The Redevelopment Option and Land Use Restrictions Over Space and Time

Draft: November 11$^{\text{th}}$, 2019

For latest draft, click here.

Abstract

We incorporate uncertainty surrounding future land-use restrictions to empirically assess the option value of redevelopment embedded in real estate prices for New York City (NYC) from 2003-2015. Using a two-stage estimation procedure, we interact predicted probabilities of land-use (re)zoning to either residential, commercial or manufacturing with an additional proxy for the property's redevelopment propensity. Over the period spanning 2003 to 2015, estimates of the average option value to redevelop in Manhattan and Brooklyn are 20% and 8.5% of total estimated property value, respectively. There is also evidence that manufacturing lots identified as likely to be rezoned by the model sell at a premium of up to 50% per square foot. Lastly, there is evidence that the option value as a percentage of total property value is counter-cyclical.

Keywords: Land Use Restrictions, Real Options, Hedonic Pricing. Real Estate

JEL Codes: D120, D460, G130, R310.

1 Introduction

Few studies focus on zoning regulation's impact on property values through the lens of an options pricing framework. Intuitively, land that is legally allowed to be developed into multiple uses should fetch a higher price than the same plot of land restricted to a single-use. Childs et al. (1996) and Geltner et al. (1996) provide a theoretical foundation in the options pricing framework to cement the above intuition. Yet, no studies have attempted to estimate how potential changes in land use restrictions impact the option value to redevelop real estate.

Zoning policy typically regulates two dimensions of property characteristics: intensity and use. Intensity corresponds to the allowed size and dimensions a structure can be built on a given plot of land. One justification for such restrictions is to allow adequate sunlight to penetrate to the street.[1]

The second dimension restricts the allowed use of the built property. The justification for regulating land use is to minimize the external costs that certain building types impose on nearby tenants and to maximize external benefits resulting from agglomeration effects.[2] For example, living next to an industrial factory creates both air and sound pollution, which detracts from the welfare of the neighboring residents. Thus it is optimal to allocate certain areas of land for industrial use far away from residential areas. Similarly, centrally located businesses within an industry can minimize the cost of travel required for face-to-face interaction.

Michael Bloomberg was the mayor of the City of New York from 2002 through 2013. Over that period, the Bloomberg Administration rezoned approximately 40% of all the properties in the city. As Dan Doctoroff, the former deputy mayor of economic development for the City of New York, noted in an interview:

> We believed New York's economy needed a massive jolt of energy and a push

[1] Bliss (2016)

[2] Strange and Rosenthal (2001)

> forward to prepare it for the 21st century. If you looked across the city, there were thousands of acres of land that were unused, largely because they were zoned for industrial or manufacturing uses for which could not compete anymore. Many were on some of the most valuable real estate in New York, like along the waterfront, on the West Side of Manhattan, or near Downtown Brooklyn.... In the end, we did 140 separate rezonings during the Bloomberg administration. 40 percent of the city was rezoned.[3]

We take advantage of the numerous rezonings that occurred during the Bloomberg administration to examine how the option value to redevelop real estate fluctuates with changes in zoning designations that correspond to the allowed land use. While many studies in the Urban Economics literature have attempted to disentangle the impact zoning has on property values,[4] the they focus on the impact of changes in the allowed intensity of the building.

To address the question of how changes in the land use component of zoning regulations impact property values, we estimate time-varying propensities of being zoned to each land use designation to proxy for the probabilities of switching zoning designation in the future. We then incorporate these probabilities into the hedonic pricing model of Rosen (1974) to identify how changes in the long-run likelihood of being rezoned impacts the redevelopment option via the framework employed in Clapp and Salavei (2010). Variation in expectations of future zoning designation around major rezoning events is employed in an hedonic estimation to identify the redevelop option component of real estate values.

We find evidence that, on average, the redevelopment option constitutes 20% of total estimated property value in Manhattan and 8% in Brooklyn. There is also evidence that manufacturing-zoned properties identified by our model as wholly residential sell for a 50% premium per square foot. Lastly, we present evidence that the redevelop option as a percent of the total property value is counter-cyclical with the real estate cycle.

[3] Florida (2017)

[4] McMillen and McDonald (1991), McMillen and McDonald (2002), Glaeser and Gyourko (2003), Ihlanfeldt (2007)

The rest of the paper is organized as follows: Section 2 presents the motivation for the research and presents a motivation model to accompany the empirical analysis. Section 3 discusses the data sources and summary statistics. Section 4 presents the empirical framework and identification strategy. Section 5 presents and discusses the results of the estimation. Finally, Section 6 concludes.

2 Motivation

The value of a property, $PV_{i,t}$, can roughly be separated into the value of the land, $V^L_{i,t}$, and the building that sits on the land, $V^B_{i,t}$. This relationship holds exactly if the construction industry is assumed to be perfectly competitive:

$$PV_{i,t} = V^{(B)}_{i,t} + V^{(L)}_{i,t}. \tag{1}$$

The value of the building, $V^{(B)}_{i,t}$, is the replacement cost to build a structure as similar as possible to the one that already exists on the land. This value tends to decrease with depreciation, $D_{i,t}$, increase with renovations, $I_{i,t}$, and size, $SIZE_{i,t}$. The main driver is the cost of construction, $C_{i,t}$, which includes labor and capital expenses.

$$V^{(B)}_{i,t} = f(I_{i,t}, D_{i,t}, C_{i,t}, SIZE_{i,t}) \tag{2}$$

Most of the variation in property values comes from the value of the land that the structure sits on (Davis and Heathcote (2005), Knoll et al. (2017)). In the urban economics literature, the value of land is thought to represent a call option on development (or redevelopment).[5] When choosing to develop land, the developer must choose both the use of the land and the intensity to which to develop the land. Childs et al. (1996) and Williams (1997) examine the option to continually redevelop a property. Childs et al. (1996) focuses on the option to reconfigure a property between two uses (as well as mixing the two uses), while

[5]Titman (1985)

Williams (1997) focuses on the option to increase the intensity or quality of the property while holding its use constant. Both Childs et al. (1996) and Williams (1997) find that the option to redevelop decreases with an increased correlation between market value of uses and increases with uncertainty regarding future rents.

Zoning restrictions limit both land use and building intensity. Thus a change in zoning designation will have implications for the price of a property. In New York City, a rezoned property can continue to operate under the current use, but significant renovations or redevelopment cannot occur without switching land use to the new designation.[6] When making the redevelopment decision, the developer must consider the expected net payoff of redeveloping the property to its highest and best use today, as well as the net payoff of delaying the redevelopment into the future.

The critical insight of this paper is to recognize that the value of the highest and best use of a property today is limited by the zoning designation today, and the *expected* value of the highest and best use tomorrow is a function of *expected* zoning designation tomorrow. In other words, the zoning designation of any property in the future is uncertain, and the value of the (re)development option should reflect this uncertainty.

Consider a simple two-period example where a landowner is deciding to develop vacant land under uncertainty with regards to the zoning designation. There are two possible zoning designations, $z_t \in \{1, 2\}$. Today the zoning designation is use i, $z_1 = i$. Tomorrow, there is some probability, $\alpha > 0$, of being rezoned to use $j \neq i$, and a probability of $(1 - \alpha) > 0$, of keeping the same zoning designation. From the perspective of the landowner, the choice to develop will depend on weighing the payoff of developing the property to its highest and best use today for property type i, HBU_i, with tomorrow's expected payoff and with the possibility of being able to develop to use j. In this situation, the value of land is

$$V_{i,t}^{(L)} = \max \left\{ NPV_i, \frac{(1-\alpha)}{(1+r)} NPV_i + \frac{\alpha}{(1+r)} NPV_j \right\}, \tag{3}$$

[6] New York City Zoning Resolution, Article 5, Chapter 2, 52-20. https://zr.planning.nyc.gov/article-v/chapter-2/52-20.

where NPV_i is the net present value of developing to the highest and best use given the zoning designation today is i, and $r > 0$ is the rate of return on the built property (assumed constant across uses). The developer will only choose to wait if

$$NPV_j - NPV_i \geq (\frac{r}{\alpha})NPV_i. \tag{4}$$

That is, the developer will wait to the later period to develop for the chance event of the zoning designation changing only if the developed property is expected to be more valuable under use j by an amount larger or equal to the rents foregone by waiting a period to develop, $r \cdot NPV_i$, divided by the probability of switching zoning designation, α. This hurdle rate, $\frac{r}{\alpha}$, is decreasing in the probability of switching zoning designations and increasing in the opportunity cost of waiting to develop. For the special case where a zoning change is certain, $\alpha = 1$, the value of the property if it were to be developed to use j must be at least as large as the opportunity cost of waiting. As the probability of being rezoned becomes extremely unlikely, the hurdle value gets larger and the developer will never choose to wait. The key takeaway from this toy model is that the probability of being rezoned, something that occurs with a non-insignificant frequency in New York City from 2003-2015, will impact the option value of redevelopment. The intuition of the toy model above is easily extended to the case of redevelopment.

In many geographies the zoning process is static. In New York City, we find that 2% of all properties in the beginning of the sample switched zoning designation at the end of the sample in 2015. Table 1 details the number of zoning changes that occurred for all lots present in the beginning of the sample. Figure 1 shows the nature of each transition and where they occurred across space.

Out of 294,040 lots, 5,758 (or roughly 2%) of lots changed land use designation. On net there was an increase in residential and mixed use zoned lots, a small decline in commercial lots, and a 20% decline in manufacturing lots. Yet, there is significant variation in changes conditioning on the initial zoning designation. For example, even though there was a steep decline in manufacturing lots overall, 84 residential and 30 commercial lots were rezoned to be manufacturing.

Out of the 16,149 commercially zoned lots at the beginning of the sample, 14,598 (90%) stayed commercial, 1,466 (9%) switched to residential, and the remaining percent is split relatively evenly between manufacturing and mixed-use manufacturing /residential. For the 13,325 lots zoned manufacturing at the beginning of the sample, where we see the most zoning changes, 10,538 (79%) stayed manufacturing, 865 (6.5%) switched to commercial, 1,922 (14.5%) switched to residential or mixed use manufacturing/residential. Most residential lots stayed residential, but a not insignificant amount of residential lots did switch. Out of the 263,365 lots zoned residential at the beginning of the sample, 595 switched to commercial and 84 switched to manufacturing. The big picture that the data presents to us is that there was a large migration from commercial and manufacturing to residential, with most zoning changes coming from manufacturing zoned lots. However, the changes are not monotonic, and in a given quarter, we may observe a single change from residential to manufacturing.

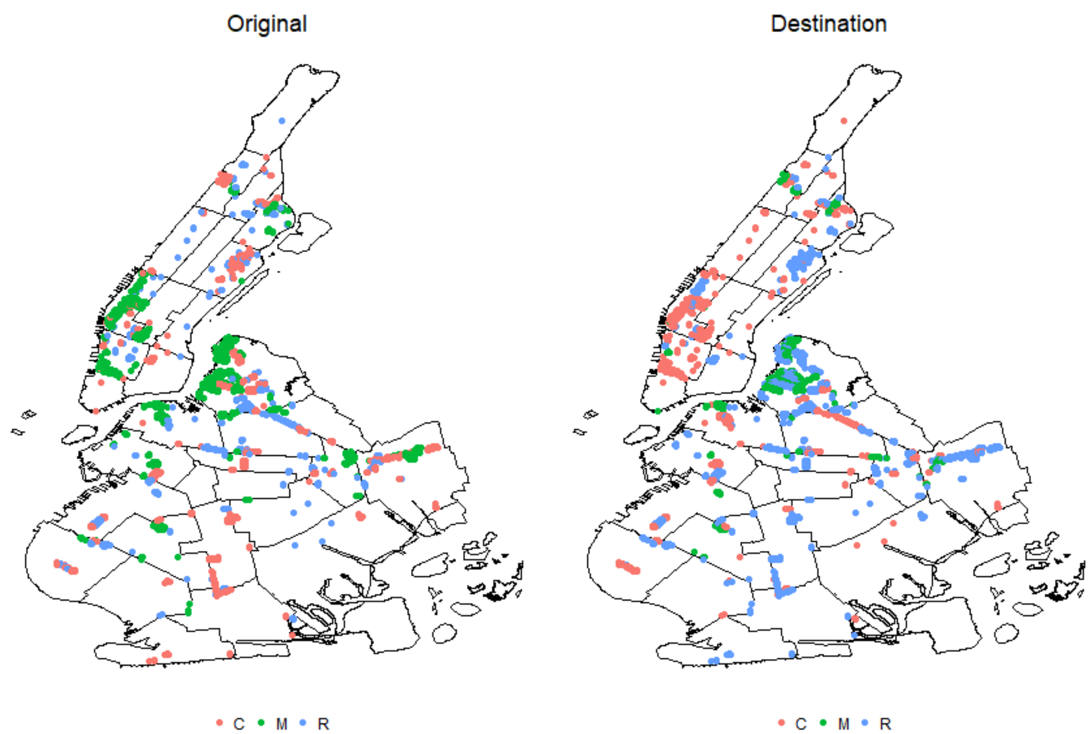
2.1 Motivational Model

In reality, the assumption of perfect competition in the construction industry doesn't hold, and we cannot separate the value of land from the structure that sits on the land so easily. A more true-to-life model must take into account uncertainty regarding when the existing structure will be redeveloped and average the net present value of cash flows from the current

Table 1: Zoning changes that occurred or did not for all lots present in the beginning of the sample from all boroughs from 2003-2017. Res/Man refers to a mixed zoning designation where the property can be used for both residential and manufacturing purposes.

Source	Commercial	Manufacturing	Residential	Res/Man	Total
Commercial	14,598	30	1,466	55	16,149
Manufacturing	865	10,538	935	987	13,325
Residential	595	84	262,683	3	263,365
Res./Man.	24	1	713	463	1,201
Net Change	(-67)	(-2,672)	(2,432)	(307)	

Figure 1: Zoning transitions in Brooklyn (bottom) & Manhattan (top) from 2003 -2015. The figure on the left shows the original zoning designation of each property, and the figure on the right shows the zoning designation each property switched to. The dots are colored by zoning designation; red being commercial, green being manufacturing, and blue being residential.



structure over various redevelopment horizons.

First, assume that macro-fundamentals, X_t, are exogenous and drive dynamics in both the rental market for real estate and the market for selling real estate. Macro-fundamentals also drive time-variation in the zoning process, $Z_{i,t} \in \{C, M, R\}$, and the zoning process follows a time-varying Markov chain conditional on vocational characteristics, $q_{i,t}$:

$$Z_{i,t} \sim \Pi(X_t, q_{i,t}), \tag{5}$$

where C represents commercially zoned lots used for office and retail space, M represents manufacturing zoned lots which are typically warehouse structures used for production, storage, or logistic centers, and R represents residential zoned lots which is any structure for which people reside, including apartments.

Each property owner either chooses to redevelop, $R_{i,t} = 0$, or to wait, $R_{i,t} = 0$, with the objective of maximizing the value of his property. If he chooses to wait, his building gets older, which will impact his expected cash flow next period. If he chooses to redevelop, the new use of the building must conform with the current zoning designation, $LU_{i,t} = Z_{i,t}$, where $LU_{i,t}$ is the operational use of the building; he must pay the cost of redevelopment, K, upfront; and he must wait d periods to begin collecting cash flow from the new building,

$$age_{i,t} = \left\{ \begin{array}{ll} age_{i,t-1} + 1, & \text{if } R_{i,t} = 0 \\ -d, & \text{if } R_{i,t} = 1, \end{array} \right\}. \tag{6}$$

$$CF_{i,t} = \mathbb{1}(age \geq 0) \exp\{\kappa(X_t, LU_{i,t})q_{i,t} + \gamma(X_t, LU_{i,t})age_{i,t} + \phi(X_t, LU_{i,t})age_{i,t}^2 + w_{i,t}\} - R_t K, \tag{7}$$

$$LU_{i,t} = Z_{R_{i,t}^{-1}} \text{ s.t. } R_{i,t}^{-1} := \{\max \tau \leq t : R_{i,\tau} = 1\}. \tag{8}$$

In Equation (7), $q_{i,t}$ represents any locational characteristics that may generate a premium or discount on cash flows, as well as building characteristics independent from the age of the structure. The coefficients κ, γ, and ϕ are the implied value of product characteristics as they relate to rents. They are time-varying because rental markets respond to macro-fundamentals, and the implied prices map into current equilibrium rents. Also there are different coefficients for each land use as there are independent rental markets for residential, commercial, and manufacturing properties. Cash flows are log-normally distributed, $w_{i,t} \sim N(0, \sigma_w^2)$.

Equation (8) simply states that the current land use at any building is the zoning designation at time of last redevelopment. This assumption is justified by New York City's zoning policies, which state that when a zoning change occurs, a building can operate under its former use until a redevelopment occurs.

Finally, the price of a property is pinned down by the assumption of no arbitrage in the market for selling real estate, implying prices are equal to risk-adjusted discounted expectations of future cash flows:

$$P_{i,t} = \max_{R(\cdot)} \sum_{s \geq t} \mathbb{E}_t^Q \big[\beta_s CF\big(X_s, Z_{i,s}, LU_{i,s}, q_{i,s}, age_{i,s}(R(\cdot)), R(\cdot)\big)\big]. \tag{9}$$

We can then break the price of a property into expected discounted cash flows before the first redevelopment and the expected value of cash flows after the first redevelopment. Let $R^1_{i,t} := \{\min \tau \geq t : R_{i,\tau} = 1\}$ be defined as the period when the property is first redeveloped after period t. Then,

$$\begin{aligned} P_{i,t} = {} & \overbrace{\mathbb{E}_t^Q \big[\sum_{t \leq s < R^1_{i,t}} \beta_s CF\big(X_s, LU_{i,s}, q_{i,s}, age_{i,s} = age_{i,s-1} + s - t, R^*(\cdot) = 0\big)\big]}^{V^{(B)}_{i,t}} \\ & \underbrace{+ \mathbb{E}_t^Q \big[\sum_{R^1_{i,t} \leq s} \beta_s CF\big(X_s, LU_{i,s}, q_{i,s}, age_{i,s}, R^*(\cdot)\big)\big]}_{V^{(L)}_{i,t}}, \end{aligned} \tag{10}$$

where

$$R^*(\cdot) := \arg\max_{R(\cdot)} \sum_{s \geq t} \mathbb{E}_t^Q \left[\beta_s CF\big(X_s, Z_{i,s}, LU_{i,s}, q_{i,s}, age_{i,s}(R(\cdot)), R(\cdot)\big)\right]. \tag{11}$$

The remaining portion of this paper focuses on empirically separating the first term from Equation (10) - henceforth *current structure value* - representing the expected discounted cash flows from prior to the first future redevelopment, and the second term - henceforth *option value to redevelop* - which is the expected future discounted cash flows post first redevelopment.

3 Data

The panel dataset used is constructed from various data sources described below. The final dataset includes 48,929,331 observations. Each observation represents a single lot-quarter from 2002q2 until 2017q4. For each observation, there are a total 109 descriptive variables representing the characteristics of the building, the geographical characteristics of the area around the building, as well as demographic information at the census tract level.

Below, we describe the sources of the data. The Primary Land Use Tax Lot Output dataset is used to track zoning designation as well as a wealth of other property and locational characteristics. It is used in both the estimation of (re)zoning propensities and the hedonic sales model. New York City's Rolling Sales dataset reports all property transactions and is used in the hedonic estimation. The policy dummies in the first stage come from the New York City Department of City Planning's Zoning Application Portal, which was scraped to compile the dataset. Lastly, both the American Community Survey data as well as Bing Maps data is used in post-estimation to decompose option value by location and demographics.

Selected summary statistics are presented in Table 1

3.1 Primary Land Use Tax Lot Output

The main dataset employed is the Primary Land Use Tax Lot Output (PLUTO), provided by the New York City Department of City Planning. It contains extensive land use and geographic data for the vast majority of the tax lots in NYC. The panel dataset was released on an annual basis from 2002-2007, but since 2009, it has been released on a semi-annual basis.

PLUTO includes over 70 fields for each tax lot. Of main interest to this paper is extensive information on zoning designations for each property. Zoning is broken down into primary zoning district, overlay district, and special purpose districts. Each category can contain multiple items ordered by the share of the tax lot from largest to smallest.

Other variables of interest include the Floor Area Ratio (FAR) of a building broken down into commercial, residential, office, retail, garage, storage, and factory. There are a number of building descriptors such as number of floors of the building (NumFloors), the area of the lot (LotArea), an identifier for an irregular lot, an identifier for special lot characteristics such as waterfront property, basement type, year built (YearBuilt), the last year the building was significantly altered (LastAlter), and the separately assessed value of the land (AssessLand) and building (AssessBldg) as estimated annually by the Department of Finance for tax purposes.

The process of merging all versions of PLUTO together is no trivial task. The variables included with each release can change from year-to-year, so special care must be taken so that all variables are comparable. The merged dataset contains 48,929,331 million rows, and then we create a quarterly time series from each observation date. For example, if a there is an observation from a property in 2002q1 and 2003q4, I duplicate the 2002q1 row to create observations for 2002q2 -2003q3.

3.2 Transaction Data

Property transactions come from the New York City Department of Finance's Rolling Sales data. The Rolling Sales dataset is released on an annual basis from 2003-present. It contains all property transactions that take place in the city. In addition to a subset of variables from the PLUTO dataset, the Rolling Sales data contains the nominal transaction price of the property (SalePrice), as well as the gross square footage (GrossSqFt) of the building and the square footage of the lot the building sits on (LandSqFt). Gross square footage is defined as the total floor area encompassed by the exterior walls of a property. The land area is then the remaining square footage of the lot.

Real estate property transactions are notoriously noisy. In order to remove outliers, for each year and borough I regress the logarithm of sale price on the logarithm of gross square ft. interacted with census tract – a granular measure of geography – along with other building and lot characteristics. I then remove all observations whose studentized-residual is greater than 3.7. Finally, I do a visual inspection for each year and borough to remove all clear outliers that may impact estimation in the hedonic model.

3.3 American Community Survey

The American Community Survey (ACS) is an annual survey performed by the US Census Bureau that collects demographic, economic, and housing data for neighborhoods across the US. The survey commenced in 2005, but for smoothing purposes, data is only available from 2009 -present. Each observation represents a 5-year moving average. The data we use is on the census tract level. To get an idea of how granular a census tract is, NYC is broken up into 1,335 individual census tracts; this is far more granular than neighborhoods or even community districts.

Demographic data includes median age (MedAge) and the racial makeup of each neighborhood broken down into white, African-American, Asian, Native American, and other. The "other" designation is largely made up of the Latin American population.

Economic data in the ACS includes the percentage of households with children, the percentage of households that report using public transportation as their main transportation to work, the percentage of those households that primarily walk to work, mean transit time to work, median household income (MedHHIncome), the percentage of those reported being unemployed (%Unemployed), the percentage of households fully covered by health insurance, the percentage of households below the poverty line, and the percentage of households with an annual income greater than $100,000.

The housing portion of the ACS includes the estimated vacancy rate, the percentage of residents who rent (%Rent), the percentage of residents without a vehicle, and the median gross rent of the area (MedGrossRent).

Finally, the education section includes the percentage of those with a less than ninth grade primary school education, the percentage with somewhere between a ninth to twelfth grade education, the percentage of those without a high school diploma, the percentage of those with a bachelor's degree of higher, and lastly, the percentage of residents born in the US.

3.4 ULURP Data

Data concerning the Uniform Land Use Review Process was collected by scraping data from the DCP's newly released Zoning Application Portal (ZAP). ZAP is a web tool that allows the public to track the progress of current and past zoning applications through the ULURP process. The information includes the date that each step of the ULURP was cleared, as well as the status of each step.

From the above information I am able to "dummy out" the ULURP process, creating six review-stage legislative dummies: one for the quarter in which the application was certified (CertDum), one for the community board recommended approval or disapproval (ResCom-Dum), one for the approval of the borough president, one for if the DCP has held its public forum, one for if the DCP approved the application, and lastly, and one for if the City

Council voted to pass the application into law.

I match each zoning application by its identifier to lot identifiers (BBLs) using the shape files released by the Department of City Planning for zoning map amendments. Using post-GIS and a shape file version of PLUTO called mapPLUTO, I find all properties that intersect the ZMA shape file for each version of MapPLUTO before and after the zoning amendment went into effect. I then check that each intersecting property did indeed change zoning designation and remove those that do not.

3.5 Bing Maps Data

Finally, we conjecture that simple network distance to a city center is not sufficient to capture the complexities of NYC's transit system. The subway system is extremely important to getting around in New York City. Mainly, we believe that residents value how easily they are able to travel to midtown Manhattan, from which you can get anywhere else in the city. Yet, transit times will also be correlated with distance to nearest subway or bus station, how many subway lines are in proximity to a property, and the idiosyncrasies of each subway line.

To capture such complexities in the data, we opt to use the Bings Map API to calculate transit time using either the subway or bus to Grand Central Station. In addition to being a central point in New York City's local subway system, it is also the meeting point of other major transportation systems such as New Jersey Transit, Long Island Rail, and the Metro North system.

Lastly, in order to handle the fact that transit times vary across time of day, I average the predicted transit times at 9am, 12pm, 5pm, and 9pm for all weekdays in a fixed week in August 2018. One implicit assumption is that estimated transit times in August of 2018, are representative of the entire sample.

3.6 Data Discussion

A brief discussion of the data is presented below.

Selected summary statistics for the full panel dataset are reported in Table 2. This sample includes all observations for all boroughs and all years.

Figure 2 presents the range of sale price to gross square feet between the 33^{rd} and 66^{th} percentiles in Manhattan and Brooklyn. There appears to be a common cyclical component between the two boroughs, but real estate prices in Manhattan have increased far more than Brooklyn over the sample period.

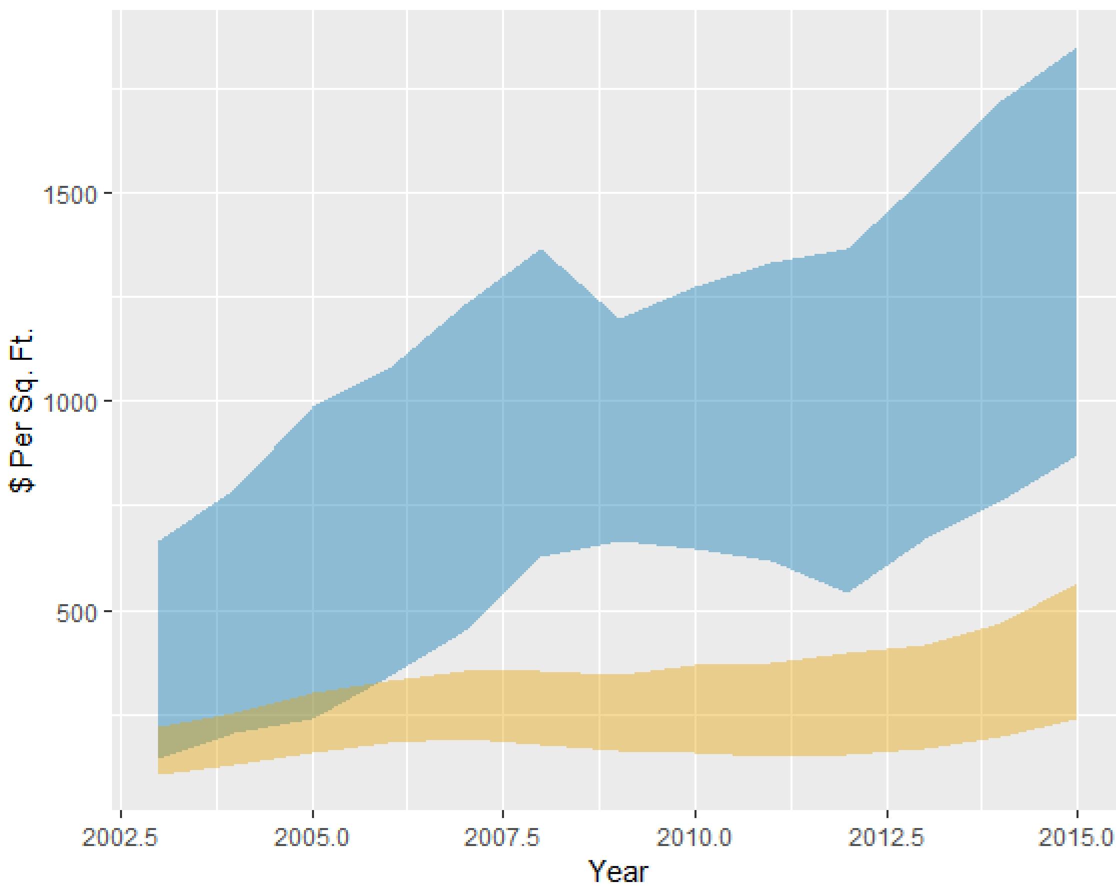


Figure 2: This figure shows how sale price per sq. ft. in Manhattan (top) and Brooklyn (bottom) has evolved through time. The lower bound of the range is the 33^{rd} percentile, and the upper bound is the 66^{th} percentile.

Table 2: Selected summary statistics.

Statistic	N	Mean	St. Dev.	Min	Max
SalePrice	470,565	1,309,629	16,215,434	10,040	4,111,111,766
GrossSqFt	470,565	4,270	35,125	10	8,942,176
LandSqFt	470,565	3,266	16,202	0	7,649,136
NumFloors	48,929,331	2.4	1.77	0	104
BldgArea	48,929,331	6,469	65,238	0	49,547,830
LotArea	48,929,331	5,165.	71,089	1	29,305,534
YearBuilt	48,861,359	1938	27.42	1661	2017
MetroTimeGCS	48,929,175	18.92	8.39	0.07	51.83
UR	21,951,046	6.286	3.359	0.000	62.700
MedHHIncome	21,939,060	62,398.91	23,743.18	8,694	247,167
%Rent	21,947,586	52.11	25.17	0	100
MedGrossRent	21,413,884	1,294.56	288.22	231	3,479
MedAge	21,949,333	37.22	5.74	17.3	94
HSGradRate	41,108,638	60.85	11.33	27	85.6

4 Empirical Framework

Clapp and Salavei (2010) develop a method to identify the option value to redevelop a property within the hedonic model. They posit that using the age of a structure as a proxy for structural depreciation produces a biased estimate for net depreciation because structure age might be positively correlated with the option value to redevelop the property. The intuition is that as a building ages, depreciation causes the value of the property to decline. In addition, as the building ages, the difference between the value of the property if it were developed for its highest and best use and the current property value increases, causing the option value to redevelop to increase as a share of total property value. These opposing effects make building age a biased proxy for net depreciation. In order to separate the two effects, Clapp and Salavei (2010) suggest using a proxy for the likelihood of redevelopment. The proxies proposed correspond to various measures of intensity. Examples are the fraction of nearby sales that are teardowns, the ratio of assessed structure value to assessed land value, and the ratio of average floor-area-ratio (FAR) of nearby new construction to FAR of the property considered.

In the context of the example above, the proxy of Clapp and Salavei (2010) captures the relationship between the built structure, and the option value to redevelop. The closer to redevelopment a property is, the greater the difference between the rate of return for the property built to its highest and best use, r_{HBU}, and the current rate of return, r. Simultaneously the value from the current structure, $V_{i,t}^{(B)}$, decreases as cash flows are relatively low and redevelopment is expected to commence soon so there are few remaining cash flows from the current structure. The redevelopment value, $V_{i,t}^{(L)}$, increases as redevelopment becomes more likely because the rents foregone by redevelopment, $r \cdot NPV_i$, decrease. In the context of the motivational model above, interacting this paper's land use uncertainty proxy with the proxy of Clapp and Salavei (2010) should increase identification of the option value to redevelopment by capturing both the opportunity cost that works through foregone rents, r, as well as through rezoning probabilities, α. As an additional contribution, identification comes from exogenous variation in the (re)zoning propensities caused by nearby zoning changes.

Munneke and Womack (2017) extend the framework in Clapp and Salavei (2010) by directly estimating the probability of redevelopment in a probit regression, and using those probabilities directly in place of the proxy used by Clapp and Salavei (2010). They make further improvements by addressing selection bias via Lee et al. (1982).

To address how the uncertainty regarding zoning designation in the future impacts the redevelopment option, I propose a two-stage approach where in the first-stage we estimate a multinomial logistic model to capture the propensity in the future to be restricted to each land use over a two-year horizon. I then interact the generated propensities from the first stages in the hedonic model with the redevelopment proxy as proposed by Clapp and Salavei (2010).

4.1 First Stage: Zoning Process

We model the long-run propensities to be (re)zoned as a multinomial logistic model at each time point. In addition, to standard locational and structure characteristics, I incorporate dummy variables in the logistic regression which correspond to the various stages of the Uniform Land Use Review and Procedure. The formal model is

$$\ln \frac{Pr(\Psi_{i,t} = \phi_z | X_{i,t})}{Pr(\Psi_{i,t} = \phi_S | X_{i,t})} = \alpha_z X_{i,t}, \tag{12}$$

where $\Psi_{i,t}$ is a categorical variable that identifies the zoning designation lead two years into the future, ϕ_z is the vector whose z^{th} element - which indexes the zoning designation - is equal to unity, and all other elements are equal to zero. $X_{i,t}$ is the vector of explanatory variables. These include individual property and locational characteristics, legislative variables from the ULURP, as well as census tract-specific variables that measure how the land use distribution of the census tract deviates from the community district and borough.

The long-run (re)zoning propensities are non-linear functions of the exogenous parameters:

$$\tilde{\eta}_{i,t,z} := Pr(\Psi_{i,t+2} = \phi_z | X_{i,t}) = \frac{\exp(\alpha_z X_{i,t})}{1 + \sum_{s=1}^{S-1} \exp(\alpha_s X_{i,t})}. \tag{13}$$

Note that I call the generated regressors long-run propensities because the model does not include current zoning designation as a regressor. The zoning process is highly persistent, and if one truly wanted to know the zoning designation a process a year from any given date the best predictor would be the current zoning designation. What we attempt to estimate is akin to the ergodic or zoning distribution unconditional of a property's use today. In the context of the presented model, we are not trying to estimate the rows of $\Pi(X_t)$, which correspond to one-step-ahead transitions probabilities, but instead what $\Pi(X_t)$ converges to in its limit,

$$\lim_{t\to\infty} \Pi(X_t, q_{i,t}) = \Pi^*(X_t, q_{i,t}).$$

It can be instructive to think of $\Pi^*(X_t)$ as the optimal land use distribution for the city that moves with underlying market and economic conditions. However, because there are significant frictions associated with rezonings and redevelopments, the city planners are always *chasing* the optimal distribution.

4.2 Second Stage: Hedonic Model

The final stage is a linear regression of the logarithm of sale price per square foot on building characteristics, locational characteristics, and the long-run zoning propensities interacted with the redevelopment proxy. The model can be viewed as a hedonic model of real estate prices where propensity to be redeveloped as well as the propensity to be zoned to each use are unobserved building characteristics that are priced.

To go from Equation (12), an infinite sum of expectations of log-normally distributed cash flows, to a log-linear equation for prices, we rely on the Fenton-Wilkinson approximation of Fenton (1960), which states that an infinite series of log-normally distributed random variables approximates a log-normal distribution.

Equation (14) is the measurement equation that is taken to the data:

$$\ln P_{i,t} = a_{i,t} + v_t' q_{i,t} + \sum_z \gamma_z \eta_{i,t,z} LSA_{i,t} + u_{i,t}, \tag{14}$$

where $LSA_{i,t}$ is the proxy for redevelopment defined as

$$LSA_{i,t} := \frac{\text{Assessed Land Value}_{i,t}}{\text{Assessed Total Value}_{i,t}}. \tag{15}$$

The idea is that redevelopment option consists of an intensity dimension and a potential rezoning dimension. By interacting the probability of being rezoned to each designation next year with the intensity proxy, we can understand how the potential to be rezoned impacts the option to redevelop.

The option value to redevelop property i is then computed as

$$OV_{i,t} := \exp\big(\hat{a}_t + \hat{v}_t' q_{i,t}^0 + \sum_{j=1}^{S} \gamma_{t,j}\hat{\eta}_{i,t,j} LSA_t + \frac{\sigma_{u,i}^2}{2}\big) - \exp\big(\hat{a}_t + \hat{v}_t' q_{i,t}^0 + \frac{\sigma_{u,i}^2}{2}\big). \tag{16}$$

I then define the option value associated with the propensity to be zoned use s as

$$OV_{i,t}^s := \exp\big(\hat{a}_t + \hat{v}_t' q_{i,t}^0 + \sum_{j}^{S} \gamma_{t,j}\hat{\eta}_{i,t,j} LSA_{i,t} + \frac{\sigma_{u,i}^2}{2}\big) - \exp\big(\hat{a}_t + \hat{v}_t' q_{i,t}^0 + \sum_{j\neq s} \gamma_{t,j}\hat{\eta}_{i,t,j} LSA_{i,t} + \frac{\sigma_{u,i}^2}{2}\big). \tag{17}$$

4.3 Identification

A primary concern with regards to estimation of the hedonic model is potential endogeneity between the land use propensities and prices. It is difficult to disentangle the effect of (re)zoning propensities on prices from the effect that prices might have on (re)zoning propensities themselves.

In order to address such concerns, we focus our estimation in the hedonic model on the sub-sample of real estate transaction that occur within a year of a nearby major zoning change. We define a zoning change to be major if more than twenty lots changed zoning designation in a given census tract for any given year. We identify 41 such zoning changes throughout the sample.

For each major zoning change event, an area is defined within a mile radius of any lot that changed zoning designation, but not within 1,000 feet of any lot that switched zoning designation. The intuition is that a major zoning change event may very well be endogenous to transaction prices in the area of the rezoning. However, when an area experiences a large rezoning, neighbors that did not experience the rezoning reevaluate their expectations about what their future zoning designation looks like. For example, if the major zoning change made a neighboring area become more like our area, we would not expect a zoning change to occur in the future. However, if the land use distribution of the area that experienced a zoning change becomes more different than ours, we should increase the probability of zoning change in our neighborhood occurring in the future.

Figure 3 illustrates the identification strategy visually for the Hudson Yards rezoning.

5 Results

The results of the two-stage estimation procedure are presented below; each is accompanied by a discussion. The entire estimation procedure is run separately for the boroughs of Manhattan and Brooklyn, and each run includes years 2003 -2015.

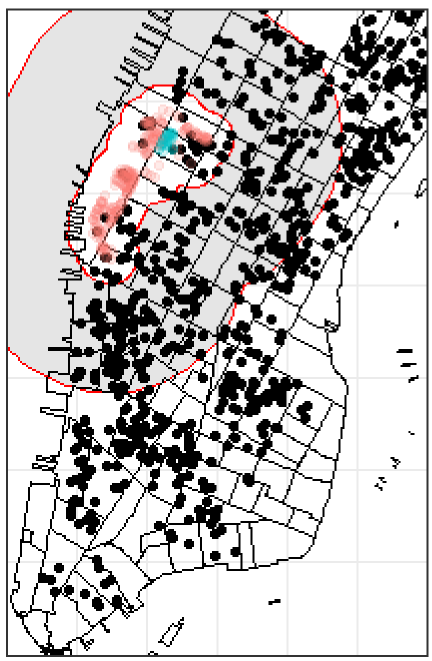
Before showing the results of the hedonic model, we present evidence that the generated proxy for land use uncertainty predicts in-sample zoning changes. In addition, we demonstrate that the proxy for redevelopment motivated by Clapp and Salavei (2010) predicts in-sample redevelopments.

5.1 First Stage: Estimating Land Use Uncertainty

In the first stage, we run a multinomial logistic regression to capture the uncertainty regarding future zoning designations. Table 3 presents summary statistics for the sample used for estimation, while Table 4 presents coefficients and significance.

The main drivers of future land use uncertainty are the distributional characteristics of a lot's census tract with respect to land use, and how it differs from the lot's community district (borough). For example, the variable $\%C|CT_{i,t} - \%C|CD_{i,t}$ is defined as the percentage of lots zoned commercial (C) in the census tract (CT) associated with lot i in year t, minus

Figure 3: Identified real estate transactions around the Hudson Yards rezoning. Each black dot represents a property transaction within a year of the Hudson Yards rezoning. The colored dots represent a properties that changed zoning designations to commercial (red) and residential (blue). The shaded area represents the identified zone: within a mile of any property that got rezoned in the Hudson Yards rezoning, but not within 1,000 ft.



the percentage of lots zoned commercial in the entire community district (CD) associated with lot i. To give a sense of how granular these groupings of geographical area are, there are two hundred and thirty eight census tracts the make up the borough of Manhattan, and fourteen community districts. Figure 4 shows the boroughs of Manhattan and Brooklyn broken into community districts (blue border), and census tracts (black border). The idea is if the localized area surrounding a property has more commercial lots than the larger geography, then the current lot is more likely to either be currently commercially zoned or rezoned to commercial in the future. Coefficients of the deviation of both the percentage of lots zoned as commercial or manufacturing in the census tract from the entire borough are significant. Results show if the census tract is either more commercial or manufacturing than the borough, then lots in that tract are more likely to be rezoned as commercial or manufacturing regardless of current zoning. This implies that commercial and manufacturing lots tend to be clumped together away from residential lots.

Table 3: Summary statistics from first-stage estimation. $\%C|CT - \%C|CD$ refers to the difference between the percentage of lots zoned to commercial (C) in a lots specific census tract (CT) and the community district (CD). The same notation holds for manufacturing zoned lots (M), and the borough (Boro). LotArea is the area of the lot. Age refers to the age of the building. LastAltered refers to the year the building had its last major alteration. MetroTimeGCS is the estimated transit time to Grand Central Station.

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
$\%C\|CT - \%C\|CD$	16,016,931	−0.00000	0.095	−0.813	−0.023	−0.004	0.940
$\%M\|CT - \%M\|CD$	16,016,931	−0.0001	0.123	−0.282	−0.039	−0.003	0.994
$\%C\|CT - \%C\|Boro$	16,016,931	−0.00005	0.123	−0.289	−0.022	−0.006	0.981
$\%M\|CT - \%M\|Boro$	16,016,931	−0.0001	0.133	−0.073	−0.041	−0.034	0.966
LotArea	16,016,931	4,132.826	29,613.590	17	1,914	2,967	9,218,615
Age	16,016,931	80.766	25.529	0.000	73.500	98.500	217.750
LastAltered	16,016,931	72.237	31.558	−0.500	52.500	94.500	216.750
MetroTimeGCS	16,016,931	15.814	5.599	0.070	12.050	19.888	40.207

Other significant variables driving the propensity to be rezoned are the age of the structure, the floor area ratio, the number of floors, the assessed value of the land, and the legislative dummies.

An interpretation can be difficult in the multinomial logit model, in which marginal effects are computed by integrating over the exogenous variables. Results for Manhattan are presented in Table 6, and for Brooklyn in Table 7.

For Manhattan, each percent deviation of commercial lots in the census tract to the borough increases the propensity to be commercially zoned by almost 0.6% and decreases the probability of being residential by nearly the same amount. It does not appear to have a statistically significant effect on the propensity to be manufacturing. Each percentage point deviation of manufacturing zoned lots from the borough increases the likelihood to be zoned commercial by 0.02%, manufacturing by 0.028%, and residential by 0.05%. Older buildings increase the likelihood of residential or manufacturing and decrease the likelihood of being commercially zoned. Taller buildings are decrease the likelihood of being zoned for manufacturing. A ten percent increase in the assessed value of the land increases the propensity to be commercial by 0.6% and decreases the likelihood of being zoned residential by 0.056%.

The legislative variables have a significant impact. Having a certified ULURP application increased the likelihood of being commercial by 32%, decreased the change of being manufacturing by roughly 3%, and decreased the likelihood of being residentially zoned by 29%. Also how the community board weighs in on the zoning application also has a statistically significant impact. Applications for which the community board is a majority in favor of the proposed zoning change are 21.9% less likely to be commercial in the future, 2.75% more likely to be manufacturing, and 19.2% more likely to be residential.

After running the estimation, the land use propensities are generated by simply taking the model prediction for each classification. Then the base case is simply unity minus the

Table 4: Results of the first-stage choice model of land use designation two years into the future. The sample includes all lots from both Manhattan & Brooklyn for years 2003-2015. $\%C|CT - \%C|CD$ refers to the deviation of % properties zoned (C)ommercial ((M)anufacturing) in the census tract from the community district $\%C|CT - \%C|Boro$ refers to the respective deviation but from the borough. CertDum is a dummy variable which is one if a ULURP application for a zoning change is in progress and it is certified. ResComDum is a dummy that is equal to one if a ULURP application is in progress and the residential community board approved of the application.

	Land Use	
	C	M
$\%C\|CT - \%C\|CD$	0.225	0.775
$\%M\|CT - \%M\|CD$	-0.449	-5.421***
$\%C\|CT - \%C\|Boro$	6.322***	4.884***
$\%M\|CT - \%M\|Boro$	4.349***	15.69***
LogLotArea	-0.527	0.327
Age	-0.00793***	-0.00341***
LastAltered	-0.0000558	0.000368
FAR	0.0468	0.0950***
LogBldgArea	-0.340	0.0641
LotFrontage	0.00160	0.000630
NumFloors	-0.0200	-0.136***
LogAssessLand	0.608***	0.187
MetroTimeGCS	0.00379	0.0333
CertDum	3.137***	1.002**
ResComDum	-2.106**	-0.215
Constant	-0.443	-9.582***
N	2,118,441	
R^2_P	0.514	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

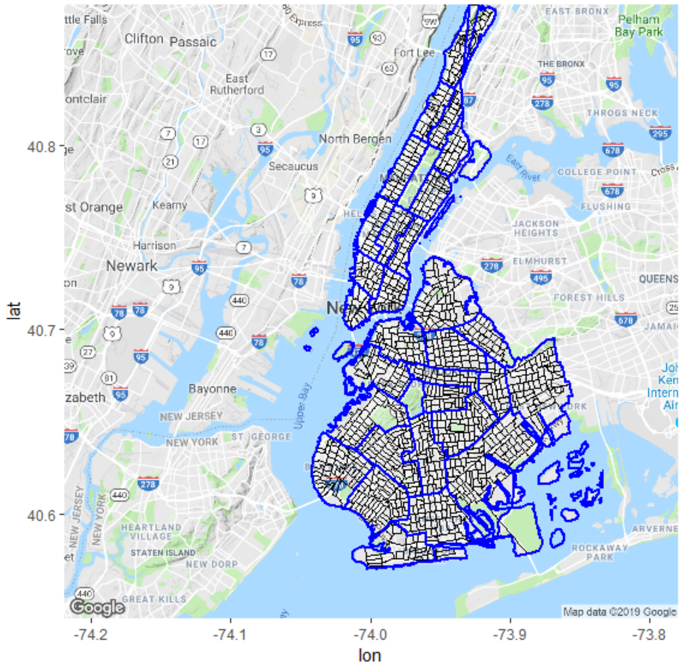


Figure 4: The boroughs of Manhattan and Brooklyn with the boundaries that community districts in blue (thick), and census tracts in black (thin). The census tracts are nested in the community districts, and the community districts within the borough. Manhattan is in the north and Brooklyn is in the south, each separated by the East River.

Table 5: Marginal effects for propensity to be rezoned commercial (C), manufacturing (M), and residential (R) in Manhattan based on the first-stage regression for years 2003-2015. Marginal effects are computed by integrating over observed values of regressors, and the t-statistics are computed using the delta-method. t-statistics in parenthesis.

	C	M	R
$\%C\|CT - \%C\|CD$	0.0119	0.0141	-0.0260
	(0.16)	(1.17)	(-0.32)
$\%M\|CT - \%M\|CD$	0.0357	-0.116***	0.0800
	(0.42)	(-4.14)	(0.90)
$\%C\|CT - \%C\|Boro$	0.593***	0.0135	-0.606***
	(9.81)	(1.35)	(-9.17)
$\%C\|CT - \%M\|Boro$	0.219*	0.288***	-0.507***
	(1.91)	(13.31)	(-4.39)
LogLotArea	-0.0607*	0.0155***	0.0452
	(-1.95)	(4.03)	(1.45)
Age	-0.000785***	0.0000444**	0.000741***
	(-5.58)	(2.12)	(5.77)
LastAltered	-0.0000115	0.00000918	0.00000236
	(-0.14)	(0.55)	(0.03)
FAR	0.00349	0.00143***	-0.00492*
	(1.29)	(2.88)	(-1.68)
LogBldgArea	-0.0369*	0.00668	0.0303
	(-1.75)	(1.61)	(1.40)
NumFloors	-0.0000150	-0.00278***	0.00280
	(-0.01)	(-3.30)	(1.56)
LogAssessLand	0.0614***	-0.00510	-0.0563***
	(7.16)	(-1.10)	(-8.46)
CertDum	0.316***	-0.0255*	-0.291***
	(5.86)	(-1.80)	(-6.44)
ResComDum	-0.219**	0.0275**	0.192**
	(-2.50)	(2.17)	(2.26)

t-statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Marginal effects for propensity to be rezoned commercial (C), manufacturing (M), and residential (R) in Brooklyn based on the first-stage regression for years 2003-2015. Marginal effects are computed by integrating over observed values of regressors, and the t-statistics are computed using the delta-method.

	C	M	R
$\%C\|CT - \%C\|CD$	-0.0533	0.0161	0.0372
	(-0.49)	(0.25)	(0.34)
$\%M\|CT - \%M\|CD$	-0.0130	-0.117***	0.130***
	(-0.41)	(-5.46)	(4.78)
$\%C\|CT - \%C\|Boro$	0.251**	0.0262	-0.278***
	(2.41)	(0.39)	(-2.68)
$\%M\|CT - \%M\|Boro$	0.0327	0.269***	-0.302***
	(1.18)	(13.25)	(-11.14)
LogLotArea	-0.0123**	0.00958***	0.00272
	(-2.51)	(4.03)	(0.40)
Age	0.0000464	0.0000970*	-0.000143**
	(1.02)	(1.93)	(-2.11)
LastAltered	-0.0000520**	-0.0000370**	0.0000889***
	(-2.43)	(-2.39)	(3.22)
LotFrontage	0.0000260***	-0.00000757	-0.0000185
	(2.90)	(-0.86)	(-1.31)
NumFloors	-0.00521***	-0.0141***	0.0193***
	(-2.83)	(-7.39)	(5.98)
LogAssessLand	0.0130***	0.00337***	-0.0164***
	(8.70)	(3.01)	(-8.08)
MetroTime	-0.000585***	-0.000137	0.000722***
	(-3.43)	(-1.22)	(3.33)
CertDum	0.00476	-0.0557***	0.0510***
	(0.75)	(-3.29)	(2.66)
ResComDum	-0.00280	-0.0141	0.0169
	(-0.97)	(-0.98)	(1.10)

t-statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

sum of all imputed values.

5.2 Do Land Use Propensities Predict Zoning Changes?

Now that land use propensities for each tax lot have been generated, we test whether our generated propensities are predictive of zoning changes. Since we have a limited sample to work with and the zoning process is highly persistent, we see if we can explain the probability that a zoning change occurs during the sample with zoning propensities at the beginning of the sample.

To determine if our generated land use propensities are actually predictive of zoning changes, we create a dummy variable, D_i, that is equal to unity if a lot has a different zoning designation in 2017 than it did in 2002, and zero otherwise. We remove the few lots that have gone through two or more rezonings over the sample to avoid the case where a lot switched to a new designation and then switched back to their original designation.

We regress $D_{i,t}$ on what we define to be the *transition propensity* at the beginning of the sample, $T_{i,2002}$, as the sum of the propensities to be a land use different from how it is currently zoned.

$$T_{i,2002} := \sum_{z \in \{C,R,M\} \setminus Z_{i,2002}} p(z)_{i,2002}, \tag{18}$$

where $Z_{i,2002}$ is the zoning designation of lot i in 2002.

For example, for a manufacturing zoned lot, the transition propensity is the sum of that lot's propensity to be commercial or residential. We run a separate regression for each zoning designation, and also include a specification with census tract fixed effects. Table 8 presents results.

Regressions (1) - (3) show the estimation results. All results show that transition propensities are positively correlated with a zoning change occurring throughout our sample, and all results are statistically significant at the 0.1% significance level. For commercial, manufacturing, and residential, respectively, a one percentage point increase in the transition propensity translates into an increased probability of being rezoned by 0.223%, 0.112%, and 0.19%. Including census tract level fixed effects does not hinder the effect, and for commercial and residential, it makes the effect much stronger. This suggests that even within a census-tract, the transition propensities help to distinguish which buildings will transition.

5.3 Does Land Share of Assessed Value Predict Redevelopments?

It is imperative that we test if the land share of assessed value, our proposed proxy motivated by Clapp and Salavei (2010), predicts redevelopments in our sample. Since redevelopments occur over a very long cycle, we opt to do an experiment similar to the prior experiment by testing whether propensities predicted observed zoning changes. We create a dummy variable, R_i, indicating whether or not a redevelopment has occurred throughout the entire sample. A redevelopment here is defined as either as a major renovation or a construction project that alters the size or egress of a structure.

Table 7: Regressions (1)-(3) show results from regressing a dummy indicator whether or not a zoning change occurred during the sample on transition propensities at the beginning of the sample defined as the sum or propensities to be a zoning designation other than the lot's current designation, for lots zoned commercial, manufacturing, and residential respectively at the beginning of the sample. Regressions (4)-(6) include census tract fixed effects.

	(1)	(2)	(3)	(4)	(5)	(6)
$T^{C}_{i.2002}$	0.223***			0.362***		
	(0.00632)			(0.0178)		
$T^{M}_{i,2002}$		0.112***			0.0832***	
		(0.00938)			(0.0118)	
$T^{R}_{i,2002}$			0.0193***			0.0554***
			(0.000844)			(0.00219)
Constant	-0.0252***	0.140***	0.00128***	-0.0973***	0.153***	-0.000256*
Tract FE	-	-	-	Yes	Yes	Yes
N	15,338	12,508	260,349	15,325	12,498	260,347
adj. R^2	0.075	0.011	0.002	0.575	0.737	0.164

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

We then regress R_i on five dummies representing which quintile of land share of assessed value the property falls into. Results are presented in Table 9.

Regression (1), the baseline, shows all coefficients are statistically significant, and there is roughly a 10 percentage point increase in the probability of redevelopment going form the bottom quintile to the top quintile. Controlling for age and the remaining FAR the property owner is legally allowed to add on to the building does not take away the effect of our proxy, although it does dampen it a bit. The difference between the top and bottom quintile decreases to roughly 7 percentage points. Age has a non-linear impact on the probability of redevelopment, reaching an inflection point at around 60 years. This implies that after 60 years of operation, the benefit from having a 'historic' building outweighs redevelopment. Instead, property owners likely invest in the upkeep of their property little-by-little each year to stave off depreciation. Remaining FAR is also statistically significant, with each additional floor that is allowed to be added on to the redeveloped building increasing the

Table 8: Regressions (1)-(4) predict redevelopment over the sample on quintile dummies of assessed land value as a percentage of assessed total value. Regression (1) presents baseline results. Regression (2) controls for a quadratic term in the age of the structure as well the remaining designated FAR permitted by the New York City Zoning Resolution (FAR Remaining). Regression (3) includes community district fixed effects, and Regression (4) includes census tract fixed effects. The bottom row F-statistics are from a test that the coefficients on all five quintiles are equal to one another.

R_i	(1)	(2)	(3)	(4)
Q1	0.0687***	0.00813*	0.225***	0.0603
	(0.00143)	(0.00340)	(0.0112)	(0.0390)
Q2	0.0733***	0.0115**	0.243***	0.0794*
	(0.00143)	(0.00366)	(0.0113)	(0.0390)
Q3	0.0793***	0.0154***	0.249***	0.0896*
	(0.00143)	(0.00374)	(0.0113)	(0.0390)
Q4	0.102***	0.0331***	0.262***	0.103**
	(0.00143)	(0.00376)	(0.0113)	(0.0390)
Q5	0.165***	0.0785***	0.290***	0.129***
	(0.00143)	(0.00382)	(0.0112)	(0.0390)
Age		0.000718***	0.00104***	0.000910***
		(0.0000928)	(0.0000943)	(0.000102)
Age^2		-0.00000161*	-0.00000823***	-0.00000771***
		(0.000000664)	(0.000000680)	(0.000000719)
CD FE	-	-	Yes	-
Tract FE	-	-	-	Yes
adj. R^2	0.111	0.120	0.141	0.153
F(4,·)	777.56***	377.48***	207.96***	190.44***

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

odds of redevelopment by 2 percentage points. Lastly, regressions (3)-(4) add community district, and census tract fixed effects respectively, and they do not seem to hinder our proxy's ability to predict redevelopment.

5.4 Hedonic Regression

Next, a hedonic model is estimated by regressing the logarithm of sale price per square feet on various property characteristics. The advantage of defining intensity as in Equation 14 is that it is increasing in option value; the greater percentage of the total value of the property is due to the land, the higher the option value should be. And when intensity is equal to zero – the land is not valued – it can be interpreted as if there is no option value to redevelop.

For years 2003-2015, I have complete data for 9,853 transactions in Manhattan, excluding condominium sales, which make up the majority of residential transactions in Manhattan. We have 105,047 transactions for Brooklyn. Summary statistics for the estimation sample are presented in Table 10 and Table 11 respectively.

Table 9: Summary statistics for hedonic regression of Manhattan; 2003-2015.

	Mean	Std. Dev.	Min	Max
SaleSqFt	680.92	947.32	1.96	35,416.67
LotArea	3,862.48	4,812.18	320	58,24
Age	95.23	19.23	1.75	119.75
p(C)	.2812	.3353	.0013	.9959
p(M)	.0703	.1952	.0001	.9859
p(R)	.6483	.3868	.0030	.9983
AssessTotalSqFt	49.73	64.94	.490929	2,163
FloorArea	3,509.80	8,691.54	130	359,251
NumFloors	5.104	4.26	1	63
LastAltered	43.94	52.95	-12.25	119.75
ExemptTotalSqFt	2.62	13.55	0	373.60
MetroTimeGCS	9.969675	6.47	.07	26.2
LSA	.4008	.2324	.0121	.9911
Observations	5,276			

Table 10: Summary statistics for hedonic regression of Brooklyn; 2003-2015.

	Mean	Std. Dev.	Min	Max
SaleSqFt	290.48	194.77	9.6153	2,854.33
LotArea	2,676.87	4,256.58	340	362,142
Age	88.11	24.04	.5	118
p(C)	.0189	.0726	1.40e-15	.9999
p(M)	.0398	.1491	1.33e-32	.9999
p(R)	.9412	.1650	3.77e-06	1
AssessTotalSqFt	11.80	12.46	.1425	499.36
FloorArea	1,570.58	2,755.36	80	146,300
NumFloors	2.4555	.8697	1	75
ExemptTotalSqFt	1.1238	4.51	0	239.59
MetroTimeGCS	16.3896	4.56	2.81	28.65
LSA	.3985	.2384	.0068	.9906
Observations	35,253			

Regression results from the hedonic model estimation for both Manhattan and Brooklyn are presented in Table 12.

The regression specification above regresses the logarithm of the transaction price per gross square foot on land share of assessed value (LSA) interacted with land use propensities, as well as the logarithm of the total assessed value per gross square foot. The LSA terms reflect the option value to redevelop into each use, while the total assessed value terms are interacted to allow for the possibility that changes in the future expected zoning of an area might also impact real estate values through the existing building characteristics.

Ignoring the interaction terms, the elasticity of prices with respect to the land share of assessed value is $\beta_{LSA} - \frac{\beta AV}{LSA}$. As a building depreciates, the land share of assessed value increases. The coefficient on LSA, β_{LSA}, is thus the impact on prices through the option value to redevelop as a building depreciates. This first-order effect is offset by the effect of the building depreciating on the total assessed value inversely proportional to the land share of assessed value. This loss of value from depreciation is most severe when the land share of assessed value is low (or when the building is most valuable).

All coefficients on the LSA term are significant at the 5% level with the exception of the p(r) term for Manhattan (which is significant at the 10% level) and the p(c) term for Brooklyn. For both boroughs, option value is most sensitive to propensities to be zoned commercial, followed by manufacturing and residential.

5.5 Discussion: Option Value to Redevelop Across Space

Table 13 and Table 14 present how option value varies across the entire sample for Manhattan and Brooklyn respectively. The mean option value in Manhattan is 20% of total value, and it is broken down into 28% related to commercial development, 8.5% manufacturing, and 66.3% residential. The option value to redevelop has a standard deviation of 10 percentage points. Lots in Brooklyn have far less option value, with a mean of 8.5% and standard deviation of 4.7 percentage points. Much of Brooklyn's option value is attributed to residential development,

Table 11: Hedonic estimation results from Manhattan and Brooklyn from the identified subsample. LSA is the land share of assessed value, and p(z) corresponds to the land use propensities. Standard errors are computed using the Huber-White sandwich estimator and are clustered across census tracts.

$\log \frac{SalePrice}{GrossSqFt}$	Manhattan β / SE	Brooklyn β / SE
p(c) × LSA	0.851***	0.439
	(0.161)	(0.446)
p(m) × LSA	0.698***	0.423**
	(0.217)	(0.166)
p(r) × LSA	0.221*	0.222***
	(0.128)	(0.036)
p(c) × $\log \frac{AssessTotal}{GrossSqFt}$	0.459***	0.735***
	(0.063)	(0.162)
p(m) × $\log \frac{AssessTotal}{GrossSqFt}$	0.590***	0.118***
	(0.124)	(0.038)
p(r) × $\log \frac{AssessTotal}{GrossSqFt}$	0.200***	0.258***
	(0.024)	(0.012)
Controls × Year	Yes	Yes
CT × Year FE	Yes	Yes
Observations	4,521	30,215
R^2	0.637	0.549

with only 1% going to commercial and 4.4% being attributed to manufacturing.

Table 12: Summary statistics for option value in Manhattan

All Properties	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
OV/Total Value	9,853	0.200	0.100	0.007	0.119	0.275	0.455
OV^C/OV	9,853	0.280	0.308	0.00000	0.022	0.438	0.995
OV^M/OV	9,853	0.085	0.207	0.000	0.004	0.037	0.987
OV^R/OV	9,853	0.663	0.364	0.001	0.306	0.976	0.999
Commercial	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
OV/Total Value	3,040	0.191	0.084	0.013	0.123	0.251	0.426
OV^C/OV	3,040	0.599	0.304	0.001	0.318	0.897	0.995
OV^M/OV	3,040	0.085	0.158	0.00001	0.014	0.064	0.969
OV^R/OV	3,040	0.352	0.331	0.001	0.060	0.685	0.993
Manufacturing	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
OV/Total Value	746	0.218	0.092	0.014	0.143	0.288	0.444
OV^C/OV	746	0.259	0.233	0.003	0.084	0.393	0.891
OV^M/OV	746	0.647	0.281	0.005	0.425	0.885	0.987
OV^R/OV	746	0.140	0.183	0.004	0.040	0.139	0.949
Residential	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
OV/Total Value	6,067	0.202	0.107	0.007	0.114	0.286	0.455
OV^C/OV	6,067	0.123	0.161	0.00000	0.013	0.184	0.965
OV^M/OV	6,067	0.016	0.063	0.000	0.003	0.011	0.950
OV^R/OV	6,067	0.883	0.162	0.019	0.840	0.985	0.999

Table 15 and Table 16 exhibit how the breakdown of option value varies across demographic factors. On the racial dimension, the higher percent Asian a census tract is, the more the redevelopment option is a fraction of total value in Manhattan. In contrast, the more African-American a neighborhood, the more the option value to redevelop constitutes total property value. Neighborhoods that are the least Asian have option values of 10.64% for commercial, 1.77% manufacturing, and 88.77% residential, while neighborhoods that are the most Asian have redevelopment rates of 57.71% for commercial, 18.75% manufacturing, and 26.43% residential. The distribution of the least African-American census tracts is 40.43% commercial, 9.17% manufacturing, and 54.25% residential. The census tracts which are most African-American have nearly all residential value; the breakdown is 9.51% commercial, 3.73% manufacturing, and 87.72% residential.

Table 13: Summary statistics for option value in Brooklyn.

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
OV/Total Value	105,047	0.085	0.047	0.002	0.042	0.125	0.305
OV^C/OV	105,047	0.010	0.044	0.000	0.002	0.007	0.793
OV^M/OV	105,047	0.044	0.139	0.000	0.007	0.021	0.994
OV^R/OV	105,047	0.948	0.144	0.001	0.975	0.990	1.000
Commercial	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
OV/Total Value	1,949	0.071	0.044	0.003	0.035	0.102	0.255
OV^C/OV	1,949	0.143	0.214	0.0001	0.012	0.172	0.793
OV^M/OV	1,949	0.073	0.116	0.000	0.015	0.081	0.981
OV^R/OV	1,949	0.792	0.243	0.017	0.712	0.969	1.000
Manufacturing	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
OV/Total Value	3,918	0.105	0.061	0.003	0.055	0.146	0.305
OV^C/OV	3,918	0.013	0.045	0.000	0.002	0.009	0.632
OV^M/OV	3,918	0.529	0.371	0.0004	0.141	0.930	0.994
OV^R/OV	3,918	0.471	0.366	0.001	0.074	0.851	0.999
Residential	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
OV/Total Value	99,180	0.084	0.047	0.002	0.042	0.125	0.274
OV^C/OV	99,180	0.008	0.027	0.000	0.002	0.006	0.766
OV^M/OV	99,180	0.025	0.071	0.000	0.007	0.019	0.974
OV^R/OV	99,180	0.970	0.075	0.025	0.977	0.990	1.000

One may think the racial breakdown may be related to income. When looking at the same breakdown but for median household income in Manhattan, it is very similar to the breakdown for Asian census tracts. The least Asian neighborhoods have a similar option value to the lowest income census tracts, while the most Asian look very close to the highest income quartile. However, the top quartile of African-American census tracts has more variation than income dynamics might suggest. The top quartile of African-American tracts have more redevelopment option value than the lowest income bracket, and the option value itself is attributed 3% more to residential development. They both have extremely little manufacturing potential.

Table 14: A breakdown of option value across various demographics for Manhattan. Q1 refers to the bottom quartile of census tracts, which contains the smallest 25% of values. Q4 is then the quartile of 25% largest of values. The first row shows the percentage of estimated property value that is attributed to total option value. The bottom three rows then break down the distribution of how optional value is related to the propensity of each property to be categorized as each land use. All demographic information is on the census tract level. GCT refers to Grand Central Station.

Percent African American	Q1	Q2	Q3	Q4
Option Value as % of Property Value	20.39	18.96	17.66	16.19
% Commercial	40.43	45.36	39.99	9.51
% Manufacturing	9.17	15.50	8.38	3.73
% Residential	54.25	42.93	54.60	87.72
Percent Asian	Q1	Q2	Q3	Q4
Option Value as % of Property Value	16.02	20.92	18.55	17.58
% Commercial	10.64	24.24	42.36	57.71
% Manufacturing	1.77	7.75	8.74	18.75
% Residential	88.77	71.84	52.23	26.73
Median Age	Q1	Q2	Q3	Q4
Option Value as % of Property Value	16.63	17.11	19.10	20.14

% Commercial	34.28	32.12	31.65	37.33
% Manufacturing	12.04	11.15	8.69	5.49
% Residential	55.95	59.35	63.10	60.37
Median Household Income	Q1	Q2	Q3	Q4
Option Value as % of Property Value	14.92	18.61	19.81	19.68
% Commercial	15.12	31.48	42.52	45.38
% Manufacturing	1.14	8.54	8.46	18.98
% Residential	84.57	62.79	53.21	39.31

Percent Renters	Q1	Q2	Q3	Q4
Option Value as % of Property Value	20.22	19.63	18.31	14.94
% Commercial	37.74	38.37	32.74	26.59
% Manufacturing	12.72	11.73	7.29	5.52
% Residential	53.14	53.58	62.61	69.53
Percent w/ Bachelor's Degree or Higher	Q1	Q2	Q3	Q4
Option Value as % of Property Value	14.38	18.52	20.50	19.62
% Commercial	15.12	38.69	37.60	43.52
% Manufacturing	1.21	8.43	15.17	12.28
% Residential	84.44	55.67	51.44	47.95
Metro Time to GCT	Q1	Q2	Q3	Q4
Option Value as % of Property Value	21.13	20.91	18.06	19.84
% Commercial	34.80	26.64	20.34	30.25
% Manufacturing	12.01	6.74	5.97	9.38
% Residential	57.05	69.29	75.42	63.41

Education also has similar option value characteristics as median household income. At the top of the income and education bracket, the quartile of tracts with the highest degree earners has 6% more option value associated with residential redevelopment and 10% more residential.

Tracts with more renters have a lower amount of option value and have less of the option value devoted to commercial use, less to manufacturing, and more to residential.

Median age of the households in the census tract seems to have a positive correlation with option value. The older the residents of the neighborhood, the more option value there is, likely due to higher incomes.

The changes to option value in Brooklyn are not as drastic as in Manhattan. There is

Table 15: A breakdown of option value across various demographics for Brooklyn. Q1 refers to the bottom quartile of census tracts, which contains the smallest 25% of values. Q4 is then the quartile of 25% largest of values. The first row shows the percentage of estimated property value that is attributed to total option value. The bottom three rows then break down the distribution of how optional value is related to the propensity of each property to be categorized as each land use. All demographic information is on the census tract level. GCT refers to Grand Central Station.

Percent African American	Q1	Q2	Q3	Q4
Option Value as % of Property Value	11.46	11.13	11.43	12.09
% Commercial	0.85	1.27	1.36	0.68
% Manufacturing	3.11	7.46	4.38	2.19
% Residential	96.37	91.74	94.60	97.42
Percent Asian	Q1	Q2	Q3	Q4
Option Value as % of Property Value	12.06	11.67	11.20	11.19
% Commercial	1.02	0.91	1.22	1.03
% Manufacturing	3.68	5.07	5.30	3.21
% Residential	95.64	94.36	93.90	96.11
Median Age	Q1	Q2	Q3	Q4
Option Value as % of Property Value	11.70	11.32	11.42	11.66
% Commercial	1.66	1.11	0.75	0.66
% Manufacturing	6.50	5.32	3.70	1.83
% Residential	92.30	93.97	95.87	97.77
Median Household Income	Q1	Q2	Q3	Q4
Option Value as % of Property Value	11.65	11.57	11.66	11.23
% Commercial	1.52	0.84	0.71	1.11
% Manufacturing	4.46	4.54	4.10	4.14
% Residential	94.44	94.99	95.52	95.08

far less variation in option value. This may have to do with the fact that the majority of option value in Brooklyn is associated with residential. More research needs to be done on what is driving the results from Brooklyn.

5.6 Discussion: Option Value Across Time

Figure 5 and Figure 6 show how the option value to redevelop as a percentage of total estimated value has changed each year for Manhattan and Brooklyn respectively. A visual inspection gives creedance to the idea that the option value to redevelop is acyclical. For both Manhattan and Brooklyn, we see option value decrease from 2003-2008, when the great financial crisis hit. From 2008-2010, there is a drastic increase in option value. Once NYC real estate starts to boom again from 2010-2015, we also see a decrease in the option value to redevelop.

This is consistent with Grenadier (1996), which finds that development cascades can occur in periods where there is a downturn in demand for real estate services. As rents from real estate decline during an economic downturn, the opportunity cost introduced from the lost rent during a redevelopment decreases. Landlords then have greater incentive to exercise the redevelopment option, as they can earn even more rents once the recession ends and the property is improved.

6 Conclusion

We ask how the option value to redevelop is related to land use restrictions, and how the redevelopment value is related to various demographic and building characteristics. A dataset is constructed that captures land use changes, property transactions, building characteristics, demographics, and economic information for all properties and census tracts in New York City from 2002-2017.

Using a two-stage estimation procedure, propensities to be zoned to either residential,

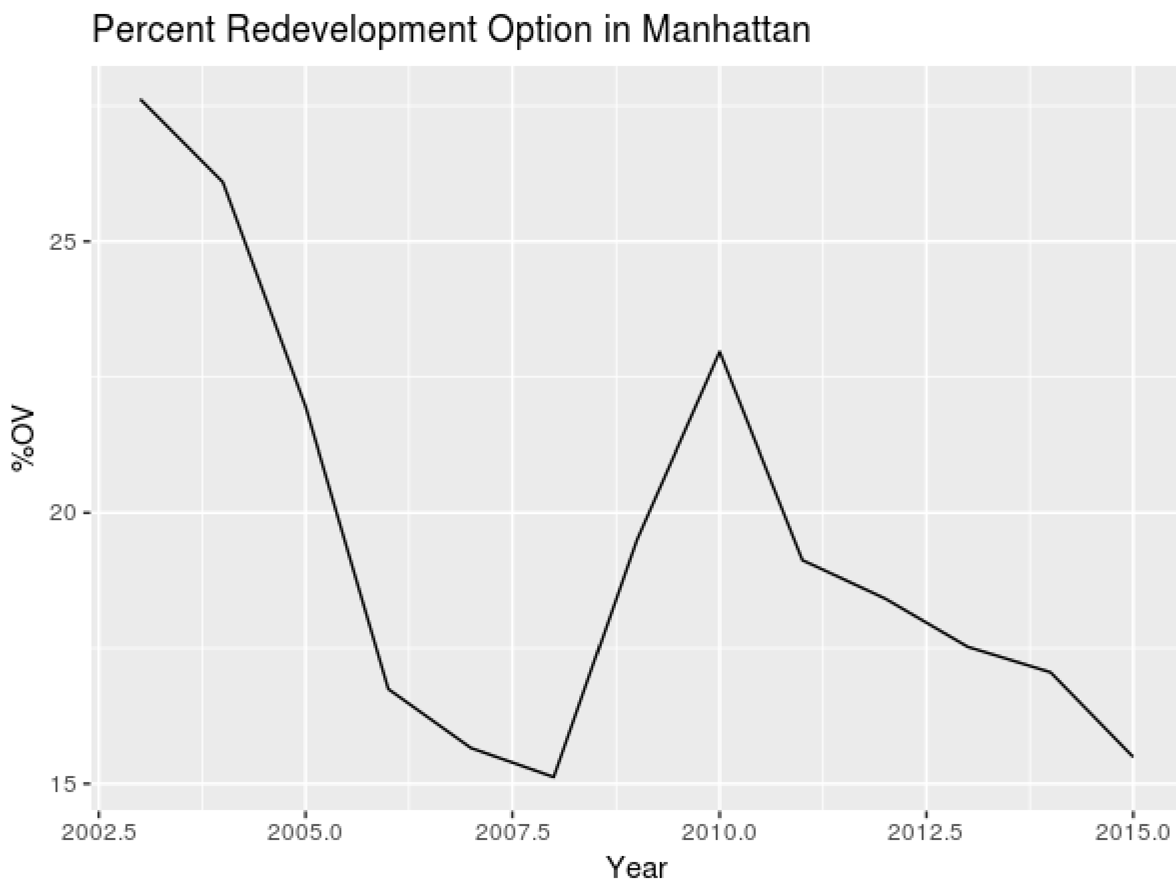


Figure 5: Option value to redevelop of a percent of total value over time for Manhattan.

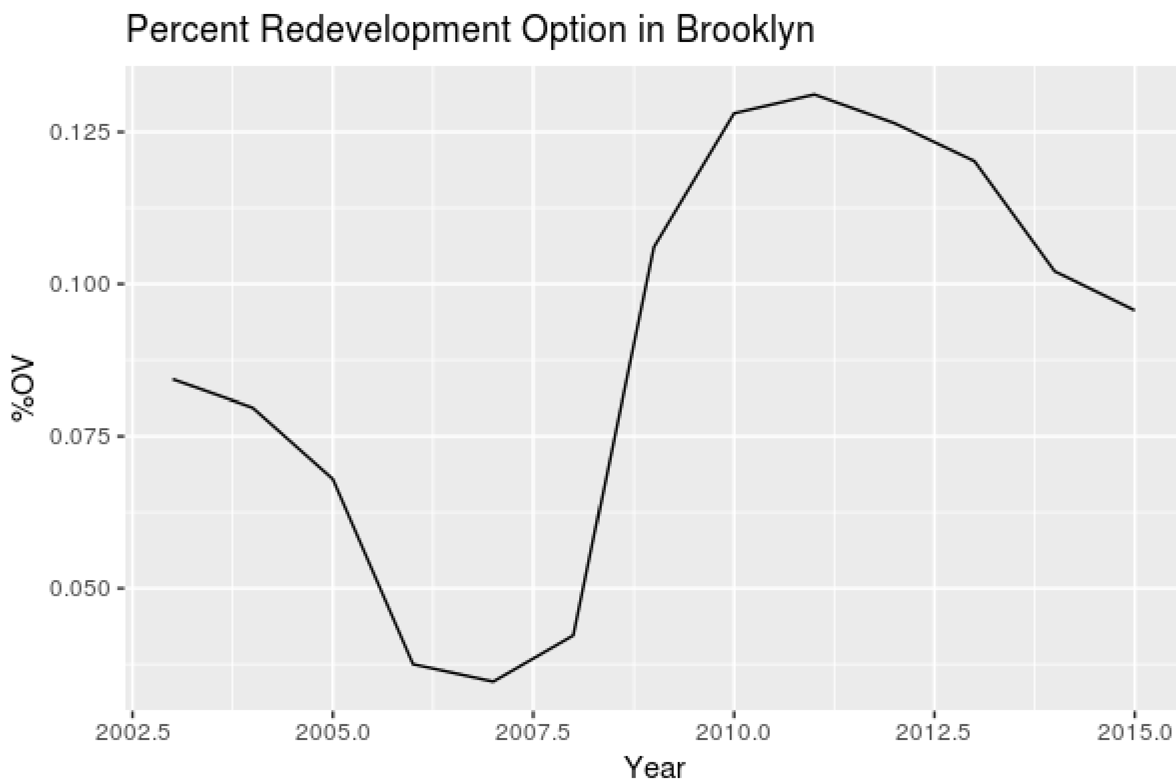


Figure 6: Option value to redevelop of a percent of total value over time for Brooklyn.

commercial, or manufacturing land uses are interacted with the ratio of assessed land to total assessed value of the property. This ratio proxies for the propensity to be redeveloped.

We find evidence that all option value terms are statistically significant. We estimate the the average option value in Manhattan for years 2003-2015 is 20% of total estimated value, and we estimate that it is 8.5% in Brooklyn. In a sense, Furthermore, the option value broken down by the potential to develop into different land uses shows clear correlations with race, income, and education. Lastly, we present evidence that the option value to redevlop is acyclical with the real estate cycle.

References

Childs, Paul D., Timothy J. Riddiough, and Alexander J. Triantis, "Mixed Uses and the Redevelopment Option," *Real Estate Economics*, 1996, *24* (3), 317–339.

Clapp, John M. and Katsiaryna Salavei, "Hedonic pricing with redevelopment options: A new approach to estimating depreciation effects," *Journal of Urban Economics*, 2010, *67* (3), 362–377.

Davis, Morris A. and Jonathan Heathcote, "Housing and the Business Cycle," *International Economic Review*, 2005, *46* (3), 751–784.

Fenton, Lawrence, "The sum of log-normal probability distributions in scatter transmission systems," *IRE Transactions on Communications Systems*, 1960, *8* (1), 57–67.

Geltner, David, Timothy J. Riddiough, and Srdjan Stojanovic, "Insights on the Effect of Land Use Choice: The Perpetual Option on the Best of Two Underlying Assets," *Journal of Urban Economics*, 1996, *39* (1), 20–50.

Glaeser, Edward, Joseph Gyourko, and Raven Molloy, "Why Is Manhattan So Expensive? Regulation and the Rise in House Prices," *Journal of Law and Economics*, 02 2005, *48*, 331–69.

Glaeser, Edward L. and Joseph E. Gyourko, "The Impact of Zoning on Housing Affordability," *FRBNY Economic Policy Review*, 2003, (June), 21–39.

Grenadier, Steven, "The Strategic Exercise of Options: Development Cascades and Overbuilding in Real Estate Markets," *The Journal of Finance*, 1996, *51* (5), 1653–1679.

Ihlanfeldt, Keith R., "The effect of land use regulation on housing and land prices," *Journal of Urban Economics*, 2007, *61* (3), 420–435.

Knoll, Katharina, Moritz Schularick, and Thomas Steger, "No Price Like Home : Global House Prices , 1870 - 2012 Moritz Schularick No Price Like Home : Global House Prices , 1870 - 2012 Abstract," *American Economic Review*, 2017, *107* (2), 331–53.

Lee, Lung-Fei, G.S. Maddala, and R.P. Trost, "Testing for Structural Change by D-Methods in Switching Simultaneous Equations Models," 1982.

McMillen, Daniel P. and John F. McDonald, "A simultaneous equations model of zoning and land values," *Regional Science and Urban Economics*, 1991, *21* (1), 55–72.

_ **and** _ , "Land Values in a Newly Zoned City," *Review of Economics and Statistics*, 2002, *84* (1), 62–72.

Munneke, Henry J. and Kiplan S. Womack, "Valuing the Redevelopment Option Component of Urban Land Values," *Real Estate Economics*, 2017, pp. 1–45.

Rosen, Sherwin, "Hedonic Prices and Implicit Markets : Product Differentiation in Pure Competition Author (s): Sherwin Rosen Published by : The University of Chicago Press Stable URL : http://www.jstor.org/stable/1830899 Accessed : 17-03-2016 13 : 46 UTC Your use of the," *Journal of Political Economy*, 1974, *82* (1), 34–55.

Strange, William and Stuart Rosenthal, "The Determinants of Agglomeration,," *Journal of Urban Economics*, 02 2001, *50*, 191–229.

Titman, Sheridan, "Urban Land Prices Under Uncertainty," *The American Economic Review*, 1985, *75* (3), 505–514.

Williams, Joseph T, "Redevelopment of real assets," *Real Estate Economics*, 1997, *25* (3), 387–407.