Socio-Economic Diversity and Its Relationship With Neighborhood Conflict in NYC

Hyungbin Park

New York University

Advisor:

Professor Amanda Geller

University of California Irvine

Professor Jacob Faber

New York University

ABSTRACT Neighborhood dynamics have been largely examined with conflict that exists among neighbors. While past research has mainly focused on ethno-racial diversity and its negative effect on neighborhood conflict, this paper addresses the complex relationship between socio-economic diversity and neighborhood conflict in New York City. This paper used 311 data which consist of noise complaint calls made to real-time government service request hotline and Neighborhood Tabulation Area (NTA) level 5-year estimates census data combined by a spatial software package, QGIS. By scrutinizing the empirical evidence of neighborhood conflict in New York City with a negative binomial fixed effects regression model, it shows that all else being equal, there are more noise complaint calls contacted through 311 in neighborhoods where the level of socio-economic diversity is higher. Similarly, the analytic approach with the tertile Gini index indicates that there are more noise complaint calls made in neighborhoods with a high Gini index and fewer calls in neighborhoods with a low Gini index, compared to neighborhoods with medium Gini index.

KEYWORDS: socio-economic diversity, neighborhood conflict, Gini index, noise complaint, neighborhood tabulation area, 311 NYC Open Data

INTRODUCTION

Socio-economic diversity is a common quality that can be noticed in this era of globalization. Especially in urban areas where the density of the population is relatively greater than in rural areas, communities are naturally formed among people with various socio-economic backgrounds. That is, diversification can have a big impact not only on the culture of the society but also on the interactions that take place between neighbors daily. While interactions among neighbors may serve a role as positive encounters and supports, they may also be bothersome and cause annoyances and conflicts. Because of such possibility of divergences, past research has mainly focused on racial heterogeneity of the neighborhoods and its negative effect on social capital and conflicts between neighbors (Legewie and Schaeffer 2016; Letki 2008).

While studying the relationship between racial diversity and neighborhood conflict allows us to understand the dynamics that exist among modern urban communities, such relationship may not fully explain the full picture of neighborhood conflict because there are other characteristics of neighborhoods that need to be taken into account to describe the inter-relationships among neighbors. Especially in urban areas where people live to seize more academic and career opportunities and therefore achieve higher economic status, it is important to look beyond the visible factor of ethno-racial diversity and also consider socio-economic diversity to better understand the dynamics of urban communities.

Knowing the importance of socio-economic diversity, this paper extended the literature of neighbor conflict and socio-economic status by conducting a statistical analysis of New York City. Analyzing the characteristics of NYC neighborhoods helps us to capture the empirical evidence of neighborhood conflicts in regard to socio-economic diversity and, therefore, understand the aggregate patterns of neighborhood dynamics.

Looking into NYC is considerably insightful because it is one of the cities where diversity is emphasized the most. Not to mention the wide range of ethno-racial diversity, it is a place for people with different socio-economic status from all over the world. Such socio-economic characteristics of NYC can be easily identified by looking into the Gini index, which is one of the most common measures of socio-economic diversity. To be specific, the Gini index is measured between zero, where the society is in total equality, and one, where all the income is earned by a single person and everyone else is earning none.

According to the statistics reported by the American Community Survey, the Gini index of income inequality for New York County was 0.60 in 2017, which is noticeably high when we consider the fact that the Gini index for Queens county, Kings county, New York State and the United States were respectively 0.45, 0.53, 0.51 and 0.48 in 2017, and New York State has the highest Gini index out of all States in the U.S. (LiveStories 2019; AHR 2019). Accordingly, looking into socio-economic diversity within NYC may provide a clear picture of the relationship it has with neighborhood conflict.

The question of interest, therefore, is, "Does socio-economic diversity influence neighborhood conflict in NYC?" With this question, this paper lays out the aggregate level of neighborhood conflict as well as its patterns in NYC in regard to socio-economic diversity. By looking into the relationship between the Gini index and the number of the noise complaint calls with a fixed effects negative binomial regression model, the present study points out that more noise complaint calls are observed in neighborhoods with a higher level of income inequality.

LITERATURE REVIEW

Ethno-Racial Diversity

In the literature of neighborhood interaction in urban areas, racial heterogeneity has been the main topic that has been discussed among scholars and policymakers. Racially diverse neighborhoods confront social tensions due to race prejudice which is derived from different aspects such as group superiority or economic competition (Blumer 1958). Blumer (1958) pointed out that the sense of group position and race prejudice is not set individually, rather it is formed as a group to group interaction. That being said, race prejudice can be seen as a form of protective device triggered by fears of racial out-group migrants. Putnam (2007) also contended that trust among residents in ethnically diverse neighborhoods is lower, and racial diversity and social solidarity are negatively correlated.

More complicated analysis regarding ethno-racial diversity in regard to neighborhood conflict looked into the boundaries between neighborhoods. By looking at poorly defined boundaries between two homogeneous communities, Legewie and Schaeffer (2016) made a significant contribution in their finding that it was at fuzzy boundaries between two distinct communities, rather than at clearly defined boundaries, where neighborhood conflict arose.

This was specifically noteworthy in their focus on the broader socio-spatial structure in New York City rather than focusing on isolated areas with predefined ethnic diversity.

Socio-Economic Status

As mentioned above, racial diversity may explain some of the overall variations in neighborhood conflict, however, it is important to consider other characteristics of neighborhoods to fully understand social conflict. Schaub, Gereke, and Baldassarri (2019) studied the effect of poverty and race on strategic cooperation and maintained that the relationship between racial diversity and cooperation may be spurious because racial diversity and socio-economic diversity are highly correlated. To be specific, the socio-economic status of ethnic minorities is often relatively low, and ethnically diverse communities are often also communities with low income level (Schaub, Gereke, and Baldassarri 2019). Letki (2008) maintained that, while cultural and racial diversities are some of the key issues that negatively affect social cohesion, the socio-economic structure is far more important because it is the main factor that undermines all types of neighborhood interactions. The results of his study indicated that when the effects of racial diversity and socio-economic status were taken into account and modeled simultaneously, racial diversity only negatively affected the attitudes of neighbors toward each other (Letki 2008). However, the low socio-economic status of the neighborhood negatively affected all types of interactions such as sociability, organizational involvement, and individual help (2008).

On top of the studies mentioned above, past studies have also shown the negative effect of socio-economic status of neighborhoods on trust and conflict among the residents (Lancee and Dronkers 2011; Méndez and Otero 2018). Because poor people are often affected by present-bias and prone to experience stress disproportionally, Schaub, Gereke, and Baldassarri (2019) showed in their experiment that cooperation was lower when participants were paired with someone with low income, and the effect of income discrimination became stronger when poor participants were paired with each other. Although this study looked into cooperation rather than conflict, we can understand that socio-economic status also plays a significant role in social interactions.

Socio-Economic Diversity

Past studies have found that at both the individual and aggregate levels, low socio-economic status has a negative impact on neighborhood dynamics. However, studying socio-economic status and the socio-economic diversity of neighborhoods are two different studies. The former examines the level of socio-economic status of each neighborhood, while the latter explores the variation of socio-economic status within each neighborhood (Lancee and Dronkers 2011). Even though these two indicators are related, as my study will focus on the socio-economic diversity of NYC, it is important to understand the differences between the definitions of these two terms.

In terms of socio-economic diversity, Alesina and La Ferrara (2002) pointed out that living in a community with a high degree of income disparity is associated with lower trust. This is largely because individuals trust those more similar to themselves. There was also a micro-level field experiment that showed individuals were “less likely to support redistributive policies in the presence of a poor person in an affluent setting (Sands 2017).” Such exposure to inequality on a daily basis may negatively affect social solidarity as a whole because people are continuously reminded of their socio-economic status relative to others. On the other hand, there is also a study that directly showed the opposite result where socio-economic diversity did not decrease trust (Lancee and Dronkers 2011). Lancee and Dronkers (2011) contended that economic differences are rather beneficial and complementary.

The fundamental differences in the outcomes of Alesina and La Ferrara's (2002) study and Lancee and Dronkers's (2011) study may be a result of having different measures of socio-economic diversity. Rather than using the Gini index, Lancee and Dronkers (2011) used a Herfindahl index, which divides income level into three groups and measures whether these three income groups are equal in size. The more equal in size these three groups are, the more diverse the neighborhood is. This paper follows how Alesina and La Ferrara (2002) used the Gini index, considering that the Gini index is the most commonly used measure of socio-economic diversity.

Socio-Structural Changes in Urban Neighborhoods

In addition to the fact that all types of neighborhood diversity are noteworthy to examine to understand the dynamics of the neighborhoods, some scholars also argue that the socio-structural process of change in urban neighborhoods, such as gentrification and densification, also needs to be taken into account. (Cheshire, Fitzgerald, and Liu 2018). Cheshire, Fitzgeral, and Liu (2018) found out that the effect of gentrification was more influential than that of densification when explaining the rise of neighborhood conflicts. By studying the changes in urban neighborhoods on top of considering neighborhood characteristics and diversities and understanding how the process of urban change is associated with different responses from residents, we can know that neighborhood nuisance is not just associated with an individual's socioeconomic circumstances. That is, neighborhood conflict has to be understood in a broader perspective of socio-spatial inequality (Méndez and Otero 2018). Moreover, collective efficacy also plays a role in neighborhood changes. Sampson (1997) points out that "the ability of neighborhoods to realize the common values of residents and maintain effective social controls" is important in the context of social characteristics of neighborhoods. Therefore, numerous geospatial analyses regarding neighborhood diversity and interactions have drawn scholars' attention.

Geospatial Analysis and Government Open Data

Past studies primarily used surveys and experiments to examine the relationship between diversity and neighborhood dynamics. While these methods allow researchers to understand the micro-level neighborhood interactions and individual's perception of conflicts, they are limited in a way to understand the objective and empirical patterns of neighborhood conflict. Thus, in recent years, the use of large-scale government open data has been growing. Such data in the form of non-emergency calls for government services and information regarding neighborhood circumstances help us grasp the overall dynamics of neighborhoods and the broader patterns of socio-spatial inequality. Also, assessing government open data, rather than survey data, is informative because the bias that stems from the unwillingness of participants to be truthful about certain questions decreases (Goerge and Lee 2002).

To take an example of a study using government open data, Liu et al. (2019) studied the patterns of neighborhood conflicts in Australia by using large-scale administrative data from Brisbane City Council's ARS dataset. The authors took a GIS-based spatial approach to map and understand the spatial distribution of urban neighborhood complaints. By doing so, they identified the categories of conflicts and the spatial distribution of conflicts over time.

The emphasis on the usage of GIS spatial analysis enables researchers to clearly recognize the objective patterns of conflicts. The detailed visualization of change in conflict patterns over time also gives more credibility to the outcomes. Even though the types of conflicts studied in Liu's paper, which were animal, building construction, property management, and visual amenity issues, were different from "neighborhood complaints," the analysis conducted in this paper indicated that using large-scale administrative data collected by the government may be highly informative.

NYC 311 Service Request

Similar to Australia's municipal government data, NYC 311 service request data from Open NYC Data has been used among researchers. Minkoff (2016) explored the social, economic, physical, and political aspects of the 311 service request data combined with Census tract-level data and explained the spatial distribution of 311 service request calling volume in New York City. He categorized the service requests into three sections: service requests regarding government goods, graffiti complaints, and noncommercial noise complaints. Then, he further studied the contacting propensity and the infrastructure age to measure respectively the frequency of service requests and actual problems in a space. By using spatial OLS regression models, Minkoff (2016) concluded that contacting 311 service requests regarding graffiti and noise complaints increased as the socio-economic status of the neighborhood became lower.

As Minkoff (2016) showed that the socio-economic status of the neighborhoods had an impact on neighborhood complaints, my research builds on his study in NYC to further examine the socio-economic diversity of the neighborhoods in NYC in the form of income inequality and examine how it deepens our understanding of neighborhood complaints. Based on the study done by Alesina and La Ferrara (2002) that suggests that socio-economic diversity decreases trust in the neighborhood, I tested the hypothesis that "There will be more neighborhood conflicts, as measured by noise complaints, in neighborhoods where the level of socio-economic diversity is higher in NYC.

311 As A Proxy for Neighborhood Conflict

One important factor to consider when using 311 neighborhood complaint calls as a proxy for neighborhood conflict is whether the volume of complaint calls closely resembles the level of neighborhood conflict. This is largely because not everyone who is annoyed by the loud noise of the neighbors will contact 311. There may also be the factor of legal cynicism in neighborhoods that potentially hinders the police from partnering with the community (Carr, Napolitano, and Keating 2014). That in mind, the neighborhood with the largest volume of noise complaint calls may not necessarily be the neighborhood with the highest level of neighborhood conflict. However, Legewie and Schaeffer (2016) point out that 311 calls are an interesting indicator of neighborhood life in the sense that it captures the "tensions and conflicts that are not resolved in a neighborly way by knocking on someone's door. Instead, residents reach out to the city as an external authority." They also go on to mention that these complaint calls, which are often more subtle forms of conflict unlike hate crimes (Lyons 2007), are not as commonly dealt with in quantitative research.

In order to reasonably consider complaint calls as a form of conflict, we need to consider the propensity to call 311 to resolve the conflict rather than knocking on the neighbor's door to ask for cooperation. Minkoff (2016) maintains that the propensity to call 311 is related to "the likelihood that people will contact government about perceived problematic conditions," and such civic engagement is "associated with the stake people have in the neighborhood and the efficaciousness of the people in the neighborhood." When people have high stakes in the neighborhood, they are naturally inclined to care more about the conditions and quality of the neighborhood. Moreover, efficacy is related to the belief people have about having an impact or making a difference (Niemi, Craig, and Mattei 1991). To control the stake that people have in the neighborhood, Minkoff (2016) included the percentage of the homes that are owner-occupied and the percentage of households with children under 18 in his model. Moreover, while efficaciousness is a subtle attribute compared to stake, he used the median income and ethnic attributes to control for efficaciousness. This is because wealthier people have more resource and time to have civic engagement. Also, there may be cultural norms that take the notion of civic engagement differently. Accordingly, this paper also includes those variables mentioned above in the models discussed in the method section.

METHOD

Data Source Description

Given the breadth of information that NYC 311 provides about neighborhood complaints, this study used 311 data to extend the study of neighborhood conflict and its relationship with socio-economic diversity. NYC 311, which is available on NYC Open Data, is a government service request hotline and online web service submission where residents in NYC can contact to report neighborhood problems ranging from noise complaints to illegal parking. The collection of 311 data has started in 2004, and it is still being updated every day to this day. In general, over two million calls are made every year. 311 data provide information on complaint type, date created, zip code, borough, latitude and longitude. While different types of complaint calls can be used to understand the patterns of neighborhood dynamics, this paper focused on the complaint calls specifically related to residential neighborhood noise, such as loud voice or partying sound. Due to the intention of keeping the analysis of this paper timely, a repeated cross-sectional study is conducted by using 311 noise complaints created in 2013, 2015, and 2017.

To analyze the relationship between the noise complaint calls made to 311 and the macro-level socio-economic diversity of the neighborhoods in NYC, I combined 311 data with five-year estimates demographic and socio-economic data from American Community Survey (ACS) of the U.S. Census Bureau. ACS is the largest household survey that the

Census Bureau administers, and it consists of detailed demographic, housing, social, and economic data. While the data from ACS are aggregated to different geographic levels such as census tract or PUMA, this study specifically looks into Neighborhood Tabulation Area (NTA) level 5-year estimates, which can be accessed on the NYC Planning webpage.

Neighborhood Size

This study examined the relationship between socio-economic diversity and conflict in NTA level neighborhoods. The reason why I specified the size of the neighborhood as NTA is that the size of the NTA neighborhood closely matches with the general definition of a neighborhood that New York City residents use in their daily bases, such as Chinatown, Upper East Side, or Hell's Kitchen. While the census tract has been commonly used as the concept of neighborhood in the quantitative literature on neighborhood effects, it may lack social meaning as each neighborhood study looks into different aspects of the residential environment (Sharkey and Faber 2014). That being said, NTA, which is larger than census tract but smaller than PUMA as shown in Figure 1, is a potentially appropriate definition of neighborhood for this study as Tract only captures a few blocks of streets and PUMA is a combination of several NTAs.

Figure 1. Neighborhood Size (Tract, NTA, PUMA from left to right)

Tract Level NTA Level PUMA Level

There are 195 NTAs in total in NYC, however, seven neighborhoods, which include five cemeteries from each borough (The Bronx, Brooklyn, Manhattan, Queens, and Staten Island), Rikers Island (NY prison), and the airport area, are excluded from this neighborhood analysis as those neighborhoods do not provide enough data on neighborhood dynamics. Therefore, this paper aggregated the noise complaint calls and grouped them by 188 NTA level neighborhoods and by years, then it looked into the socio-economic diversity indicators.

Outcome Variable

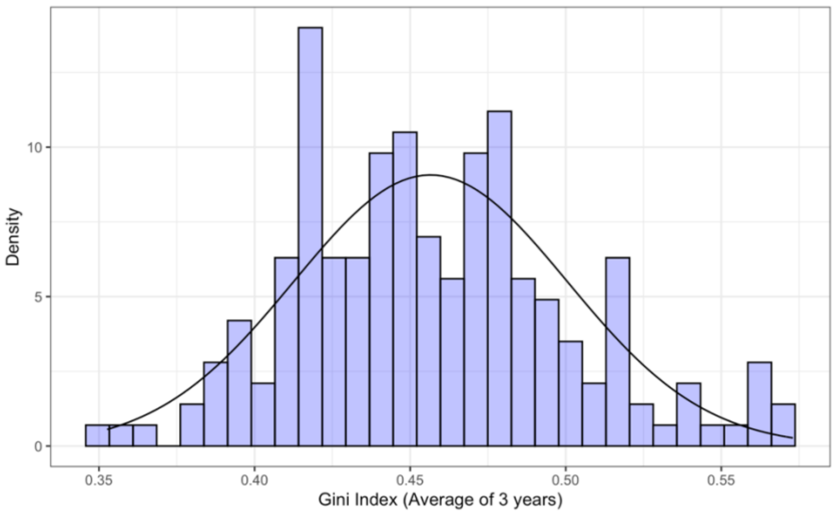
The outcome variable is the number of residential noise complaint calls made to 311 obtained from NYC Open Data. The number of noise complaint calls was further aggregated to the NTA level neighborhoods using QGIS. 1,837,049, 2,296,079, and 2,485,809 calls were made to 311 in total in 2013, 2015, and 2017 respectively. Moreover, among those calls, 151,093, 207,390, and 229,470 calls were residential noise complaint calls in 2013, 2015, and 2017 respectively. The notable trend in the increasing number of total calls and complaint calls may be explained in several ways. 1) The level of neighborhood conflict may have increased. 2) 311 may have become more popular among residents in NYC. 3) The trust level of residents towards the government has become higher than before. 4) The tendency of resolving neighborhood conflict has become more indirect in the sense that neighbors call 311 rather than knocking on each other's doors. Despite all the possibilities that may have contributed to the increase in contacting 311, such a pattern cannot be fully understood as the intention behind calling 311 is not recorded in the dataset. To ensure that the number of noise complaint calls is properly used as a proxy of neighborhood conflict as mentioned above in the literature review, extra steps are taken, which will be further discussed in the control variable section.

Predictor Variable

The predictor variable is the Gini index by NTA, which is the measure of socio-economic diversity or income inequality. ACS provides the Gini index by census tract, and there are more than 2,000 tracts in NYC. Because each NTA is formed with multiple tracts, I calculated the weighted average of tract-level Gini index to derive the Gini index for each NTA using a query language, MySQL.

On top of taking the Gini index as a continuous variable for one model, the Gini index was further categorized into tertiles (Low, Medium, High Gini) in the later modeling approaches to assess and compare the different levels of socio-economic diversity, which is an approach used in past research as well (Huynh et al. 2005). As seen in Figure 2, the Gini index presents a normal distribution, which justifies the approach to take the tertiles of the Gini index. The study uses two different groupings of the tertiles: one divides the NTAs equally into 33%, 33%, 33%, while the other divides the NTAs into 25%, 50%, 25% for Low, Medium, High Gini respectively. This tertile Gini is used as a categorical predictor, and the comparison of the modeling results between the continuous Gini index and tertile Gini is thoroughly discussed in the results section.

Figure 2. Density plot of NTA-level Gini index with normal distribution



Data Source: ACS 5-year Estimates

Control Variables

As mentioned in the outcome variable section, the increasing trend in the number of noise complaint calls cannot be fully understood. Therefore, I controlled the propensity to call 311 by counting the complaint calls not related to noise complaints as well. To be specific, complaint calls related to the heating system of the building, which is a type of complaint call that is not related to neighborhood dynamics, is controlled in the analytic models to clearly establish the number of noise complaint calls as the proxy for neighborhood conflict and avoid interpreting the number of calls as the propensity to call 311. Also, I included the percentage of owner-occupied housing units, the percentage of households with people under 18 years old, the median household income level, and the racial composition (the percentages for Black, Hispanic, and Asian residents) to control the stake and efficacy people have in the neighborhood. The percentage of white residents is used as the comparison group to prevent multicollinearity.

On top of the variables listed above, many of the control variables Minkoff (2016) included in his analysis of 311 are taken into account in this study: population of the NTA, percentage of male residents, the percentage of residents with less than high school degrees, the percentage of residents with more than bachelor’s degrees, the percentage of foreign-born residents, the percentage of residents who lived in the same house for over one year, the percentage of households with people over 65 years old, the median gross rent, and the rental vacancy rate. These are demographic and economic variables that may potentially have spurious relationships with the variables of interest.

Table 1. Descriptive Statistics of the average of 2013, 2015, 2017 NTAs (N = 188)

	Mean	SD	Source
			NYC Open Data
Noise Complaint Call	1037.04	880.90	
Non-noise Complaint			
Heat Complaint Call	1113.35	1217.01	
			ACS 5-year Estimates
Gini Index	0.46	0.04	
Population	44716.21	21907.07	
Male (%)	47.54	2.35	
Race/ Ethnicity			
Hispanic (%)	28.19	21.23	
White (%)	34.15	28.07	
Black (%)	21.62	25.38	
Asian (%)	13.40	14.64	
Educational Attainment			
Less than HS (%)	39.39	3.42	
More than BA (%)	33.64	18.82	
Housing Stability			
Lived in Same House >1 year (%)	89.67	4.42	
Rental Vacancy Rate	3.60	1.85	
Owner Occupied units (%)	35.77	22.34	
Foreign-born (%)	36.18	13.33	
Younger than 18 (%)	32.48	9.84	
Older than 65 (%)	26.67	6.80	
Median Household Income	58682.56	25115.52	
Median Gross Rent	1319.84	344.40	
N	188		

Most of the covariates are in the form of percentages within NTAs. This is largely because, for example, having one more Asian resident would not be as important as having one percent more Asian residents in a certain NTA level neighborhood in terms of social meaning. As a result, having a macro-level analysis on NTA neighborhood rather than a micro-level study on an individual resident, I aimed to understand the overall patterns of neighborhood conflict in regard to socio-economic diversity. Moreover, controlling for composition as percentages allows me to scale the data and efficiently interpret the coefficients. Table 1 lists all the descriptive statistics of the average of the years used in the study.

Statistical Analysis

For comparison, I created four different models with the Gini index in the form of a continuous variable. I first estimated a bivariate OLS model (1) using only the count of noise complaint calls and the Gini index to look into the unadjusted relationship between the outcome and predictor variables.

After confirming the unadjusted relationship, I used a year fixed effects regression model (2) to take advantage of the multiple years presented in the data. With the fixed effects model, I controlled for the omitted variable bias and heterogeneity throughout the years. Many unobserved macro-level characteristics of NYC and the effects of the changes and development that constantly happen in the city are fixed with a fixed effects model. Thus, the within-year relationship between noise complaint calls and socio-economic diversity can be clearly observed.

Next, I used a fixed effects Poisson regression model (3) to take into account the fact the outcome variable, the count of noise complaint calls, is a count variable. Also, I set the population of each NTA as an exposure in order to weight the complaint calls based on the population of the NTAs. This is done largely because, for example, looking into a thousand noise complaint calls made in a neighborhood with a thousand residents is different from observing a thousand calls made in a neighborhood with two thousand residents.

While Poisson regression model is used for datasets with count variables, it has a highly restrictive assumption that the mean of the outcome variable equals the variance. As it is shown in Table 1, however, the mean of the count of noise complaint calls does not equal to the variance. Rather, an over-dispersion can be observed with variance higher than the mean. Thus, I further used fixed effects negative binomial model (4), which is a generalization of the Poisson model that takes over-dispersed count variables. The comparison of Poisson and negative binomial models is done with a Chi-square test while estimating the fixed effects negative binomial model.

Lastly, I included a sensitivity analysis that looks at leads and lags. By doing so, I can examine whether the observed associations reflect a causal or spurious relationship. Considering the temporal order, it is unlikely that the socio-economic diversity in 2015 will predict the noise complaint calls in 2013. If so, the relationship may be spurious.

After looking into the four models described above, I took the model that best explained the relationship of interest. Then, I further compared its results with fixed effects negative binomial models that used the tertile Gini index. The tertile Gini (Low, Medium, High Gini) is taken in the models as a categorical variable and the population is used as an exposure just as the fourth model. The reason for further studying the relationship of interest with the tertile Gini index is that it helps us better grasp the substantive meaning of the relationship. To be specific, the continuous Gini index is a scale between 0 and 1, which signifies that looking into a unit change in the Gini index in the case of the OLS model would not result in a realistic interpretation as no neighborhood would score 0 and have an increase of 1 in Gini index. Therefore, using the tertile Gini index resolves this scaling issue and gives us a thorough understanding of the relationship between noise complaint calls and neighborhoods with different levels of Gini index.

The tertiles are divided into two different proportions. By observing the density plot (Figure 2) and confirming that the Gini index is normally distributed, the proportions are chosen empirically to compare the neighborhoods with different Gini levels. At first, the proportion is divided equally into 33%, 33%, 33% for Low, Medium, High Gini respectively. Following the first choice of the proportions, the tertiles were again divided into 25%, 50%, 25% so that the neighborhoods with High and Low Gini groups can be compared more clearly with the Medium Gini group.

RESULTS

The overall idea of how the Gini index and the count of noise complaint calls differ by NTAs can be assessed with Figure 3, which was created using a visualization tool, Tableau. With the 3-year average NYC neighborhood data, the NTAs with the highest Gini index were Hudson Yards, Upper West Side, Morningside Heights, Upper East Side, and East Harlem. The NTAs with the lowest Gini index were Laurelton, Springfield Gardens North, Cambria Heights, Rosedale, and Springfield Gardens South. The NTAs with the highest Gini were all in Manhattan, while the NTAs with the lowest Gini were all in Queens. In contrast to the Gini index, the number of noise complaint calls varied across the counties and did not show any pattern. The NTAs with the highest volume of noise complaint calls were Central Harlem North, Crown Heights North, Washington Heights South, Washing ton Heights North, and Flatbush. And, the NTAs with the lowest volume of noise complaint calls were Starrett City, West Brighton, Douglaston, Rossville, and Glen Oaks.

Figure 3. Maps of Gini index and noise complaint calls in NYC (Average of 3 years)

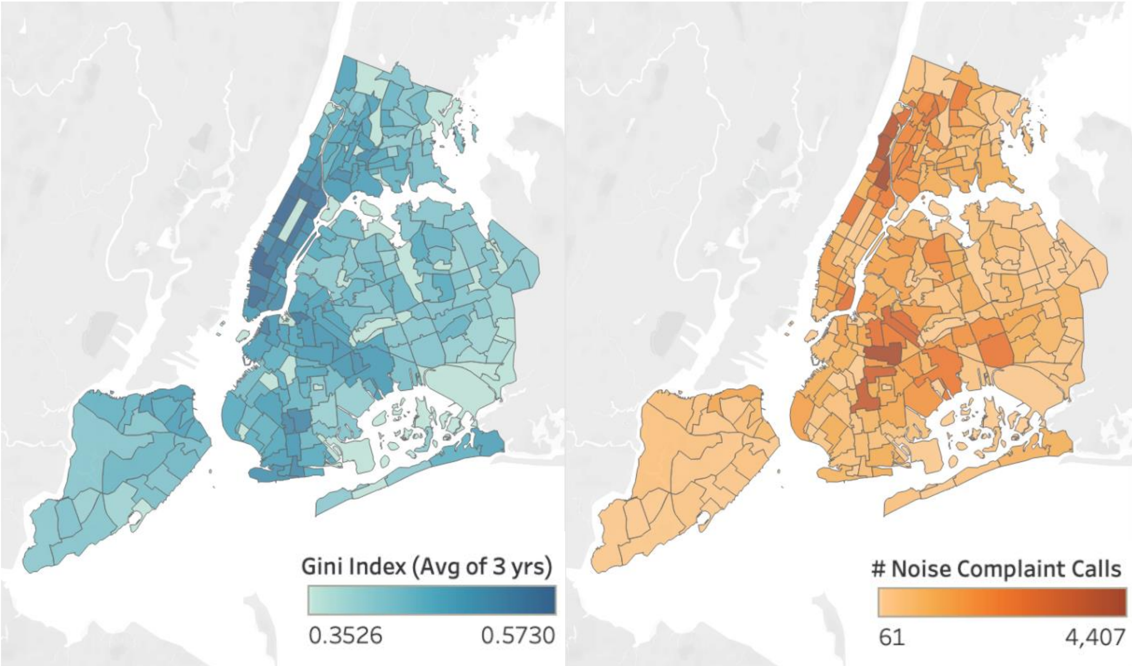


Table 2. Comparison of four regression models

	(1)	(2)	(3)	(4)
Models	Bivariate	FE	FE Poisson	FE Neg.Bin.
Variables	Outcome: Number of Noise Complaint Calls			
Gini index (NTA-level)	6,025.578***	1,464.019*	1.450***	1.540**
	(800.481)	(605.247)	(0.053)	(0.490)
# of Non-noise comp. calls (heat)		0.368***	0.000***	0.000***
		(0.020)	(0.000)	(0.000)
Population		0.014***		
		(0.001)		
Male (%)		11.893	0.003**	0.004
		(9.823)	(0.001)	(0.008)
Foreign-born (%)		1.570	0.011***	0.011***
		(2.083)	(0.000)	(0.002)
Race/ Ethnicity				
Hispanic (%)		11.002***	0.011***	0.011***
		(1.522)	(0.000)	(0.001)
Black (%)		6.179***	0.007***	0.006***
		(1.151)	(0.000)	(0.001)
Asian (%)		0.734	-0.006***	-0.005**
		(1.935)	(0.000)	(0.002)
Education Attainment				
Less than HS (%)		4.093**	-0.000	-0.001
		(1.567)	(0.000)	(0.001)
BA degree or higher (%)		-2.746	-0.003***	-0.000
		(3.475)	(0.000)	(0.003)
Housing Stability				
Same house >1 yr (%)		-3.529	-0.023***	-0.018**
		(7.039)	(0.001)	(0.006)
Rental vacancy rate		3.170	0.010***	0.012
		(8.516)	(0.001)	(0.008)
Owner-occupied units (%)		4.568*	-0.000	0.002
		(2.307)	(0.000)	(0.002)
HH with people <18 (%)		-21.309***	-0.018***	-0.017***
		(3.616)	(0.000)	(0.003)
HH with people >65 (%)		-20.589***	-0.033***	-0.033***
		(4.397)	(0.000)	(0.004)
Median gross rent ($)		-0.689***	-0.001***	-0.001***
		(0.119)	(0.000)	(0.000)
Logged median HH income		327.970*	0.421***	0.267*
		(162.014)	(0.015)	(0.136)
Constant	-1,713.901***	-3,076.569		-8.953***
	(367.294)	(1,836.489)		(1.429)
Observations	564	564	564	564
Year Fixed Effects	N	Y	Y	Y
R-squared	0.092	0.849		
Number of years		3	3	3
Exposure: Population	N	N	Y	Y

Standard errors in parentheses *** p<0.001, ** p<0.01, * p<0.05

Statistical analyses with different regression models furthered the understanding of the relationship between the predictor variable, Gini index, and the outcome variable, the number of noise complaint calls. Table 2 shows the regression results from four modeling approaches. We can confirm from all four models that NTAs with a higher Gini index have a higher number of noise complaint calls with significance.

The bivariate model (1) shows the unadjusted relationship between the Gini index and noise complaint calls without any control variables, and the coefficient of the predictor is 6,026 with significance ($p<0.001$), which indicates that the relationship between the variables of interest is positive. Thus, there are more noise complaints in NTAs where the Gini index is higher.

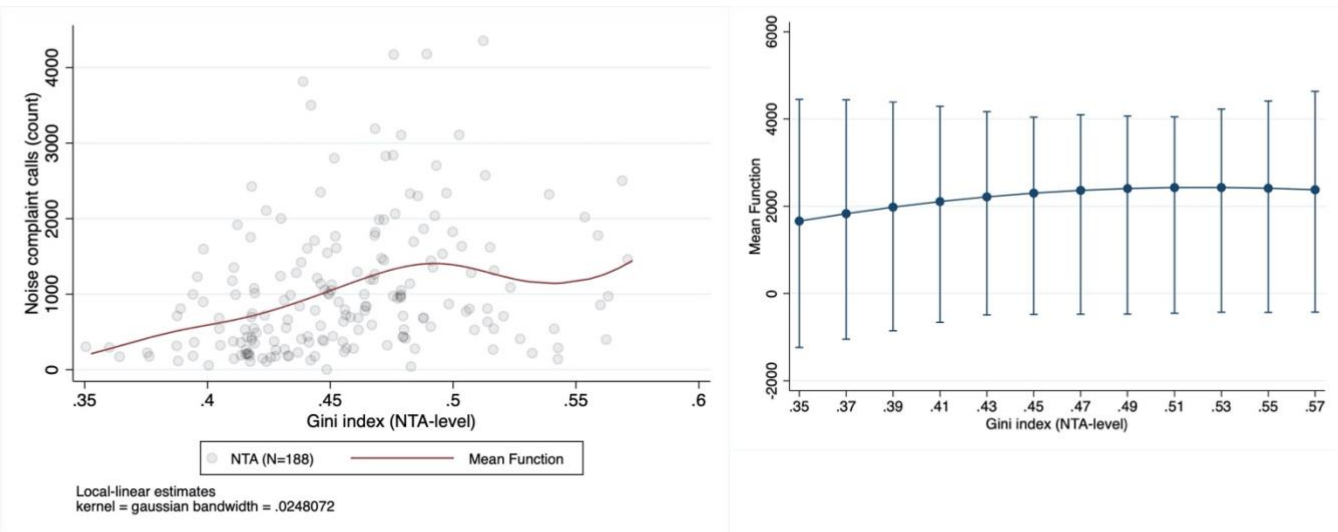
After adding all the covariates and using a fixed effects model (2), we can still observe the positive relationship between the Gini index and the number of noise complaint calls while holding all else constant. Despite the decrease in the magnitude of the coefficient from 6,026 to 1,464, the positive relationship is yet significant at the ($p<0.05$) level. This shows that the Gini index is significantly associated with the number of noise complaint calls throughout the given years 2013, 2015, and 2017.

The relationship between the Gini index and the number of noise complaint calls can be better assessed with models that take non-negative integer outcome values that count something such as the noise complaint calls made to 311 in this study. The fixed effects Poisson model in model (3) shows that the Gini index has a positive relationship with the number of noise complaint calls throughout the years with significance ($p<0.001$). The magnitude of the coefficient of the Gini index decreased tremendously because the population was used as an exposure variable rather than a control variable, which adjusts each observation so that the number of noise complaint calls is weighted based on the total population of each NTA.

The overdispersion of the noise complaint calls was taken into account by using a fixed effects negative binomial regression in model (4). The population is also used as an exposure variable like it was in model (3). While holding all else constant, model (4) shows the positive relationship between Gini index and the number of noise complaint calls with an increased magnitude of the coefficient and a decreased level of significance (p <0.01) from the fixed effects Poisson model to the fixed effects negative binomial model.

After confirming the significant relationship between the socio-economic variables and neighborhood complaint calls, I went through a sensitivity analysis using leads and lags. Based on the results of the models with leads and lags shown in the appendix, the relationship between the socio-economic heterogeneity and the number of noise complaint calls is significant for both the leads (2013) and lags (2017). This signifies that the relationship of interest is spurious, and further research is needed to confirm if there is any causal relationship.

Figure 4. mean function of noise complaint calls created with Kernel regression (left) and Margins plot of adjusted predictions of 95% Cis (right)



In terms of visualizing the relationship between the Gini index and the number of noise complaint calls, I used the npregress function in Stata to create a non-parametric plot with Gaussian kernel regression (Figure 4). The reason for using a non-parametric technique for visualization is that the scatter plot of the NTAs by Gini index and the noise complaint calls shows some varying patterns across the Gini index, which is a hint for the non-linear relationship between the variables of interest. Using kernel regression and without considering covariates, the number of noise complaint calls reaches the max when the Gini index is about 0.49, then a decreasing slope is presented with an increasing slope at the end. When the covariates are controlled and set to their means, the margins plot can be shown in Figure 4 (right).

While the positive relationship between the variables of interest is evidently shown above with the four models, the interpretations of the coefficient are rather challenging as it is hard to look into a unit change in the Gini index. Gini index is a scale from 0 to 1, thus a unit change in the Gini index would not be realistic. As an alternative approach, I further transformed the Gini index into a tertile to compare the different levels of the Gini index. Table 3 shows the modeling results of the tertile Gini created in two different ways, 33%,33%,33% and 25%,50%,25%. Each percentage represents Low, Medium, High Gini groups respectively. Moreover, the yearly overall Gini index is controlled as Gini index is a relative measure.

Table 3. Comparison among continuous Gini and tertile Gini

Models	(4)	(5)	(6)
Variables	Outcome: Number of Noise Complaint Calls		
Gini index (NTA-level)	1.540**		
	(0.490)		
33% 33% 33% Tertiles			
Low Gini		-0.075	
		(0.041)	
High Gini		0.032	
		(0.036)	
25% 50% 25% Groups			
Low Gini			-0.129**
			(0.045)
High Gini			0.110**
			(0.036)
Overall Gini index		5.945	1.091
		(15.565)	(15.626)
# of Non-noise comp. calls (heat)	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)
Male (%)	0.004	0.006	0.005
	(0.008)	(0.008)	(0.008)
Foreign-born (%)	0.011***	0.010***	0.011***
	(0.002)	(0.002)	(0.002)
Race/ Ethnicity			
Hispanic (%)	0.011***	0.011***	0.011***
	(0.001)	(0.001)	(0.001)
Black (%)	0.006***	0.006***	0.006***
	(0.001)	(0.001)	(0.001)
Asian (%)	-0.005**	-0.005**	-0.005**
	(0.002)	(0.002)	(0.002)
Education Attainment			
Less than HS (%)	-0.001	-0.000	-0.001
	(0.001)	(0.001)	(0.001)
BA degree or higher (%)	-0.000	0.002	0.000
	(0.003)	(0.003)	(0.003)
Housing Stability			
Same house >1 yr (%)	-0.018**	-0.017**	-0.017**
	(0.006)	(0.006)	(0.006)
Rental vacancy rate	0.012	0.016	0.011
	(0.008)	(0.008)	(0.008)
Owner-occupied housing units (%)	0.002	0.002	0.002
	(0.002)	(0.002)	(0.002)
HH with people <18 (%)	-0.017***	-0.017***	-0.016***
	(0.003)	(0.003)	(0.003)
HH with people >65 (%)	-0.033***	-0.032***	-0.033***
	(0.004)	(0.004)	(0.004)
Median gross rent ($)	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)
Logged median HH income	0.267*	0.223	0.315*
	(0.136)	(0.142)	(0.136)
Constant	-8.953***	-10.790	-9.428
	(1.429)	(7.356)	(7.362)
Observations	564	564	564
Number of years	3	3	3
Exposure: Population	Y	Y	Y

Standard errors in parentheses *** p<0.001, ** p<0.01, * p<0.05

Model 5 shows the regression modeling result of the 33%, 33%, 33% tertile Gini index, and model 6 shows the result of 25%, 50%, 25% three group Gini. The coefficients of Low Gini and High Gini are not significant for model 5, indicating that the tertile groups are not significantly different when it comes to noise complaint calls. On the other hand, Model 6 with 25%, 50%, 25% Gini groups shows clearer differences between the three Gini groups. Compared to the NTAs with Medium Gini index, the NTAs with Low Gini index resulted in fewer noise complaint calls with a statistically significant coefficient of -0.129 ($p < 0.01$) and the NTAs with High Gini index resulted in more noise complaint calls with a statistically significant coefficient of 0.110 ($p < 0.01$). The magnitudes of the coefficients are larger in model 6 compared to model 5, indicating that NTAs at the top (25%) and bottom (25%) of the Gini index are much more different than NTAs in the 50% Gini group.

DISCUSSION

This study examines the socio-economic diversity of NTAs measured with the Gini index and its relationship with neighborhood conflict in NYC. Given the fact that neighborhood study has been largely revolving around the topic of ethno-racial diversity and socio-economic status, this paper fills the gap that exists in the literature by emphasizing the importance of looking into socio-economic diversity. Moreover, by taking the noise complaint calls made to 311, a non-emergency government hotline, as the proxy for neighborhood conflict, the present study looks into the empirical evidence of neighborhood conflict in the form of noise complaints and overcomes the weakness of survey data that may only explain the idea of neighborhood conflict rather than actual real-time conflict happened in the neighborhood. Therefore, this paper provides a new viewpoint on how to approach neighborhood studies.

This study tested the hypothesis that “There will be more neighborhood conflicts, as measured by noise complaints, in neighborhoods where the level of socio-economic diversity is higher in NYC.” After conducting multiple statistical analyses with the fixed effects negative binomial regression models, I found that the results of the modeling coincide with the hypothesis, thus there is a positive relationship between the Gini index and the number of noise complaint calls. That is, the higher the level of socio-economic diversity of an NTA is, the more neighborhood conflicts occur within that NTA. Moreover, comparing the NTA groups with low, medium, high Gini index also showed that there exist significant differences between the neighborhoods with different levels of socio-economic diversity. This finding has a thread of connection with the findings from previous literature that living in a community with a high degree of income disparity is associated with lower trust (Alesina and La Ferrara 2002). In a sense that neighbors will call 311 rather than knocking on the other neighbor's door, we can assume that there is a lack of trust between neighbors, and the findings of this paper suggest that there are more noise complaints in neighborhoods with a high degree of income disparity.

Despite the contribution this paper made to the field of neighborhood studies, there are some limitations in this study that need to be addressed. The fundamental limitation of 311 data is the fact that it does not provide any personal information about the caller besides the time, location, and type of the complaint calls (Minkoff 2016). There is no such information as age or gender that might help researchers study the individual tendency to call 311. While 311 data provide rich information and insight into macro-level neighborhood dynamic and empirical evidence of neighborhood conflict, to look into neighborhood conflict at the micro-level, it is encouraged for future studies to incorporate survey data that show the individual experience of neighborhood conflict (Liu et al. 2019).

Moreover, NYC 311 data do not address the objective level of the noise in the neighborhood. Thus, there is a possibility that noise complaints are reflecting the levels of noise rather than levels of conflict. However, this study is focusing on the evident fact that there were complaint calls made which signifies that neighborhood conflict was present at that moment regardless of the objective intensity of the noise. Looking into objective noise level is encouraged for future studies as it may advance the use of government service request open data and potentially bolster the analysis of 311 data and neighborhood noise complaints.

Furthermore, the role that gentrification plays in neighborhood dynamics is not explicitly analyzed in this paper. Yet, variables related to housing stability were included in the analysis. Moreover, the changes in the neighborhoods throughout the years have been ruled out by the fixed effects model to clearly look into the picture of the relationship between the Gini index and noise complaint calls. Thus, future studies may incorporate datasets that address gentrification and development of the neighborhoods to further their understandings of changes in neighborhood dynamics.

CONCLUSION

This paper offers a comprehensive analysis of the socio-economic diversity of a city that builds upon the study of socio-economic status using 311 (Minkoff 2016). It also enhances the understanding of neighbor conflict in NYC in light of the rich information 311 provides. The use of government open data enabled this study to objectively assess the empirical evidence of neighborhood conflict in the form of noise complaints that prevail in NYC. The main finding of the paper, therefore, is that conflict among neighbors is more prevalent in neighborhoods with a higher degree of socio-economic diversity. Therefore, it is necessary to resolve this issue of income inequality to create social solidarity or trust among neighbors in NYC. Despite some limitations the paper has, it is encouraged for future studies to further their use of 311 to understand neighborhood dynamics and incorporate other types of datasets to understand the nature of micro-level neighborhood characteristics.

APPENDIX

Table 4. Sensitivity analysis using lead and lag

	(7)	(8)
Models	Neg. Bin. Lag	Neg. Bin. Lead
Outcomes	Noise Calls in 2017	Noise Calls in 2013
Variables in 2015		
Gini index (NTA-level)	3.236**	4.350***
	(1.032)	(0.937)
# of Non-noise comp. calls (heat) -17	0.000**	
	(0.000)	
# of Non-noise comp. calls (heat) -13		0.000***
		(0.000)
Male (%)	0.002	0.011
	(0.017)	(0.015)
Foreign-born (%)	0.014***	0.016***
	(0.003)	(0.003)
Race/ Ethnicity		
Hispanic (%)	0.015***	0.014***
	(0.003)	(0.002)
Black (%)	0.007***	0.010***
	(0.002)	(0.002)
Asian (%)	-0.006	-0.005
	(0.003)	(0.003)
Education Attainment		
Less than HS (%)	-0.004	-0.004
	(0.008)	(0.008)
BA degree or higher (%)	-0.006	-0.004
	(0.006)	(0.005)
Housing Stability		
Same house >1 yr (%)	-0.023*	-0.015
	(0.011)	(0.010)
Rental vacancy rate	0.002	-0.006
	(0.014)	(0.013)
Owner-occupied units (%)	0.000	0.007
	(0.004)	(0.004)
HH with people <18 (%)	-0.018**	-0.026***
	(0.006)	(0.006)
HH with people >65 (%)	-0.036***	-0.048***
	(0.008)	(0.007)
Median gross rent ($)	-0.001***	-0.001***
	(0.000)	(0.000)
Logged median HH income	0.685**	0.264
	(0.258)	(0.230)
Constant	-8.588**	-6.331*
	(2.855)	(2.516)
Exposure: Population	Y	Y
Observations	188	188

Standard errors in parentheses *** p<0.001, ** p<0.01, * p<0.05

REFERENCE

Alesina, A. and La Ferrara, E. (2002) ‘Who Trusts Others?’ *Journal of Public Economics* 85(2): 207-34.

America’s Health Rankings 2019. “Annual Report.” Retrieved December 10, 2019 (https://www.americashealthrankings.org/explore/annual/measure/gini/state/NY?edition-year=2018)

Anderson, Mary R. 2010. “Community Psychology, Political Efficacy, and Trust.” Political Psychology 31 (1): 59–84.

Blumer, Herbert. 1958. “Race Prejudice as a Sense of Group Position.” *Pacific Sociological* Review 1:3–7.

Brady, Henry E., Sidney Verba, and Kay L. Schlozman. 1995. “Beyond SES: A Resource Model of Political Participation.” American Political Science Review 89 (2): 271–94.

Carr, P. J., Napolitano, L., Keating, J. (2007). We Never Call The Cops And Here Is Why: A Qualitative Examination Of Legal Cynicism In Three Philadelphia Neighborhoods. *Criminology, 54(2).* 445-480.

Cheshire, Lynda, Robin Fitzgerald, and Yan Liu. 2018. “Neighbourhood Change and Neighbour Complaints: How Gentrification and Densification Influence the Prevalence of Problems between Neighbours.” *Urban Studies* 56(6):1093–1112.

DiPasquale, Denise, and Edward L. Glaeser. 1999. “Incentives and Social Capital: Are Homeowners Better Citizens?” Journal of Urban Economics 45 (2): 354–84.

Goerge, Robert M., and Bong Joo Lee. 2002. “Matching and Cleaning Administrative Data.” *New Zealand Economic Papers* 36(1):63-64.

Huynh, M., Parker, J. D., Harper, S., Pamuk, E., & Schoendorf, K. C. (2005). Contextual effect of income inequality on birth outcomes. *International Journal of Epidemiology, 34*, 888–895.

Kingston, Paul W., and John C. Fries. 1994. “Having a Stake in the System: The Sociopolitical Ramifications of Business and Home Ownership.” Social Science Quarterly 75 (3): 679–86.

Lancee, Bram, and Jaap Dronkers. 2011. “Ethnic, Religious and Economic Diversity in Dutch Neighbourhoods: Explaining Quality of Contact with Neighbours, Trust in the Neighbourhood and Inter-Ethnic Trust.” *Journal of Ethnic and Migration Studies* 37(4):597–618.

Legewie, Joscha, and Merlin Schaeffer. 2016. "Contested Boundaries: Explaining Where Ethnoracial Diversity Provokes Neighborhood Conflict.” *American Journal of Sociology* 122(1):125–61.

Letki, Natalia. 2008. “Does Diversity Erode Social Cohesion? Social Capital and Race in British Neighbourhoods.” *Political Studies* 56(1):99–126.

Liu, Yan, Lynda Cheshire, Siqin Wang, and Xuanming Fu. 2019. “A Socio-Spatial Analysis of Neighbour Complaints Using Large-Scale Administrative Data: The Case in Brisbane, Australia.” *Cities* 90:168–80.

LiveStories 2019. “New York County Gini Index of Income Inequality.” Retrieved December 10, 2019 (https://www.livestories.com/statistics/new-york/new-york-county-gini-index-income-inequality)

Lyons, Christopher J. 2007. “Community (Dis)Organization and Racially Motivated Crime.” American Journal of Sociology 113:815–63.

Méndez, María Luisa, and Gabriel Otero. 2018. "Neighbourhood Conflicts, Socio-Spatial Inequalities, and Residential Stigmatisation in Santiago, Chile.” Cities 74:75–82.

Minkoff, Scott L. 2016. “NYC 311: A Tract-Level Analysis of Citizen–Government Contacting in New York City.” *Urban Affairs Review* 52(2):211–46.

Niemi, Richard G., Stephen C. Craig, and Franco Mattei. 1991. “Measuring Internal Political Efficacy in the 1988 National Election Study.” American Political Science Review 85 (64): 1407–13.

Putnam, Robert D. 2007. "E Pluribus Unum: Diversity and Community in the Twenty-First Century." Scandinavian Political Studies 30:137–74.

Sampson, R. J. (1997). Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy. Science, 277(5328), 918-924.

Sands, Melissa L. 2017. “Exposure to Inequality Affects Support for Redistribution.” Proceedings of the National Academy of Sciences114(4):663–68.

Schaub, Max, Johanna Gereke, and Delia Baldassarri. 2019. “Does Poverty Undermine Cooperation in Multiethnic Settings? Evidence from a Cooperative Investment Experiment.” *Journal of Experimental Political Science* 1–14.

Sharkey, Patrick, and Jacob Faber. 2014. "Where, When, Why, and For Whom Do Residential Contexts Matter? Moving Away from the Dichotomous Understanding of Neighborhood Effects." *Annual Review of Sociology* 40: 559-579.