Determinants of New York City Residential Rental Prices

Larissa Marco[1]
Washington University in St. Louis

Abstract

Economists and financial analysts have invested a tremendous amount of energy in an attempt to explain housing sale and rental prices. Many different approaches have been taken to try to understand residential prices, taking into account qualities such as neighborhood characteristics and housing units' physical characteristics. This paper considers the physical and demographic characteristics of New York City neighborhoods to better understand residential rental prices in this region. The empirical analysis reveals that premiums are charged for rental properties based on location in the following order: Staten Island, The Bronx, Queens, Brooklyn, and Manhattan, with Manhattan commanding the largest premium. High rental prices are also correlated with high median household incomes and high crime rates. Finally, increases in rental prices are correlated with increases in rent-regulated or rent-subsidized housing. At a certain point, however, the level of rent-regulated housing reaches its maximum and rental prices fall thereafter. Ultimately, this study offers a bundle of attributes that may be used to understand housing prices and establishes the attributes most relevant to the pricing equation. With this study, renters, landlords, and real estate developers can better understand the real estate market and make more informed decisions as players in the market.

---

1 Larissa Marco is a fourth year Economics student at Washington University in St. Louis. She would like to thank Robert Parks for his support, guidance, and dedication, all of which were instrumental to the writing of this paper. She would also like to thank her friends and family for their love and encouragement. Please direct any questions or comments to larissa.marco@gmail.com.

I. Introduction

The real estate market is a highly complex and challenging one to understand. "Housing is a multidimensional good differentiated into a bundle of attributes that vary in both quantity and quality."[2] Because each piece of property is unique, it is often difficult to identify the appropriate variables that will explain residential sale and rental prices. This paper's objective is twofold: to ascertain the bundle of attributes that can be used to understand housing prices in New York City and to determine the attributes within this bundle that are most relevant to the pricing equation.

II. Literature Review

Economists and financial analysts have invested a tremendous amount of energy and effort in an attempt to explain residential sale and rental prices. One approach that has been used to model residential prices evaluates properties according to their neighborhood characteristics. These characteristics include quality of schooling systems, level of noise pollution, air quality, proximity to parks, proximity to bodies of water and quality of transportation systems. Other research has tried to explain residential prices on the basis of housing units' physical characteristics, which include number of bedrooms, number of bathrooms, square footage and age. Additional surveys model residential prices as functions of apartment amenities such as indoor pools, gyms, and covered parking. Palmon, Smith and Sopranzetti (2004) offer an alternative approach in "Clustering in Real Estate Prices."[3] The authors look at transactions themselves as the units of analysis, investigating the impacts that listing price, closing price, number of days on the market and number of properties listed by a given real estate agent have on the resulting real estate prices.

A review of the available literature on real estate prices reveals the many lenses through which these prices can be studied. In a housing market such as New York City, the high wages and cultural attractions available in NYC translate into a particularly high demand for space in the region, making the wide variance in NYC residential prices all the more interesting. An analysis of these prices must therefore reflect New York City's unique characteristics. erhaps a more thorough analysis could come from a study that incorporates all sets of potentially relevant data; future research could

2 Can, Ayse. "The Measurement of Neighborhood Dynamics in Urban House Prices." Economic Geography (66.3, 1990).

3 Palmon, Oded, Barton A. Smith and Ben J. Sopranzetti. "Clustering in Real Estate Prices: Determinants and Consequences." JRER (26.2, 2004) 115-136.

involve such an analysis. This paper, however, will be limited to a select set of physical and demographic attributes of New York City neighborhoods in the hopes of solving the mystery of NYC residential rental prices.

Compelling results from a study of the determinants of residential rental prices can have useful real-world implications. With knowledge of the variables relevant to the rental price equation, landlords can learn how to extract maximum rents from their tenants. A refined rental price equation can also help landlords and developers refine their estimates of the expected returns on their investments during their cost/benefit analysis of each residential project. From the consumer side, prospective tenants can break down rental asking prices into their relevant components and negotiate along these lines. Furthermore, pricing information can give economists and financial analysts a better understanding of the real estate market and its supply and demand functions.

III. Data

The data for this research were obtained from *The Furman Center for Real Estate and Urban Policy: State of New York City's Housing and Neighborhoods 2005*;[4] this publication represents the results of the NYU Furman Center's annual collection of data on housing and demographics in New York City. The data are cross-sectional, drawing from the 59 community districts within the five boroughs of New York City. The unit of observation for this data is a community district.

The dependent variable is *Median Monthly Rent*, which was measured in 2005. Data on this variable were obtained from the *New York City Housing and Vacancy Survey*, excluding subsidized units. Hence, the rents are 'gross' rather than 'net.' Using *median* monthly rents relieves the model of potential bias; median rents are less influenced by the highest and lowest rent levels than are mean rents.

4 Been, Vicki, Caroline K. Bhalla, Ingrid Gould Ellen, Solomon J. Greene., Andrew E. Schnizel and Ioan Voicu. "The Furman Center for Real Estate and Urban Policy: State of New York City's Housing and Neighborhoods 2005." The Furman Center (2005).

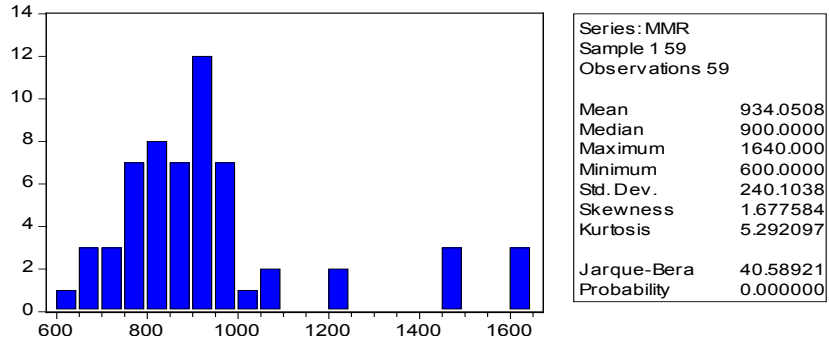


Figure 1: Descriptive Statistics of *MEDIAN MONTHLY RENT*

Variables

Table 1: Names of Variables

Name	Description	Unit of Measurement
Mmr	Median Monthly Rent	(in 2005 dollars)
Brook	Brooklyn	
crimerate	Felony Crime Rate	(per 1,000 residents)
Idr	Income Diversity Ratio	
Manhatt	Manhattan	
medhouinc	Median Household Income	(in 2004 dollars)
Bronx	Bronx	
numhouse	Number of Housing Units	
percentimig	% Immigrant Households	
Percpublic	% Public Housing	
Percrent	% Rent-Regulated/Other subsidized	
Pop	Population (2000)	
Povrate	Poverty Rate	
Queens	Queens	
Racdiv	Racial Diversity Index	
Rentalvac	Rental Vacancy Rate	
Staten	Staten Island	

Independent Variables for Monthly Rent Regression: *Bronx, Brooklyn, Crime Rate, Income Diversity Ratio, Manhattan, Median Household Income, Number of Housing Units, Percentage of Immigrant Households, Percentage of Public Housing, Percentage of Rent-Regulated Housings, Population,*

Poverty Rate, Queens, Racial Diversity Index, Rental Vacancy Rate and Staten Island.

Bronx, brook, Manhattan, Queens, and *Staten Island* are dummy variables to indicate the borough associated with each observation. Each dummy takes on a value of 1 when the observation is located in that borough.

Crime Rate is the number of crimes per 1,000 people living in the area as reported by the New York City Police Department. The crime rate data reflect reports of burglary, larceny and motor vehicle theft, murder, rape, robbery, and assault. The rates are calculated for each community district through a population weighting formula.

Income Diversity Ratio represents the 80th percentile income divided by the 20th percentile income. A larger number for this variable indicates a larger range of incomes for a given community.

Median Household Income is measured for all members of a household who are fifteen years of age and older measured in 2004 dollars.

Number of Housing Units in a given area includes houses, apartments, and any other spaces intended for occupancy as separate living quarters.

Percentage of Public Housing is calculated by dividing the total number of public housing units by the number of housing units in the area. A public housing unit is a unit that is owned and maintained by the NYC Housing Authority.

Percentage of Immigrant Households is calculated against total households in a given area.

Percentage of Rent-Regulated Housing includes housing units that are rent-stabilized, rent-controlled, or city-owned.

Population reflects estimates as reported by the United States Census Bureau.

Poverty Rate is calculated by measuring household income data against a yardstick of poverty, which is adjusted for the number of dependents in a household, number of members of household and age of head of household.

Racial Diversity Index measures the probability of two heads of a household being of different race. A higher number indicates a neighborhood that is more racially diverse.

Rental Vacancy Rate reflects the number of vacant rental units divided by the total number of rental units in the area.

Table 2: Expected Signs of Independent Variables

Name	Expected Sign
Bronx	?
Brooklyn	?
Crime Rate	?
Income Diversity Ratio	?
Manhattan	?
Median Household Income	+
Number of Housing Units	-
Percentage of Immigrant Households	-
Percentage of Public Housing	-
Percentage of Rent-Regulated Housing	-
Population	+
Poverty Rate	-
Queens	?
Racial Diversity Index	-
Rental Vacancy Rate	-
Staten Island	?

IV. Explanation of Expected Signs

Prior to running the regressions, the direction of the effect of the boroughs is uncertain, although it is expected that the prices of land vary across locations even within a given city.[5] The signs of *Crime Rate* and *Income Diversity Ratio* are also indeterminate. An increase in the crime rate may decrease the rental prices being asked and paid for housing units in that area because crime is undesirable. However, crime may be committed more in high-income areas because it is most lucrative in these areas.

Median Household Income and *Population* are expected to have positive signs. It can be assumed that an increase in the median household income of a community will drive rental prices up. Wealthy renters can not only afford more expensive properties but may also prefer them as luxuries and status symbols. Increases in population will also drive up rental prices because they increase demand.

5 Cheshire, Paul and Stephen Sheppard. "Estimating the Demand for Housing, Land, and Neighbourhood Characteristics." Oxford Bulletin of Economics and Statistics (60.3, 1998) 357-382.

It can be hypothesized that the *Number of Housing Units, Percentage of Immigrant Households, Percentage of Public Housing, Percentage of Rent-Regulated Housing, Poverty Rate,* and *Rental Vacancy Rate* will all have negative coefficients. Areas with a higher number of housing units will have lower rents according to the laws of supply and demand. It can be hypothesized that there is a negative relationship between *Percentage of Immigrant Households* and *Racial Diversity Index* and rental prices. According to David Harris, "[a]mong nonpoor whites, at least 70 percent of movers select new neighborhoods that are in nonpoor areas where no more than 30 percent of residents are black. The probability of moving to a neighborhood that is at least 60 percent black is almost zero for this group."[6] High levels of public housing are typically undesirable. Since rent controls lower rents, increases in the number of rent-regulated units are expected to lower rents.

Table 3: Descriptive Statistics as Common Sample of Variables

	Mean	Median	Maximum	Minimum	Std. Dev.	Skewness	Kurtosis	Jarque-Bera	Probability	Observations
BRONX	0.189655	0	1	0	0.395452	1.583278	3.50677	24.85275	0.000004	58
BROOK	0.310345	0	1	0	0.466675	0.819892	1.672222	10.75872	0.004611	58
CRIMERATE	36.42241	29.4	226.9	13.1	31.37464	4.517984	25.85549	1459.72	0	58
HOMERATE	31.23103	28.35	73.7	2.2	18.50228	0.687601	2.741312	4.732077	0.093852	58
IDR	5.553448	5.4	8.3	3.5	1.062968	0.430242	2.809253	1.877312	0.391153	58
MANHATT	0.206897	0	1	0	0.408619	1.447136	3.094203	20.26541	0.00004	58
MEDHOUINC	40796.03	39150	76010	15544	15880.7	0.529162	2.654835	2.994708	0.223721	58
MMR	936.3621	900	1640	600	241.5379	1.651107	5.193751	37.98315	0	58
NUMHOUSE	58720.97	52033.5	135015	40426	18404.27	2.106314	8.206526	108.3975	0	58
PERCENTIMIG	33.64138	36.6	60.9	10.9	13.97684	-0.027852	1.91547	2.849994	0.240509	58
PERCPUBLIC	8.808621	4.4	45.6	0	10.16534	1.352254	4.724076	24.85977	0.000004	58
PERCRENT	55.09828	58.35	94.6	8	20.64179	-0.296197	2.427811	1.6393	0.440586	58
POP	135582.7	127793.5	242952	34420	46199.91	0.125836	2.64288	0.461279	0.794026	58
POVRATE	18.97759	15.65	45.6	4.4	10.54978	1.023351	3.383098	10.47807	0.005305	58
QUEENS	0.241379	0	1	0	0.431657	1.208734	2.461039	14.82537	0.000604	58
RACDIV	0.686552	0.705	0.96	0.26	0.184786	-0.368448	2.196199	2.87369	0.237676	58
RENTALVAC	2.846552	2.7	6.3	0	1.614758	0.14625	2.322838	1.314921	0.518165	58
STATEN	0.051724	0	1	0	0.223404	4.048195	17.38788	658.6929	0	58

A brief review of the descriptive statistics reveals that Brooklyn has the greatest representation in this sample. Staten Island, on the other hand, is least represented. An ideal study would feature equal representation of community districts from each borough. Variance in representation may have implications for the results of this analysis. Thus,this inequality must be accepted as a limitation of this study.

V. Models

The goal of this paper is to identify the independent variables that

6 Harris, David R. "'Property Values Drop When Blacks Move In Because…': Racial and Socioeconomic Determinants of Neighborhood Desirability." American Sociological Review (64.3, 1999) 461-479.

explain median monthly rents of New York City residential properties and to assess the robustness of these variables as explanatory instruments. An OLS method was used to regress median monthly rent of 2005 (*Median Monthly Rent*) against 15 independent variables.

Table 4: Models

OLS Regressions
Dependent Variable: Median Monthly Rent (Median Monthly Rent) for 2005

	Equation 1	**Equation 2**	**Equation 3**	**Equation 4**
C	367.149	5.781196	6.099663	6.213655
BRONX	-101.561	-0.10323	-0.12406	*-0.12395*
BROOKLYN	-91.6838	**-0.12274**	**-0.10937**	*-0.11455*
STATEN ISLAND	*-319.851*	*-0.27219*	-0.30777	*-0.31885*
QUEENS	-97.8713	-0.12705	-0.11571	**-0.09738**
CRIME RATE	**1.549827**	**0.001499**	0.001481	*0.001387*
INCOME DIVERSITY RATIO	7.725554	0.01996		
MEDIAN HOUSEHOLD INCOME	*0.010915*	*1.34E-05*	1.14E-05	*1.09E-05*
NUMBER OF HOUSING UNITS	0.002816	6.34E-07		
PERCENTAGE OF IMMIGRANT HOUSEHOLDS	1.459037	0.002175		
PERCENTAGE OF PUBLIC HOUSING	1.492536	-0.0002		
PERCENTAGE OF RENT-REGULATED HOUSING	-0.20112	*0.011469*	0.009567	*0.009304*
PERCENTAGE OF RENT-REGULATED HOUSING^2		*-0.00011*	-9.14E-05	*-8.93E-05*
POPULATION	-0.00093	-2.94E-07		
POVERTY RATE	-0.28494	0.001893		
RACIAL DIVERSITY INDEX	53.52097	**0.155106**	0.126587	
RENTAL VACANCY RATE	-3.19515	-0.01131		
R-squared	0.875718	0.898504	0.881725	0.87396
Adjusted R-squared	0.831332	0.858896	0.859548	0.853382
S.E. of regression	99.19781	0.084959	0.084762	0.086603
Sum squared resid	413288.6	0.295938	0.344863	0.367503
Log likelihood	-339.571	70.76494	66.32808	64.48408
Durbin-Watson stat	2.12305	2.287611	2.038685	2.050118
Mean dependent var	936.3621	6.814738	6.814738	6.814738
S.D. dependent var	241.5379	0.226172	0.226172	0.226172
Akaike info criterion	12.26106	-1.85396	-1.94235	-1.91324
Schwarz criterion	12.82946	-1.25004	-1.5871	-1.59352
F-statistic	19.72942	22.68486	39.75932	42.47073

Prob(F-statistic)	0	0	0	0
Ramsey RESET Test:				
F-statistic Probability	0.010812	0.13001	0.065406	***0.301108***
Log likelihood Probability	0.00141	0.04812	0.03211	***0.227363***
Italics = Significant at 1% level			Bold = Significant at 5% level	

VI. Discussion of Models

EQUATION 1

For the first regression, *Manhattan* is removed from the set of independent variables and used as a baseline due to the fact that the data involved five classifications in terms of geographical location and six descriptors. Research from "Why is Manhattan So Expensive? Regulation and the Rise in Housing Prices" by Glaeser, Gyourko and Saks (2005) supports the use of *Manhattan* as a baseline.[7] According to the paper, data on condominium sales in the five boroughs suggest two distinct housing markets: the first market being that of Manhattan, which is associated with high values for *Median Monthly Rent*, and the second being that of the remaining boroughs, which are associated with lower values for *Median Monthly Rent*.

The data make this dual market distinction readily apparent:

Table 5: *Median Monthly Rent* by Borough

Borough	Median Monthly Rent
Bronx	768
Brooklyn	850
Manhattan	**1186**
Queens	950
Staten Island	850

EQUATION 2

The second regression incorporates modifications of the functional forms of the variables to determine whether additional independent variables are relevant to the *Median Monthly Rent* regression. Squaring all of the terms produces a statistically significant combination of *Percentage of Rent-Regulated Housing* and *Percentage of Rent-Regulated Housing^2*.

7 Glaeser, Edward L., Joseph Gyourko and Raven Saks. "Why Is Manhattan So Expensive? Regulation and the Rise in Housing Prices." Journal of Law and Economics (XLVII, 2005) 331-369.

Because of this significance, the second equation includes a squared term for *Percentage of Rent-Regulated Housing*.

The second regression is in log-linear form for two reasons. First, the literature addressing the subject of housing prices supports the use of log-linear form. Furthermore, only this form passed the Ramsey test for goodness of fit.

EQUATION 3

Having identified an appropriate model for the data in terms of the functional form of the dependent and independent variables, Equation 3 reflects the elimination of statistically insignificant variables *Income Diversity Ratio, Number of Housing Units, Percentage of Immigrant Households, Percentage of Public Housing, Population, Poverty Rate* and *Rental Vacancy Rate*. The elimination of these variables is supported by the WALD test, which tests for a statistically significant relationship between independent and dependent variables.

EQUATION 4

In the fourth regression, *Racial Diversity Index* is eliminated on account of the fact it is significant at the 10% level and the model is operating at the 95% confidence level. Further support for the elimination of this variable comes from the fact that a report entitled, "Does Federally Subsidized Housing Depress Neighborhood Values," suggests that *Median Household Income* captures some of the effects *Racial Diversity Index* may have on New York City rental prices.[8] The report states that public housing developments are commonly sited in neighborhoods comprising low-income families with the highest poverty rates and highest proportion of non-Hispanic Black and Hispanic populations. This being the case, the statistically significant inclusion of *Median Household Income* may be a sufficient stand-in for the other variables omitted for their statistical insignificance (namely, *Poverty Rate, Percentage of Public Housing* and *Racial Diversity Index*).

VII. Interpretation of Results

Bronx, *Brooklyn*, *Queens*, and *Staten Island* all command lower

8 Ellen, Ingrid Gould, Michael H. Schill, Amy Ellen Schwartz and Ioan Voicu. "Does Federally Subsidized Rental Housing Depress Neighborhood Property Values?" (Working Paper. Furman Center for Real Estate and Urban Policy, 2003).

rental prices than *Manhattan.* As the *Median Household Income* of residents in a particular community district increases, so too do the median monthly rents charged. Renters with greater amounts of disposable income can afford to pay higher rents and may even look for more expensive properties for the status and luxuries that these properties offer their occupants.

Together, *Percentage of Rent-Regulated Housing* and *Percentage of Rent-Regulated Housing^2* produce a cap; initial increases in the percentage of properties that are rent-regulated increase the median monthly rent, but only to a point. Before reaching a peak, increases in rent-regulated properties limit the supply of non-subsidized housing units. This limitation puts upward pressure on the prices of remaining residential units. When enough rental units become regulated, increases in *Percentage of Rent-Regulated Housing* make the lower rents associated with subsidized housing the predominant rental values in the market.

The crime rate data used in this study reflect New York City Police Department reports of burglary, larceny and motor vehicle theft, murder, rape, robbery and assault. While data indicating the breakdown of reported crimes would be useful, the positive coefficient assigned to this variable links crime to rental areas. The association of high rental prices with high rates of crime suggests that more crimes are committed in these areas on account of their relative value as targets. Based on the positive relationship between rents and median household income, it can be argued that higher crime rates are also associated with higher income areas.

VIII. Implications

Judging from the results of the final model for this data, landlords can ascertain how to price their rental properties and renters can break down the prices they are charged to understand how certain attributes are valued. More specifically, premiums are charged based on location in the following order: Staten Island, The Bronx, Queens, Brooklyn, and Manhattan, with Manhattan commanding the largest premium. Prospective renters with higher median household incomes will typically pay higher median monthly rents and live in areas with higher crime rates. Districts where the percentage of subsidized housing is below 51.06% will see higher rents as subsidized housing increases to a certain point, after which rent-regulated prices will come to dominate.

With an understanding of the relationships between these variables and monthly rents, renters can make more informed decisions about where to reside according to their willingness to pay for housing and negotiate along the relevant dimensions of the pricing equation. An understanding of how to

break down rents into their relevant components also helps landlords and real estate developers approximate the expected returns on their investments with more accuracy. Knowing that the borough, median household income, crime rate and percentage of subsidized housing in a district influence monthly rents, resources can flow to investments that will best suit investors' strategic desires and expectations. A matching of expectation and returns of this kind can ultimately reduce volatility in the real estate market as developers and landlords' secure their ability to collect the rents necessary for subsistence and profitability.

With regard to economic efficiency, the results of this study make public the inputs of the residential pricing formula. With more information available to the general public, transactions costs should decrease, and consumers and sellers should be able to make more informed decisions and realize an increase in social welfare.

IX. Areas of Future Research

There are multiple aspects of this project that could be further developed in future studies. Among these is a review of data on the sales market to see whether the factors relevant to rental prices are similarly relevant to sales prices. Any discrepancies between these two sets of prices could be explored in an attempt to model the difference between the sale and rental markets. It would also be interesting to see whether this model would be transferable to rental prices of other cities.

It is also possible to look at the data used in this project over a period of time. Such a panel study could reveal the relative importance of these independent variables over time. Perhaps some understanding of their relationship with general changes and trends in the real estate market could be developed.

Most interesting, however, is the prospect of conducting a comprehensive study of New York City rental pricing that would incorporate every kind of attribute. From neighborhood characteristics to apartment amenities to individual unit dimensions and features, there is an incredibly wide range of variables that may be relevant to this pricing question. Including data that represent all of these variables would illustrate which bundles of attributes and which attributes within those bundles are most significant in explaining the rental prices of New York City residences. This comprehensive data set would then yield results that would paint the most detailed portrait of rental prices as a function of their many different determinants.

Appendix A

Discussion of Equation 1

In Equation 1, very few of the variables result in compelling p-values. However, the Ramsey RESET test, which tests the model for misspecification of functional form, suggests that the regression is correctly specified. The regression passes the White Test, indicating that the model does not suffer from heteroskedasticity, corroborating the assumption that the error term has a constant variance.

To determine whether borough (as an indicator for geographical location) is significant in explaining housing prices or rents empirically, a WALD Test is run. This test would reveal if the coefficients for the borough dummy variables were simultaneously equal to zero and could consequently be omitted. The poor results for the WALD test lend empirical support for the inclusion of these dummies.

Theory also supports the notion that borough is significant in explaining housing rental prices. Oftentimes residential units of seemingly identical qualities demand much higher rents in one region than in another. Cheshire and Sheppard (1998) discusses this notion of rental disparity in "Estimating the Demand for Housing, Land, and Neighborhood Characteristics."[9] Their analysis of residential prices operates on the premise that rental prices are largely location-dependent.

9 Cheshire, Paul and Stephen Sheppard. "Estimating the Demand for Housing, Land, and Neighbourhood Characteristics." Oxford Bulletin of Economics and Statistics (60.3, 1998) 357-382.

Equation 1: Rental price regressed on all of the variables (with Manhattan as a baseline)

Dependent Variable: MMR
Method: Least Squares
Date: 11/14/06 Time: 21:49
Sample: 1 59
Included observations: 58
Excluded observations: 1

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	367.149	310.7906	1.181339	0.2441
BRONX	-101.561	60.70606	-1.67299	0.1018
BROOK	-91.6838	60.67182	-1.51114	0.1382
STATEN	-319.851	92.32771	-3.4643	**0.0012**
QUEENS	-97.8713	80.984	-1.20853	0.2336
CRIMERATE	1.549827	0.668143	2.319603	**0.0253**
IDR	7.725554	20.94101	0.36892	0.714
MEDHOUINC	0.010915	0.002844	3.837305	**0.0004**
NUMHOUSE	0.002816	0.001961	1.436259	0.1583
PERCENTIMIG	1.459037	1.863143	0.783105	0.438
PERCPUBLIC	1.492536	2.353582	0.634155	0.5294
PERCRENT	-0.20112	1.061202	-0.18952	0.8506
POP	-0.00093	0.000705	-1.32145	0.1935
POVRATE	-0.28494	3.532521	-0.08066	0.9361
RACDIV	53.52097	89.40825	0.598613	0.5526
RENTALVAC	-3.19515	12.69394	-0.25171	0.8025

R-squared	0.875718	Mean dependent var	936.3621
Adjusted R-squared	0.831332	S.D. dependent var	241.5379
S.E. of regression	99.19781	Akaike info criterion	12.26106
Sum squared resid	413288.6	Schwarz criterion	12.82946
Log likelihood	-339.571	F-statistic	19.72942
Durbin-Watson stat	2.12305	Prob(F-statistic)	0

***Bold** indicates statistical significance at the 5% level.

WALD Test on Equation 1:

Wald Test:			
Equation: REGRESSION1			
Null Hypothesis:	C(2)=0		
	C(3)=0		
	C(4)=0		
	C(5)=0		
F-statistic	3.226243	Probability	**0.021373**
Chi-square	12.90497	Probability	**0.011750**

RAMSEY Test on Equation 1:

Ramsey RESET Test:			
F-statistic	5.080364	Probability	0.010812
Log likelihood ratio	13.12847	Probability	0.00141

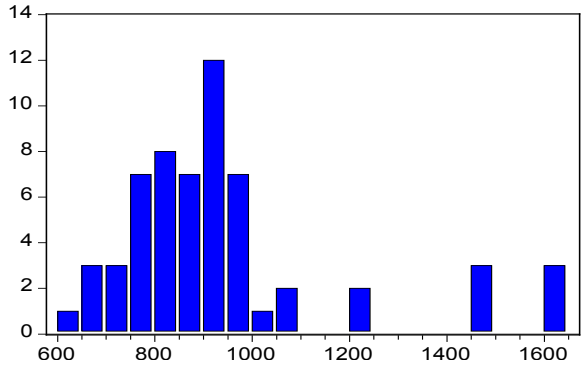
WHITE Test on Equation 1:

White Heteroskedasticity Test:			
F-statistic	1.062293	Probability	0.432327
Obs*R-squared	27.32767	Probability	0.392311

Discussion of Equation 2

Taking the log of *Median Monthly Rent* significantly reduces the skew-factor and improves the normality of the variable, as measured by the Jarque-Bera Test. Using log-linear form produces better R-squared, Akaike and Schwarz values, all of which measure the goodness of fit of the model. The equation also produces coefficient estimates of greater statistical significance and passes the Ramsey and White tests.

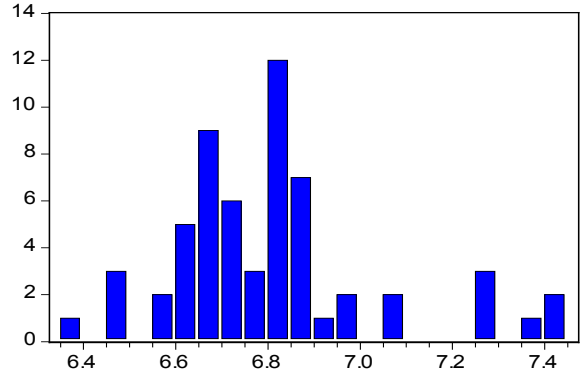
Normalizing Median Monthly Rent changed the distribution from Median Monthly Rent:



Series: MMR	
Sample 1 59	
Observations 59	
Mean	934.0508
Median	900.0000
Maximum	1640.000
Minimum	600.0000
Std. Dev.	240.1038
Skewness	1.677584
Kurtosis	5.292097
Jarque-Bera	40.58921
Probability	0.000000

Figure 2: Un-normalized *Median Monthly Rent*

To log(Median Monthly Rent):



Series: LOGMMR	
Sample 1 59	
Observations 59	
Mean	6.812533
Median	6.802395
Maximum	7.402452
Minimum	6.396930
Std. Dev.	0.224853
Skewness	1.110127
Kurtosis	4.120632
Jarque-Bera	15.20563
Probability	0.000499

Equation 2: Log(rental price) regressed on all of the variables with the addition of percrent^2 (with Manhattan as a baseline)

Dependent Variable: LOG(MMR)
Method: Least Squares
Date: 11/20/06 Time: 20:48
Sample: 1 59
Included observations: 58
Excluded observations: 1

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	5.781196	0.280541	20.60729	0
BRONX	-0.10323	0.052344	-1.97211	***0.0554***
BROOK	-0.12274	0.055243	-2.22179	**0.0319**
STATEN	-0.27219	0.079078	-3.442	**0.0013**
QUEENS	-0.12705	0.07171	-1.77166	***0.0839***
CRIMERATE	0.001499	0.000574	2.61272	**0.0125**
IDR	0.01996	0.01796	1.111306	0.2729
MEDHOUINC	1.34E-05	2.54E-06	5.284622	**0**
NUMHOUSE	6.34E-07	1.78E-06	0.357063	0.7229
PERCENTIMIG	0.002175	0.001597	1.362029	0.1806
PERCPUBLIC	-0.0002	0.002105	-0.0949	0.9249
PERCRENT	0.011469	0.003396	3.376777	**0.0016**
PERCRENT^2	-0.00011	3.13E-05	-3.59481	**0.0009**
POP	-2.94E-07	6.20E-07	-0.47364	0.6383
POVRATE	0.001893	0.003133	0.604191	0.549
RACDIV	0.155106	0.076635	2.023946	**0.0495**
RENTALVAC	-0.01131	0.01123	-1.0072	0.3197

R-squared	0.898504	Mean dependent var	6.814738
Adjusted R-squared	0.858896	S.D. dependent var	0.226172
S.E. of regression	0.084959	Akaike info criterion	-1.85396
Sum squared resid	0.295938	Schwarz criterion	-1.25004
Log likelihood	70.76494	F-statistic	22.68486
Durbin-Watson stat	2.287611	Prob(F-statistic)	0

***Bold** indicates statistical significance at the 5% level.
****Bold and italicized*** indicates statistical significance at the 10% level.

RAMSEY Test on Equation 2:

Ramsey RESET Test:			
F-statistic	2.150691	Probability	0.13001
Log likelihood ratio	6.068129	Probability	0.04812

WHITE Test on Equation 2:

White Heteroskedasticity Test:			
F-statistic	1.134003	Probability	0.36725
Obs*R-squared	29.29569	Probability	0.346709

WALD Test on Equation 2:

Wald Test:			
Equation: Untitled			
Null Hypothesis:	C(7)=0		
	C(9)=0		
	C(10)=0		
	C(11)=0		
	C(14)=0		
	C(15)=0		
	C(17)=0		
F-statistic	0.9683	Probability	**0.466883**
Chi-square	6.778098	Probability	**0.452344**

Discussion of Equation 3

After eliminating *Income Diversity Ratio, Number of Housing Units, Percentage of Immigrant Households, Percentage of Public Housing, Population, Poverty Rate* and *Rental Vacancy Rate*, all of the remaining variables become significant at the 5% level except for *Racial Diversity Index*, which is significant at the 10% level. The Ramsey Test results, however, strongly indicate misspecification of some type, which calls for testing for multicollinearity (as measured by the Variance Inflation Factor—VIF) and heteroskedasticity (as measured by the White Test). The

results from the VIFs identify *Percentage of Rent-Regulated Housing* and *Percentage of Rent-Regulated Housing^2* as being somewhat collinear, which indicates that there is a high degree of linear correlation between the two variables. On account of the fact that the two variables are highly statistically significant and that they are derived from the same data, the multicollinearity is understandable. The variables are therefore kept in the regression. The regression passes the White Test but fails the Ramsey Test, suggesting that the regression is incorrectly specified.

Equation 3: Equation 2 having omitted those whose elimination was supported by WALD

Dependent Variable: LOG(MMR)
Method: Least Squares
Date: 11/20/06 Time: 20:55
Sample: 1 59
Included observations: 58
Excluded observations: 1

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	6.099663	0.121152	50.34712	**0**
BRONX	-0.12406	0.044692	-2.77594	**0.0078**
BROOK	-0.10937	0.04129	-2.64876	**0.0109**
STATEN	-0.30777	0.069253	-4.4441	**0.0001**
QUEENS	-0.11571	0.042964	-2.69317	**0.0097**
CRIMERATE	0.001481	0.000418	3.545903	**0.0009**
MEDHOUINC	1.14E-05	9.62E-07	11.83738	**0**
PERCRENT	0.009567	0.002792	3.425969	**0.0013**
PERCRENT^2	-9.14E-05	2.53E-05	-3.61691	**0.0007**
RACDIV	0.126587	0.07131	1.775179	***0.0822***

R-squared	0.881725	Mean dependent var	6.814738
Adjusted R-squared	0.859548	S.D. dependent var	0.226172
S.E. of regression	0.084762	Akaike info criterion	-1.94235
Sum squared resid	0.344863	Schwarz criterion	-1.5871
Log likelihood	66.32808	F-statistic	39.75932
Durbin-Watson stat	2.038685	Prob(F-statistic)	0

RAMSEY Test on Equation 3:

Ramsey RESET Test:			
F-statistic	2.895408	Probability	0.065406
Log likelihood ratio	6.877144	Probability	0.03211

MULTICOLLINEARITY in Equation 3:

Equation	R-squared	VIF
BRONX BROOK STATEN QUEENS CRIMERATE MED	0.596459	2.478063
BROOK STATEN QUEENS CRIMERATE MEDHOUINC	0.660529	2.94576
CRIMERATE MEDHOUINC PERCRENT PERCRENT^2	0.266117	1.362615
MEDHOUINC PERCRENT PERCRENT^2 RACDIV C BI	0.459384	1.849742
PERCRENT PERCRENT^2 RACDIV C BRONX BROOK	0.962063	**26.35949**
PERCRENT^2 RACDIV C BRONX BROOK STATEN Q	0.959424	**24.64511**
QUEENS CRIMERATE MEDHOUINC PERCRENT PER	0.63352	2.728662
RACDIV C BRONX BROOK STATEN QUEENS CRIME	0.274066	1.377536
STATEN QUEENS CRIMERATE MEDHOUINC PERCR	0.473415	1.899029

WHITE Test on Equation 3:

White Heteroskedasticity Test:			
F-statistic	1.24845	Probability	**0.27993**
Obs*R-squared	15.62899	Probability	**0.269746**

Discussion of Equation 4

Eliminating *Racial Diversity Index* from the regression yields coefficient estimates for the independent variables that are all statistically significant at the 1% level save *Queens*, which is significant at the 3% level. The regression passes the Ramsey and White Tests, confirming that there is neither misspecification nor heteroskedasticity.

The regression is then tested for dummy interactions to eliminate the potential of each dummy having its own distinct slope and intercept. The resulting coefficient estimates are statistically insignificant and their omission from the regression is supported by the WALD Test. The interaction results confirm that there is no difference in slope between each borough—the only difference is in the intercept.

Equation 4: Equation 3 with the elimination of racdiv

Dependent Variable: LOG(MMR)
Method: Least Squares
Date: 11/21/06 Time: 00:51
Sample: 1 59
Included observations: 58
Excluded observations: 1

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	6.213655	0.104966	59.19697	**0**
BRONX	-0.12395	0.045662	-2.71448	**0.0091**
BROOK	-0.11455	0.042081	-2.72215	**0.009**
STATEN	-0.31885	0.070469	-4.52465	**0**
QUEENS	-0.09738	0.042609	-2.28529	**0.0267**
CRIMERATE	0.001387	0.000423	3.276336	**0.0019**
MEDHOUINC	1.09E-05	9.46E-07	11.54418	**0**
PERCRENT	0.009304	0.002849	3.265603	**0.002**
PERCRENT^2	-8.93E-05	2.58E-05	-3.46195	**0.0011**

R-squared	0.87396	Mean dependent var	6.814738
Adjusted R-squared	0.853382	S.D. dependent var	0.226172
S.E. of regression	0.086603	Akaike info criterion	-1.91324
Sum squared resid	0.367503	Schwarz criterion	-1.59352
Log likelihood	64.48408	F-statistic	42.47073
Durbin-Watson stat	2.050118	Prob(F-statistic)	0

RAMSEY Test on Equation 4:

Ramsey RESET Test:

F-statistic	1.231469	Probability	**0.301108**
Log likelihood ratio	2.962411	Probability	**0.227363**

WHITE Test on Equation 4:

White Heteroskedasticity Test:

F-statistic	1.181531	Probability	**0.325902**
Obs*R-squared	12.77724	Probability	**0.308134**

Discussion of Crime Rate

Crime Rate's distribution is highly skewed:

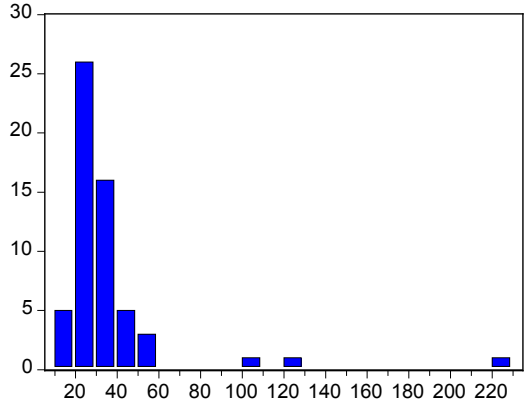


Figure 5: *Crime Rate* Descriptive Statistics

The mean is 30, while the maximum is 226.9. Three outliers may be driving the non-normality of this variable: 106.4, 124.9 and 226.9. These outliers are related to observations 1, 4, and 5, all of which fall within the borough of Manhattan. Omitting these three observations, however, produces insignificant results. Normalizing *Crime Rate* by using *log(Crime Rate)* in place of *Crime Rate* does not improve the outcome either. As a result, *Crime Rate* remains in its original form.

References

Been, Vicki, Caroline K. Bhalla, Ingrid Gould Ellen, Solomon J. Greene., Andrew E. Schnizel and Ioan Voicu. "The Furman Center for Real Estate and Urban Policy: State of New York City's Housing and Neighborhoods 2005." The Furman Center 2005.

Can, Ayse. "The Measurement of Neighborhood Dynamics in Urban House Prices." Economic Geography 66.3 (1990): 254-272.

Cheshire, Paul and Stephen Sheppard. "Estimating the Demand for Housing, Land, and Neighbourhood Characteristics." Oxford Bulletin of Economics and Statistics 60.3 (1998): 357-382.

Ellen, Ingrid Gould, Michael H. Schill, Amy Ellen Schwartz and Ioan Voicu. "Does Federally Subsidized Rental Housing Depress Neighborhood Property Values?" Working Paper. Furman Center for Real Estate and Urban Policy, 2003.

Glaeser, Edward L., Joseph Gyourko and Raven Saks. "Why Is Manhattan So Expensive? Regulation and the Rise in Housing Prices." Journal of Law and Economics XLVII (2005): 331-369.

Harris, David R. "'Property Values Drop When Blacks Move In Because…': Racial and Socioeconomic Determinants of Neighborhood Desirability." American Sociological Review 64.3 (1999): 461-479.

Palmon, Oded, Barton A. Smith and Ben J. Sopranzetti. "Clustering in Real Estate Prices: Determinants and Consequences." JRER 26.2 (2004): 115-136.

Sirmans, G. Stacy, C. F. Sirmans and John D. Benjamin. "Determining Apartment Rent: The Value of Amenities, Services and External Factors." The Journal of Real Estate Research 4.2 (1989): 33-43.