

Data Sources and Analysis for New York City Population Growth and Housing Prices

Jason M. Barr
Rutgers University-Newark

I. Data Sources

1. *Index of Real Housing Prices from 1974 to 2016*: I use the same housing price index as given in the 2008 NYU Furman Center report, "[Trends in New York City Housing price Appreciation](#)." I updated this index up to 2016.
2. *Annual NYC City Unemployment Rate*: <https://labor.ny.gov/stats/nyc/index.shtm>. I take average of monthly values from 1974 to 2016. (1973-1975 was obtained directly from NYS Dept of Labor.)
3. *U.S. Unemployment Rate*: [Table LNU04000000](#). Annual 1974-2016.
4. Annual number of Legal Immigrants to U.S.: Department of Homeland Security: <https://www.dhs.gov/immigration-statistics/yearbook/2016/table1>.
5. *Annual Number of Birth and Deaths in NYC*: 1910-1940: Table A-4. 1941-1967; Table A-5 in Rosenwaike, I. (1972). *Population History of New York City*. Syracuse University Press. 1968- 2016: Table PC1: "[Summary of Vital Statistics 2016: The City of New York](#)."
6. *Annual NYC Population*. Annual population was estimated by interpolating from census population counts. First, for two census years, an average annual growth rate was calculated from the formula $g_c = \ln(\text{NYC Census Pop}_c / \text{NYC Census Pop}_{c-10}) / 10$, where c is a census year. Then, for a given inter-census period, population was calculated as $\text{Pop}_t = \text{Pop}_{t-1} \exp(g_c)$. For 2010 to 2016, the growth rate was as above, but calculated by dividing by 6 to get an annual rate. Census population figures are from the Wikipedia page on New York City.
7. *Annual Net Migration Estimates*: For each year, I first took the change in population from the year prior, then I subtracted from this the number of births minus deaths. That is annual migration for a given year was given by the formula: $\text{Net Migration}_t = (\text{Pop}_t - \text{Pop}_{t-1}) - (\text{Births}_t - \text{Deaths}_t)$.

I. Descriptive Statistics & Graphs

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Avg Ann.Net Migration	43	-35030.77	41312.55	-116408	9461
Ln(price index)	44	.6041353	.4630859	-.1036864	1.305937
year	45	1995	13.13393	1973	2017
US immigration	44	836746	318084.4	385378	1826595
births	44	122116.5	9108.43	106021	139630
deaths	44	66352.36	9631.493	52455	82319
birth - deaths	44	55764.11	14844.58	28320	74888
NYC population	45	7730016	465040.3	7071639	8622698

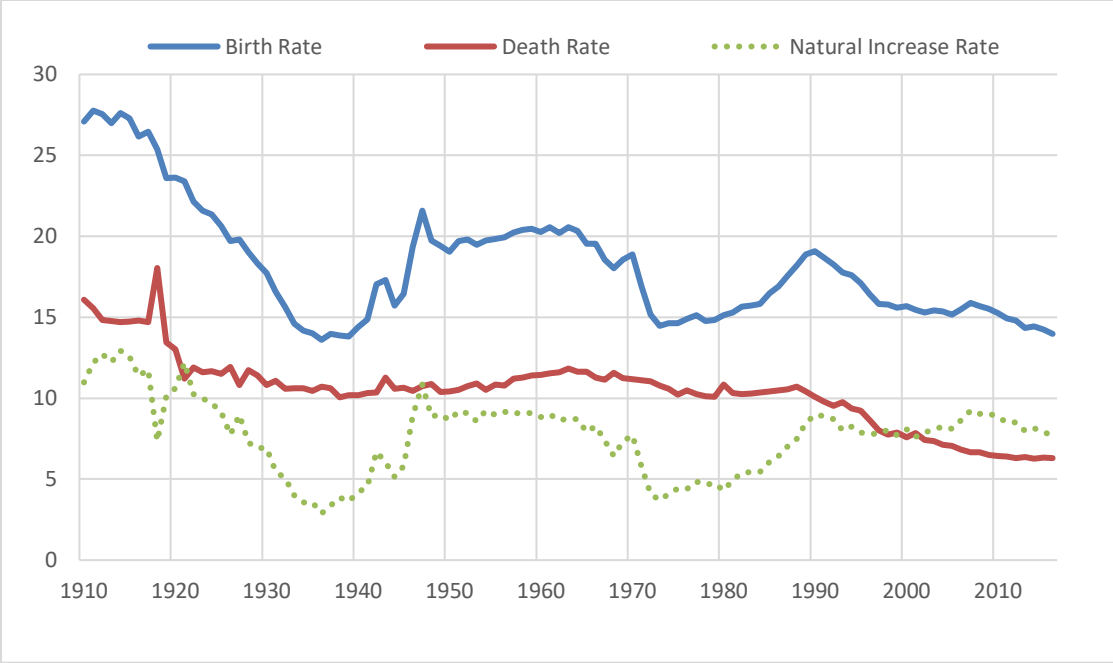


Figure 1: Birth Rates, Death Rates, and Natural Increase Rates for New York City, 1910-2016. Each rate is per 1000 people. Sources: See above.

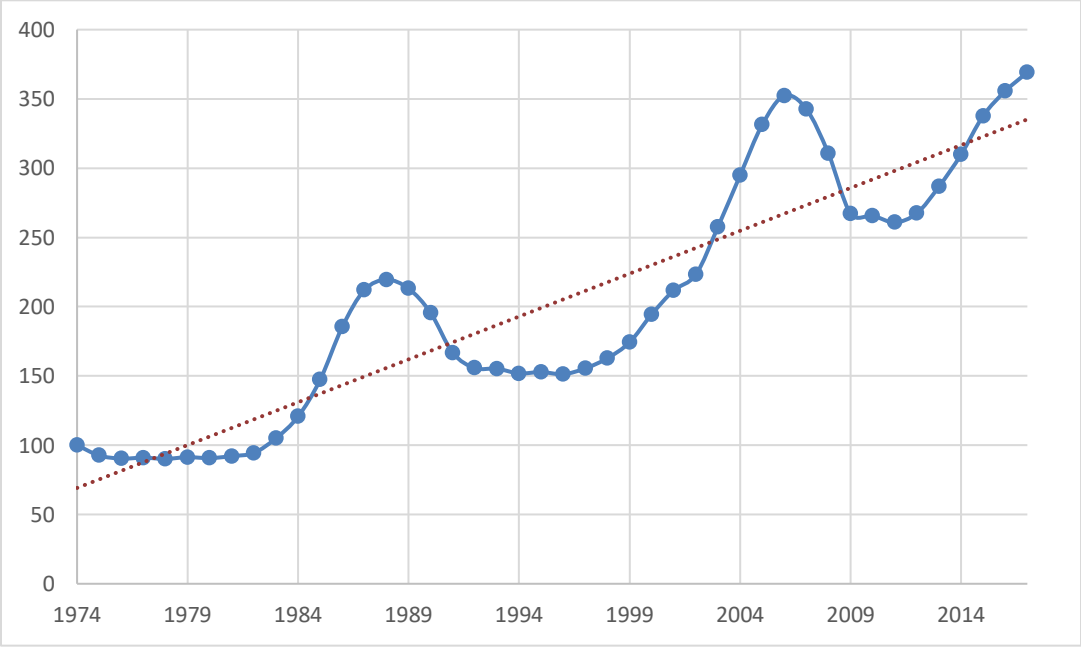


Figure 2: NYC Real Housing Price Index from 1974 (100) to 2017. Brown line is the trend line. Source: See above.

II. Regression Results

Each regression has annual net migration as the dependent variable. The independent variables include the log of the housing price index, log of the total migration to the U.S., log of the ratio of NYC unemployment to U.S. unemployment, and the log of New York City population. These variables are each lagged one period, to help mitigate endogeneity and because they likely provide a better fit of the data.

These control variables are included to better isolate a price effect, if one exists. First is the ratio of the New York City unemployment rate to the national rate. When it's high it suggests New York a relatively worse place to work relative to the rest of the country; thus there should be an negative relationship between this variable a net migration. The international immigrants is included to measure the degree to which inflow are affected by foreign migrations, since, presumably, many of them will choose to come to New York. Thus there should be a positive relationship between international migration and New York City net migration.

All specifications also include a year (trend) variable and decade dummy variables. Some specifications include a lag dependent variable. Note that Equations (1) – (3) are estimated using ordinarily least squares; Equation (4) is estimated via Newey-West because of the presence of serial correlation in the residuals, as indicated by the Breusch-Godfrey LM test for autocorrelation. Equations (1) - (3) are include robust standard errors. P-values are given below the coefficient estimates. P-values less than 0.1 suggest statistical significance. Note that Equation (4) includes the lag of the population which is likely endogenous, but inclusion of it does not change the results.

The regression results show that there is a negative and statically significant relationship between housing prices and net migration, controlling for other factors driving net migration. The results suggest that for example a 10% increase in housing prices reduces annual net migration between 4,000-5,000 people per year. Also note that the migration and unemployment measure have the expected signs, though they are not statistically significant.

Variable	eq1	eq2	eq3	eq4
Ln(index)				
L1.	-41946.67 0.01	-57236.00 0.02	-38226.20 0.07	-37832.11 0.09
year	3191.04 0.05	3742.61 0.14	1573.68 0.31	2233.31 0.12
decade2	74977.37 0.00	69314.20 0.03	43111.79 0.16	34267.32 0.10
decade3	84463.99 0.01	84250.63 0.04	56341.65 0.10	51808.39 0.01
decade4	24680.34 0.62	21329.09 0.73	23807.28 0.56	22705.03 0.34
decade5	34715.47 0.58	23833.78 0.78	34477.62 0.52	32968.11 0.32
Ln(Migration)				
L1.		2591.43 0.87	5529.45 0.59	133.38 0.99
Ln(NY_USEMP)				
L1.		-28175.58 0.57	-24726.82 0.38	-23845.17 0.44
netmigration				
L1.			0.47 0.06	0.47 0.00
lnPop				-156252.02 0.47
_cons	-6426503.01 0.05	-7543312.14 0.14	-3239417.66 0.31	-2001885.23 0.55
N	42	42	42	42
aic	953.23	956.04	936.36	.
r2	0.81	0.81	0.89	