Data Sources for Why Has Gotham Been in a "Housing Emergency" for 80 Years? (Part I)

Sources for NYC Vacancy 1909 to 1923

Variable	Details	Source
vancancy_int_ann	Annual Vacancy Interpolated	Various reports on NYC Housing, missing years were obtained via interpolation command in Stata
netnewper1000housing	Net new housing per 1000 of existing units	Various reports on NYC Housing
lnPop	log of annual NYC Pop	Census years and interpolated to make an annual series
lnHH	log annual # of households	Census years and interpolated to make an annual series
lnPop_lnHH	Interaction of log pop and log # HHs	
lnRGDPcap	log per cap US RGDP	Measuringworth.com
year	Trend	
lnTaxRate	Tax rate on Class 2 Buildings	Various reports of NYC Dept of Finance
worldwar	dummy if year of World War (1917-19, 1940-45)	
inflation	US CPI Inflation Rate	Measuringworth.com
vr	Residuals from vacancy regression (to aid with OVB and autocorrelated errors)	
lnRealRents	Index of real asking rents in US	From Shertzer et al (2024)
rcYears	Dummy from 1943-68	
rsYears	Dummy from 1969-2023	
dlnMataerialCosts	Inflation Rate for Building Materials Costs	See Barr (2010) and FRED Index WPUSI012011
post1964	Dummy for 1964-2023 (when 1961 zoning codes were in effect)	
Note: More details about the variables are available upon request		

Regressions for Annual Vacancy Rates

	(1) OLS	(2) OLS	(3) OLS	(4) 2SLS	(5) 2SLS	(6) 2SLS
L2.netnewper1000housing	0.0361***	0.0313***	0.0301***	0.0747**	0.0762**	0.0725**
	(5.80)	(6.52)	(6.97)	(2.21)	(2.24)	(2.30)
L2.lnPop	53.85*	74.16***	85.27***	-124.6	-120.4	-146.8
	(1.67)	(3.57)	(2.93)	(-0.92)	(-0.88)	(-1.35)
L2.lnHH	64.03*	82.79***	91.28***	-98.76	-94.03	-125.3
	(1.91)	(3.84)	(3.08)	(-0.71)	(-0.67)	(-1.07)
L2.lnPop_lnHH	-3.918*	-5.214***	-5.860***	7.636	7.323	9.248
	(-1.79)	(-3.70)	(-2.98)	(0.83)	(0.79)	(1.22)
L.lnRGDPcap	-2.776**	-0.647	-0.510	-16.28***	-16.75***	-17.59***
	(-2.35)	(-0.71)	(-0.56)	(-6.21)	(-6.97)	(-7.75)
year	0.0485	0.0192	0.0156	0.290***	0.299***	0.297***
	(1.62)	(0.92)	(0.72)	(3.72)	(3.88)	(3.86)
L2.lnTaxRate	0.410	-0.0689	-0.0626	-3.625	-3.481	-3.339
	(0.61)	(-0.18)	(-0.13)	(-1.27)	(-1.26)	(-1.18)
L.vacancy_int_ann	0.904***	0.923***	0.944***			
	(21.31)	(28.46)	(23.96)			
worldwar		-0.950***	-0.775**	-0.480		
		(-3.14)	(-2.39)	(-0.54)		
L.inflation		-0.0273*	-0.0302**	-0.0415	-0.0460	-0.0833
		(-1.92)	(-2.18)	(-0.69)	(-0.79)	(-1.45)
L.vr		0.370**	0.295*	0.908**	0.924**	
		(2.57)	(1.89)	(2.26)	(2.33)	
L2.lnRealRents			0.354			
			(0.99)			
L3.rcYears			0.530	-5.938***	-5.620***	-5.007***
			(1.28)	(-3.56)	(-3.95)	(-5.03)
L3.rsYears			0.705	-2.106	-1.863	-0.327
			(1.42)	(-0.95)	(-0.94)	(-0.23)
_cons	-950.2*	-1208.4***	-1355.5***	1251.1	1174.5	1614.4
	(-1.87)	(-3.74)	(-3.03)	(0.59)	(0.55)	(0.92)
N	114	113	113	113	113	115
R-sq				0.734	0.732	0.706
adj. R-sq				0.702	0.703	0.678

t statistics in parentheses * p<0.10, ** p<0.05, *** p<0.01. Newey-West s.e. for OLS equations. Robust se for IV equations. Note "L" is short for lagged variable, with number being years lagged, i.e. L2 is variable is lagged two years prior.

First Stage Regression for Equation (4)

```
                                     Number of obs =    113
                                     F(15, 97)     =  59.58
                                     Prob > F      = 0.0000
                                     R-squared     = 0.9021
                                     Adj R-squared = 0.8870
                                     Root MSE      = 5.1238
```

L2. netnewper1000housing	Coefficient	Std. err.	t	P>|t|	[95% conf. interval]	
year	-.6456778	.2098652	-3.08	0.003	-1.062202	-.2291534
lnTaxRate L2.	13.01084	6.622386	1.96	0.052	-.1327663	26.15444
inflation L1.	-.0626128	.1295382	-0.48	0.630	-.3197102	.1944846
lnRGDPcap L1.	22.14333	7.788738	2.84	0.005	6.684838	37.60181
lnPop L2.	-203.7034	380.527	-0.54	0.594	-958.9441	551.5374
lnHH L2.	-305.7811	403.1273	-0.76	0.450	-1105.877	494.3151
lnPop_lnHH						

```
              L2. |   17.17438    26.26319     0.65    0.515    -34.95079    69.29956
                  |
               vr |
              L1. |  -.9345795    1.025318    -0.91    0.364    -2.969552    1.100393
                  |
          rcYears |
              L3. |   1.380455    3.231468     0.43    0.670    -5.033114    7.794025
                  |
          rsYears |
              L3. |   1.914937    4.816121     0.40    0.692     -7.64373     11.4736
                  |
  dlnMataerialCosts |
              L3. |  -11.37578    6.745591    -1.69    0.095    -24.76391    2.012349
                  |
         worldwar |
              L2. |  -8.889769    3.243776    -2.74    0.007    -15.32777   -2.451771
                  |
       lnRealRents |
              L3. |   8.283898    5.011094     1.65    0.102    -1.661737    18.22953
                  |
netnewper1000housing |
              L3. |   .6987407    .0535709    13.04    0.000     .5924174     .805064
                  |
          post1964 |
              L3. |  -6.998026    3.205939    -2.18    0.031    -13.36093   -.6351241
                  |
            _cons |   4713.958    6080.734     0.78    0.440    -7354.617    16782.53
```

Data Sources for Zip Code Regressions

Variable	Details	Source
lnResunits15	ln # residential units in zip code in 2015	2015 NYC PLUTO file
avgMaxResidFAR15	Avg max. allowable FAR in zip code in 2015	2015 NYC PLUTO file
lnMedHHY19	ln of Median HH income in 2019	ACS 5-year Avg, 2015-19
lnMedHHY_sq19	ln of Median HH income squared in 2019	ACS 5-year Avg, 2015-19
lnHH19	Log of number of Households in zip code in 2019	ACS 5-year Avg, 2015-19
lnPop	Log of pop in zip code ca 2019	ACS 5-year Avg, 2015-19
vacancyrate_2015	Avg zip code vacancy rate in 2015	ACS 5-year avg, 2011-15
AvgMaxResidFAR6plus	Avg. max allowable FAR of 6 or higher in zip code	2015 NYC PLUTO file
avgMaxReFAR15_4	Dummy if avg. max allowable far is bt 4 and <5	2015 NYC PLUTO file
avgMaxReFAR15_5	Dummy if avg. max allowable far is bt 5 and <6	2015 NYC PLUTO file
avgMaxReFAR15_6	Dummy if avg. max allowable far is bt 6 and <7	2015 NYC PLUTO file
avgMaxReFAR15_7	Dummy if avg. max allowable far is bt 7 and <8	2015 NYC PLUTO file
avgMaxReFAR15_8	Dummy if avg. max allowable far is bt 8 and <9	2015 NYC PLUTO file
avgMaxReFAR15_9	Dummy if avg. max allowable far is bt 9 and <10	2015 NYC PLUTO file
avgMaxReFAR15_10	Dummy if avg. max allowable far is 10	2015 NYC PLUTO file
BK	Brooklyn Dummy	
QN	Queens Dummy	
BX	Bronx Dummy	
SI	Staten Island Dummy	

Dep var.: Avg zip code vacancy for 2015-19.

	(1)	(2)	(3)	(4)	(5)
lnResunits15	2.293**	1.945**	2.066**	1.596*	2.317**
	(2.39)	(2.23)	(2.31)	(1.81)	(2.06)
avgMaxResidFAR15	0.298**	0.267**			
	(2.62)	(2.63)			
lnMedHHY19	37.26**	37.55***	37.98***	38.68***	39.38***
	(2.65)	(3.00)	(3.75)	(3.38)	(3.58)
lnMedHHY_sq19	-1.683**	-1.696***	-1.735***	-1.769***	-1.803***
	(-2.62)	(-2.97)	(-3.75)	(-3.38)	(-3.57)
lnHH19	-0.455	-0.246	0.0404	0.217	-0.686
	(-0.46)	(-0.25)	(0.04)	(0.21)	(-0.58)
lnPop	-2.358***	-2.045***	-2.414***	-2.205***	-2.449***
	(-2.95)	(-2.91)	(-3.75)	(-3.59)	(-2.71)
BX	-0.615**	-0.483**	-0.546***	-0.446**	0.0231
	(-2.66)	(-2.24)	(-2.80)	(-2.11)	(0.05)
SI	3.317**	2.413**	2.295**	2.173**	3.719***
	(2.61)	(2.39)	(2.39)	(2.37)	(2.86)
vacancyrate_2015		0.288***	0.278***	0.327***	
		(2.69)	(2.81)	(3.42)	
AvgMaxResidFAR6plus			2.340***		
			(5.27)		
avgMaxReFAR15_4				0.546	1.298**
				(1.13)	(2.48)
avgMaxReFAR15_5				1.063**	2.002***
				(2.01)	(3.88)
avgMaxReFAR15_6				2.792***	3.775***
				(4.94)	(5.96)
avgMaxReFAR15_7				2.625***	3.533***
				(3.67)	(3.30)
avgMaxReFAR15_8				4.212***	2.782***
				(4.21)	(3.12)
avgMaxReFAR15_9				1.094	3.070***
				(1.11)	(5.24)
avgMaxReFAR15_10				1.220	1.376*
				(1.63)	(1.98)
BK					0.994*
					(1.95)
QN					0.422
					(0.69)
_cons	-196.0**	-200.5***	-200.0***	-203.2***	-201.9***
	(-2.60)	(-2.96)	(-3.68)	(-3.28)	(-3.44)
N	170	170	170	170	170
R-sq	0.443	0.522	0.551	0.575	0.505
adj. R-sq	0.415	0.495	0.526	0.533	0.453
AIC	646.3	622.3	611.5	612.4	640.2
BIC	674.5	653.6	642.8	659.4	690.3

Dep. Var.: Avg. vacancy rate from 2015-2019 at the zip code. t statistics in parentheses. s.e.'s clustered by community district. * p<0.10, ** p<0.05, *** p<0.01. All equations estimated via OLS.