Gentrification and Crime in the Recession Era: Evidence from New York City neighborhoods using bayesian linear mixed modeling

Master's Thesis

Quantitative Methods in the Social Sciences

Columbia University

Jessica Kim

December 31, 2016

Acknowledgments

I thank my thesis advisor, Lance Freeman, and the QMSS thesis course instructor, Elena Krumova, for their guidance and welcome critiques throughout my research process. I also thank Jordan Holliday for his programming help in the data collection process. I finally thank my peers and instructors throughout my experience in the QMSS program for the academic and social support to pursue a skill set that culminates in this independent study.[1]

[1] All code used for this thesis can be found at `https://github.com/jesskim/gentrification_thesis`

Abstract

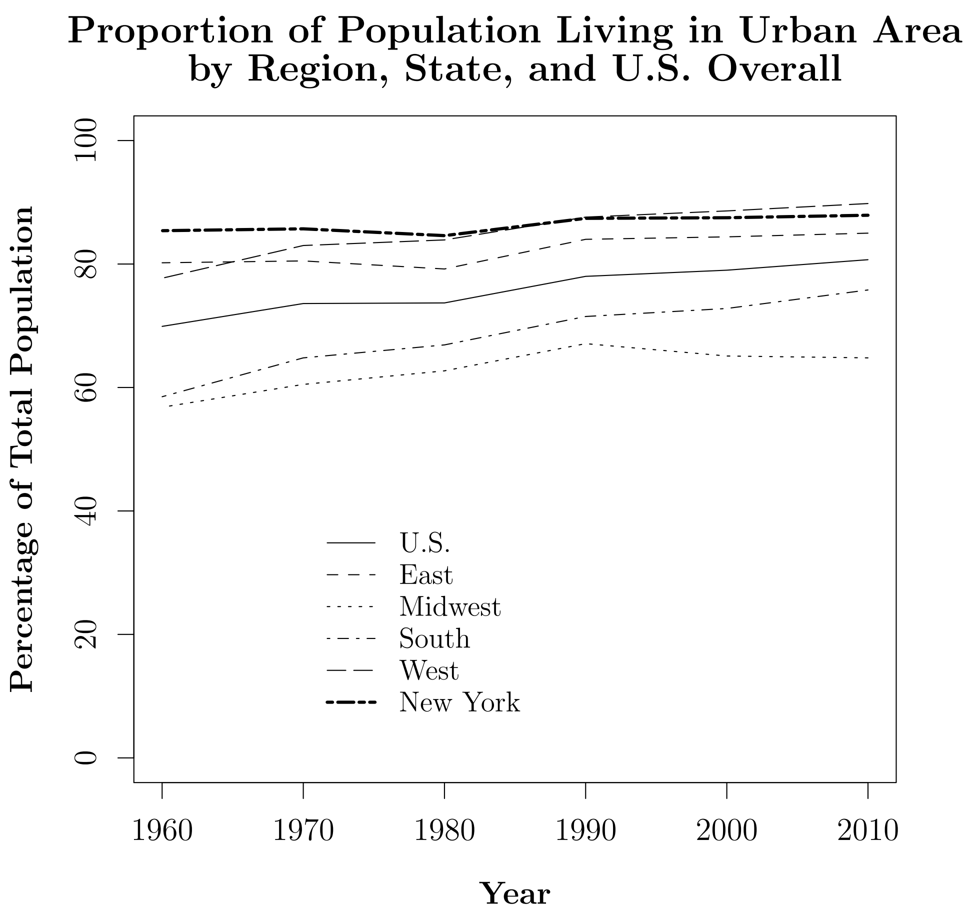
Is gentrification good for New York City? While I do not aim to provide a direct answer to that question, this paper tests a contentious finding in the contemporary debate surrounding gentrification across U.S. cities. I account for neighborhood heterogeneity and time trend in linear mixed models to contribute to the research on gentrification's effect on crime rates. I used a Bayesian approach to provide estimates with a relatively small panel of 34 New York City neighborhoods and annual data from 2004 to 2014. I found that there was no probable effect of gentrification on either overall felonies per 1,000 residents or on property felony rates specifically. I did however find evidence of increases in violent felony incident rates as well as increases in crime among neighborhoods that already had relatively high investment at a baseline level and experienced greater increases in richer and more educated residents. Uncertainty in these findings may be reduced as municipal data is more publicly-available in both the time and unit directions. Either way, a Bayesian approach is advantageous for further research at the local level.

Introduction

The term 'gentrification' has become a source of civil controversy in discussions about the growth of modern metropolises in the Contemporary era. Indeed, Figure 1 shows the time trends in the growth of the proportion of urbanized residents in the U.S. and in its four major regions. Since the 1960s, this proportion has been steadily increasing while the proportion in New York state remains at the highest of those aggregated trends. As increasingly more of the U.S. population chooses to reside in urban areas, the tension surrounding gentrification is unlikely to abate. Yet despite the popular recognition that gentrification has happened and will linger so long as cities develop, formal definitions of the phenomenon and its consequences differ among scholars across and within disciplines. This paper seeks to parse out the particular consequence of gentrification on neighborhood public safety, building upon prior efforts to provide a meaningful understanding of that relationship.

First, it is important to recognize that the mechanism for gentrification in the U.S. experienced a marked change through the decades surrounding the 2000s. Hackworth (2002) distinguishes gentrification that occurred before and after the recession of the late 1980s and early 1990s, also known as

Figure 1:



Source: U.S. Census Bureau

the savings and loan crisis. In that pre-recession era, individuals tended to initiate gentrification, whereby individual developers rehabilitated homes in underdeveloped neighborhoods for personal consumption. After a neighborhood reached a certain threshold of these new individual developers, corporate developers would follow. Working-class communities were occasionally displaced as a result of this development.

In contrast, Hackworth (2002) argues that after urban communities experienced the recession's economic shock, corporate firms were no longer accelerants for gentrification initiated by wealthier individual or small-scale investors, but had become the initiators themselves. Property values for both corporate and owner-occupier real estate investors fell considerably. In New York City, the larger, corporate developers responded by consolidating with financial capital sources. In the post-recession era of the 1990s and 2000s, these consolidated firms had the means to continue to invest in the more expensive neighborhoods that remained after the ones that were more accessible to owner-occupiers were gobbled-up.

Local governments contributed to this shift in the mechanism for gentrification by providing incentives through deregulation of the development process (e.g., relaxing fire code regulations and re-zoning land to accommodate development). Additionally, consumer demand for urban spaces grew, indicating a change from previous land economics that prized suburban neighborhoods. The gradual increase in the average cost of land further prohibited non-corporate property purchases.

Given the conditions that the 1990s savings and loan crisis created for corporate development, there is reason to believe that the recession in 2008 led to similar neighborhood conditions that were amenable to gentrification. While the economic impact of the 2008 recession was markedly greater than that of the 1990s, there are similarities in their potential impact on urban development. The impetus for corporate consolidation from the 1990s was arguably even stronger in the post-recession era following the 2008 subprime mortgage crisis. Furthermore, the nation as a whole continues to urbanize and demand for housing in cities like New York remain strong.

The movement to cities has been associated with changes to neighborhood characteristics that scholars in various disciplines have attempted to understand, including the impact on crime rates. There are several common theories that posit a causal effect of gentrification on crime. Although there is disagreement about the type of impact that gentrification has on crime, there is largely consensus in the literature that a relationship exists. Barton and

Gruner (2016) summarizes the major theories that attempt to explain the mechanism behind gentrification leading to increases or decreases in crime.

The main distinctions between these theories are whether disruptions in informal social order or changes to incentives for criminal behavior are the root causes. Criminological theories known as social disorganization and civic communities argue that gentrification inhibits the ability of poorer neighborhoods to utilize community organizations and feelings of kinship to prevent crime. Alternatively, the broken windows, routine activities, and defended communities theories posit that gentrification alters opportunities for crime.

While I did not conduct research to directly determine which of these theories was the most plausible, my aim in this study was to test the post-recession gentrification theory proposed by Hackworth (2002) and improve upon the operationalization of gentrification. I explain the 2008 recession's effect on gentrification and evaluate the effects of gentrification on felony crimes. Understanding public safety outcomes of gentrification is an important consideration for development, public safety, and housing policy-makers. In particular, there may be implications regarding economic and urban planning strategies for improving the equity of human capital opportunity and safety of neighborhoods.

Literature Review

Gentrification and Crime

Previous literature on the relationship between gentrification and crime show some consensus that gentrification tends to decrease violent crime specifically, but there is substantial variation across studies on how to define gentrification and its relationship to crime rates generally. Some studies incorporate cultural icons into operationalizing gentrification while others rely solely on neighborhood investment measurements. Additionally, studies that have looked specifically at the impact on crime levels find both positive and negative effects from gentrification. The lack of consensus among relevant studies make it necessary to further test different ways that gentrification has been operationalized and what its effect has been on crime overall and by crime type.

Hipp, Butts, Acton, Nagle, and Boessen (2013) evaluated the effect of neighborhood social network attachment on crime, using census data on household location, neighborhood population density, and household member count from the five cities of Buffalo, Cincinnati, Cleveland, Sacramento, and Tucson. Greater social cohesion predicted a decrease in violent crime and an even greater decrease in property crime. Although the study was limited to a two-year period from 2000 to 2002, this finding suggests that neighborhoods with stronger social ties have lower crime incidence.

Most other studies compare neighborhoods within one city or metropolitan area and tend to find a differential effect of gentrification by crime type and baseline neighborhood characteristics. Findings from Barton (2014) find that New York City sub-boroughs that gentrified between the 1980s and early 2000s experienced lower levels of violent crime, but some noticeable increase in theft. Papachristos, Smith, Scherer, and Fugiero (2011) found that there were different effects of gentrification on crime in Chicago between 1991 and 2005 based on the racial make-up of the neighborhood; increases in coffee shops were associated with decreases in crime in predominantly White and Hispanic neighborhoods but increases in robbery in Black neighborhoods. Consistent with Papachristos et al. (2011), Smith (2014) found that demographic changes and increases in private investment (i.e. coffee shops) in Chicago between 1994 and 2005 were associated with decreases in gang-related homicides.

However, Smith (2014) also concluded that public housing demolition led to increases in gang homicides, suggesting a negative effect from gentrification. Criminology theory attributed to Becker (1968) and furthered by Ehrlich (1973), supports this negative relationship; new wealth begets new opportunities for crime. More recently, Freedman and Owens (2016) found that in neighborhoods where only some residents experienced increased income, crime increased between 2000 and 2010. They did a difference-in-difference analysis looking at the levels of crime before and after a government program that provided construction jobs was implemented discriminately in San Antonio, Texas. Neighborhoods that experienced a greater influx of beneficiaries had a greater increase in felony charges for property crime and domestic violence compared to counterpart neighborhoods with fewer beneficiaries. While Freedman and Owens (2016) exploited a natural experiment, simply measuring economic development in terms of changes in individual wealth may not capture the principal mechanism behind gentrification in cities of corporate housing development.

In addition to a dispute on the direction of the relationship between gentrification and crime, an important and general limitation in the literature is the issue of endogeneity. Lee (2010) and Kreager, Lyons, and Hays (2011) point out that the disparate conclusions may be due to insufficiently addressing directional endogeneity. Lee (2010) used an instrumental variable approach to understanding the effect of gentrification on crime. Following the 1994 Northridge Earthquake, the Los Angeles city government incentivized buying homes in low- and moderate-income neighborhoods, which were largely bought by middle and upper income people. They used LAPD data with home mortgage disclosure act data at the census tract level to assess crime trends between 1990 and 2000. They found that in the short term, gentrification increases assaults, robberies, car theft, and theft from cars.

Kreager et al. (2011) found a similar relationship in exploring the effect of urban revitalization in Seattle on crime between 1982 and 2000 in Seattle. They used two measures of gentrification and used census tracts as their unit of analysis. They looked at demographic changes and yearly housing investments and also found a curvilinear association between gentrification and crime where housing investments in the 1980s predicted a slight increase in crime while those in the 1990s decade predicted a decrease. This suggests that there may be an increase in crime in the short term with a decrease in crime in the long term during gentrification.

Operationalizing Gentrification

Over the past century, the understanding of gentrification has become increasingly fragmented. Barton and Gruner (2016) broadly distinguishes between qualitative and quantitative approaches to specifying gentrification. Qualitative studies have typically focused on describing a process of displacement and have defined gentrification as a change from a non-white, ethnic culture to a homogeneous white and middle-class culture. Quantitative research has instead used multiple neighborhood features, including racial make-up, that were measurable over time for large samples. Operationalizing gentrification has therefore involved using multiple indicators in relation to the distribution of these indicators across a given city. Studies have typically used a baseline status to identify neighborhoods with the potential to become gentrified. They then compare outcomes among those 'gentrifiable' neighborhoods based on whether they gentrified or not.

The quantitative literature on gentrification tends to agree that it is a noticeable change in the iconic features of a metropolitan neighborhood that occurs over about a decade, usually associated with increases in the wealth of businesses or residents. Papachristos et al. (2011) and Smith (2014) attempt to measure cultural changes and used increases in coffee shops as an indicator for private investment and neighborhood wealth. While Smith (2014) also included demographic changes and public housing demolitions, the central limitations of using cultural icons are the lack of the generalizability of neighborhood effects and the possibility that these changes are actually outcomes of gentrification that appear ex post.

Alternatively, Lee (2010) and Kreager et al. (2011) use home mortgage changes to indicate gentrification, estimating effects in Los Angeles and Seattle respectively. Albeit a broader operationalization of the term, gentrification was tailored toward cities where home ownership may be more common than in a city like New York where renting may be more commonplace. Furthermore, Hackworth (2002) points out that home sales may not be indicative of the type of corporate gentrification following a recession. Rather, property investment is more commonly from corporate entities.

The variation in quantifying gentrification largely comes from how to include popular conceptions of gentrification. However the inclusion of cultural indicators such as coffee shops in Papachristos et al. (2011) and Smith (2014) or social cohesion in Hipp et al. (2013) have been supplementary features to measures of real estate investment. How exactly to measure that investment depends on the type of housing economics for particular metropolitan areas and therefore requires knowledge about the type of property investment that is common for the area of interest. I therefore exclude home mortgage measures as utilized in Lee (2010) and Kreager et al. (2011) and opt for a measure of property sales to capture the more corporate nature of neighborhood development in New York City.

Across the literature on gentrification, outcomes are compared among areas that are poorer for a given city at the start of the period of analysis as richer neighborhoods are seen as already gentrified. Barton (2014), for example, identified fifty-five New York City sub-boroughs as either 'gentrifiable' or 'gentrified' using changes relative to median income and education levels as measurements. Gentrifiable neighborhoods were defined as those with lower average household income, educational attainment, and proportion of old houses than those of the rest of the city at the start of the decade. Gentrified neighborhoods had an above-average increase in educational attainment and housing prices as well as a recognition of being gentrified by a New York Times article. Similarly Kreager et al. (2011) only compared outcomes for among neighborhoods that had mean family incomes below the mean family income across neighborhoods in Seattle, and Covington and Taylor (1989) used Baltimore neighborhoods with relatively high changes in home values to indicate gentrification. Gentrification therefore is restricted to neighborhoods that are initially poor and have low investment.

Research Question

Few studies apart from Lee (2010) and Hackworth (2002) have considered the effect of an economic shock that was endogenous to crime and may have augmented the process of gentrification. Therefore, I will contribute to the literature on gentrification and crime with more recent data than used previously and include the effect of the 2008 recession.

The primary questions I would like to answer are:

1. What has post-Recession gentrification done to the safety and quality of New York City neighborhoods?

2. What was the effect of the 2008 Recession on gentrification in New York City?

Hypotheses

The three hypotheses of my study are:

H1. Gentrification led to a decrease in violent crime.

H2. Gentrification led to a short term increase in non-violent crime followed by a longer term decrease.

H3. The 2008 Recession had an amplifying effect on Gentrification.

Variables

Gentrification

The primary explanatory variable of my analysis is gentrification and I opt for a parsimonious operationalization of the term to capture changes in property investment and neighborhood wealth. While other studies have used supplementary measures such as public housing demolitions or the presence of cultural icons, the widely accepted mechanism for gentrification has been changes in housing and land economics. Furthermore, studies that used cultural indicators were using a very subjective and possibly spurious measure of gentrification and have not honed in on the central issues of income disparity and displacement due to increased property values. Those studies that have used cultural indicators have tended to use them in addition to measures of income and property development.

Consistent with the literature on operationalizing gentrification, the process in this study is measured in relative terms. That is, neighborhood characteristics must be substantially different from other city neighborhoods at a baseline status and then experience changes that are substantially different from other neighborhoods over time. First, gentrification is a process that manifests in neighborhoods that have lower household income and corporate investment compared to the rest of the city. Wealthier neighborhoods may experience increased investment but they would not be considered gentrifying areas because the baseline conditions would not be met.

Those baseline conditions I use to identify 'gentrifiable' neighborhoods are median household income and median property sales values that are below the median of the distributions across neighborhoods. Gentrified neighborhoods are ones that experience an increase in average income and education level that is greater than the median of increases across gentrifiable neighborhoods.

Data for property sales are publicly-available from 2003 to 2015 and come from the New York City Department of Finance (NYCDF). The NY-

CDF has published public data on property sales every since 2003, known as the Annualized Sales Update Data, a population dataset with all properties sold in New York City.

Crime

The outcome variables are felony crime rates in terms of incidents per 1,000 residents for each UHF neighborhood. These include total felonies, grand larceny, rape, burglary, assault, vehicular grand larceny, homicide, and robbery. The data is publicly available yearly starting from 2000 through New York City Police Department (2014) and includes latitude and longitude coordinates of the nearest intersection of the crime occurrence, with the exception of rape incidents, which are geo-located to the nearest police station. Crime definitions are in accordance with New York State Penal Law (New York City Police Department, 2016).

Covariates

As both Papachristos et al. (2011) and Smith (2014) find that racial make-up and immigration influences crime, I include proportions of Black and foreign-born residents. This data is drawn from the New York City Community Health Survey (CHS) that is run by the Department of Health and Mental Hygiene. The data are cross-sectional and collected annually through a randomized landline and cell phone survey of approximately 8,500 adults from all five New York City boroughs. Respondents are asked to report about their own ecological characteristics and health status, including healthcare use, health status, and behavioral risk factors. They also report their racial or ethnic background and whether or not they were born in the U.S.

The CHS is publicly-available for each year from 2002 to 2014. Spatial identifiers include zip code, borough, and a custom neighborhood classification developed by the United Hospital Fund (UHF), a nonprofit organization that collaborates with the NYCDOHMH for research on public health. According to United Hospital Fund (2010) the UHF neighborhoods were developed to analyze health care services in geographic areas that mirrored the New York City Community Planning Districts. For the most part, zip codes are nested within UHF neighborhoods and neighborhood boundaries were created to join zip codes with residents of similar characteristics. The final delineation of these 34 neighborhoods were created after consulting with representatives from several NGOs and local government departments in urban planning and public health. While prior literature on gentrification tend to rely on census tracts, a neighborhood delineation such as the UHF may reflect a more accurate conception of neighborhood identities than census tracts and a more useful one for the purposes of community planning.

Analytical Approach

I first identified neighborhoods that have the potential to become gentrified in the first three years (i.e. 2004 to 2006) of the available data and determined which neighborhoods did gentrify in the remaining years (i.e. 2007 to 2014). I then employed a Bayesian linear mixed model to produce coefficient estimates within credible intervals. Using a Bayesian approach is especially advantageous in this study as the sample size of cross-sectional units is quite small (i.e. 34 neighborhoods). Sorensen and Vasishth (2015) note that while frequentist approaches to linear mixed models require large sample sizes for well-performing estimators, Bayesian models enable calculations of true probabilities of parameter values with uncertainties explicitly included in posterior distributions.

The two primary challenges for estimating the coefficient for gentrification ceteris paribus are omitted variable bias from unobserved heterogeneity across neighborhoods and the potentially endogenous relationship between gentrification and crime rates. Regarding the former, it is not reasonable to assume that there are neighborhood-specific and unobservable factors that are uncorrelated with the explanatory variables. That is, a given precinct may have factors, particularly spatial ones, that would impact crime apart from gentrification such as access to public transportation and proximity to Manhattan where many people commute for work. To account for this unobserved, cross-sectional heterogeneity, I allowed for neighborhood variation in the intercept as specified here:

$$y_{i,j} = \beta_0 + \beta_{Gentrified} x_i + ... + \epsilon_i, \text{for i = 1,...,n}$$

$$where$$

$$\beta_0 \sim N(0, \sigma)$$

I additionally accounted for time trend by allowing for a random slope on the year indicator conditional on the neighborhood. Finally, whether or not the neighborhood was gentrifiable was interacted with the variable that indicates a gentrified neighborhood. For brevity, the model equation below omits the interacted variables in the non-interacted forms. The final model specification was according to the set of parameters here:

$$y_{i,j} = \beta_0 + \beta_{Gentrified} * \beta_{Gentrifiable} + \beta_{Prop.Male} +$$

$$\beta_{Prop.LessthanHS} + \beta_{Prop.Black} + u_{0,i} + u_{1,i,j} + \epsilon_{i,j},$$

$$where$$

$$\beta_0 \sim N(0, \sigma)$$

The final model above accounts for the variation by neighborhood i with the random intercept term $u_{0,i}$ that vary the intercept for neighborhood. I also model that neighborhoods will experience different rates of change in felony incidence across neighborhoods and within the same ones over years j by including a varying slope as $u_{1,i,j}$.

I implemented these models using the Rstanarm package developed by Gabry and Goodrich (2016) package in R. The Rstanarm package uses a Markov Chain Monte Carlo algorithm to draw samples from the posterior distribution of the model that is conditioned on the data. I used a weakly informative standard normal prior distribution as it is possible for parameter values to be less than or greater than zero.

Identifying Gentrifiable Neighborhoods

I classified neighborhoods as 'gentrifiable', or with the potential to become gentrified, if the following conditions were met:

1. Neighborhood median property sales were less than the median across neighborhoods in a given year between 2004 and 2006.

2. The proportion of neighborhood residents living at or below 100% of the Federal poverty level (FPL) was greater than the median across neigh-

borhoods in a given year between 2004 and 2006.

Figures 2 and 3 provide graphical explanations of the criteria for identifying gentrifiable neighborhoods. Figure 1 shows that the median of the median neighborhood property sales for gentrifiable neighborhoods varied considerably across the entire time period between 2004 and 2014. To establish a baseline status from which percentage changes can be analyzed, I limited the time period for identifying gentrifiable neighborhoods to 2004 to 2006. Between these years, the median of the median property sales was roughly $175 million to $225 million among gentrifiable neighborhoods while that figure was about $275 million to $350 million among non-gentrifiable neighborhoods.

To incorporate the condition of gentrifiable neighborhoods as lower-income, I used the proportion of residents at or below the Federal poverty level (FPL). Data on neighborhood income came from the CHS and is only in the form of household income relative to the FPL. Gentrifiable neighborhoods were those that had a proportion of poor residents that was greater than the median proportion across neighborhoods for a given year. Figure 2 shows the difference in the median of these proportions over time by the potential of a neighborhood to become gentrified. Between the years of 2004 and 2006, the median proportion of poor residents was around 25% among gentrifiable neighborhoods while it was closer to 15% among non-gentrifiable neighborhoods.

Results

Descriptive Results

Table 1 shows summary statistics on covariates, gentrification indicators, and outcome variables across all neighborhoods and years. There is considerable variation across neighborhoods and years for the covariate variables of percentages of the total population that are Black, Foreign-born, have less than a high school education, and are at or below the Federal poverty level (FPL). The standard deviations are particularly high for the proportion of Black residents and foreign-born residents at 23.2 percentage points and 12.9 percentage points, respectively. For the proportion of residents with less than a high school education the standard deviation is 9.5 percentage points, and for the proportion of residents with a combined household income at or below the FPL it is 10.2 percentage points. These variables therefore should be at least tested for their ability to confound the causal relationship between

Figure 2:

Median Neighborhood Property Sales By Gentrification Potential

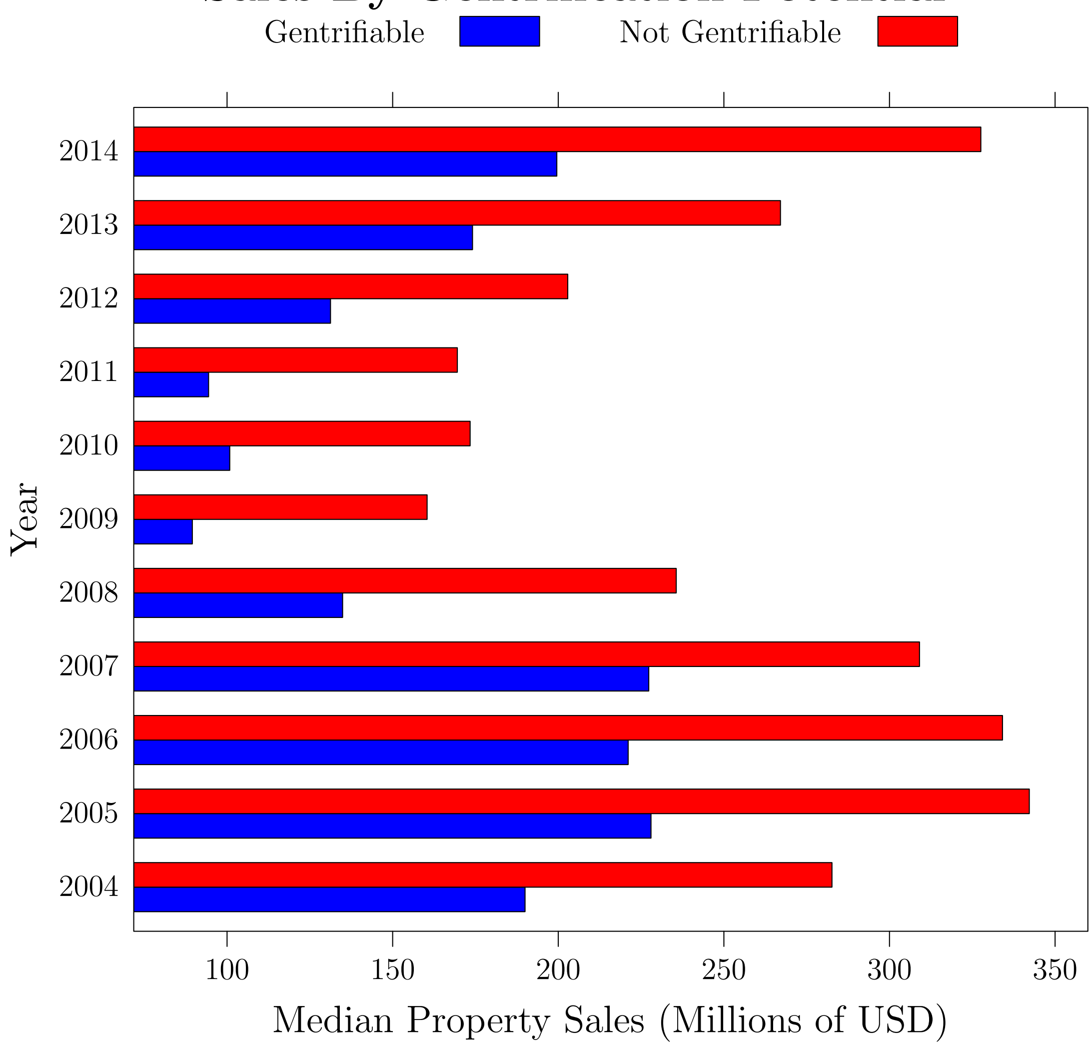
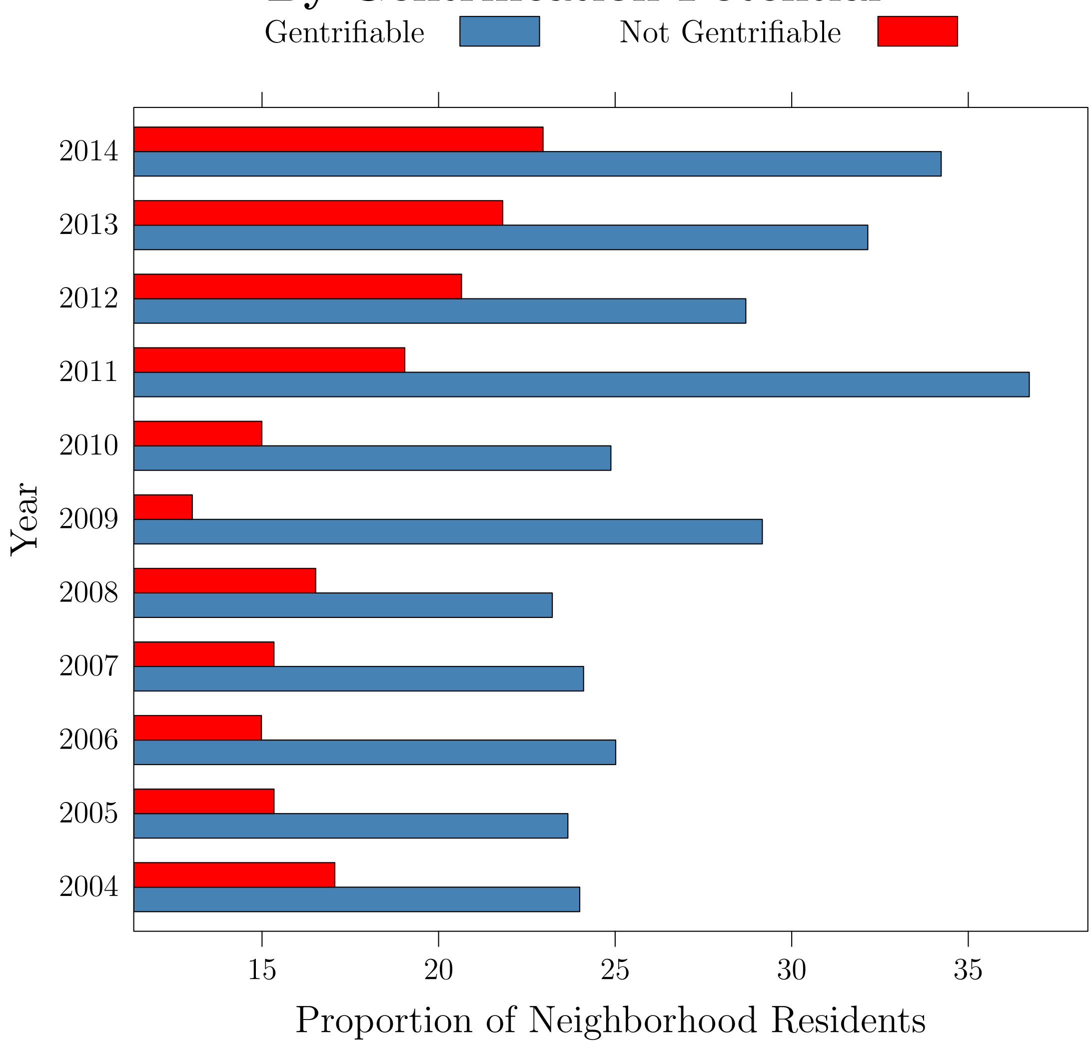


Figure 3:

Proportion of Residents At Or Below FPL
By Gentrification Potential



gentrification and crime rates. They were included in unrestricted models during inferential analyses.

Crimes were predicted as total felony incidents per 1,000 residents and then by two crime types as listed in Table 1. Property crimes included grand larceny, grand larceny of a vehicle, robbery, and burglary while violent felony crimes consisted of rape and assault. Once again there is noticeable variation in these crime rates across neighborhoods, with the standard deviations for felony rates overall being 19.9; they were 16.8 for property felonies, and 4.0 for violent felonies.

Table 1: Summary Statistics

Variable	Mean	St. Dev.	Min	Max
% Black	24.2	23.2	0.0	80.1
% Foreign-born	43.8	12.9	14.6	78.8
% Less than high school education	17.7	9.5	1.1	50.0
% Residents at or below FPL	21.9	10.2	4.2	53.2
Property sales (millions)	$2,369.6	$3,906.5	$201.6	$29,272.4
Median property sales (millions) over years	$333.0	$394.4	$36.1	$2,636.1
% change median property sales (millions)	7.1	38.1	−67.0	316.2
% change % residents at or below FPL	9.1	40.0	−57.3	251.2
Felony incidents per 1,000	23.7	19.9	0.03	128.3
Property Felonies per 1,000	19.4	16.8	0.02	117.3
Violent Felonies per 1,000	4.3	4.0	0.0	20.7

Inferential Results

Figure 4 shows the credible intervals from the joint posterior distributions of all the parameters for each of the three models that predicted all felony offenses per 1,000 residents as the outcome variable. The first plot shows the intervals from the restricted model, the second shows those from the unrestricted model excluding the recession indicator, and the third plot shows parameter probability intervals from the unrestricted model with an indicator for the recession in 2008 and 2009. The x-axes are standardized across the three plots. Table 2 compared the three models using the Leave-One-Out Information Criterion (LOOIC), which evaluates the predictive capability of the models. LOOIC is similar to the Aikaike Information Criterion (AIC) in that it estimates the expected log predicted density (ELPD)[2] of the out-of-sample posterior distribution. Table 2 shows that the unrestricted models are better out-of-sample predictors of the outcome, but there is little difference in the LOOIC from including the recession indicator variable. For the sake of evaluating its influence, it is included in the summary credible interval table.

[2] Vehtari, Gelman, and Gabry (2016) explain: "expected log predictive density (ELPD) for a new dataset = $\sum_{i=1}^{n} \int p_t(\tilde{y}_i) \log p(\tilde{y}_i|y) d\tilde{y}_i)$, where $p_t(\tilde{y}_i)$ is the distribution representing the true data-generating process for $\tilde{y}_i$."

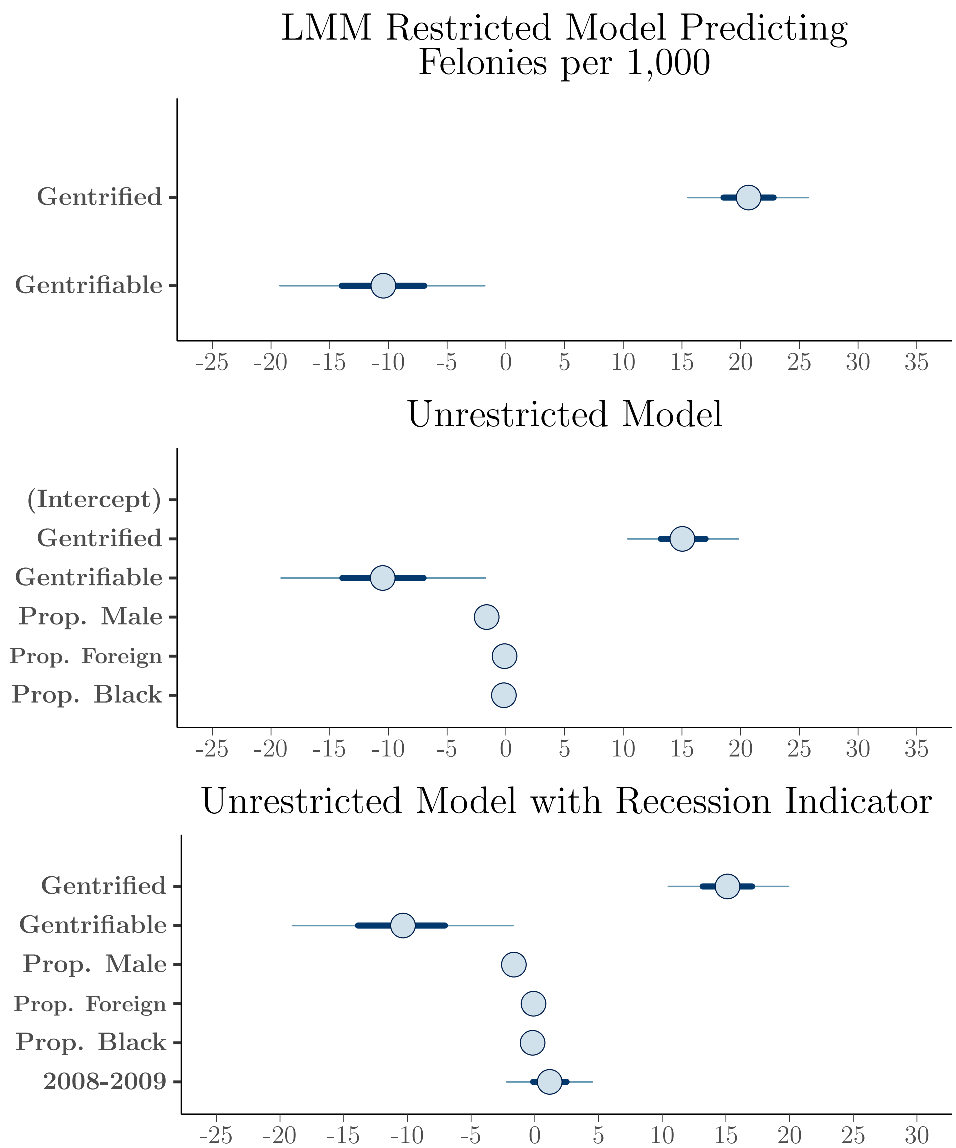
Table 2: LOOIC Comparison of Models Predicting Rates of All Felonies

	LOOIC
Restricted	$2,866.26$
Unrestricted	$2,798.59$
Unrestricted with recession indicator	$2,799.89$

The credible intervals in Figure 4 suggest that there is a positive effect from being gentrified on felony crime incidents per 1,000 residents. The interval range is slightly moderated moving from the restricted to the unrestricted models, but they remain well above zero. The credible intervals for the interacted gentrifiable indicator variable display do not include zero and have relatively wide intervals below zero. There is some evidence that the intervals on the gentrifiable effect parameter moderated the effect of becoming gentrified. However, there is uncertainty with that conclusion as the combined posterior probability densities for being gentrifiable and gentrified include zero. I am therefore reluctant to conclude that there is a meaningful gentrification effect on felony incident rates from a baseline status of being gentrifiable in this model.

The samples of the posterior probability distribution for the parameters are summarized in Tables 3, 4, and 5. In addition to summary statistics on the posterior distributions of the parameter estimates, Tables 3, 4, and

Figure 4:



LMM Restricted Model Predicting
Felonies per 1,000
Gentrified
Gentrifiable
-25
-20
-15
-10
-5
0
5
10
15
20
25
30
35
Unrestricted Model
(Intercept)
Gentrified
Gentrifiable
Prop. Male
Prop. Foreign
Prop. Black
-25
-20
-15
-10
-5
0
5
10
15
20
25
30
35
Unrestricted Model with Recession Indicator
Gentrified
Gentrifiable
Prop. Male
Prop. Foreign
Prop. Black
2008-2009
-25
-20
-15
-10
-5
0
5
10
15
20
25
30

5 also include the posterior distribution sampling metrics, Rhat, N Eff., and MCSE. Rhat is a statistic developed by Gelman and Rubin (1992) that indicates convergence of the Markov chains to a common distribution. It is a ratio of the average variance of the samples within a given chain to the variance of the samples that are pooled across chains, and therefore should be one. The MCMC sampling method to estimate the posterior distributions of the parameters generate samples that are correlated in each chain and are therefore not as accurate as independent samples. N Eff. is therefore an estimate of the number of effective, or independent, samples in a markov chain that would provide the same resulting distribution. MCSE is the Monte Carlo standard error and is the posterior standard deviation divided by the square root of the N Eff. MCSE is thus an estimate of the error from attempting to estimate the posterior mean (Carpenter et al., 2016).

The actual values of the credible interval bounds for the posterior probability density of the parameter value samples on being gentrified after a baseline gentrifiable status consistently showed that the combined intervals between the variables encompass zero and were therefore too uncertain to conclude there was an effect. The credible intervals for the proportion of the population that is foreign-born and the binary indicator for the 2008 to

2009 recession also included zero and did not report probable effects. The proportion of the neighborhood population that is male did however have a 90% credible interval in the negative direction with a mean value of -1.65 percentage points, indicating a slight negative effect on felony rates. There was also a probable negative effect from the proportion of Black residents, but the 90% credible interval was between -1 and 0 and was therefore too small to conclude a practical effect on the outcome.

Table 3: LMM Restricted Model Predicting the Effect of Gentrification on All Felonies per 1,000 Residents

	Mean	MCSE	SD	5%	95%	N. Eff.	Rhat
(Intercept)	23.85	0.06	3.00	18.93	28.77	2,368	1.00
Gentrified	20.66	0.05	3.09	15.49	25.77	4,000	1.00
Gentrifiable	-10.50	0.12	5.35	-19.25	-1.80	2,021	1.00
Gentrified*Gentrifiable	10.16	0.17	8.41	-3.76	23.97	6,021	1.00

Figure 5 shows the inferential results from sequential models with the outcome variable being only violent felonies per 1,000 residents. The x-axis ranges are the same as the interval plots for the models predicting total felony rates. Table 9 employs the LOOIC to compare the three models and similarly show that the unrestricted models improve the predictability of the outcome, with little difference from including the recession indicator.

Table 4: LMM Unrestricted Model Predicting the Effect of Gentrification on Felonies per 1,000 Residents

	Mean	MCSE	SD	5%	95%	N. Eff.	Rhat
Gentrified	15.10	0.05	2.86	10.39	19.82	4,000	1.00
Gentrifiable	-10.46	0.12	5.29	-19.15	-1.73	2,111	1.00
Gentrified*Gentrifiable	5.36	0.19	8.15	-8.76	18.09	6,111	1.00
Proportion male	-1.63	0.003	0.19	-1.94	-1.31	4,000	1.00
Proportion foreign-born	-0.11	0.002	0.12	-0.30	0.09	4,000	1.00
Proportion Black	-0.17	0.002	0.10	-0.33	-0.01	2,321	1.00

Table 5: LMM Unrestricted Model with Recession Indicator Predicting the Effect of Gentrification on Felonies per 1,000 Residents

	Mean	MCSE	SD	5%	95%	N. Eff.	Rhat
Gentrified	15.14	0.05	2.91	10.48	19.92	4,000	1.00
Gentrifiable	-10.45	0.10	5.30	-19.03	-1.71	2,701	1.00
Gentrified*Gentrifiable	4.69	0.15	8.21	-8.55	18.21	6,701	1.00
Proportion male	-1.65	0.003	0.19	-1.96	-1.32	4,000	1.00
Proportion foreign-born	-0.10	0.002	0.12	-0.28	0.10	4,000	1.00
Proportion Back	-0.17	0.002	0.10	-0.34	-0.02	2,764	1.00
Year = 2008 or 2009	1.18	0.03	2.02	-2.21	4.53	4,000	1.00

The intervals for the coefficient samples from being gentrifiable still include zero and therefore indicate an improbable effect from a neighborhood with the potential to become gentrified. The credible intervals for the effect of gentrification without a baseline status of gentrifiable however remain above zero and provide further evidence that increased neighborhood investment in non-gentrifiable neighborhoods led to an increase in property crime rates over time; Table 8 shows that the intervals were specifically 6.86 and 14.69.

Table 6: LMM Unrestricted Model Predicting Treatment Effect on Property Felonies per 1,000

	Mean	MCSE	SD	5%	95%	N. Eff.	Rhat
(Intercept)	20.28	0.06	2.68	15.80	24.59	1,900	1.00
Gentrified	15.37	0.04	2.66	10.99	19.71	4,000	1.00
Gentrifiable	-9.69	0.10	4.67	-17.58	-1.97	2,004	1.00
Gentrified*Gentrifiable	5.68	0.14	7.33	-6.59	17.74	6,004	1.00

Table 7: LMM Unrestricted Model Predicting Treatment Effect on Property Felonies per 1,000

	Mean	MCSE	SD	5%	95%	N. Eff.	Rhat
Gentrified	10.85	0.04	2.50	6.76	14.96	4,000	1.00
Gentrifiable	-9.51	0.10	4.56	-17.08	-1.99	2,170	1.00
Gentrified*Gentrifiable	1.34	0.14	7.16	-10.32	12.97	6,170	1.00
Proportion Male	-1.32	0.003	0.16	-1.58	-1.06	4,000	1.00
Proportion foreign-born	-0.13	0.002	0.10	-0.28	0.04	4,000	1.00
Proportion Black	-0.17	0.002	0.09	-0.31	-0.03	2,650	1.00

Figure 5:

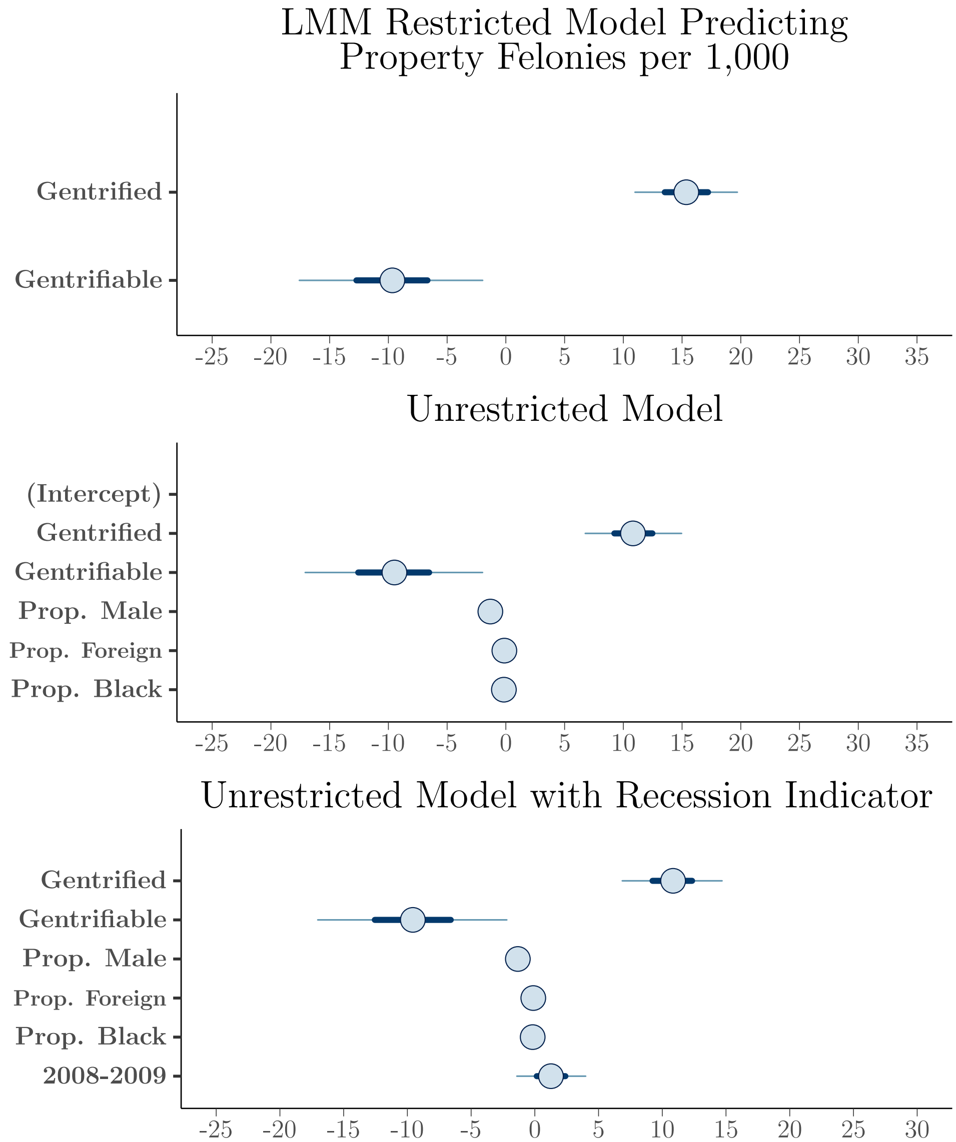

LMM Restricted Model Predicting
Property Felonies per 1,000
Gentrified
Gentrifiable
-25 -20 -15 -10 -5 0 5 10 15 20 25 30 35
Unrestricted Model
(Intercept)
Gentrified
Gentrifiable
Prop. Male
Prop. Foreign
Prop. Black
-25 -20 -15 -10 -5 0 5 10 15 20 25 30 35
Unrestricted Model with Recession Indicator
Gentrified
Gentrifiable
Prop. Male
Prop. Foreign
Prop. Black
2008-2009
-25 -20 -15 -10 -5 0 5 10 15 20 25 30

Table 8: LMM Unrestricted Model 2 Predicting Treatment Effect on Property Felonies per 1,000

	Mean	MCSE	SD	5%	95%	N. Eff.	Rhat
Gentrified	10.81	0.04	2.37	6.86	14.69	4,000	1.00
Gentrifiable	-9.59	0.10	4.52	-17.03	-2.20	2,006	1.00
Gentrified*Gentrifiable	1.22	0.14	6.89	-10.17	12.49	6,006	1.00
Proportion Male	-1.33	0.003	0.16	-1.60	-1.06	4,000	1.00
Proportion foreign-born	-0.12	0.002	0.10	-0.28	0.05	4,000	1.00
Proportion Black	-0.17	0.002	0.08	-0.31	-0.04	2,187	1.00
Year = 2008 or 2009	1.27	0.03	1.65	-1.41	3.99	4,000	1.00

Table 9: LOOIC Comparison of Models Predicting Rates of Property Felonies

	looic
Restricted	2,744.26
Unrestricted	2,681.53
Unrestricted with recession indicator	2,683.33

Figure 6 shows the same sequence of restricted and unrestricted credible interval plots for the final outcome variable of violent crimes per 1,000 residents. Table 13 provides LOOIC results for model comparison. The unrestricted models are better predictors out of sample and therefore indicate that the additional covariates should be included.

Looking at the results in Figure 6, the intervals are noticeably smaller in these plots than in prior ones predicting other outcome variables, but the direction of the coefficient sample intervals are relatively unchanged. Tables 10, 11, and 12 show the posterior distribution statistics for the parameter estimates predicting violent felony rates in the same sequence of restricted and unrestricted models as the tables for previous outcome variables. The estimates from the final unrestricted model are shown in Table 12 and show that gentrifiable neighborhoods that gentrified led to an increase in violent felony rates. The mean parameter estimate for that interaction term was 3.45 violent felonies per 1,000 with a 90% credible interval between 0.86 and 4.53. In contrast, gentrifiable neighborhoods that did not gentrify were associated with a mostly negative effect on violent crime rates as the 90% credible interval for the gentrifiable variable parameter estimate alone is between -2.52 and 0.72.

Although there is once again a positive effect from gentrification for non-gentrifiable neighborhoods, the 90% credible interval for the size of the effect on violent felonies per 1,000 is between about 3 and 5. The proportion of male residents do seem to have had a negative effect on violent felony rates, but the credible intervals on the coefficient samples are between -1 and 0 and are too small to justify there being a substantial relationship.

Table 10:

	Mean	MCSE	SD	5%	95%	N. Eff.	Rhat
(Intercept)	3.64	0.01	0.61	2.63	4.67	2,550	1.00
Gentrified	5.38	0.01	0.59	4.40	6.37	4,000	1.00
Gentrifiable	-0.83	0.02	1.07	-2.58	0.95	2,433	1.00
Gentrified*Gentrifiable	4.55	0.03	1.68	1.82	7.32	6,433	1.00

Table 11: LMM Unrestricted Model Predicting Treatment Effect on Violent Felonies per 1,000

	Mean	MCSE	SD	5%	95%	N. Eff.	Rhat
Gentrified	4.32	0.01	0.56	3.39	5.26	4,000	1.00
Gentrifiable	-0.91	0.02	1.03	-2.64	0.75	1,865	1.00
Gentrified*Gentrifiable	3.41	0.03	1.59	0.75	6.01	5,865	1.00
Proportion Male	-0.31	0.001	0.04	-0.37	-0.25	4,000	1.00
Proportion foreign-born	0.02	0.0004	0.02	-0.01	0.06	4,000	1.00
Proportion Black	-0.004	0.0004	0.02	-0.04	0.03	1,858	1.00

Figure 6:

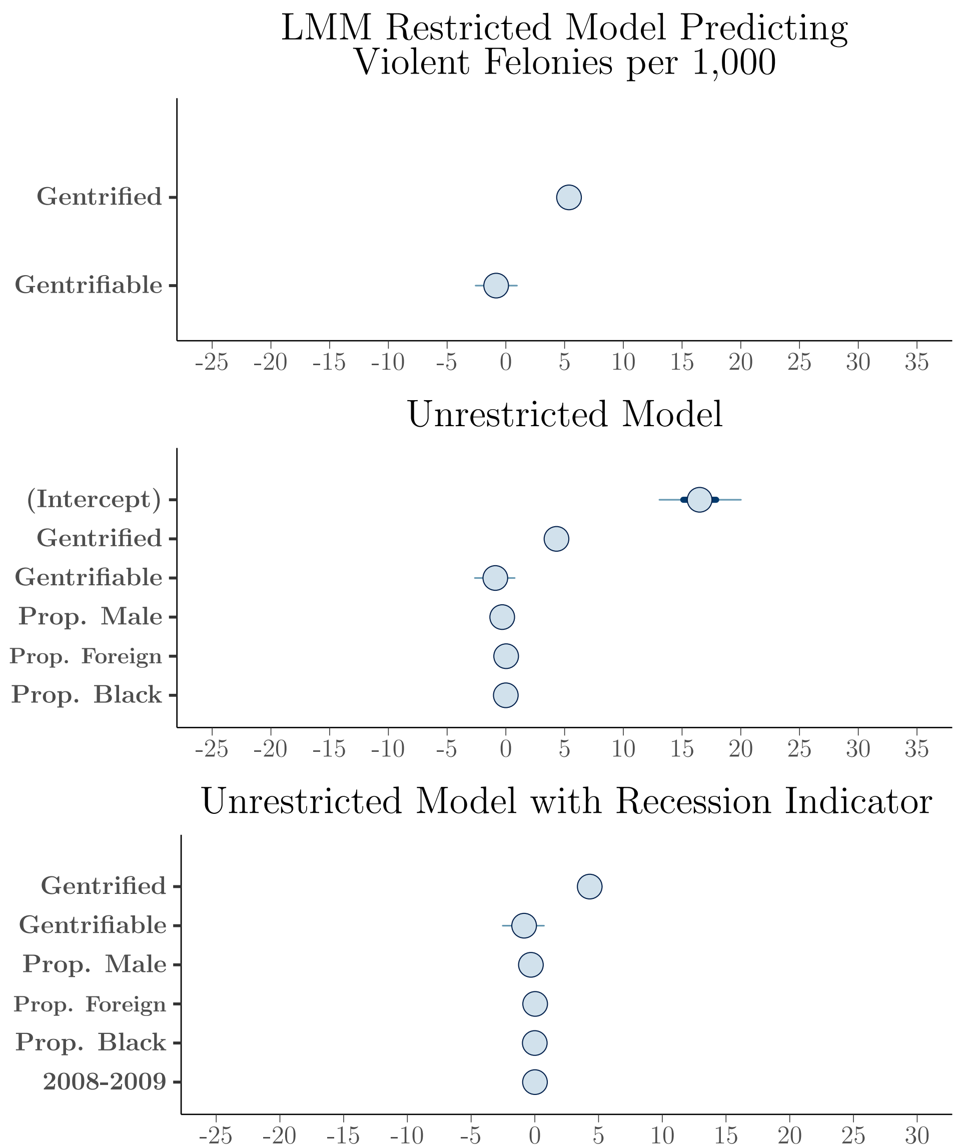

LMM Restricted Model Predicting
Violent Felonies per 1,000
Gentrified
Gentrifiable
-25 -20 -15 -10 -5 0 5 10 15 20 25 30 35
Unrestricted Model
(Intercept)
Gentrified
Gentrifiable
Prop. Male
Prop. Foreign
Prop. Black
-25 -20 -15 -10 -5 0 5 10 15 20 25 30 35
Unrestricted Model with Recession Indicator
Gentrified
Gentrifiable
Prop. Male
Prop. Foreign
Prop. Black
2008-2009
-25 -20 -15 -10 -5 0 5 10 15 20 25 30

Table 12: LMM Unrestricted Model 2 Predicting Treatment Effect on Violent Felonies per 1,000

	Mean	MCSE	SD	5%	95%	N. Eff.	Rhat
Gentrified	4.31	0.01	0.56	3.38	5.25	4,000	1.00
Gentrifiable	-0.86	0.02	1.00	-2.52	0.72	2,197	1.00
Gentrified*Gentrifiable	3.45	0.03	1.56	0.86	4.53	6,197	1.00
Proportion Male	-0.31	0.001	0.04	-0.37	-0.25	4,000	1.00
Proportion foreign-born	0.02	0.0004	0.02	-0.01	0.06	4,000	1.00
Proportion Black	-0.004	0.0004	0.02	-0.04	0.03	2,447	1.00
Year = 2008 or 2009	0.01	0.01	0.37	-0.60	0.63	4,000	1.00

Table 13: LOOIC Comparison of Models Predicting Rates of Violent Felonies

	looic
Restricted	1,732.77
Unrestricted	1,663.92
Unrestricted with recession indicator	1,666.86

Discussion

Using a Bayesian approach to linear mixed modeling, the inferential results provide mixed evidence for evaluating my three hypotheses. Regarding H1, that gentrification had a negative impact on violent crime, the credible intervals on the gentrification indicator variable in fact suggest the opposite for gentrifiable neighborhoods that gentrified; gentrified neighborhoods experienced increases in violent crime rates after accounting for any time trend and confounding demographic characteristics.

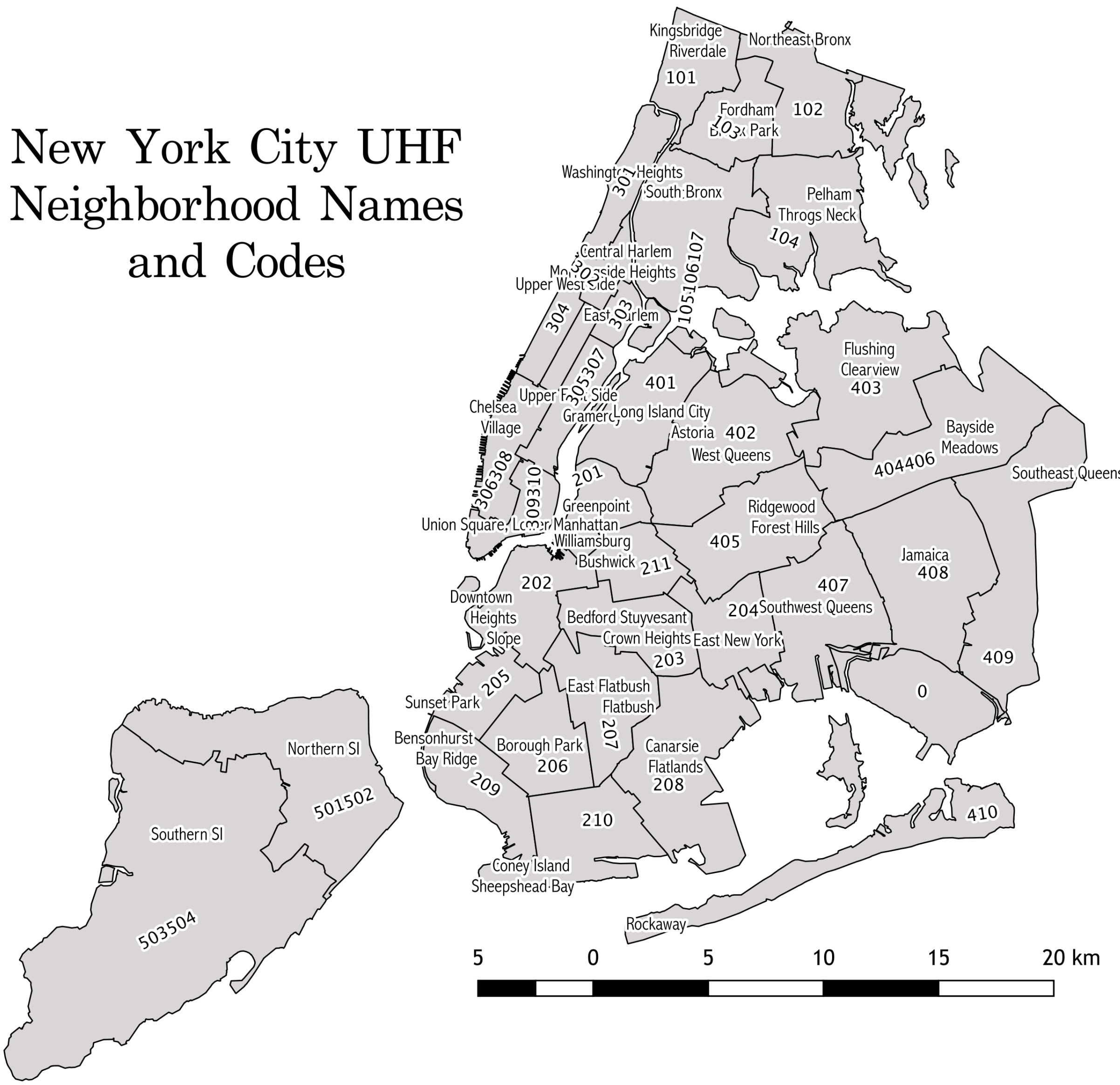
H2 posited that there would be short term increases in non-violent crime that would be preceded by longer term decreases and H3 argued that the recession would amplify gentrification. These hypotheses are rejected as none of the models provided 90% credible interval ranges for the recession indicator variable coefficient samples fully above or below zero. There was no noticeable impact from the years of 2008 or 2009, which differentiate the short and long time periods as well as indicate any effect from the recession.

There were several limitations regarding the availability of appropriate data as well as caveats to note for the assumptions in this paper. First, as I used the CHS for covariate variables and for constructing the gentrifiable and gentrification indicators, I was limited to the 34 UHF neighborhood delineations used in the CHS. I attempted to use a neighborhood classification that would be useful for the New York City planning community by using the CHS over demographic data provided by the Census Bureau's American Community Survey. The data provided are similar with the main distinction being the geographic boundaries for neighborhoods as the American Community Survey does not include UHF neighborhood indicators. While these boundaries were beneficial for providing estimates for locally-meaningful neighborhoods, they limited the number of cross sectional units. Some of the UHF neighborhoods are also larger areas than they are popularly considered and the UHF delineation may have inhibited capturing only portions of these neighborhoods that were gentrifying.

Although Bayesian approaches do not require minimum sample sizes, a small number of cross-sectional units likely contributed to the high uncertainty in the gentrifiable variable coefficient probabilities. There were also limitations in the time direction to 10 time periods. In particular, attempting to estimate the effect of the recession may require a longer stretch of time post-recession than was included in this study.

Future research on the effect of gentrification on crime can employ Bayesian linear mixed models as was done here, but should still use a greater amount of neighborhoods. This may be more feasible in the near future as data at the municipal level becomes increasingly more publicly accessible. As many cities in the U.S. may still not provide many years or geographically precise figures on neighborhood characteristics, researchers will benefit from Bayesian modeling to estimate coefficient values.

Appendix



New York City UHF
Neighborhood Names
and Codes
Kingsbridge
Riverdale
101
Northeast Bronx
Fordham
Bronx Park
102
103
Washington Heights
301
South Bronx
Pelham
Throgs Neck
104
Central Harlem
Morningside Heights
Upper West Side
105106107
East Harlem
303
304
305307
Upper East Side
Gramercy
Long Island City
401
Astoria
402
West Queens
Flushing
Clearview
403
Bayside
Meadows
404406
Southeast Queens
Chelsea
Village
306308
309310
201
Greenpoint
Union Square, Lower Manhattan
Williamsburg
Bushwick
211
Ridgewood
Forest Hills
405
Jamaica
408
407
204
Southwest Queens
202
Downtown
Heights
Slope
Bedford Stuyvesant
Crown Heights
203
East New York
409
205
0
Sunset Park
East Flatbush
Flatbush
207
Bensonhurst
Bay Ridge
Borough Park
206
Canarsie
Flatlands
208
209
210
410
Coney Island
Sheepshead Bay
Rockaway
Northern SI
501502
Southern SI
503504
5
0
5
10
15
20 km

Figure 7:

References

Barton, M. S. (2014). Gentrification and violent crime in new york city. *Crime & Delinquency*, 0011128714549652.

Barton, M. S., & Gruner, C. P. (2016). A theoretical explanation of the influence of gentrification on neighborhood crime. *Deviant Behavior*, *37*(1), 30–46.

Becker, G. S. (1968). Crime and punishment: An economic approach. In *The economic dimensions of crime* (pp. 13–68). Springer.

Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., ... Riddell, A. (2016). Stan: A probabilistic programming language. *J Stat Softw*.

Covington, J., & Taylor, R. B. (1989). Gentrification and crime robbery and larceny changes in appreciating baltimore neighborhoods during the 1970s. *Urban Affairs Review*, *25*(1), 142–172.

Ehrlich, I. (1973). Participation in illegitimate activities: A theoretical and empirical investigation. *The Journal of Political Economy*, 521–565.

Freedman, M., & Owens, E. G. (2016). Your friends and neighbors: localized economic development and criminal activity. *Review of Economics and Statistics*, *98*(2), 233–253.

Gabry, J., & Goodrich, B. (2016). rstanarm: Bayesian applied regression modeling via stan [Computer software manual]. Retrieved from `https://CRAN.R-project.org/package=rstanarm` (R package version 2.13.1)

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical science*, 457–472.

Hackworth, J. (2002). Postrecession gentrification in new york city. *Urban Affairs Review*, *37*(6), 815–843.

Hipp, J. R., Butts, C. T., Acton, R., Nagle, N. N., & Boessen, A. (2013). Extrapolative simulation of neighborhood networks based on population spatial distribution: Do they predict crime? *Social Networks*, *35*(4), 614–625.

Hlavac, M. (2015). stargazer: Well-formatted regression and summary statistics tables [Computer software manual]. Cambridge, USA. Retrieved from `http://CRAN.R-project.org/package=stargazer` (R package version 5.2)

Kling, J. R., Ludwig, J., & Katz, L. F. (2005). Neighborhood effects on crime for female and male youth: Evidence from a randomized housing voucher experiment. *The Quarterly Journal of Economics*, 87–130.

Kreager, D. A., Lyons, C. J., & Hays, Z. R. (2011). Urban revitalization and seattle crime, 1982–2000. *Social problems*, *58*(4), 615–639.

Lee, Y. Y. (2010). Gentrification and crime: Identification using the 1994 northridge earthquake in los angeles. *Journal of Urban Affairs*, *32*(5), 549–577.

Ludwig, J., Duncan, G. J., Hirschfield, P., et al. (2001). Urban poverty and juvenile crime: Evidence from a randomized housing-mobility experiment. *QUARTERLY JOURNAL OF ECONOMICS-CAMBRIDGE MASSACHUSETTS-*, *116*(2), 655–680.

New York City Department of Finance. (2014). *NYCDF Annualized Sales Update Data 2004 to 2014.* http://www1.nyc.gov/site/finance/taxes/property-annualized-sales-update.page.

New York City Police Department. (2014). *NYPD Historical Crime - Incident Level Data.* http://www.nyc.gov/html/nypd/html/crime_prevention/crime_statistics.shtml.

New York City Police Department. (2016). *NYPD Incident Level Data Footnotes.* http://www.nyc.gov/html/nypd/html/crime_prevention/crime_statistics.shtml.

Papachristos, A. V., Smith, C. M., Scherer, M. L., & Fugiero, M. A. (2011). More coffee, less crime? the relationship between gentrification and neighborhood crime rates in chicago, 1991 to 2005. *City & Community*, *10*(3), 215–240.

Pearsall, H. (2013). Superfund me: a study of resistance to gentrification in new york city. *Urban Studies*, *50*(11), 2293–2310.

QGIS Development Team, O. S. G. F. (2009). Qgis geographic information system [Computer software manual]. Retrieved from `http://qgis.osgeo.org`

R Core Team. (2016). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `https://www.R-project.org/`

Sarkar, D. (2008). *Lattice: Multivariate data visualization with r*. New York: Springer. Retrieved from `http://lmdvr.r-forge.r-project.org` (ISBN 978-0-387-75968-5)

Sharpsteen, C., & Bracken, C. (2016). tikzdevice: R graphics output in latex format [Computer software manual]. Retrieved from `https://github.com/yihui/tikzDevice` (R package version 0.10-1)

Smith, C. M. (2014). The influence of gentrification on gang homicides in chicago neighborhoods, 1994 to 2005. *Crime & Delinquency*, *60*(4), 569–591.

Sorensen, T., & Vasishth, S. (2015). Bayesian linear mixed models using stan: a tutorial for psychologists, linguists, and cognitive scientists. *arXiv preprint arXiv:1506.06201*.

United Hospital Fund. (2010). *Sources, Methods, and Definitions - United Hospital Fund.* https://www.uhfnyc.org/assets/592.

Vehtari, A., Gelman, A., & Gabry, J. (2016). Practical bayesian model evaluation using leave-one-out cross-validation and waic. *arXiv preprint arXiv:1507.04544*.

Wickham, H., & Francois, R. (2016). dplyr: A grammar of data manipulation [Computer software manual]. Retrieved from https://CRAN.R-project.org/package=dplyr (R package version 0.5.0)