Data Sources and Regression Results for Housing Gotham II

Jason Barr, Rutgers University-Newark
December 13, 2021

I. Data Sources and Variable Names

A panel data set was created using data from 2003, 2006, 2010 and 2019. For each, year I collected the following variables from the following sources for each of the 59 Community Districts in New York.

NYC PLUTO File

Source: https://www1.nyc.gov/site/planning/data-maps/open-data/dwn-pluto-mappluto.page

- Number of housing units
- Per cent of land zoned for residential
- Number and per cent of structures that are one- or two-family homes
- Th total land area for residential use
- The average age of residential buildings
- The average allowable residential floor area ratio value
- Number of properties that are in a historic district

NYC Department of Finance

Source: https://www1.nyc.gov/site/finance/taxes/property-annualized-sales-update.page
For the panel years, I downloaded all real estate transactions for the respective years (plus the quarter after). I then calculated the sale price per square foot of building area. Prices were adjusted for inflation to prices using the NYC CPI without shelter (set to 2020 prices). Median real prices per square foot for each year and CD were taken, after excluding outliers in the bottom and top one percentile, respectively.

NYU Furman Center New York Neighborhood Data Profiles

Source: https://furmancenter.org/neighborhoods
Note: Their data for each CD and Borough is given for 2000, 2006, 2010, and 2019. For median real monthly rent, I adjusted the rents to 2003 based on CD growth rates from between 2000 and 2006. For the other variables, I simply used 2000.
CD level:

- Median monthly rents adjusted for inflation
- %White, %Black, %Hispanic
- Unemployment rate
- % of people who are foreign born.

Borough Level:

- % of residents in each borough with a college degree or higher
- Fraction of households with child less than 18 years of age.

BEA County Data

Source: https://www.bea.gov/data/gdp/gdp-county-metro-and-other-areas
Note: Each county in NYC is the same as the borough.

- Borough GDP
- Borough population
- Borough personal income
- Borough population

FRED

Source: https://fred.stlouisfed.org/series/CUURA101SA0L2.

- NYC CPI less shelter

CD Shapefile

Source: https://data.cityofnewyork.us/City-Government/Community-Districts/yfnk-k7r4

- CD area in square feet
- Latitude and longitude of CD centroids.
- Given the latitude and longitude coordinates, I also created the distance of each centroid to the Empire State Building (in degrees).

Table of Variable Names

Variable	Full Name	Units
medrppsf	Real Median sales price per square foot	$
medrent	Real Median monthly rent	$
totalResUnits	Total Residential Units	
b_pipc	Borough personal income per capita	$
b_hhch	Borough fraction households with children <18	
b_firegdp	borough GDP from FIRE sector	$
b_ba	Borough fraction with college degree or higher	
b_gdp	Borough GDP $	
fb	Fraction residents foreign born	
black	Fraction Black	
white	Fraction White	
hisp	Fraction Hispanic	
unemp	Unemployment rate	
numonetwofam	Number one- or two-family homes	
Rzone_pct	% of CD zoned for residential	
totalResArea	Total CD Residential area	square feet
cdarea_sqft	Total CD area	square feet
totallotarea	Total area of all lots in CD	square miles
%onetwofamily	% of structures that are one- or two- family	
HDProps	Number of structures in CD in a historic district	
B_Pop	Borough population	
DistESB	Distance of CD Centroid to ESB	degrees
AvgMaxFAR	Avg. Max. allowable Residential FAR for CD (unweighted)	

II. Descriptive Statistics

Table: Variable in levels, by CD unless otherwise noted.

Variable	Obs	Mean	Std. dev.	Min	Max
medrppsf	236	333.2943	156.0365	98.29156	848.7324
medrent	236	1340.72	400.1903	590	2840
totalResUnits	236	56974.72	21730.26	16016	151223
b_pipc	236	54042.1	41101.51	22504	187213
b_hhch	236	.3209322	.0771237	.17	.44
b_firegdp	236	5.52e+07	7.30e+07	2379361	2.20e+08
b_ba	236	.3209746	.1375388	.15	.62
b_gdp	236	1.53e+08	1.86e+08	1.07e+07	6.35e+08
fb	236	.3533517	.1244568	.117	.668
black	236	.2317966	.236169	.001	.895
white	236	.3363432	.2581583	.004	.891
hisp	236	.2896314	.2050553	.049	.743
unemp	236	.0905903	.046251	.0131	.2363
numonetwofam	236	9499.784	10820.51	14	44237
Rzone_pct	236	85.60486	20.20603	.2582311	99.52338
totalResArea	236	5.26e+07	2.56e+07	5231555	1.62e+08
cdarea_sqft	236	1.29e+08	1.19e+08	3.77e+07	5.99e+08
totallotarea	236	9.57e+07	9.58e+07	2.41e+07	4.92e+08
%onetwofamily	236	55.30301	29.83313	1.611047	99.44028
numHDProp	236	7493.559	10949.02	0	49965
b_pop	236	1925246	569488.6	443728	2559903
distESB	236	.1209472	.0626103	.003723	.2917135
AvgMaxFAR	227	2.403292	1.896169	.4837194	10.59374

Log Differences

Variable	Obs	Mean	Std. dev.	Min	Max
dlnMedRppsf	177	.225191	.3700811	-.616796	1.180819
dlnResUnits	177	.0126795	.0311611	-.1306314	.2320061
dlnBHH	177	-.0755551	.0634374	-.2170645	0
dUnemp	177	-.0178102	.0477344	-.1574	.0946
dBlack	177	-.0110395	.0355037	-.1799999	.091
dWhite	177	-.0078475	.04451	-.123	.148
dHisp	177	.0066667	.0339062	-.121	.128

III. The Regression Model

The regression model starts with log median sales price at time $t = \{2003, 2006, 2010, 2019\}$ in neighbored $i = \{CD_{101}, \ldots . CD_{503}\}$, which are the 59 Community Districts across the five boroughs, as a function of number of units and other controls that affect the price:

$$\ln(price_{it}) = \beta_0 + \beta_1 ln(units_{it}) + \beta_2 X_{it} + \gamma t + \delta d_i + \varepsilon_{it},$$

where β_1 is the percent change in price with a 1% change in units. d_i are the neighborhood fixed effects. *t* is year variable (year dummy). X_{it} are vector of control variables.

This specification is problematic on two fronts. First is that the interpretation is in comparing two CDs, holding everything constant, β_1 gives the difference in price. But we want to know how new construction affects prices within (and across) CDs. For this reason, I estimate the model in differences:

$$\Delta\ln(price_{it}) = \alpha_0 + \beta_1 \Delta ln(units_{it}) + \beta_2 \Delta X_{it} + \gamma \Delta t + \varepsilon_{it},$$

which estimates the effect of a change of units within a CD.

Secondly, $\Delta ln(units_{it})$ is endogenous and for this reason, I estimate two types of models—instrumental variables and three stage regressions.

In the IV model:

$$\Delta ln(units_{it}) = \theta_0 + \theta_1 Z_{it} + \mu_{it},$$

Where Z_{it} are controls that affect the changes in units and obey the condition, $corr\big(Z_{it}, \Delta\ln(price_{it})\big) = 0$.

The related three-stage model is based on the simulateoss equation model:

$$\Delta\ln(price_{it}) = \alpha_0 + \beta_1 \Delta ln(units_{it}) + \beta_2 \Delta X_{it} + \gamma \Delta t + \varepsilon_{it},$$

$$\Delta ln(units_{it}) = \theta_0 + \theta_1 Z_{it} + \theta_2 \Delta\ln(price_{it}) + \mu_{it}$$

IV. Regression Results

1. Instrumental Variables Regressions

This section gives the results of several regressions. Because we are looking at the effect of unit growth on prices, all the regressions are performed in differences, with a change in housing prices over time for each CD as determined by a set of controls, which are changes in the CD and borough variables. In some specifications, year and borough dummies are also included. (*Boro1=MN*, *Boro2=BX*, *Boro3=BK*, *Boro4=QN*, *Boro5=SI*). Note that "*_L1*" means the variable was lagged one period.

Table 1 gives results of an instrumental variable regression of the changes in the log of real median price per square foot of residential buildings (of all kinds), `dlnMedRppsf`, on the

change in the log of residential units (which is instrumented), change in unemployment, the lag of the fraction of one- or two-family units, the change in the % of residents with college degrees or higher at the borough level, change of percent of white residents (as a proxy for gentrification), the total area of the CD devoted to residential land, lagged one period, and the log of the CD area.

All regression specification show that *overid* test suggest valid instruments and that the instruments are moderately strong. See Table 2 for first stage regressions. The results show the coefficient for *dlnResUnits* is between -1.1 and -1.73 across specifications.

Table 1: IV Regression.

	(1) dlnMedRppsf	(2) dlnMedRppsf	(3) dlnMedRppsf	(4) dlnMedRppsf
dlnResUnits	-1.641***	-1.131***	-1.373***	-1.725***
	(-7.19)	(-5.29)	(-6.33)	(-18.15)
dUnemp	-1.342**	-1.316**	-1.273**	-1.421**
	(-2.20)	(-2.03)	(-2.12)	(-2.39)
onefamily_L1	-0.000851***	-0.000820***	-0.00109***	-0.00137***
	(-3.17)	(-3.05)	(-3.85)	(-4.93)
dB_BA	0.521	0.749	0.835*	0.387
	(0.80)	(1.28)	(1.84)	(0.55)
dWhite	1.349***	1.318***	1.253***	1.400***
	(6.73)	(7.02)	(9.87)	(6.80)
lnTotalResLArea			-0.0585	
			(-1.43)	
lnCDArea				0.0289**
				(2.18)
_cons	0.361***	0.342***	1.361**	0.359***
	(9.44)	(7.66)	(1.99)	(8.01)
N	177	177	177	177
R-sq	0.753	0.763	0.764	0.752
adj. R-sq	0.743	0.753	0.752	0.741
First	14.7	14.4	9.8	14.4
Overid (p-value)	0.28	0.26	0.76	0.277
Endog (p-value)	0.046	0.11	0.06	0.04

dlnResUnits are instrumented. See Table 2 for first stage. t statistics in parentheses. * p<0.10, ** p<0.05, *** p<0.01t statistics in parentheses. Year dummies in all equations. Standard errors clustered by borough

Table 2 shows a first-stage OLS regression od change in log of residential units on several instruments. The change in the log of households with children; the fraction of the CD area zoned for residential land, lagged one period; the log of income generated in the CD by the Finance, Insurance, and Real Estate (FIRE) industry, lagged one period, and a year dummy for 2019 (the other years were not stat. sig.). The second stage controls are given in Table 3.

Table 2: First-stage regressions for IV regressions in Table 1

	(1) dlnResUnits	(2) dlnResUnits	(3) dlnResUnits	(4) dlnResUnits
dlnBHH	0.200***	0.198***	0.103*	0.200***
	(5.85)	(5.42)	(2.22)	(5.83)
Rzone_pct_L1	-0.000647***	-0.000586***	-0.000630***	-0.000645***
	(-11.20)	(-9.68)	(-15.36)	(-11.32)
lnbb_firegdp_L1	-0.0106***	-0.0116***	-0.232***	-0.0106***
	(-6.09)	(-6.75)	(-5.59)	(-6.35)
year3	-0.0143**	-0.0150**	0.0375***	-0.0143**
	(-4.56)	(-3.52)	(4.94)	(-4.32)
year4	0.0248	0.0252	0.0985***	0.0247
	(1.33)	(1.42)	(6.59)	(1.32)
dUnemp	-0.126	-0.133	-0.119	-0.128
	(-1.12)	(-1.24)	(-1.19)	(-1.12)
%onefamily_L1	0.00000334	-0.0000784	0.00000306	-0.0000129
	(0.04)	(-1.31)	(0.02)	(-0.14)
dB_BA	-0.407*	-0.448*	-0.214*	-0.410
	(-2.17)	(-2.13)	(-2.64)	(-2.12)
dWhite	0.0999	0.110	0.0842	0.101
	(1.52)	(1.47)	(1.21)	(1.49)
lnLotArea_L1		0.0610		
		(1.69)		
boro2			-0.644***	
			(-5.14)	
boro3			-0.484***	
			(-4.99)	
boro4			-0.499***	
			(-5.01)	
boro5			-0.928***	
			(-5.13)	
lnTotalResArea_L1			-0.00457	
			(-1.87)	
lnCDArea				0.000841
				(0.29)
_cons	0.274***	-0.733	4.517***	0.274***
	(9.71)	(-1.22)	(5.63)	(10.24)
N	177	177	177	177
R-sq	0.298	0.341	0.409	0.298
adj. R-sq	0.260	0.301	0.358	0.256

t statistics in parentheses. * p<0.10, ** p<0.05, *** p<0.01. Standard errors clustered by borough.

Table 3 repeats the IV regression exercise but with different controls and instruments. The results are broadly similar.

Table 3: Instrumental Variable Regressions

	(1) dlnMedRppsf	(2) dlnMedRppsf	(3) dlnMedRppsf	(4) dlnMedRppsf
dlnResUnits	-0.889***	-1.641***	-1.596***	-1.377***
	(-3.57)	(-7.19)	(-7.18)	(-7.19)
year3	-0.437***	-0.451***	-0.450***	-0.444***
	(-6.20)	(-7.80)	(-7.40)	(-7.00)
year4	0.142***	0.147***	0.146***	0.144***
	(3.84)	(3.39)	(3.45)	(3.64)
dUnemp	-1.304**	-1.342**	-1.339**	-1.400**
	(-1.98)	(-2.20)	(-2.20)	(-2.24)
dB_BA	0.857	0.521	0.541	0.547
	(1.64)	(0.80)	(0.87)	(0.80)
%onefamily_L1	-0.000805***	-0.000851***	-0.000849***	-0.00132***
	(-2.89)	(-3.17)	(-3.09)	(-4.96)
dWhite	1.303***	1.349***	1.346***	1.376***
	(7.38)	(6.73)	(6.96)	(6.80)
lnCDArea				0.0276**
				(2.08)
_cons	0.333***	0.361***	0.359***	0.346***
	(6.86)	(9.44)	(8.38)	(7.00)
N	177	177	177	177
R-sq	0.766	0.753	0.754	0.760
adj. R-sq	0.757	0.743	0.744	0.748
F-stat	11.3	14.0	12.4	11.9
Overid (p-value)	0.23	0.13	0.46	0.07
Endgo (p-value)	0.09	0.02	0.04	0.01

t statistics in parentheses. * p<0.10, ** p<0.05, *** p<0.01. Standard errors clustered by borough.

Table 4: First stage IV regression for Table 3 results

	(1) dlnResUnits	(2) dlnResUnits	(3) dlnResUnits	(4) dlnResUnits
dlnBHH	0.0971	0.200***	0.197***	0.167**
	(1.86)	(5.85)	(6.94)	(4.27)
Rzone_pct_l1	-0.000663***	-0.000647***	-0.000650***	-0.000636***
	(-16.81)	(-11.20)	(-10.56)	(-13.79)
lnbb_firegdl_L1	-0.243***	-0.0106***	-0.00986***	-0.0792
	(-4.93)	(-6.09)	(-4.97)	(-1.18)
boro2	-0.672**			
	(-4.58)			

boro3	-0.506**			
	(-4.43)			
boro4	-0.521**			
	(-4.48)			
boro5	-0.969**			
	(-4.56)			
year3	0.0372**	-0.0143**	-0.0115***	-0.00463
	(3.93)	(-4.56)	(-4.75)	(-0.50)
year4	0.0987***	0.0248	0.0277	0.0319
	(6.10)	(1.33)	(1.39)	(1.34)
dUnemp	-0.121	-0.126	-0.130	-0.130
	(-1.26)	(-1.12)	(-1.32)	(-1.24)
dB_BA	-0.202*	-0.407*	-0.380	-0.303**
	(-2.18)	(-2.17)	(-2.13)	(-2.84)
%onefamily_L1	0.00000698	0.00000334	0.0000184	-0.000125
	(0.05)	(0.04)	(0.18)	(-1.01)
dWhite	0.0883	0.0999	0.134	0.0857
	(1.18)	(1.52)	(1.77)	(1.49)
dFB			0.144*	
			(2.32)	
lnBGDP_l1				0.0732
				(1.01)
lnCDArea				-0.00180
				(-1.61)
_cons	4.640***	0.274***	0.258***	0.108
	(5.00)	(9.71)	(7.97)	(0.63)
N	177	177	177	177
R-sq	0.406	0.298	0.316	0.316
adj. R-sq	0.359	0.260	0.275	0.270

t statistics in parentheses. * p<0.10, ** p<0.05, *** p<0.01. Standard errors clustered by borough

Table 5 gives IV regressions for change in log of median monthly rents. The `dlnResUnits` coefficient is between -0.75 and -1.16 across spedfications.

Table 5: IV regressions for dln(Median Monthly Rents).

	(1)	(2)	(3)	(4)	(5)
	dlnRents_adj	dlnRents_adj	dlnRents_adj	dlnRents_adj	dlnRents_adj
dlnResUnits	-1.155***	-1.108***	-1.087**	-0.746**	-0.738***
	(-3.13)	(-3.07)	(-2.12)	(-2.15)	(-3.62)
year3	0.0710***	0.0707***	0.0609***	0.0631***	0.0673***
	(3.16)	(3.17)	(2.77)	(2.85)	(3.33)
year4	0.0464	0.0465	0.0385	0.0478	0.0547**
	(1.49)	(1.51)	(0.98)	(1.29)	(2.04)
%onefamily_L1	-0.000312	-0.000308	0.000125	0.000199	0.000102
	(-0.92)	(-0.92)	(0.30)	(0.58)	(0.27)
lnTotalResArea_L1	-0.0546***	-0.0537***	-0.0529***	-0.0462***	-0.0472***
	(-8.86)	(-8.75)	(-6.41)	(-8.71)	(-11.75)

dlnBPIPC	0.329***	0.326***	0.387**	0.317**	0.267***
	(4.43)	(4.28)	(2.37)	(1.98)	(3.54)
dWhite	0.418***	0.419***	0.459***	0.429***	0.411***
	(5.11)	(5.07)	(7.22)	(4.93)	(4.06)
dB_BA			-0.453	-0.255	
			(-1.53)	(-0.79)	
distESB			-0.224**	-0.281***	-0.229***
			(-2.36)	(-6.02)	(-3.95)
_cons	0.984***	0.968***	0.968***	0.851***	0.865***
	(9.25)	(9.13)	(6.58)	(8.89)	(12.63)
N	177	177	177	173	177
R-sq	0.182	0.197	0.221	0.306	0.299
adj. R-sq	0.148	0.164	0.179	0.268	0.266
First stage stat	6.65	8.27	6.46	5.82	9.14
Overid (p-value)	0.55	0.62	0.41	0.33	0.32
Endog (p-value)	0.002	0.004	0.01	0.05	0.023

t statistics in parentheses * p<0.10, ** p<0.05, *** p<0.01. Standard errors clustered by borough.

Table 6: First stage of IV regressions

	(1)	(2)	(3)	(4)	(5)
	dlnResUnits	dlnRents_adj	dlnResUnits	dlnResUnits	dlnResUnits
dlnBHH	0.196*	-0.344*	0.0549	0.0883	0.183*
	(2.14)	(-2.24)	(0.92)	(1.05)	(2.48)
Rzone_pct_l1	-0.000450***	0.000242	-0.000563***	-0.000543**	-0.000493***
	(-9.21)	(0.84)	(-6.57)	(-4.56)	(-7.75)
lnBPop_l1	-0.00347				
	(-0.55)				
dBlack	0.271	-0.291	0.205	0.245	0.353
	(1.54)	(-1.82)	(1.61)	(1.69)	(1.64)
dHisp	0.342	-0.528*	0.274	0.299	0.401
	(1.40)	(-2.70)	(1.41)	(1.60)	(1.52)
year3	-0.00306	0.0790***	-0.00576	0.00704	-0.0000849
	(-0.34)	(5.21)	(-0.60)	(0.32)	(-0.01)
year4	0.0433	-0.0338	-0.0103	0.0152	0.0444
	(1.09)	(-0.59)	(-0.52)	(0.38)	(1.47)
%onefamily_L1	0.000137	-0.000268	0.000118	0.000213	0.0000746
	(1.08)	(-0.62)	(0.79)	(1.28)	(0.58)
lnTotalResArea_L1	-0.0145*	-0.0346**	-0.0132*	-0.0112**	-0.0120
	(-2.24)	(-4.31)	(-2.69)	(-3.18)	(-2.09)
dlnBPIPC	-0.0434	0.458**	0.162	0.126	-0.0279
	(-0.50)	(3.17)	(1.62)	(1.34)	(-0.46)
dWhite	0.317	0.0427	0.316	0.377	0.438
	(1.71)	(0.28)	(1.98)	(1.84)	(1.89)
dlnResUnits		0.0892			
		(0.39)			
dB_BA			-0.498*	-0.408*	
			(-2.68)	(-2.69)	

distESB			0.101**	0.0323	0.0523
			(3.71)	(1.08)	(1.21)
dFB				0.206*	0.236
				(2.36)	(2.02)
lnHDProps_L1				-0.00192	
				(-0.66)	
_cons	0.362***	0.562**	0.272**	0.247**	0.264**
	(4.78)	(3.83)	(3.84)	(4.40)	(2.91)
N	177	177	177	173	177
R-sq	0.255	0.390	0.326	0.370	0.308
adj. R-sq	0.205	0.349	0.276	0.315	0.258

t statistics in parentheses. * p<0.10, ** p<0.05, *** p<0.01. Standard errors clustered by borough.

Table 7: Additional IV Specifications for Monthly rental income

	(1)	(2)	(3)	(4)	(5)
	dlnRents_adj	dlnRents_adj	dlnRents_adj	dlnRents_adj	dlnRents_adj
dlnResUnits	-1.012***	-0.546***	-0.697***	-1.138***	-0.465**
	(-3.97)	(-3.71)	(-3.36)	(-3.34)	(-2.49)
%onefamily_L1	-0.000228	-0.000180	0.000310	0.000282	0.000370
	(-0.66)	(-0.58)	(1.01)	(0.76)	(1.45)
lnTotalResArea_L1	-0.0218**	-0.0139	-0.0223***	-0.0303***	-0.0156
	(-2.13)	(-1.47)	(-6.34)	(-10.89)	(-1.43)
dlnBPIPC	0.291***	0.267***	0.371***	0.425***	0.343***
	(21.07)	(10.68)	(5.29)	(5.27)	(7.74)
dWhite	0.391***	0.402***	0.476***	0.483***	0.461***
	(6.84)	(5.70)	(4.60)	(5.18)	(4.09)
dB_BA			-0.707***	-0.923***	-0.595***
			(-4.52)	(-6.73)	(-5.28)
distESB			-0.264***	-0.280***	-0.300***
			(-3.84)	(-3.49)	(-6.24)
_cons	0.448***	0.304*	0.464***	0.612***	0.346*
	(2.62)	(1.94)	(8.42)	(11.09)	(1.81)
N	177	177	177	173	173
R-sq	0.120	0.223	0.239	0.150	0.270
adj. R-sq	0.094	0.201	0.207	0.114	0.239
First	5.7	6.7	9.1	6.6	7.9
Overid	0.20	0.34	0.29	0.021	0.001
Endog	0.065	0.41	0.27	0.020	0.33

t statistics in parentheses. * p<0.10, ** p<0.05, *** p<0.01. Standard errors clustered by borough.

Table 8: First stage regressions for results in Table 7.

	(1)	(2)	(3)	(4)	(5)
	dlnResUnits	dlnResUnits	dlnResUnits	dlnResUnits	dlnResUnits
dlnBHH	0.0783**	0.0713*	0.0688*	0.0580	0.0557
	(3.16)	(2.29)	(2.57)	(1.74)	(1.56)
Rzone_pct_L1	-0.000461***	-0.000488***	-0.000561***	-0.000581**	-0.000981***
	(-6.52)	(-8.06)	(-5.75)	(-4.45)	(-5.23)
lnBPop_L1	-0.0107*				
	(-2.51)				

%onefamily_L1	0.000190	0.000188	0.0000778	0.000197	-0.000153
	(1.16)	(1.22)	(0.50)	(1.27)	(-1.01)
lnTotalResArea_L1	-0.0149**	-0.0148*	-0.0174**	-0.0116**	-0.00575
	(-2.92)	(-2.74)	(-4.48)	(-3.54)	(-1.14)
dlnBPIPC	0.0743	0.0679	0.136*	0.157**	0.149**
	(1.28)	(1.10)	(2.20)	(2.81)	(3.03)
dWhite	0.0590**	0.0405	0.128**	0.156*	0.154*
	(2.96)	(1.46)	(3.23)	(2.57)	(2.44)
dB_BA			-0.481	-0.538*	-0.533*
			(-1.93)	(-2.17)	(-2.25)
distESB			0.107**	0.0590	0.0775
			(3.27)	(1.38)	(1.67)
dFB				0.141**	0.168**
				(2.89)	(3.08)
lnHDProps_L1				-0.00178*	-0.00237*
				(-2.56)	(-2.51)
AvgMax FAR					-0.00899***
					(-7.44)
_cons	0.448**	0.296**	0.345***	0.254**	0.230**
	(3.38)	(3.33)	(5.25)	(4.44)	(2.99)
N	177	177	177	173	173
R-sq	0.184	0.167	0.281	0.327	0.371
adj. R-sq	0.150	0.138	0.246	0.285	0.328

t statistics in parentheses. * p<0.10, ** p<0.05, *** p<0.01. Standard errors clustered by borough.

2. Three stage least squares regression

The three stage regressions also allow for an estimatin of the housing supply elasticity—i.e. the % change in supply with a given % change in price. Based on the data set, for median sales prices, the elasticity is hovers around 0.03 across specifications. Even larger estimates (in absolute value) are found for response of supply on prices.

Table 9: Three-stage least square results for dlnUnits and dlnMed PPSF

	(1)	(2)	(3)	(4)	(5)
	dlnResUnits	dlnResUnits	dlnResUnits	dlnResUnits	dlnResUnits
dlnResUnits					
dlnMedRppsf	0.0289***	0.0282***	0.0284***	0.0401***	0.0326***
	(3.21)	(3.17)	(3.38)	(3.52)	(3.71)
Rzone_pct_l1	-0.000323***	-0.000310***	-0.000726***	-0.000311***	-0.000631***
	(-2.91)	(-2.81)	(-5.05)	(-2.78)	(-4.89)
dlnBHH	0.0880*	0.0970**	0.101**	0.142**	0.226***
	(1.91)	(2.14)	(2.34)	(2.52)	(4.27)
dFB	0.184***	0.175***	0.121*	0.231***	0.141**
	(2.70)	(2.62)	(1.92)	(3.18)	(2.14)
lnAvgMaxFAR_L1			-0.00903**		
			(-2.23)		
year2				-0.0124**	

				(-1.96)	
dYear					0.00310***
					(2.72)
dHisp					0.154**
					(2.40)
boro2					0.0361***
					(4.63)
boro3					0.0263***
					(3.70)
boro4					0.0262***
					(3.42)
boro5					0.0363***
					(3.23)
_cons	0.0399***	0.0396***	0.0808***	0.0444***	0.0348***
	(4.07)	(4.06)	(5.54)	(4.43)	(3.48)
dlnMedRppsf					
dlnResUnits	-1.461	-2.791**	-1.587	-1.665	-2.230**
	(-1.17)	(-2.39)	(-1.47)	(-1.41)	(-2.41)
dUnemp	-1.677***	-1.643***	-1.647***	-1.332**	-1.375**
	(-3.21)	(-3.18)	(-2.98)	(-2.46)	(-2.52)
dPov	-0.380				
	(-0.85)				
dlnBGDP	0.836***	0.873***	0.922***	0.0648	
	(2.97)	(3.09)	(2.97)	(0.12)	
dWhite	1.362***	1.169***	1.172***	1.405***	1.559***
	(4.09)	(3.34)	(3.33)	(3.93)	(4.77)
year3	-0.447***	-0.477***	-0.459***	-0.612***	-0.469***
	(-9.48)	(-9.95)	(-9.26)	(-6.68)	(-9.91)
onefamily_~1		-0.00116**	-0.00104*	-0.000894*	
		(-2.21)	(-1.88)	(-1.75)	
year2				-0.146*	
				(-1.91)	
year4					0.154***
					(3.79)
_cons	0.256***	0.348***	0.313***	0.527***	0.346***
	(5.82)	(6.50)	(5.63)	(4.35)	(11.17)
N	177	177	168	177	177

t statistics in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table 10 gives three-stage least squares regression for the dln(monthy rents) and dln(residential units). Similar are results are found, but with a higher supply elasticity (may be due to more sticky nature of rents?)

Table 10: Three-stage least square regressions for monthly rents and units

	(1)	(2)	(3)	(4)	(5)
	dlnRents_adj	dlnRents_adj	dlnRents_adj	dlnRents_adj	dlnRents_adj
dlnRents_adj					
dlnResUnits	-1.314***	-1.332***	-1.278***	-0.525	-0.723**

	(-3.13)	(-3.20)	(-2.63)	(-1.29)	(-2.10)
year3	0.0740***	0.0742***	0.0654***	0.0601***	0.0671***
	(4.56)	(4.62)	(3.93)	(3.82)	(4.45)
year4	0.0470	0.0449	0.0415	0.0478	0.0548**
	(1.62)	(1.57)	(1.29)	(1.59)	(1.99)
onefamily_~1	-0.000362*	-0.000364*	0.000110	0.000244	0.000103
	(-1.74)	(-1.78)	(0.33)	(0.78)	(0.32)
lnTotalRes~1	-0.0543***	-0.0536***	-0.0519***	-0.0437***	-0.0472***
	(-3.83)	(-3.82)	(-3.65)	(-3.15)	(-3.73)
dlnBPIPC	0.333***	0.339***	0.369**	0.325**	0.267**
	(2.73)	(2.81)	(2.28)	(2.19)	(2.31)
dWhite	0.387***	0.379***	0.403***	0.508***	0.415***
	(3.05)	(3.02)	(3.08)	(4.12)	(3.51)
dB_BA			-0.352	-0.367	
			(-0.89)	(-1.05)	
distESB			-0.258	-0.239	-0.227
			(-1.62)	(-1.59)	(-1.48)
_cons	0.980***	0.967***	0.955***	0.801***	0.864***
	(4.00)	(3.99)	(3.87)	(3.33)	(3.96)
dlnResUnits					
dlnRents_adj	0.188**	0.150**	0.128*	0.291***	0.166**
	(2.37)	(2.03)	(1.82)	(3.42)	(2.29)
dlnBHH	0.0694	0.0575	0.0517	0.0558	0.0422
	(1.56)	(1.33)	(1.22)	(1.23)	(0.99)
Rzone_pct_l1	-0.000278**	-0.000312***	-0.000322***	-0.000313**	-0.000340***
	(-2.31)	(-2.68)	(-2.79)	(-2.55)	(-2.92)
lnBPop_l1	-0.0122*				
	(-1.83)				
dBlack	0.102	0.107	0.0976	0.153*	0.0974
	(1.34)	(1.42)	(1.32)	(1.93)	(1.31)
dHisp	0.244***	0.250***	0.236***	0.282***	0.223**
	(2.72)	(2.80)	(2.71)	(2.99)	(2.54)
dFB				0.194**	0.177**
				(2.48)	(2.43)
lnHDProps_l1				-0.00245***	
				(-2.60)	
_cons	0.198**	0.0288**	0.0313***	0.0331**	0.0280**
	(2.15)	(2.33)	(2.58)	(2.49)	(2.28)
N	177	177	177	173	177

t statistics in parentheses. * p<0.10, ** p<0.05, *** p<0.01

V. Information about the Figures in the Blog Post

1. Maps: Look at the percent change in units and medial sales prices between 2003 and 2019 for each CD.
2. First Set of Scatter Plots: % Change in real median sale prices per square foot and real median monthly rents for each CD from 2003 to 2019.

3. Second Set of Scatter Plots.
    a. First a regression was run for real median sales prices per square foot on all controls except the *dlnResUnits* and the residuals were obtained, using the specification in Table 1, Equation (3).
    b. Next *dlnResUnits* was regressed on all IVs and controls, and a predicted values was gotten.
    c. Next the predicted value was regressed on the controls in the price equation (in (a)), and the residuals were obtained.
    d. The scatter plot is all residuals for the price equation (a) and dln(housing units) equation in (c) for all years and all CDs.
    e. The above steps were run for the median monthly rents using Table 5, Equation (3).