficients are omitted below.

	Means Ratio		Odds Ratio	
	95% CI	P-Value	95% CI	P-Value
Higher Fee	[0.3, 4.2]	0.77	[0.8, 3.7]	0.16
			AIC	27943.58
			BIC	28417.5
			Log-Likelihood	-13,924
			R^2_M	0.65

A likelihood ratio test confirms that a negative binomial hurdle model better fits the data than a Poisson hurdle model. (See **Table 6** in the Appendix.) Further, Voung's test indicates that a hurdle model - whether the count is modelled as a Poisson or as a negative binomial - significantly outperforms the corresponding count data model without the hurdle. (See **Table 7** in the Appendix.)[7]

The results of this model are robust to different specifications of the threshold h for selecting parcels near the fee boundary. (See **Table 8** in the Appendix.) This robustness check illustrates that a larger h is associated with a lower variance (which explains the narrower confidence intervals), albeit at the cost of increased bias.[26] Yet, no matter how large one makes h, no confidence interval excludes the null hypothesis that there is no effect. This means the results of the model are robust to different specifications of h.

5 Robustness Check

The hurdle model assumes each variable contributes linearly to a parcel's probability of development and to a parcel's expected unit count, if developed. While the selected hurdle model includes six interaction effects, a critic could claim this is insufficient to control for the non-linear effects of confounding variables on housing production.

To confirm that results do not hang on a parametric assumption about the effect of any confounding variable or, say, the compounding treatment of zoning, I perform a robustness check using Double/Debiased Machine Learning, which uses flexible machine learning methods to control for the effect of confounders or the compounding treatment on hous-

[7]Zero-inflated models were not considered because there are no sampling zeros given that I defined development as a pulled permit that provides net new units. If permits with zero new units were included in the scope of the study, there would be sampling zeros, and one should then compare the goodness of fit of the hurdle model and the zero-inflated model.

ing production. This robustness check confirms the same story that a 50% reduction in fees does not register as a statistically significant constraint removal.

In the framework of Double/Debiased Machine Learning, the partially linear model for the local average treatment effect of F is:

$$log(1 + Y) = \theta F \mathbb{I}(x^d < h) + g_0(R) + \epsilon_1 \tag{5}$$

where Y is the count of units built, F is the higher fee treatment, x^d is the distance to the nearest boundary, $h > 0$ is the distance cutoff, R are the control variables, $g_0(R)$ is an unknown, non-linear function of R, and ϵ_1 is an unobserved error term with conditional expectation equal to zero. Here, I use the log(1 + Y) transformation of the response data because doing so, for this dataset, ensures the transformation back onto the original scale yields nonnegative predictions. As before with the hurdle model, I estimate the local average treatment effect of F by estimating the effect of F locally near the RDD boundary - that is, for parcels where $x^d < h$.

Further, the propensity score model is:

$$F = m_0(R) + \epsilon_2 \tag{6}$$

where $m_0(R)$ is an unknown, non-linear function of R, and ϵ_2 is an unobserved error term with a conditional expectation of zero.

There are three major modelling advantages to using Double/Debiased ML. First, with Double/Debiased ML, one can train cutting-edge machine learning algorithms to approximate the non-linear nuisance functions $m_0(R)$ and $g_0(R)$. As a result of leveraging machine learning, the outlook for controlling for complex, non-linear effects caused by confounding variables is much improved.

Second, Double/Debiased ML is doubly robust in that only *one* of the nuisance functions $m_0(R)$ and $g_0(R)$ needs to be well-modelled for the estimate to converge to the correct parameter. In contrast, the RDD hurdle model - and, indeed, most econometric methods - is only singly robust, which is to say that its estimate is only unbiased if the model is correctly specified.[27]

Third, machine learning excels with high-dimensional datasets, and so leveraging machine learning improves the prospects for approximating $g_0(R)$ and $r_0(R)$. The dataset under study contains nearly a hundred covariates, many of which are highly collinear and have only a weak relationship with the treatment and the outcome. As a result, most variables available were excluded *ex ante* or by the feature selection algorithm for the hurdle model. In contrast, I can leverage the entire dataset when using Double/Debiased ML.

To separately model $m_0(R)$ and $g_0(R)$, I train the gradient boosting algorithm CatBoost, which iteratively builds decision trees as weak learners. The initial algorithm predicts a constant value, and, at each learning iteration, is augmented with a weak learner to minimize an RMSE loss, denoted $L(y, x)$. The algorithm computes the derivatives of the loss for each sample,

$$\forall i \in \{1, .., n\}, r_{i,m} = -\frac{\partial L(y_i, f_{m-1}(x_i))}{\partial f_{m-1}(x_i))} \tag{7}$$

where predictor f_{m-1} is a predictor built using $m-1$ trees. CatBoost then builds a decision tree t_m to predict the negative derivatives of the loss using the samples $\{(x_i, r_{i,m})\}_{x_i \in [1,n]}$. The step size towards the direction of gradient descent is controlled by the learning rate hyperparameter γ. This iterative process is continued until the algorithm either converges or exhausts allotted compute.[20] Empirically, a gradient boosting approach achieves excellent results on medium-sized datasets like the one under study.[3]

Another benefit of CatBoost - so-named because it is a boosting algorithm designed for categorical data - is its handling of categorical data, which predominates in my dataset. The conventional approach of one-hot encoding categorical variables poses curse-of-dimensionality issues for my dataset as some variables, such as one for the property class code, have over a hundred possible values. Rather than one-hot encoding, CatBoost calculates target statistics for categorical values and does so without introducing a prediction shift; specifically, CatBoost randomly permutes the dataset and then computes a target statistic for observation i by calculating the target statistic for i based on its categorical value among observations $j \in \{1, 2, ..., i-1\}$, thereby avoiding a prediction shift.[20]

After tuning hyperparameters for CatBoost to separately fit models of $m_0(R)$ and $g_0(R)$ that minimize expected validation set error, this Double/Debiased ML approach yields a bootstrapped p-value of 0.18, which implies a failure to reject the null hypothesis. The 95% confidence interval for θ is [-0.02, 0.12] and covers zero, so there is no evidence that subjecting a parcel to the higher impact fee reduces housing production.

6 Discussion

Two methods with very different assumptions arrive at the same conclusion: after controlling for parcel-level variables, there is no evidence that the higher impact fee reduces housing production. The corroboration of two methods should increase our confidence in this finding.

This paper's negative findings agree with the negative finding in Mayer and Somerville. Mayer and Somerville studied forty-four metropolitan areas over a decade; while their study found significant effects from other regulations, they found no statistically significant effect for impact fees on housing production. Their finding and this paper's finding run contrary to earlier work by Skidmore and Peddle in 1998, which, unlike this paper, benefited in terms of external validity from studying 29 cities, thereby averaging out city-specific peculiarities that interact with the treatment. On the other side of the ledger, however, is internal validity: because Skidmore and Peddle lack a rich dataset and sophisticated identification strategy, there are many plausible ways in which their findings could go wrong. For example, if local opposition to housing varies jointly over time and by city, and if local opposition to housing both leads to increases in impact fees and reductions in housing production, then Skidmore and Peddle's estimate would be biased.

Yet, this paper's null finding is somewhat perplexing. If the relevant market actors are informed and rational, one would assume they would respond to a market incentive to build on parcels with the lower impact fee, all else equal.

There are at least three explanations for why this paper's finding might be null. First, the fee under investigation may simply not be large enough for its effect to be apparent. Though the neighboring parcels under study have an impact fee that differs by 50%, this difference only increases total development costs for a prototypical project - namely, a 100-unit apartment building - by around 1.2%, after accounting for carrying costs. It's possible that if all processing fees and in-lieu fees were likewise reduced by 50% that there would be a statistically discernible effect. Future research could replicate the identification method in this paper to alternative datasets that contain larger fee disparities. Nonetheless, if this is the reason for why the finding is null, then the negative finding at least would be informative for policymakers: it would imply that reducing total development costs for prototypical projects by 1.2% is negligible. Put positively, on this view, pro-supply policymakers would be well-advised to pursue reforms that reduce total development costs by more than 1.2% for the effect to register in the market.

Alternatively, perhaps there is no apparent effect from the impact fee because the study period simply lacks any parcels that are near the margin of being economically feasible to develop. On this view, there is such a wide gap between the financially feasible parcels and the infeasible parcels that reducing impact fees don't tip any parcels over the edge into feasibility. This theory, while cogent, runs contrary to the data available. For one, according to the hurdle model's predictions, there are no parcels that are likely to be developed, which undercuts the notion that the feasible parcels are so financially lucrative to redevelop that increasing impact fees wouldn't make a difference on the margin. Secondly, the occurrence of abandoned projects in San Francisco suggests that there are parcels whose viability to redevelop changes over time.

A third possible reason for the lack of impact on development could be the unpredictability and opacity of fees, making it difficult for developers to factor them into pro formas. Research shows that such fees are often non-transparent in California.[21] However, this explanation seems less likely for two reasons. First, there's no indication that this specific fee was poorly publicized. Second, even if the fee structure wasn't clear initially, its eventual imposition should still affect the likelihood of project abandonment, thus leading to fewer new housing developments overall.

A limitation of my study is its focus on per-square-foot impact fees. While the Mayer and Somerville dataset includes various fee structures, my conclusions only apply to fees based on square footage. This doesn't rule out the possibility that other fee structures, like per-unit fees, could affect housing production differently. For example, per-unit fees might lead to fewer, larger homes. Therefore, my findings don't negate the potential for reforming impact fees to increase housing production; they simply indicate that halving per-square-foot impact fees is unlikely to significantly boost it.

Halving impact fees may seem like a dramatic reduction in impact fees. The evidence provided in this paper suggests that such a policy change would not produce meaningfully more housing in cities like San Francisco. Policymakers motivated to end regional housing shortages should prioritize exercising other levers to produce housing and aim for deeper cuts to the total cost of developing housing.

References

[1] Brian Asquith, Evan Mast, and Davin Reed. "Supply shock versus demand shock: The local effects of new housing in low-income areas". In: *Available at SSRN 3507532* (2019).

[2] Vicki Been. "Impact fees and housing affordability". In: *Cityscape* (2005), pp. 139–185.

[3] Candice Bentéjac, Anna Csörgő, and Gonzalo Martínez-Muñoz. "A comparative analysis of gradient boosting algorithms". In: *Artificial Intelligence Review* 54 (2021), pp. 1937–1967.

[4] Sara C Bronin. "Zoning by a thousand cuts". In: *Pepp. L. Rev.* 50 (2023), p. 719.

[5] Alex Casey et al. *Modeling New Housing Supply in Los Angeles: Simulations from the Terner Housing Policy Dashboard*. Tech. rep. Terner Center, Oct. 2022.

[6] Planning Department City of San Francisco. *San Francisco Eastern Neighborhoods Nexus Study*. May 2008. URL: https://sfplanning.s3.amazonaws.com/archives/documents/1467-SFEastern_NexusStudy_5_21_08.pdf.

[7] Jennifer S Evans-Cowley and Larry L Lawhon. "The effects of impact fees on the price of housing and land: A literature review". In: *Journal of Planning Literature* 17.3 (2003), pp. 351–359.

[8] Edward Glaeser and Joseph Gyourko. "The economic implications of housing supply". In: *Journal of economic perspectives* 32.1 (2018), pp. 3–30.

[9] Kyle F Herkenhoff, Lee E Ohanian, and Edward C Prescott. "Tarnishing the golden and empire states: Land-use restrictions and the US economic slowdown". In: *Journal of Monetary Economics* 93 (2018), pp. 89–109.

[10] Luke J Keele and Rocio Titiunik. "Geographic boundaries as regression discontinuities". In: *Political Analysis* 23.1 (2015), pp. 127–155.

[11] Xiaodi Li. "Do new housing units in your backyard raise your rents?" In: *Journal of Economic Geography* 22.6 (2022), pp. 1309–1352.

[12] Evan Mast. "The effect of new market-rate housing construction on the low-income housing market". In: *Upjohn Institute WP* (2019), pp. 19–307.

[13] Sarah Mawhorter, David Garcia, and Hayley Raetz. "It all adds up: The cost of housing development fees in Seven California cities". In: *Terner Center for Housing Innovation at UC Berkeley. Retrieved from: https://ternercenter. berkeley. edu/development-fees* (2018).

[14] Christopher J Mayer and C Tsuriel Somerville. "Land use regulation and new construction". In: *Regional Science and Urban Economics* 30.6 (2000), pp. 639–662.

[15] Daniel McFadden. "Quantitative methods for analysing travel behaviour of individuals: some recent developments". In: *Behavioural travel modelling*. Routledge, 2021, pp. 279–318.

[16] Andreas Mense. "The impact of new housing supply on the distribution of rents". In: (2020).

[17] Ben Metcalf et al. "Will Allowing Duplexes and Lot Splits on Parcels Zoned for Single-Family Create New Homes?" In: *Terner Center for Housing Innovation. https://ternercenter. berkeley. edu/research-and-policy/duplexes-lot-split-sb-9* (2021).

[18] Kate Pennington. "Does building new housing cause displacement?: the supply and demand effects of construction in San Francisco". In: *The Supply and Demand Effects of Construction in San Francisco (June 15, 2021)* (2021).

[19] Shane Phillips. "Reducing Development Costs with Impact Fee Deferral". In: (2021).

[20] Liudmila Prokhorenkova et al. "CatBoost: unbiased boosting with categorical features". In: *Advances in neural information processing systems* 31 (2018).

[21] Hayley Raetz et al. "Residential impact fees in California: Current practices and policy considerations to improve implementation of fees governed by the Mitigation Fee Act". In: (2019).

[22] San Francisco Planning Department. *San Francisco Citywide Development Impact Fee Register*. Updated as of December 1, 2022, rates effective as of January 1, 2023. 2022. URL: https://sfplanning.org/sites/default/files/forms/Impact_Fee_Schedule.pdf.

[23] SF Planning Department. *Appendix B: Sites Inventory and Rezoning Program*. In ADOPTED HOUSING ELEMENT 2022 UPDATE, adopted January 31, 2023. City and County of San Francisco. 2023. URL: https://sfplanning.org/sites/default/files/documents/citywide/housing-choice/housingchoice_element_appendixB_sites_inventory.pdf.

[24] Mark Skidmore and Michael Peddle. "Do development impact fees reduce the rate of residential development?" In: *Growth and change* 29.4 (1998), pp. 383–400.

[25] Jake Wegmann. "Death to single-family zoning... and new life to the missing middle". In: *Journal of the American Planning Association* 86.1 (2020), pp. 113–119.

[26] Jeffrey M Wooldridge. *Econometric analysis of cross section and panel data*. MIT press, 2010.

[27] Paul N Zivich and Alexander Breskin. "Machine learning for causal inference: on the use of cross-fit estimators". In: *Epidemiology (Cambridge, Mass.)* 32.3 (2021), p. 393.

Appendix

Table 5: Description of key explanatory variables in panel dataset.

Category	Variable Name	Description
Location attributes	X Coordinate	The horizontal coordinate of the parcel
	Y Coordinate	The vertical coordinate of the parcel
	Neighborhood	One of 97 neighborhoods
	District	Supervisorial district
	Local Home Price	Zillow neighborhood estimate
Lot attributes	Lot Area	Square footage of lot
	Partly Steep Lot	Part of the lot is on a 20 degree slope
	Entirely Steep Lot	Entire lot is on a 20 degree slope
	Lot Shape	Either square, rectangular, or ’other’
	Lot Depth	Distance from street to the back of the lot
	Lot Frontage	Length of lot along the street
Zoning	Buildable Envelope	Square feet (in 1000) allowed per zoning
	Upzone Ratio	Buildable sq ft divided by existing sq ft
	Residential 1	Indicates one home is permitted on lot
	Residential 2	Indicates two homes is permitted on lot
	Residential 3	Indicates three homes is permitted on lot
	Residential 4+	Indicates 4+ homes is permitted on lot
	Form-Based	Indicates city regulates building size, not # of homes built
	Public	Indicates park space or public facility
	Office/Commercial	Indicates office or commercial zoning

Continued on next page

Table 5 – continued from previous page

Category	Variable Name	Description
	Industrial	Indicates zones for industrial uses
	Redevelopment	Indicates zones targeted for revitalization
Existing Use	# of Units	Number of units if residential
	# of Stories	Number of stories in existing building
	# of Bedrooms	Number of bedrooms in existing building
	# of Bathrooms	Number of bathrooms in existing building
	Property Area	Square footage of building
	General Use Code	Indicates hotel, retail, residential, office, industrial, or government use.
Property Tax	Land Value	Value of land on its own
	Improvement Value	Value of building(s) on top of land
	Fixtures Value	Value of permanent add-ons to structure
	Personal Property Value	Value of personal items
	Homeowner Exemption Value	Value of property exempt from tax
Property History	Year Built	Year the building was built (or renovated)
	Years Since Last Sale	Year of last sale
	Historical Status	Indicates if parcel is a historical resource
Lag Permits[a]	New Construction	# of permits for new construction
	Additions/Alterations	# of permits for large changes
	Signs	# of permits to add a sign
	Excavate	# of permits to excavate
	Demolitions	# of permits to demolish space

Continued on next page

Table 5 – continued from previous page

Category	Variable Name	Description
	OTC Alterations	# of over-the-counter permits
		[a] Refers to permit applications filed for a parcel in the preceding 5 years.

Table 6: Likelihood Ratio Tests

Model	df	LogLik	χ^2	χ^2 P-Value
Poisson	25	-52615		
Negative Binomial	26	-11222	82786	< 2.2e-16
Hurdle (Poisson)	37	-14678		
Hurdle (Negative Binomial)	38	-10831	7694.5	< 2.2e-16

Table 7: Evidence for Hurdle Model per Vuong Test

	Vuong	p-value
Negative Binomial vs NB Hurdle	-11.8	< 2.22e-16
Poisson vs Poisson Hurdle	-10.3	< 2.22e-16

Table 8: Robustness Check: Varying Threshold h for RDD

		Odds Ratio		Means Ratio	
h	m	95% CI	p-value	95% CI	p-value
125m	5314	[0.704, 2.012]	0.516	[0.142, 11.694]	0.823
100m	4707	[0.720, 2.088]	0.452	[0.179, 8.096]	0.849
75m	4058	[0.712, 2.095]	0.468	[0.153, 8.004]	0.921
50m	3139	[0.788, 2.438]	0.257	[0.027, 29.633]	0.953
25m	2014	[0.741, 2.651]	0.299	[0.070, 13.690]	0.986
15m	1413	[0.839, 3.374]	0.143	[0.094, 9.836]	0.975
10m	1243	[0.840, 3.812]	0.131	[0.011, 78.242]	0.966
5m	1037	[0.679, 3.263]	0.320	[0.005, 148.214]	0.945
1m	984	[0.670, 3.229]	0.337	[0.005, 148.085]	0.945