Spring 5-2023

Gentrification and Crime in the Twin Cities: Insights and Challenges through a Statistical Lens

Erin G. Franke
efranke@macalester.edu

Gentrification and Crime in the Twin Cities: Insights and Challenges through a Statistical Lens

Erin Franke

Vittorio Addona, Advisor

Kelsey Grinde, Reader

Daniel Trudeau, Reader

May 2023

Macalester College

Department of Mathematics, Statistics, and Computer Science

Abstract

Gentrification is a complex process of urban redevelopment that typically involves an in-migration of educated people to neighborhoods experiencing a period of disinvestment. While gentrification is widely regarded for its potential to displace long-time businesses and residents of the neighborhood, its impact on crime is highly controversial. There is not a consensus on the relationship between gentrification and crime across criminological theory and past statistical studies have also shown contradictory results. Measuring gentrification on the tract level with census data, we seek to understand gentrification's relationship with violent crime and theft in the Twin Cities. Using a Poisson model with spatial components, our results show no indication that gentrification results in reduced rates of violent crime or theft. Broader crime patterns and implications of gentrification are also discussed.

Contents

List of Figures

List of Tables

Acknowledgements

There are many people that have supported me throughout my time at Macalester but I would first like to thank my honors project advisor, Victor Addona. You have helped me become a smarter student and better person and I am so grateful for your support. Thank you for spending part of your spring break in Olin-Rice answering all my paper questions and for always responding to my emails (even if it is just to maintain your happy sun.. just kidding) with smiley faces and exclamation points included! And of course, thanks for brightening my day by waving at me through the window every morning! :)

I would also like to thank the rest of my committee, Kelsey Grinde and Daniel Trudeau, for your thoughtful comments and questions. I appreciate you helping me to look at this topic from various perspectives. Thank you, Dan, for contributing your gentrification expertise and committing time and energy to this project without ever having met me prior to the defense. Thank you, Kelsey, for helping me improve my presentation skills over the past few semesters and supporting me academically/career-wise/athletically. And for the occasional morning walk waves!

Third, I want to thank my academic advisor, Brianna Heggeseth. We may have first met accidentally on a sidewalk (boo COVID), but you have taught me so much and have been an integral part to helping me become passionate about working with data. I also want to thank you for helping me learn to map with sf and make spatial models in Correlated Data. Don't forget to tell me when you discover the next cool RStudio trick! :)

Fourth, I want to thank Suzanne Burr and the MSCS department as a whole for creating a welcoming community where I have formed many friendships. And last but not least, thank you to my family, friends, teammates, and coaches for being so fantastic and supporting me throughout this project and at Macalester. I'm so lucky to have you all in my life!

1 Introduction

Dating back to 2010, the Twin Cities have shown evidence of rapid growth. According to decennial U.S. census counts, the Twin Cities seven-county metropolitan region population gained 314,000 residents between 2010 and 2020, bringing the total population to 3.16 million in 2020 [Metropolitan, 2021]. Demographers from the Metropolitan Council[1] estimate the region gained 116,000 residents from migration during the 2010s, compared with a net loss of 26,000 during the 2000-2010 decade. Minneapolis and Saint Paul saw population growth rates of 12.4% and 9.3% between 2010 and 2020, respectively. The share of residents in the two cities who identify as Black, Indigenous, and People of Color (BIPOC) in 2020 was estimated to be 31%, up from 24% in 2010 [Metropolitan, 2021].

With rapid population growth it is not unreasonable to expect parts of the Twin Cities to undergo gentrification. *Governing* magazine looked at demographic change data for the United States' 50 largest cities in population and found 50.6% of eligible census tracts in Minneapolis to have gentrified between 2000 and 2015[2], the third highest in the nation behind only Portland, Oregon and Washington, DC [Maciag, 2015]. With gentrification comes change in residential mobility, housing renovation and reconstruction rates, and commerce among other categories. Whether or not such changes are "good" or "bad" for neighborhood residents is heavily debated by developers, city planners, politicians, and academics alike [Papachristos et al., 2011]. Proponents of gentrification often cite the reduction of crime as one benefit, using the logic that underprivileged neighborhoods "upgrade" with the inflow of more prosperous residents and as a result a reduction in crime is observed [Papachristos et al., 2011]. For example, the City of Baltimore created a program called *Crime Prevention Through Environmental Design*, which aims to reduce violence by cleaning, creating green spaces, upgrading lighting and demolishing or repurposing vacant homes and spaces [Mendez, 2022]. However, prior studies on gentrification's relationship with crime have displayed greatly mixed results. A number of such studies show a positive relationship between gentrification and crime ([Covington and Taylor, 1989], [Smith, 2012]) while others show a negative relationship ([McDonald, 1986], [Papachristos et al., 2011], [Smith, 2012], [Barton, 2016]) and still others show a more complex relationship that varies with time [Kreager et al., 2011]. These contradicting results are likely the result of a number of factors, including the difficult task of measuring complex gentrification processes and the diverse crime outcomes studied [Smith, 2012], as well as the methodology used.

Given the degree to which the Twin Cities have experienced urban change over the past two decades

[1]The Metropolitan Council is a policy-making body, planning agency, and provider of essential services for the Twin Cities metropolitan region.

[2]*Governing* came up with this number by first denoting if a tract was *eligible to gentrify*. To be eligible, tracts had to have a median household income and a median home value in the bottom 40th percentile metrowide. A tract was then marked as *gentrified* if it recorded increases in home value and percent of the population with bachelor's degrees in the top one-third of all tracts in the metro area.

[Maciag, 2015], we believe it is important to understand the impacts these changes have had on crime. Thus, in this study, we explore the relationship between gentrification and neighborhood violent crime and theft in Minneapolis and Saint Paul, MN over the period 2010-2020. The paper is structured as follows. We begin by conceptualizing gentrification and discussing the challenges that come with trying to measure and understand it. We then review relevant criminological theory in addition to statistical studies on crime and gentrification. Next, we present our analysis, highlighted by a series of Poisson regression models. After a review of the results, we conclude with a brief discussion of the broader impacts of gentrification in the Twin Cities.

2 Defining Gentrification

Gentrification is a multifaceted, difficult to define process involving political, corporate, and independent actors and, as a result, its precise definition and political utility creates intense academic debate [Papachristos et al., 2011]. The United States Department of Housing and Urban Development defines gentrification as "a form of neighborhood change that occurs when higher-income groups move into low-income areas, potentially altering the cultural and financial landscape of the original neighborhood" [States Department of Housing and Urban Development, 2016]. Urban scholars commonly define gentrification as "the class transformation of those parts of the city that suffered from systematic outmigration, disinvestment, or neglect in the midst of rapid economic growth and suburbanization" [Kreager et al., 2011, Wyly and Hammel, 1999]. Additional research has conceptualized gentrification as "a churning process that involves the in-migration of wealth and the out-migration of poverty, most often resulting in over time increases in median household incomes, property values, and presence of lifestyle amenities that appeal to the tastes—and meet the demands of—the wealthier residents" [Papachristos et al., 2011, Lees et al., 2008]. While these represent a small sample of definitions of gentrification and they are not fully interchangeable, there are consistent ideas across them. The neighborhoods that have the potential to undergo gentrification are urban neighborhoods comprised largely of low-income residents that have previously experienced disinvestment. Additionally, the process of gentrification itself is centered around an inflow of wealthier residents and reinvestment [Freeman, 2005].

Encapsulating the complexity of the definition of gentrification in a quantitative study is not easy. Prior studies have most commonly utilized census data on the census tract level to track changes in neighborhood characteristics over a number of years. While census data covers several of the ideas included in the gentrification definitions (e.g. increased median household income, increased home values), it certainly has its limitations. First, these census measures typically cannot pick up on changes in local culture, an important aspect of the gentrification process [Barton, 2016]. Second, the quality and type of data necessary to do these analyses is typically only captured in the decennial census. As a result, measures of neighborhood change must be interpolated using 10-year time intervals, essentially estimating any change associated with gentrification as a linear trend [Papachristos et al., 2011]. Past research has shown the relationship between gentrification and crime may not be linear [Kreager et al., 2011]. Third, relying on aggregate census indicators at the neighborhood level implies an assumption that only individual residents drive gentrification processes. This assumption fails to consider the role that corporate and political actors play in the process [Papachristos et al., 2011]. The forces that gentrify neighborhoods vary as private economic development can gentrify a neighborhood but so can state intervention. A study of gentrification's effect on gang homicides in Chicago found that private investment gentrification has a marginally significant and negative effect on gang homicide, while in contrast, state-based gentrification has a positive effect on gang homicide [Smith, 2012]. Thus, the way gentrification is occurring has an important context and it may be difficult for census data to pick up on this. In a similar manner, these census metrics cannot identify subtle neighborhood changes such as condominium sales, repainting, putting up new signage, or simply better landscaping [Kreager et al., 2011].

To account for these limitations of census data, recent studies have attempted to find other ways to measure gentrification. A study on the relationship between gentrification and neighborhood crime rates in Chicago utilizes coffee shops as a measure of gentrification, stating, "Measuring the number of coffee shops in a neighborhood has the distinct advantage over the more commonly employed census and survey indicators in that coffee shops provide an on-the-ground and visible manifestation of a particular form of gentrification—the increased presence of an amenity often associated with gentrifiers' lifestyles" [Papachristos et al., 2011]. Mixed methods approaches have also been used. Researchers Daniel J. Hammel and Elvin K. Wyly utilized historical records and income indicators from the 1960 census to identify urban tracts which underwent substantial disinvestment post WWII and then visited these "at-risk of gentrifying" tracts between 1994 and 1998 to record evidence of housing stock development (e.g., condominium and housing construction) and renovations (e.g., exterior paint, signage). A study through the University of Minnesota, *The Diversity of Gentrification: Multiple Forms of Gentrification in Minneapolis and Saint Paul* [Goetz et al., 2019], uses a similar mixed methods process (but does not focus on the relationship between gentrification and crime). Instead, their report relies on the methods of three peer reviewed publications [Freeman, 2005, Bates, 2013, Ding et al., 2016] to identify tracts eligible for gentrification and denote whether or not these tracts gentrified using census data. Goetz et al. then conduct an in-depth qualitative analysis through interviews with public officials, community leaders, and neighborhood residents to best understand the residents' impression of the neighborhood change that occurred and its implications. Finally, gentrification has been measured through the use of annual home mortgage investment data. Kreager et al. utilized this yearly data as a proxy for urban revitalization to test for a curvilinear relationship between gentrification and crime [Kreager et al., 2011].

While these methods certainly have benefits, they also have their faults. Even though using coffee shops as a measure of gentrification has the benefit of being an "on-the-ground visible manifestation of gentrification," the location of coffee shops is influenced by city planning efforts, individual tastes, and residential preferences. Furthermore, the coffee shops in this study were clustered in the central business district, potentially excluding a significant number of areas that underwent similar neighborhood change just without a coffee shop [Barton, 2016]. The annual home mortgage data used by Kreager et al. (2011) does not have a neighborhood exclusion issue, however the data is directly related to property value and therefore may be more likely to result in changes in property crime as opposed to violent crime [Barton, 2016]. Perhaps having the fewest analytical faults are the mixed methods studies, however these can be time-consuming and costly to carry out.

Due to cost and time constraints, we utilize census data to identify gentrifiable areas and measure gentrification. Despite its limitations, census metrics do allow us to identify gentrifying areas in a way that aligns with society's notions about gentrification: neighborhoods seeing an influx in educated residents and increased home values or rent. Following our analysis, we rely heavily on the findings and in-depth qualitative interviews from a study by Goetz et al. to understand *how* gentrification occurred in different parts of the Twin Cities and its implication on residents.

Before getting into prior research, it is important for us to acknowledge that some of the prominent studies we review include race as a component in their definition of gentrification or in their models. A study by a Stanford sociologist has shown that the negative effects of gentrification are felt disproportionately by minority communities, whose residents have fewer options of neighborhoods they can move to compared to their white counterparts [Feder, 2020]. However, we find it inappropriate to include race as a metric for identifying an area as eligible to gentrify or to denote it as having gentrified. Moving forward, we review these studies for the purpose of explaining some of the most eminent past work, but as a reader please understand we do not endorse these methods and refuse to do this in our own study. Instead, following our results, we choose to discuss how gentrification may disproportionately affect certain groups and what steps can be taken to prevent this.

3 Gentrification and Crime: Theory and Prior Analysis

Research on gentrification and crime is largely built upon the routine activities and social disorganization/collective efficacy theories [Barton, 2016]. The routine activities theory represents the idea that crime is more likely to occur when a high value target lacking in capable guardianship converges in time and space with a motivated offender [Barton, 2016]. With this in mind, we expect gentrification to prompt crime as middle/upper class people move into comparatively disadvantaged neighborhoods populated by residents that may be unhappy about the changes occurring in their neighborhood. Social disorganization theory is built upon the idea that gentrifying areas might experience an increase in crime rates as neighborhood social structures undergo a period of flux and socioeconomic heterogeneity, which decreases a community's ability to control crime internally. However, as the community continues through the gentrification process and residential mobility stabilizes, disorder should decline and social organization should increase [Papachristos et al., 2011, Kirk and Laub, 2010]. This is based on the proposition that with a more populated area (perhaps busy with new businesses), there should be more eyes on the street and fewer opportunities for crime to occur. Thus, social disorganization theory suggests a curvilinear relationship between gentrification and crime.

While both social disorganization and routine activities theories support at least in part some kind of positive relationship between gentrification and crime, advocates of gentrification often make other claims. These include that the influx of more affluent residents and improvements in resources, institutions, and amenities lower crime rates and improve overall safety in gentrifying neighborhoods [Papachristos et al., 2011, McDonald, 1986]. Another argument asserts that gentrification reduces crime and delinquency through an increase in law enforcement efforts, new economic and social opportunities, or the displacement of criminal residents [Papachristos et al., 2011, Kirk and Laub, 2010]. All together, theories about gentrification's relationship with crime exhibit competing thoughts. Keeping this in mind, we will dive deeper into a couple of the more prominent studies on gentrification and crime in the past few decades to learn more.

3.1 More Coffee, Less Crime? The Relationship between Gentrification and Neighborhood Crime Rates in Chicago, 1991 to 2005

In September 2011, Andrew V. Papachristos, Chris M. Smith, Mary L. Scherer, and Melissa A. Fugiero published the study *More Coffee, Less Crime? The Relationship between Gentrification and Neighborhood Crime Rates in Chicago, 1991 to 2005.* This work brought a novel way to quantify gentrification to past research by measuring annually the growth and geographic spread of coffee shops, one of gentrification's most prominent symbols [Papachristos et al., 2011]. According to the authors, this measurement is able to provide an on-the-ground and visible manifestation of an amenity often associated with gentrifiers' lifestyles, something that the more commonly employed census and survey indicators cannot. With the goal of assessing the influence of gentrification on homicide and street robbery in Chicago, the coffee shop variable is represented as a three-year average count of coffee shops for the 15-year period between 1991 and 2005 (yielding a total of five time periods). The study includes 341 neighborhood clusters, which together are composed of 847 census tracts. The dependent variables of the study are the annual counts of homicides and street robberies, which too are collapsed into three-year counts given that crime data, especially homicide data, are zero-inflated (meaning they exhibit a relatively high frequency of zeros). The authors present a series of overdispersed longitudinal Poisson models with neighborhood fixed effects, which allow for modeling gentrification as a process over time rather than a singular event or singular change. These models regress the total number of neighborhood homicides and robberies (separately) on census factors, a lagged measure of coffee shops, and lagged levels of neighborhood crime[3]. Due to several census indicators being highly correlated, the authors used principal component analysis to create two census factors used in the models: *Neighborhood Change*, which represents a level of change in the proportion of the population with bachelor's degrees, recently moved individuals, new housing, and overall mean family income, as well as *Non-Black Hispanic Neighborhoods*, which stresses the racial segregation of Chicago neighborhoods.

The results of this study highlight that those neighborhoods experiencing gentrification also experience a greater than expected decline in homicide. Even when controlling for census factors, the authors find the lagged number of coffee shops to retain statistical significance, implying when the number of coffee shops increases, the number of subsequent homicides decreases. Robbery results differ from homicide results in that when the coffee shop variable is considered in conjunction with the census factors, the coffee shop variable is positive. This indicates that when other factors of gentrification are considered, an increase in the number of coffee shops is associated with an increase in robberies. According to the authors, this can be for a number of reasons. Gentrification's effect on crime may vary by crime type—in this case, gentrifiers may be more tolerant of nonlethal crimes like robbery versus something like homicide. It is also possible that as the routine activities theory suggests, the increased wealth associated with gentrifiers provides an opportunity for criminal activity. After further analysis, the authors find that this observed effect of the increase of coffee shops on robbery may relate to important ways in which gentrification unfolds across neighborhoods with different racial composition. More specifically, they find gentrification to be associated with declines in robbery in White and Hispanic neighborhoods, but increases in robbery in Black neighborhoods[4]. According to the authors, this result underscores the qualitative research finding that gentrification in Black neighborhoods takes a different form than gentrification in White neighborhoods - coffee shops are not opening in gentrifying Black neighborhoods, a result that is consistent with research on other neighborhood-level resources such as a 2006 study by Mario L. Small and Monica McDermott [Papachristos et al., 2011].

[3]Population was used as an exposure term in all models in order to effectively compare counts across census tracts

[4]The authors categorized a neighborhood as either "White", "Black", or "Hispanic" by the dominate race residing there.

3.2 Urban Revitalization and Seattle Crime, 1982–2000

Shortly following the study published by Papachristos et al. came *Urban Revitalization and Seattle Crime, 1982–2000* by Derek A. Kreager, Christopher J. Lyons, and Zachary R. Hays in November 2001. This analysis is completed in two parts, the first examining whether gentrification in Seattle relates to changes in crime in the 1990s and the second seeking to understand the *process* of gentrification within Seattle's poor tracts. Prior to the analysis, the authors hypothesize that gentrification in the 1980s was "spotty" and incomplete, and therefore positively related to crime change, but then reversed in the 1990s when gentrification became more consolidated and complete [Kreager et al., 2011]. In this study, tracts are classified as gentrifying, appreciating, or non-gentrifying poor using a mixed-methods system developed by Elvin K. Wyly and Daniel J. Hammel. Hammel and Wyly utilized historical records and income indicators from the 1960 census to identify urban tracts which underwent substantial disinvestment post WWII and then visited these "at-risk of gentrifying" tracts between 1994 and 1998 to record evidence of housing stock development (e.g., condominium and housing construction) and renovations (e.g., exterior paint, signage) to classify them [Wyly and Hammel, 1999, Kreager et al., 2011].

To first understand the relationship between gentrification and crime in Seattle, the authors use two-panel difference models, or change-score regression analyses, which regress changes in a dependent variable between two time points (here 1990 and 2000) on changes in independent variables or on a treatment occurring between the two panels. In this study, the authors treat the dependent variable as the difference in average annual crime between 1999-2001 and 1989-1991, creating separate models for total, property, and violent crime. While the timing of the gentrification measure is appropriately placed between the two endpoints, gentrification is not randomly assigned to the observed tracts. In order to reduce endogenous sources of bias, the authors control for 1990 crime (a strong predictor of 1999-2001 crime) as well as several time-varying covariates commonly thought to relate to changes in crime including total tract population (measured in 100s), percent foreign-born, percent black, percent college educated, percent residential mobility in the past five years, percent of homes built in the past five years, mean home mortgage value (measured in \$100k), and mean family income (measured in \$1k). The models also include a significant spatial error term. Initial results show that gentrification predicts significant reductions in total and property crime (but not violent crime), however when including average 1989-1991 crime in the models these results become statistically insignificant. According to the authors, this may be due to regression to the mean, as areas denoted as "gentrified" are more likely to have higher crime at the beginning of the decade.

The authors next complete an analysis to understand gentrification *as a process*. To test their hypothesis that the relationship between gentrification and crime is curvilinear (with a positive slope in the 1980s and a negative slope in the 1990s), the authors use yearly home mortgage investment as a proxy for urban revitalization [Kreager et al., 2011]. They use the total dollar amount of home loans originated in each tract per year from 1981 to 2000 in order to understand yearly tract-level housing activity. This analysis is centered around three conditional fixed effects negative binomial models of within-tract change in total crime, property crime, and violent crime counts reported by the Seattle Police Department between the years 1982–2000 in previously poor Seattle tracts. The fact that the census doesn't capture yearly changes in demographic characteristics prevent these controls from being included in the model. However, the authors use linear, quadratic, and cubic terms for time in order to control for time-varying characteristics that occur above the city level, such as changes in state and national economies, incarceration rates, and city-wide ordinances. They also include tract population in these fixed-effects models by linearly interpolating values for the between-census years. Mortgage investment is included in the model and interacted with the linear and quadratic time terms in order to test for non-linear relationships between housing activities over time. The mortgage investment variable and time trends are lagged to (t-1) in order to estimate the effects of past mortgage investments on future tract-level crime. Results of total and property crime models show a significant negative term for mortgage investment, indicating tracts that have large amounts of mortgage investment are expected to have less crime in the following year. Mortgage investment's interactions with time and time-squared are also significant, suggesting that the mortgage effects vary over the two decades and that there is a curvilinear relationship between gentrification and crime (with increases in crime in earlier years and decreases in crime in later years). Mortgage investment and its interactions are not significant for violent crime models. Despite the evidence of the curvilinear relationship they hypothesized, the authors note that the magnitude of this relationship is modest at best. When combined with findings from their decennial change-score analyses, there is evidence that gentrification alone explains less than ten percent of the change in total crime between 1990 and 2000 with unobserved factors playing a larger role.

3.3 Gentrification and Violent Crime in New York City

The final and most recent analysis we will highlight is Michael S. Barton's *Gentrification and Violent Crime in New York City*, which seeks to understand the relationship between gentrification and violent crime in 55 New York sub-boroughs[5] between 1980 and 2009. To do this, Barton uses a hybrid fixed-effect technique that allows for the inclusion of time-invariant measures, which help determine whether changes in the sampled neighborhoods are due to stationary characteristics such as access to transportation or if the changes are due to variation in population composition [Barton, 2016]. Barton operationalizes gentrification using a technique developed by Raphael W. Bostic and Richard W. Martin [Bostic and Martin, 2003]. With

[5]A sub-borough is composed of roughly 40 census tracts.

this method, tracts are identified as gentrified using information on median family income, proportion of residents with college degrees, home ownership rates, proportion of residents aged 30 to 44, proportion of White non-family households, proportion of managerial and administrative workers, proportion of residents with some college, percentage of residents in poverty, and percentage of Black residents. Barton creates separate models for aggravated assault, robbery, and homicide, with the dependent variable in each being a rate per 1,000 residents. Each model includes independent variables of the percentage of the sub-borough that gentrified, its concentrated disadvantage index [6], its residential stability[7], its percentage of foreign born residents, its percentage of tracts with subway entrances, and an indicator variable representing if the decade was the 1990s or 2000s[8] [Barton, 2016]. Results find a negative and statistically significant association between gentrification and all three violent crimes. Given Kreager et al.'s finding of the curvilinear relationship between gentrification and crime, Barton chose to investigate if a similar pattern occurred in New York by doing a fixed effect analysis with time interactions. Here, the curvilinear pattern was not found as non-significant associations of the time interactions indicate that the relationship of gentrification and each crime type was stable over the period of the analysis.

3.4 Summary and our contribution

From the three studies discussed we learn that the relationship between gentrification and crime is complex. Papachristos et al. finds that those neighborhoods experiencing gentrification also experience a greater than expected (statistically significant) decline in homicide. Similarly, Barton's study finds a significant and negative relationship between gentrification and three types of violent crime (aggravated assault, robbery, and homicide). On the other hand, Kreager et al.'s study focuses on the relationship between gentrification and three types of crime: total, property, and violent crime. They find there to be a negative relationship between gentrification and each crime type, though this relationship is insignificant. However, a highlight from their study is evidence of a curvilinear relationship between gentrification and crime, in which a tipping point of neighborhood investment is reached and crime rates decline after initially rising. After testing for a similar relationship, Barton denoted the relationship between gentrification and crime to be stable. It is important to keep in mind that each analysis is set in a different city and uses its own unique definition of gentrification in addition to different modeling techniques.

Our analysis will add to the current literature in two ways. First, the focus of our analysis will be assessing the relationship between gentrification and crime in the Twin Cities specifically, on which there doesn't

[6] An index the authors created using the percentage of residents receiving public assistance, percentage living below poverty, unemployment rate, percentage female-headed households, and percentage Black.

[7] Barton measures residential stability through change in the population who lived in the same home for at least 5 years.

[8] Barton includes this variable due to the fact that crime rates increased between the 1960s and 1980s in most major cities but then declined dramatically during the 1990s.

appear to be any prior literature. While *The Diversity of Gentrification: Multiple Forms of Gentrification in Minneapolis and Saint Paul* dives deeply into the recent history of gentrification in the Twin Cities, the impact on crime is not the focus. Second, unlike the Chicago and New York City analyses, our study will seek to understand gentrification's relationship with theft in addition to violent crime.

4 Data and Methods

4.1 Units of Analysis

The units of analysis of our study are census tracts in the Minneapolis and Saint Paul city limits across the period 2010-2020. A census tract is a small and relatively permanent statistical subdivision of a county and has an optimal population size of 4,000, though this can vary between 1,200 and 8,000 residents [Bureau, 2022]. Census tract boundaries are created with the goal of being permanent so that statistical comparisons can easily be made between decennial censuses, but due to population growth or decline, tracts are split or merged on occasion. In the Saint Paul city limits, three census tracts split over the decade. In Minneapolis, six census tracts split between 2010 and 2020 while one set of two tracts merged. Four of the six splitting census tracts are located in Northeast Minneapolis and shown in Figure 1.

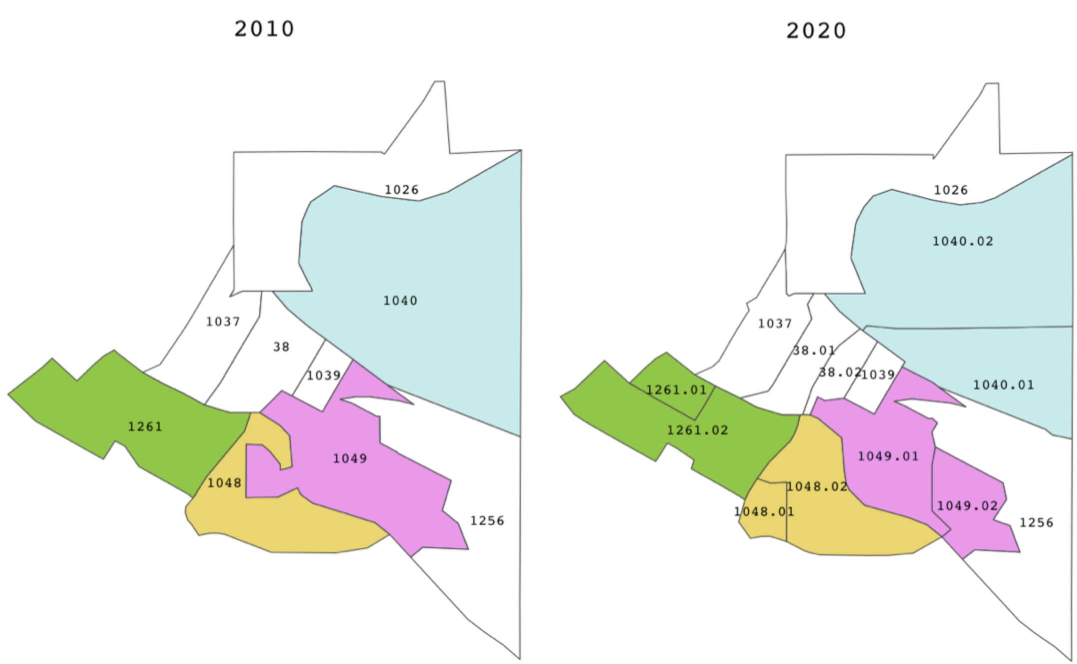

Figure 1: An example of redistricting in northeast Minneapolis. Census tracts 1261.00, 1048.00, 1040.00, and 1049.00 split between 2010 and 2020.

Redistricting poses problems for our analysis, as we are trying to measure whether or not a particular census tract gentrifies over the course of the decade. We do not have the data to determine, for example, if only the 1040.02 area of census tract 1040.00 was eligible to undergo gentrification in 2010. As a result, we default to always using the largest area of a particular census tract across our study. This means continuing to treat census tracts that split over the decade as one tract in 2020 (summing the crimes occurring in the split tracts and taking a weighted average based on population for census metrics such as the proportion of residents with a bachelors degree or median family income). Similarly, we treat our single pair (1023.00 and 1029.00) of Minneapolis tracts that merged over the decade as one census tract in 2010.

4.2 Operationalization of Gentrification

As previously discussed, encapsulating the difficult to define process of gentrification in a quantitative study is not a simple task with a correct answer. We measure gentrification using information from the 2010 and 2020 U.S. decennial censuses, accessed through the `TidyCensus` R package. This package allows users to interact with a select number of the U.S. Census Bureau's data APIs and return `tidyverse`-ready data frames, optionally with simple feature geometry included [Walker and Herman, 2022]. We first begin by denoting whether or not each census tract is *eligible to gentrify* in 2010. Given that we are choosing to follow past studies in measuring gentrification as a binary process, this step is necessary to separate out new growth in communities previously experiencing a period of disinvestment from continued development of wealthier neighborhoods. We define a census tract as eligible to gentrify in 2010 if it has a median household income that is less than the citywide median household income, as specified in the 2016 publication *Gentrification and residential mobility in Philadelphia* [Ding et al., 2016]. Figure 2 shows a map of the Twin Cities with each census tract designated as either eligible to gentrify or not. Major roads and highways are drawn in light grey.

Of census tracts eligible to gentrify, we next identify tracts that underwent gentrification over the course the decade. Again, we follow the methods of Ding et al., who define a tract as having gentrified if (a) the tract observes a change in the share of adults with college degrees greater than the city-level change *and* (b) sees a change in median rents above the citywide change *or* a change in median home value greater than the citywide change [Ding et al., 2016]. With these conditions, we denote 25 out of 58 eligible census tracts (43.1%) to have gentrified in Minneapolis and 11 out of 44 (25.0%) in Saint Paul. 28 out of 94 *ineligible to gentrify* tracts across both cities also meet these standards, though again our analysis attempts to focus on areas that have previously seen periods of disinvestment - thus these 28 tracts will not be marked as gentrifying. Figure 3 presents a visualization of these categorizations.

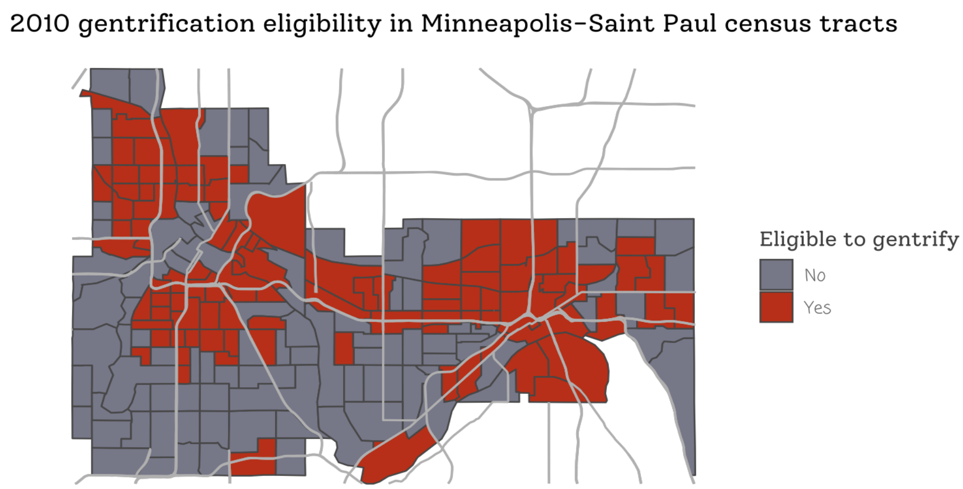

Figure 2: Census tract eligibility to undergo gentrification in 2010. The majority of census tracts eligible to gentrify are located in the northern half of Minneapolis and are clustered around the downtown area in Saint Paul.

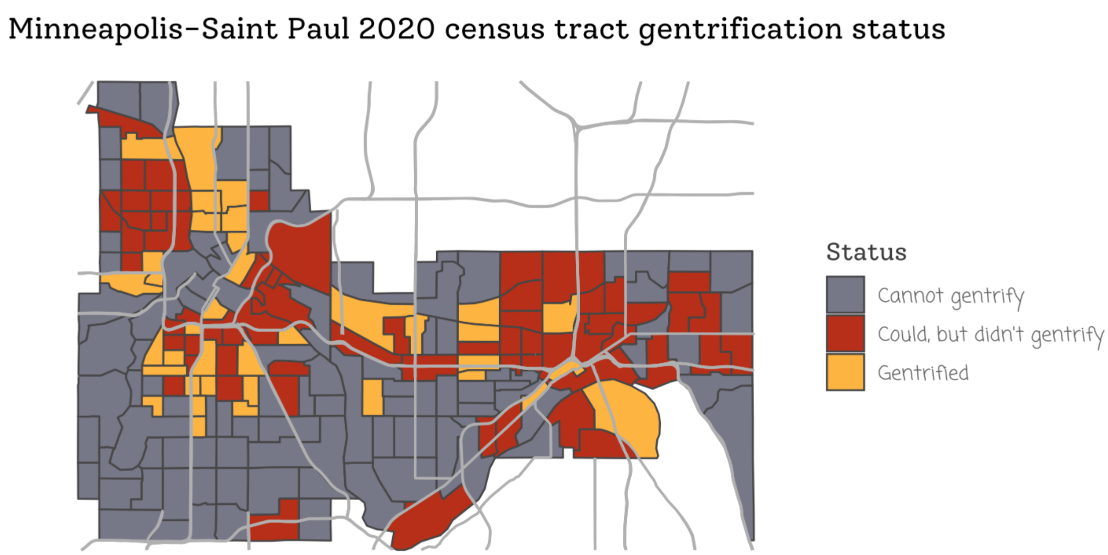

Figure 3: A map building off Figure 2 to identify census tracts that underwent gentrification over the period 2010-2020.

4.3 Dependent Variables: Violent Crime and Theft

4.3.1 Data

We are interested in understanding gentrification's relationship with both violent crime and theft in the Twin Cities. We acquire crime data for Minneapolis from the public data portal *Open Minneapolis*, which has recorded police incidents in the city limits dating back to 2010. Similarly, we obtain Saint Paul crime data from *Open Information Saint Paul*, which has incident level data from the Saint Paul Police Department dating back to August 14, 2014. While we would have preferred to include additional years in our study, the lack of publicly available incident level crime data is the primary reason that leads us to center our study around the period 2010-2020. To see how we classify each incident as violent crime or theft, please refer to Appendix 8.1.

With the crimes on the incident level, we aggregate them to the census tract annual level. *Open Minneapolis* provides an approximate latitude and longitude of each incident, making the aggregation process straightforward. On the other hand, *Open Information Saint Paul* provides an approximate address or intersection of the incident. Using string manipulation techniques, we are able to yield an approximate latitude and longitude for 88.35% of the crimes occurring in Saint Paul. However, it is likely that incidents missing coordinates are not missing at random as numbered streets (typically downtown) are especially difficult for the geocoder. Thus, we do not feel entirely comfortable drawing strong conclusions from the model results using the Saint Paul data and therefore place the Saint Paul results in Appendix 8.2. The complete data cleaning process for Saint Paul will be described in Appendix 8.2 as well[9].

It is important to note that even for the Minneapolis data, we are not entirely confident about the process in which the data is recorded. While unlikely, it is possible that the Minneapolis Police Department recorded each crime at an address or intersection and then passed them through a geocoder to get approximate coordinates. In such a case, it is unlikely all crimes would be recorded in the final dataset and we may have Minneapolis data similar to the geocoded Saint Paul data. Moving forward, however, we assume the data represent the true crime patterns in Minneapolis.

[9] Note that we break the rule of placing the appendix after the references. This is because we feel it contains some pretty neat stuff, and as a result we want it to feel like a fully included part of the paper!

5 Modeling Framework

5.1 Model Specification

We utilize Poisson regression to model the relationship between gentrification and crime in Minneapolis. Given that our outcome - the number of violent crimes or thefts per census tract in 2020 - is a count, a Poisson generalized linear model is an appropriate choice. Figure 4 displays the right skewed nature of our data, which the Poisson distribution is suited for modeling.

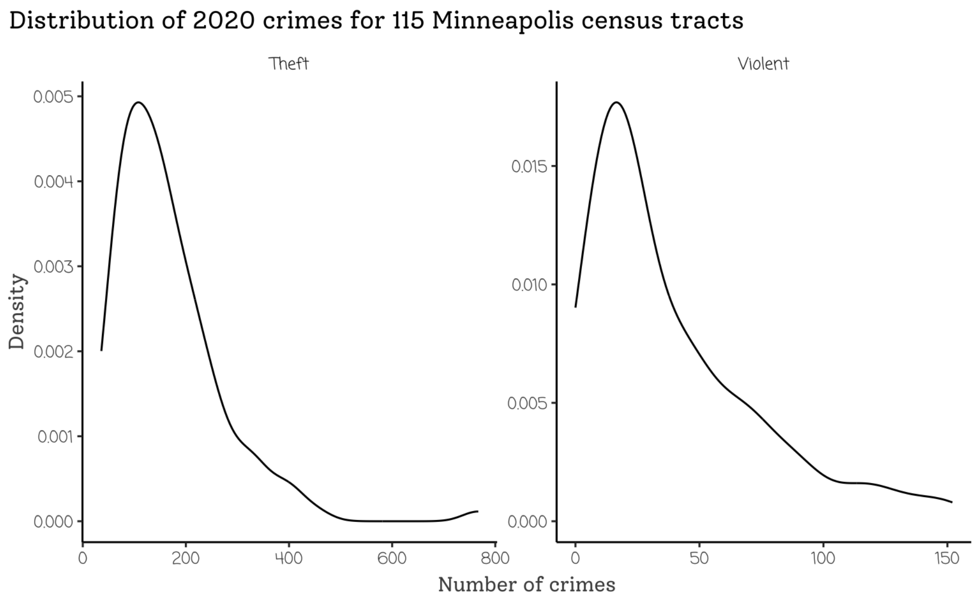

Figure 4: The distribution of both violent crimes and thefts by census tract in Minneapolis is right skewed, making the Poisson a good modeling choice.

The predictor variables of our Poisson regression include:

- Tract 2020 population size [10]
- Tract number of violent crimes or thefts in 2010
- Tract gentrification status: our independent variable of primary interest. This categorical variable takes one of three levels: Cannot gentrify; Could, but didn't gentrify; and Gentrified.

The population size of a census tract is positively correlated with both the number of violent crimes and thefts committed in that tract, and thus is important to include in our model. We do this through an offset, including the log of a census tract's 2020 population size on the right side of our model. The rationale for an offset is as follows. If we think of λ as the mean number of violent crimes per census tract in 2020, then we

[10] Technically, this variable is not necessarily an independent variables as we include it as an offset.

can make this count comparable across tracts of different population sizes by converting them to rates, or per capita quantities, i.e. $\frac{\lambda}{Population2020}$. The concept of adjusting the count of violent crimes by population is equivalent to adding $log(Population2020)$ to our Poisson regression as a so-called "offset" term, as illustrated in Equation (1).

$$\begin{aligned}
log(\frac{\lambda}{\text{Population2020}}) &= \beta_0 + \beta_1(\text{ViolentCrime2010}) + \beta_2(\text{CannotGentrify}) + \beta_3(\text{Gentrified}) \\
log(\lambda) - log(\text{Population2020}) &= \beta_0 + \beta_1(\text{ViolentCrime2010}) + \beta_2(\text{CannotGentrify}) + \beta_3(\text{Gentrified}) \\
log(\lambda) = \beta_0 + \beta_1(\text{ViolentCrime2010}) &+ \beta_2(\text{CannotGentrify}) + \beta_3(\text{Gentrified}) + log(\text{Population2020})
\end{aligned} \tag{1}$$

With the offset, we essentially add a predictor with a fixed coefficient of 1 to our model and gain the ability to interpret our coefficients as acting on crime rates, as opposed to raw counts. Specifically, each exponentiated coefficient represents the multiplicative change in a crime rate for a 1-unit increase in a predictor, assuming that the other predictors remain constant. Alternatively, we could have included population size as a traditional predictor in the model, but we chose to pursue the route of including an offset given its advantage of allowing us to interpret our outcome as a rate.

Aside from population, the other control variable in our model is a tract's prior (2010) crime count. By controlling for initial between-tract crime differences we get a more conservative estimate of gentrification's relationship with crime. It may be that revitalization largely occurred in areas with above average crime rates and that prior crime, not gentrification, is a significant predictor of crime change [Kreager et al., 2011].

To summarize, under the Poisson distribution we specify our two models as:

$$\begin{aligned}
\text{Model 1: } log(Violent2020) &\sim \beta_0 + \beta_1(\text{Violent2010}) + \beta_2(\text{CannotGentrify}) \\
&+ \beta_3(\text{Gentrified}) + \text{offset}(log(\text{Population2020})) \\
\text{Model 2: } log(Theft2020) &\sim \beta_0 + \beta_1(\text{Theft2010}) + \beta_2(\text{CannotGentrify}) \\
&+ \beta_3(\text{Gentrified}) + \text{offset}(log(\text{Population2020}))
\end{aligned}$$

The reference category for gentrification status is *Could, but didn't gentrify.* This allows us to compare gentrifying census tracts to those with similar demographic characteristics and spatial locations within the city that did not gentrify.

5.2 Assumptions

A Poisson model makes the following four assumptions:

1. The response variable is a count per unit of space or time.

2. The log of the mean rate, $log(\lambda)$, is a linear function of independent variables.

3. The mean of the response variable is equal to its variance.

4. Observations are independent from one another.

Assumptions 3 and 4 deserve some further discussion. Assumption 3 is centered around the idea that the mean of the number of violent and theft crimes in 2020 is equal to its variance. It is not uncommon for this assumption to be violated in practice, a phenomenon referred to as overdispersion. We test for overdispersion using the R function *check_overdispersion()* of the performance package, which returns an approximate estimate of an overdispersion parameter for a generalized linear model based on code from Gelman and Hill (2007) [Lüdecke et al., 2021, Gelman and Hill, 2009]. This test returns an overdispersion ratio significantly greater than 1, leading us to believe Assumption 3 is tenuous. To account for this, we will fit a Quasi-Poisson in addition to our regular Poisson generalized linear model. Quasi-Poisson models are used when there is more variation in the response than the model assumes. Not correcting for this overdispersion can yield artificially small standard errors, which can create artificially small p-values for model coefficients [Roback and Legler, 2021]. The Quasi-Poisson model uses an estimated dispersion factor $\hat{\phi}$ to inflate standard errors, $\hat{\phi} = \frac{\sum(\text{Pearson residuals})^2}{n-p}$, where $n =$ the number of observations and $p =$ the number of estimated parameters. Standard errors in the Quasi-Poisson model are represented by $SE_Q(\hat{\beta}) = \sqrt{\hat{\phi}} * SE(\hat{\beta})$, where $SE(\hat{\beta})$ represents the standard error from the traditional Poisson model.

The fourth assumption is also dubious. To meet this assumption, the crime in one census tract must be unrelated to the number of crimes in surrounding tracts. It seems plausible that the number of crimes in neighboring tracts are correlated, as displayed by the seeming clusters of higher and lower crime neighborhoods in Figure 5.

A violation of the independence assumption will result in our Poisson model systematically over- or under-predicting crime rates in certain areas of the Twin Cities. Additionally, similar to the consequences of failing to correct for overdispersion, failing to account for correlation in crime between census tracts can lead to artificially small p-values. We will address these concerns with a spatial Poisson model.

5.3 A Spatial Poisson Model

The first step of estimating a spatial model is to specify a correlation structure, or in other words, to define what is meant by the neighbors of a census tract. Figure 6 presents a few different options for the definition of "neighbor."

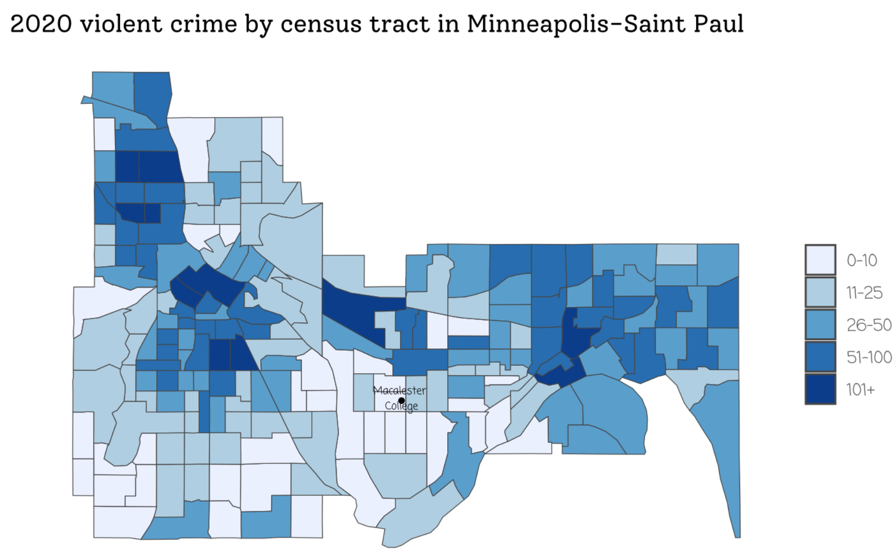

Figure 5: A choropleth map of Minneapolis and Saint Paul depicting the 2020 violent crime count by census tract.

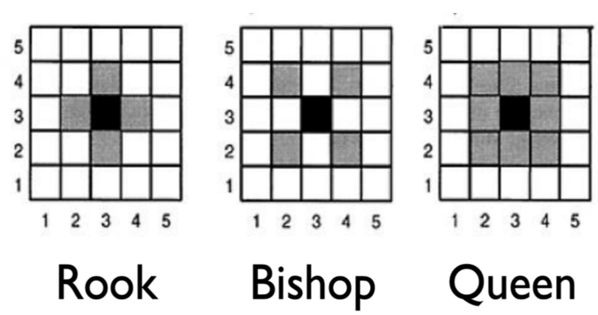

Figure 6: Three types of neighborhood structure: Rook, Bishop, and Queen [Heggeseth, 2022].

We can specify a census tract is neighbors with another if the two tracts share an edge (Rook), if they share a single point (Bishop), or if they touch in any matter (Queen). There is also the option to use K Nearest Neighbors, denoting a neighborhood to be formed from the K nearest polygons where distance is based on the centroid of each tract. For our analysis, we use the Queen structure. This structure is codified in our model using a spatial proximity, or weighting matrix, W. W is an $n \times n$ matrix with values w_{ij} of either 0 or 1 that reflect whether or not the ith area is a neighbor of the jth area [Heggeseth, 2022].

To fit our spatial Poisson model we use the `CARBayes` R package, which is suited for spatial areal unit modelling with conditional autoregressive priors [Lee, 2013]. More specifically, we use the `S.CARdissimilarity()` function to fit a spatial generalised linear mixed model to areal unit data with a Poisson response variable. It includes inference in a Bayesian setting using Markov chain Monte Carlo (MCMC) simulation. As a whole, this function allows us to fit our original Poisson model with a few extra parameters:

- The spatial proximity matrix W.
- A dissimilarity matrix, Z. This required parameter allows users to specify when neighboring polygons might be more different than expected, often (but not always) because of some physical barrier such as a river or highway. We do not feel there to be any significantly notable physical barriers or policies in Minneapolis that may impact crime patterns, and thus elect to fill this matrix with small (0.01) values in the off diagonal elements and zeros along the diagonal.
- The number of MCMC samples to generate. We choose to generate 100,000 samples.
- The number of MCMC samples to discard as the burn-in period. We choose to discard 30,000 samples.
- The level of thinning to apply to the MCMC samples to reduce their temporal autocorrelation. We choose level 20.

6 Results

6.1 Crime: Spatial and Temporal Trends

Over the past decade, Minneapolis has seen an increase in crime. While between 19,338 and 22,068 incidents were logged annually by the Minneapolis Police Department between 2010 and 2018, this number has hovered between 22,733 and 26,511 from 2019-2022. As displayed in Figure 7, we are seeing this increase largely as a result of a spike in auto theft, a trend that is not unique to Minneapolis. According to a study by the Council on Criminal Justice, motor vehicle thefts across 30 major U.S. cities have increased by 59% from 2019 to 2022 [Sganga, 2023].

While we feel it is important to give context about the increase in auto theft rates, the remainder of our analysis will classify all auto theft incidents in the *theft* category. Figure 8 shows census tracts of high theft are generally located in central Minneapolis and around the University of Minnesota.

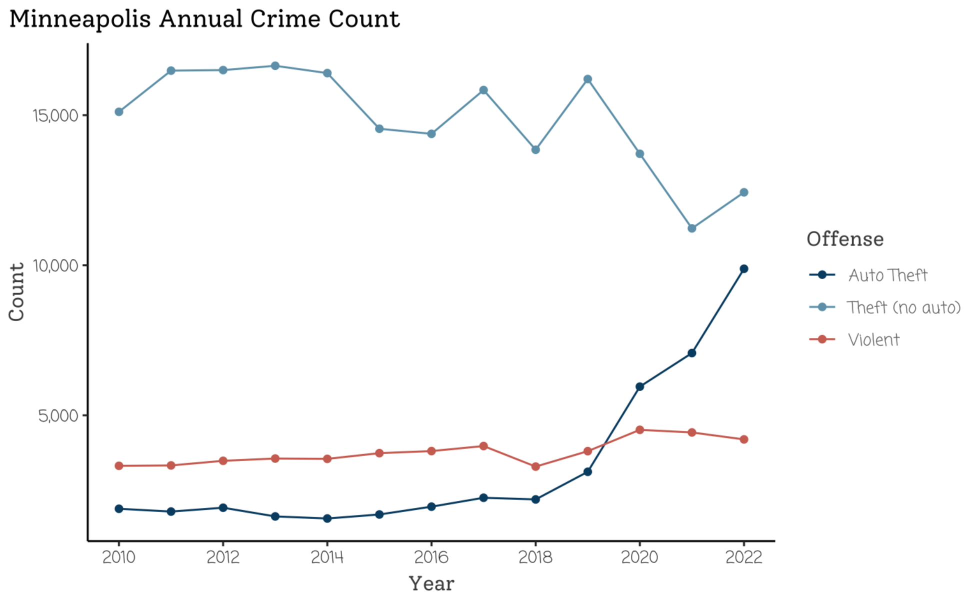

Figure 7: Auto theft in Minneapolis has been dramatically on the rise since 2019, while violent crime rates have slightly increased over the last decade and theft has declined. This graph classifies crimes listed as auto theft, theft of motor vehicle parts, carjacking, and motor vehicle theft as *Auto Theft*.

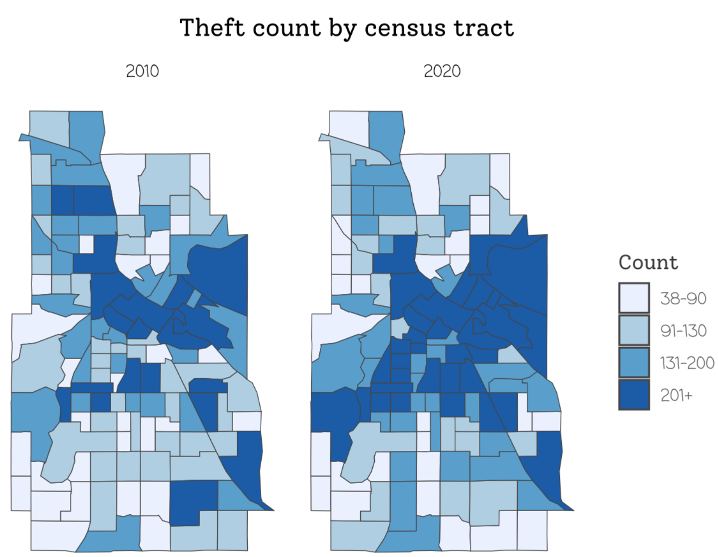

Figure 8: A choropleth map of the count of thefts in Minneapolis in 2010 and 2020.

On the other hand, high areas of violent crime are located in central and northwest Minneapolis. As seen in Figure 9, it is clear that areas that observed high levels of violent crime in 2010 also saw high levels in 2020, and areas observing low levels of violent crime in 2010 similarly observed low levels in 2020.

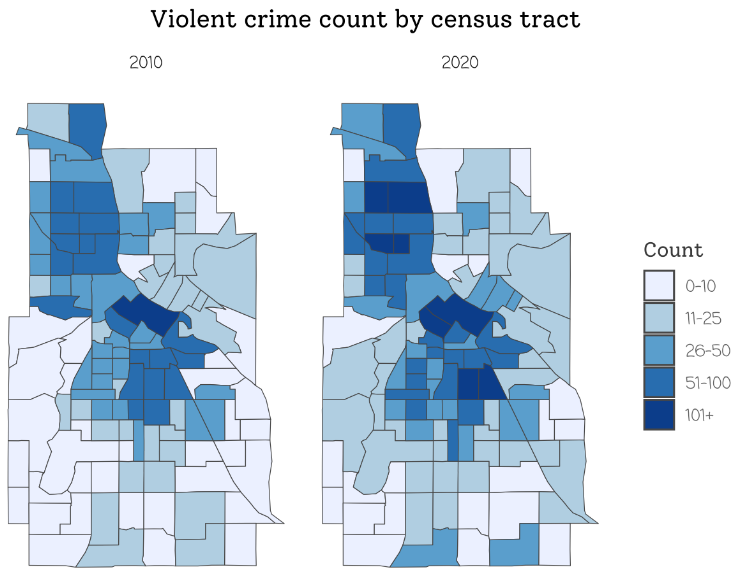

Figure 9: A choropleth map of the count of violent crimes in Minneapolis in 2010 and 2020.

Connecting theft back to gentrification, Table 1 displays descriptive statistics for tracts ineligible to undergo gentrification, tracts that were eligible but did not gentrify, and census tracts undergoing gentrification for the years 2010 and 2020.

Variable	Cannot gentrify	Could, but didn't gentrify	Gentrified
Violent Crimes/1000 residents	5.4	13.4	11.5
	6	16.1	14.6
Thefts/1000 residents	49.3	54.7	46.9
	45.2	57.7	57.4
Population	3,360	3,461	2,954
	3,751	3,978	3,176
Bachelor's %	35	16	18
	39	17	29
Median Income($)	67,409	28,531	33,493
	93,704	41,773	53,371
Median Home Value ($)	284,316	202,200	203,519
	326,643	213,995	250,008
Median Contract Rent ($)	875	681	668
	1,185	901	941

Table 1: Statistics for census tracts based on gentrification status, computed using a weighted average on tract population. For each variable, the 2010 metric is listed first, followed by 2020.

There are few key findings from Table 1. First, in both 2010 and 2020 tracts ineligible to undergo gentrification have a significantly lower average rate of violent crimes per 1000 residents than those eligible for gentrification in 2010. Second, across the three gentrification statuses we notice that the census tracts that underwent gentrification over the course of our study have the lowest rate of theft per 1000 residents in 2010. These tracts have a comparable proportion of residents with a bachelor's degree, mean rent, and mean home value as tracts that did not complete the gentrification process, thus we wonder if educated people viewed these neighborhoods as "safer" and were drawn to them as a result. Interestingly, by 2020 theft rates between gentrifying census tracts and tracts that did not complete the gentrification process are essentially identical.

6.2 Violent Crime Modeling

Table 2 presents three models of 2020 violent crime counts in Minneapolis census tracts, with Model 1 being our original Poisson model, Model 2 the Quasi-Poisson model, and Model 3 the spatial Poisson.

	Poisson	Quasi-Poisson	Spatial-Poisson
	(1)	(2)	(3)
Cannot Gentrify	0.417^{***}	0.417^{***}	0.624^{***}
	(0.389, 0.447)	(0.3, 0.571)	(0.495, 0.795)
Gentrified	0.916^{*}	0.916	0.921
	(0.85, 0.986)	(0.648, 1.292)	(0.716, 1.225)
Violent Crime 2010	1.0095^{***}	1.0095^{***}	1.0096^{***}
	(1.009, 1.01)	(1.0066, 1.012)	(1.006, 1.013)
Note:		*p<0.05; **p<0.01; ***p<0.001	

Table 2: Our three Minneapolis violent crime models, each having the dependent variable being the 2020 violent crime count in each of 115 census tracts. Coefficients and respective 95% confidence intervals are exponentiated.

Beginning with our Poisson model (Model 1), we see from the *Cannot Gentrify* coefficient that tracts ineligible to undergo gentrification have a statistically significant lower violent crime rate than census tracts that were eligible to but did not gentrify (our reference category). This coefficient represents the multiplicative scaling of crime *rate* (due to our offset of population size) of tracts ineligible to gentrify compared to the reference level, holding prior crime constant. Our *Gentrified* coefficient is also negative, though borderline significant ($p = 0.0199$). This provides evidence that gentrifying census tracts see a drop in violent crime compared to similar tracts that do not undergo gentrification. Finally, our coefficient on prior violent crime is significant and positive, indicating past violent crime is a strong indicator of future violent crime holding constant gentrification status.

As discussed in Section 5.2, there is evidence of overdispersion in our original Poisson model. Thus, a Quasi-Poisson can provide more conservative estimates by inflating coefficient standard errors. Results from Model 2 continue to show significance for *Cannot Gentrify* and *Violent Crime 2010* coefficients. On the other hand, the *Gentrified* coefficient is no longer significant, indicating that there is not evidence that census tracts undergoing gentrification over the period of our study see a meaningful reduction in violent crime compared to tracts that don't, at least at this point of the study (2020).

Models 1 and 2 treat violent crime rates in neighboring census tracts as independent of one another. A quick look at Model 1's residual plot (left map of Figure 10) reiterates there is correlation in the violent crime rates between neighboring census tracts. There are seeming clusters of tracts where the model is over- or under-predicting violent crime counts. The presence of spatial autocorrelation is confirmed with a Moran's I test (p-value = 1.922×10^{-6}).

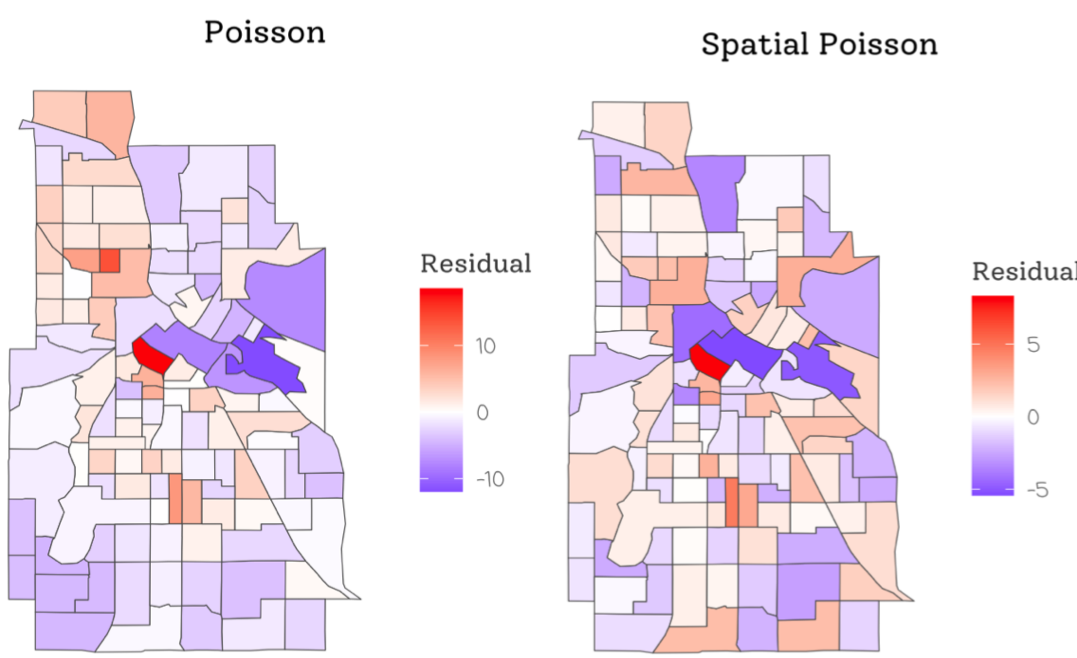

Figure 10: Residuals for our original Poisson model (Model 1) versus spatial Poisson model (Model 3). Units are in number of violent crimes. Positive residuals (red) indicate higher observed counts than predicted, while negative residuals (blue) indicate lower observed counts than predicted.

The differing scales of the legends of Figure 10 indicate the size of our residuals decreases on average with the spatial model. Slight spatial autocorrelation may still be present in the spatial Poisson model as the Moran's I p-value is 0.013, though it has been greatly reduced from the original and Quasi-Poisson model. Overarching conclusions remain the same as the Quasi-Poisson model. There is no evidence that census tracts undergoing gentrification over the course of decade see a meaningful drop in violent crime by 2020, accounting for 2010 crime rate. Violent crime in 2010 continues to be a strong indication of violent crime in 2020, and census tracts ineligible to gentrify have a statistically significant lower violent crime rate than census tracts that were eligible to but did not gentrify.

6.3 Theft Modeling

Table 3 presents our three models for 2020 theft counts in Minneapolis census tracts. Again, Model 1 represents our original Poisson model, Model 2 the Quasi-Poisson model, and Model 3 the spatial Poisson model.

	Poisson	Quasi-Poisson	Spatial-Poisson
	(1)	(2)	(3)
Cannot Gentrify	0.79***	0.79**	0.993
	(0.764, 0.815)	(0.663, 0.942)	(0.842, 1.143)
Gentrified	1.1***	1.1	0.992
	(1.056, 1.142)	(0.887, 1.355)	(0.832, 1.183)
Theft 2010	1.0013***	1.0013***	1.0011***
	(1.0012, 1.0014)	(1.0009, 1.0017)	(1.0006, 1.0016)
Note:		*p<0.05; **p<0.01; ***p<0.001	

Table 3: Our three Minneapolis theft models, each having the dependent variable being the 2020 theft count in each of 115 census tracts. Coefficients and respective 95% confidence intervals are exponentiated.

Most intriguing from Model 1 is the *positive* significant *Gentrified* coefficient. This coefficient represents the theft rate in a gentrifying tract is estimated to be approximately 1.09 times what the theft rate would be in an "eligible-but-didn't" tract, assuming that they had the same prior crime value in 2010. Moving to the Quasi-Poisson model, this coefficient remains positive but loses its significance as a result of the inflated standard errors (p-value of 0.39). With an exponentiated *Gentrified* coefficient of 0.9932, our spatial Poisson model essentially shows there to be no relationship between a tract undergoing gentrification and theft rate when accounting prior theft rate.

Out of the three models, the spatial Poisson provides the most trustworthy coefficients given it accounts for dependence between observations. Similar to the original Poisson model for violent crime, as seen in Figure 11 the original Poisson model for theft has parts of the city where it is consistently over- or under-predicting theft. Using the spatial Poisson model, spatial autocorrelation is greatly reduced.

Across all three models, 2010 theft is a significant and positive predictor of 2020 theft. Unlike the violent

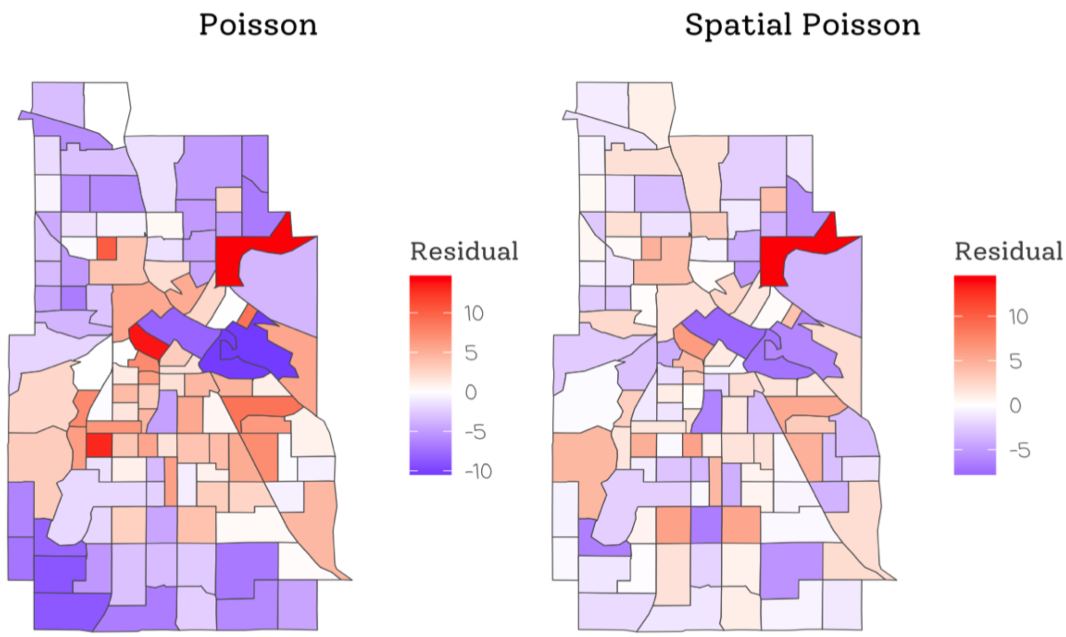

Figure 11: Residuals for our original Poisson theft model (Model 1) versus spatial Poisson theft model (Model 3). Units are in number of thefts. Positive residuals (red) indicate higher observed counts than predicted, while negative residuals (blue) indicate lower observed counts than predicted.

crime models, there is *not* consistent evidence that census tracts ineligible to undergo gentrification see a significantly lower 2020 theft rate than "eligible-but-didn't" tracts, accounting for 2010 theft. While the *Cannot Gentrify* coefficient is significant and negative in our original Poisson model, it is less so in the Quasi-Poisson model and essentially 0 in our spatial Poisson model.

7 Discussion and Conclusion

7.1 Key Takeaways

This analysis contributes to the current literature by introducing a new geographic area of the study (the Twin Cities) and placing emphasis on understanding gentrification's relationship with theft in addition to violent crime. We measure gentrification over the 2010-2020 decade with census data, first using tract median family income to identify tracts eligible to undergo gentrification and then using college education, home value, and rent to determine what tracts completed the process. Using Poisson, Quasi-Poisson, and spatial-Poisson models, we come to understand that gentrification's relationship with violent crime and theft in the Twin Cities is murky. Unlike findings from Barton (2016) and Papachristos et al. (2011), we do not find gentrification to be related to significant drops in violent crime in the Twin Cities. We also see no evidence of gentrification being associated with reductions in theft. Thus, our results most closely match those of Kreager et al. (2011), whose *gentrification-as-outcome* analyses did not find gentrification to be associated with increases or decreases in total, violent, or property crime when including prior crime rate as a model predictor.

It is plausible that gentrifying neighborhoods are still experiencing a period of residential turnover and social adjustment, and that checking crime rates in a few years (e.g. 2025) will show significant reductions in crime - violent crime specifically. This would align with the social disorganization theory, which advocates for a curvilinear relationship between gentrification and crime rate. Additionally, given that neighborhoods ineligible to gentrify see significantly lower violent crime rates than "eligible-but-didn't" tracts across all models, this would make intuitive sense as gentrified tracts become more similar to that "ineligible-to-gentrify" category as their median incomes rise. We start to investigate this possible phenomenon by running our models on 2022 violent crime and theft counts in Minneapolis. Results are comparable to the 2020 models. Additionally, in case COVID-19 has impacted crime patterns in Minneapolis, we run our models with 2019 data. Again, we reach the same conclusions as with 2020 data. Given these findings, we currently do not believe there is substantial evidence to use reductions in crime to advocate for gentrification, at least in Minneapolis.

Our violent crime results may differ from the conclusions reached by Papachristos et al. (2011) and Barton (2016) for a wide variety of reasons, the first being location. Minneapolis is a much smaller city than both Chicago and New York City, both in square mileage and in population, which can create different cultural dynamics. Second, each study defines gentrification in a different way - Papachristos et. al use coffee shops while Barton uses a different set of census measures than our analysis. Third, these studies all take place at different times. Barton's study takes place over the 29 year period 1980-2009, while Papachristos et. al set their study between 1991-2005. As discussed above, we may need to wait several more years to see gentrification's implications on crime fully play out.

7.2 Limitations and Future Steps

Gentrification is a complex process that can play out in many different ways, and thus our first limitation is that our use of census indicators to measure gentrification captures some but certainly not all gentrification processes occurring in the Twin Cities. We miss out on aspects including residential mobility, demolition of housing, city planning, and signage for and presence of new businesses. Our second major limitation comes as a result of the crime data we have available. The three studies reviewed have lengths of 14 years, 19 years, and 29 years. Unfortunately, we only have access to incident level crime data dating back to 2010 in

Minneapolis and August of 2014 in Saint Paul, resulting in a shortened period of analysis. Additionally, the quality of the Saint Paul data is lacking to a degree that prevents us from feeling comfortable drawing strong conclusions from it. Third, no model is without limitations. While we were able to account for overdispersion of our data with the Quasi-Poisson model and spatial autocorrelation with the spatial Poisson model, we did not account for both of these assumptions together in one final model.

Future research could be beneficial by seeking to understand *why* gentrifying census tracts see changes in violent crime and/or theft overtime. For tracts that see increased levels of crime, is it a result of increased social tension and what steps can be taken to improve neighborhood dynamics? Additionally, could an increase in crime be due an increased willingness of residents to call the police and fewer instances of crime going unreported? For tracts seeing decreased levels of crime, it would be interesting to understand if crime is displaced to census tracts.

7.3 Greater Impacts of Gentrification in the Twin Cities

Gentrification is a highly controversial phenomenon that can have serious consequences for neighborhood residents. Before concluding our paper, we would like to highlight some of the greater impacts of gentrification felt in the Twin Cities. To do so, we rely heavily on the study *The Diversity of Gentrification: Multiple Forms of Gentrification in Minneapolis and Saint Paul* (2019), which combines a statistical analysis of neighborhood-level data with an in-depth qualitative analysis of interviews with public officials, community leaders, and neighborhood residents [Goetz et al., 2019]. Over the period 2000-2015, this study identifies gentrifying neighborhoods that are fairly similar in location to our study. These neighborhoods are displayed in Figure 12.

First, gentrification is often associated with displacement of residents from their homes as a result of unsustainable housing costs. Almost all renters interviewed in the Hamline-Midway neighborhood of Saint Paul expressed increasing housing costs have continued to push them further and further east, away from Highway 280 toward the east side of St. Paul [Goetz et al., 2019]. However, displacement does not only have to be physical - it can also be political and cultural as residents that once had a strong say in the neighborhood no longer feel their voices are heard by local neighborhood officials and community organizations that once spoke on their behalf [Hyra, 2017, Goetz et al., 2019]. For example, residents interviewed in a group of census tracts in North Minneapolis collectively expressed a feeling of powerlessness, a notion that change was happening to them rather than with them. These feelings were especially strong for minority residents.

Goetz et al. find additional consequences of gentrification in the Twin Cities. Gentrification in the Arts

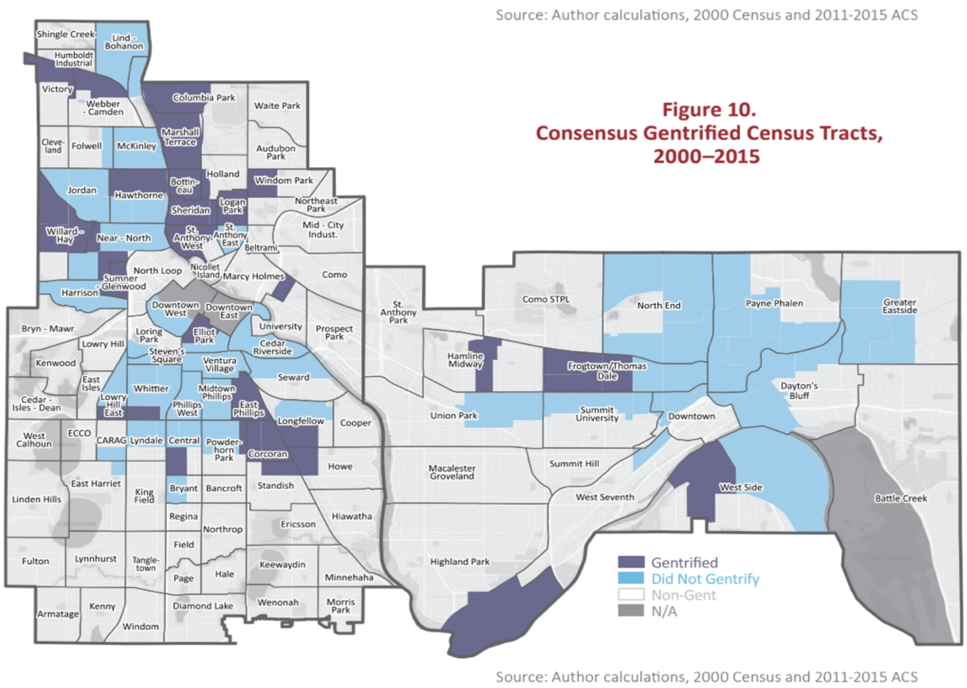

Figure 12: Census tract gentrification status in 2015 according to the study *The Diversity of Gentrification: Multiple Forms of Gentrification in Minneapolis and Saint Paul* [Goetz et al., 2019]. Gentrification was measured by combining methods of Freeman (2005), Bates (2013), and Ding et al. (2016).

District in Northeast Minneapolis brought in "makers" who create art in bulk at cheaper prices to meet ever-increasing demand. Long-term resident-artists in the area, most of whom identified as "raw" artists (produce fewer pieces annually at higher prices), felt replaced and struggled financially. Residents living just south of downtown Minneapolis struggled with the concept they called "Uptowning", which represents an influx of young white families bringing new businesses that do not cater to the needs of the long-time neighborhood residents. Those interviewed in the Thomas-Dale neighborhood of Saint Paul expressed frustration about the new affordable housing developments, which are offered at a level beyond their ability to pay [Goetz et al., 2019].

There are several local organizations in the Twin Cities with a mission to fight these harms of gentrification and others. Common themes of these organizations include empowering long-term residents to have a say in the neighborhood change happening around them, redirecting resources toward long term affordable housing, and shifting the discourse about low-wealth communities and their residents. Goals of these organizations range from preventing gentrification before it happens to assisting residents struggling with cultural displacement coming as a result of the process [Goetz et al., 2019].

It is clear that consequences of gentrification can be widespread. While gentrification's relationship in the Twin Cities with crime is murky, the process does bring funding and development to communities previously experiencing periods of financial neglect. Bringing investment and development to neighborhoods in a manner that empowers long-time community residents and preserves neighborhood culture can reduce the harms of gentrification. With less social tension, neighborhoods that see development in this manner may even see a meaningful decline in crime.

8 Appendix

8.1 Crime Classification

We classify the following offenses as **violent crimes in Minneapolis**: Assault in the 1st-4th degrees, murder, domestic assault in the 1st-3rd degrees, domestic assault - strangulation, robbery of person, robbery of business, aggravated robbery, adulteration/poison, arson, rape.

We classify the following offenses as **theft in Minneapolis**: Theft, theft from motor vehicle, burglary of a dwelling, bike theft, theft of motor vehicle part, theft by swindle, shoplifting, theft from person, theft from building, carjacking, motor vehicle theft, theft/coinop device, pocket-picking, petty theft, package theft, looting.

We classify the following offenses as **violent crimes in Saint Paul**: Simple domestic assault, aggravated assault, aggravated domestic assault, homicide, arson, rape, robbery.

We classify the following offenses as **theft in Saint Paul**: Theft, burglary, vandalism, auto theft, criminal damage, graffiti.

8.2 Saint Paul Analysis

A major limitation to this analysis is the fact that we are unable to render coordinates for 12% of the recorded Saint Paul incidents. Given that these 21,336 incidents over the period 2015-2022 may not be missing at random (causing certain areas of Saint Paul to have lower crime rates than reality), we decide to place our results and discussion of the relationship between gentrification and crime in Saint Paul here in Appendix 8.2.

8.2.1 Data Cleaning

Table 4 displays a small subset of the Saint Paul incident data.

	DATE	INCIDENT	BLOCK
1	2014/08/14 12:10:00+00	Theft	160X OLDHUDSON RD
2	2014/08/14 13:27:00+00	Theft	ASHLAND AV & CLEVELAND
3	2021/08/16 03:30:00+00	Theft	SELBY AVE AND SNELLING
4	2014/09/25 15:00:00+00	Theft	106X BANDANA BD E
5	2015/07/25 02:06:00+00	Theft	7XX 6 ST E

Table 4: A small subset of Saint Paul crime incident data.

Unlike the Minneapolis incident data, an approximate latitude and longitude of the incident are not recorded. Thus, we use *tidygeocoder*, an R package that can convert from addresses to coordinates and vice-versa [Cambon et al., 2021], to approximate a latitude and longitude for each indcident. Before running the data through the geocoder, however, there are several issues with this data that we identify and fix.

- From Table 4, first notice that final number or two of street numbers are recorded with an X for privacy reasons. In certain instances, the *geocode()* function can produce approximate coordinates with these Xs included, but in the majority instances we test it cannot. As a result, we replace all Xs with a 0. While it is likely this process creates a viable address in the correct census tract, it is not guaranteed, especially for street numbers containing multiple Xs such as Entry 5.

- Second, note that multiple word streets are recorded as one word. For example, the street name of Entry 1 is recorded as *OLDHUDSON* instead of *OLD HUDSON*. The geocoder does not recognize this as a street. We correct this issue by extracting all unique street names from the data, manually identifying streets that should be two words, and using string manipulation techniques to insert a space where appropriate for each instance of the street in the data.

- Carefully inspecting Table 4 allows one to see that the street identifiers (avenue, street, road, etc) of incidents are not consistently recorded one way. For example, *avenue* is recorded as *AV* in Entry 2 and *AVE* in Entry 3. Interestingly, in certain cases (but not always) the geocoder will only take *AVE*. Avenue is not the only identifier that can be problematic. The geocoder will provide an approximate address for *140 MOUNDS BLVD* but not *140 MOUNDS BD*. Similarly, the geocoder accepts *PKWY* but not PA. Understanding what the geocoder does and does not accept is the difficult part; replacing these mistaken identifiers with the appropriate version is a simple process.

- Entry 1 of Table 4 is missing a street direction of E, W, N, or S. For several thousand incidents, this prevents coordinates from being geocoded. For these crimes, we use string manipulation to insert a direction following the street number and rerun them through the geocoder, using trial and error until the correct direction is specified.

- The geocoder especially struggles with numbered streets, most of which are located downtown. Entry 5 of Table 4 will not be recognized by the geocoder for two reasons. First, *6 ST* should be recorded as *6TH ST*. Second, the street direction is located at the end of the string. While certain addresses in Saint Paul list direction after the street identifier (Entry 4), other addresses, both downtown and throughout Saint Paul, require the direction to be placed after the street number (Entry 5). Using string manipulation techniques, we are able to capture coordinates of over 12,000 more incidents by reformatting them.

- Lastly, not all incidents are addresses - 10.2% are recorded as an intersection (e.g. *ASHLAND AV CLEVELAND*). The geocoder cannot recognize intersections, and as a result we run them in ArcGIS Pro [11]. However, before doing this these intersections need cleaning:

  - The second street of the intersection is lacking an identifier such as AVE or PKWY as well as a street direction N, E, W, or S. ArcGIS cannot render coordinates for these intersections without this key information. As a result, we find the most common identifier and direction for a particular street and append this to the end of the intersection. Unfortunately, certain streets in Saint Paul have different street directions depending on where you are. Parts of Bandana Blvd are "E" and parts are "W". Similarly, parts of Snelling Ave are "N" and parts are "S". Thus, specifying the most common identifier and street direction will not always yield the correct one. This is likely the primary reason for a low success rate of identifying coordinates for intersections.

Using the tidygeocoder package and ArcGIS Pro, we are able to yield approximate coordinates for 40.73% of intersections and 93.76% of addresses for a total success rate of 88.35%. Missing data may not be missing at random, so as we proceed the results our Saint Paul analysis should be taken with reservations.

8.2.2 Crime: Spatial and Temporal Trends

Figure 13 displays the annual counts of theft, violent crime, and auto theft in Saint Paul over the period 2015-2022.

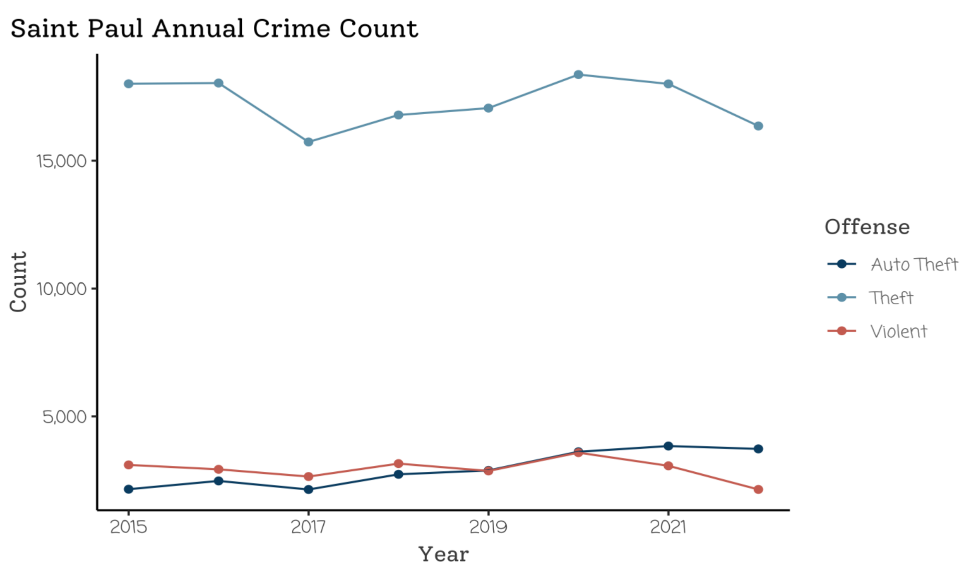

Figure 13: Line graph of annual counts of theft, violent crime, and auto theft in Saint Paul over the period 2015-2022.

[11] Special thanks to 2022 Macalester graduate Claire McHenry for running these addresses through ArcGIS Pro for us.

Annual counts of theft and violent crime appear to be fairly similar to Minneapolis, which is somewhat surprising given Saint Paul's 2020 population was 310,942 while Minneapolis had a population of 429,014. In recent years, however, theft counts have declined in Minneapolis while remained fairly constant in Saint Paul. What we do not see in Figure 13 that we see in Minneapolis (Figure 7) is a sharp increase in auto theft over the last four years. It is not clear if this is a result of how the data is recorded in Saint Paul (it is plausible that auto thefts could be consistently recorded as "theft") or if auto theft has truly skyrocketed in Minneapolis while just mildly increased in Saint Paul. Note that because this data is not spatial, this graph is an accurate depiction of the full data that has been provided by the Saint Paul Police.

Figure 14 displays theft counts by census tract in 2020 for Saint Paul. This map is on the same scale as Minneapolis. A large portion of the census tracts have 201+ thefts recorded in 2020. We are unsure if Saint Paul's theft counts are so high (given its smaller population) compared to Minneapolis as a result of different incident recording policies between the two cities or differences in criminal activity between the two cities.

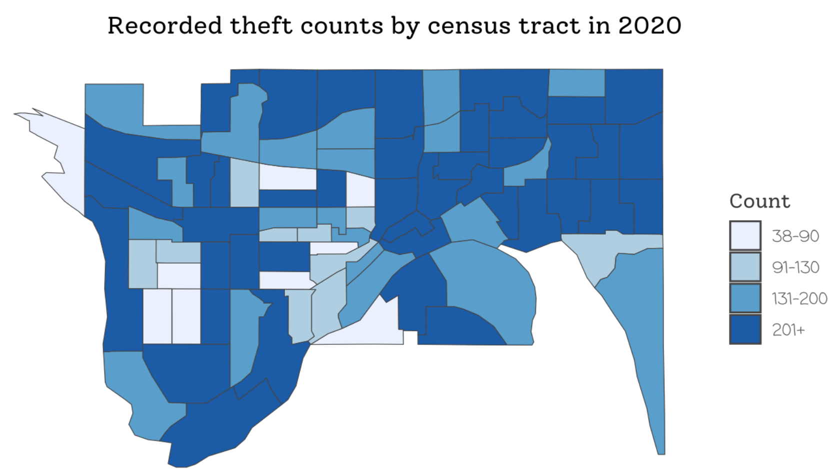

Figure 14: A choropleth map of the count of thefts in Saint Paul by census tract in 2020.

In a similar manner, Figure 15 displays violent crime counts by census tract in 2020. It is clear that counts of violent crime are much higher on the eastern half than the southwestern portion of the city.

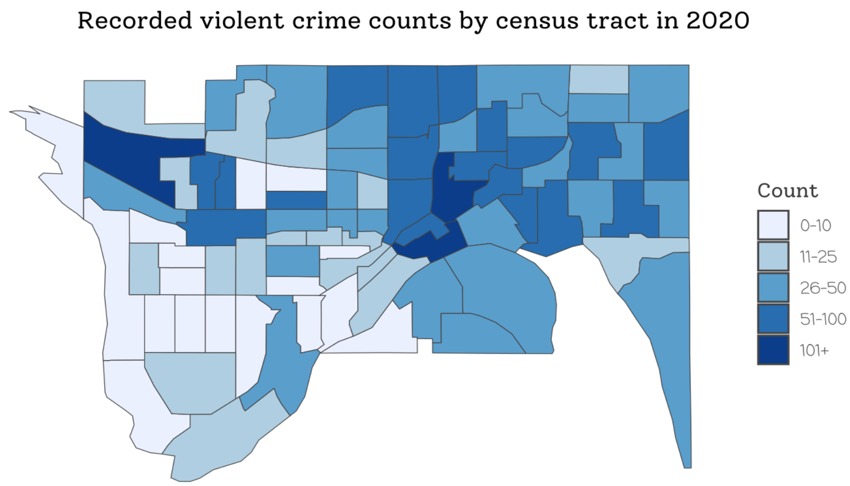

Figure 15: A choropleth map of the count of thefts in Saint Paul by census tract in 2020.

Table 5 introduces summary statistics for Saint Paul tracts that were in eligible to gentrify in 2010, were eligible to but did not gentrify, and underwent gentrification.

Variable	Cannot gentrify	Could, but didn't gentrify	Gentrified
Violent Crimes/1000 residents	5.2	12.6	14.9
	5.9	14.1	15.4
Thefts/1000 residents	41.9	69.4	129.2
	51.2	71.6	113.8
Population	3,780	3,746	2,729
	3,883	3,981	2,802
Bachelor's %	29	11	19
	31	13	25
Median Income ($)	61,379	34,359	36,934
	77,289	47,305	55,062
Median Home Value ($)	256,405	175,305	200,930
	284,649	172,867	228,746
Median Contract Rent ($)	748	648	635
	1,027	859	905

Table 5: Statistics for census tracts based on gentrification status, computed using a weighted average on tract population. For Violent Crime, Theft, and Population, the 2015 metric is listed first followed by 2020 metric. For the remaining census variables, the 2010 metric is shown first followed by 2020.

Similar to Minneapolis, census tracts ineligible to gentrify have noticeably lower theft and violent crime rates than census tracts eligible to gentrify. Unlike Minneapolis, census tracts that underwent gentrification have higher rates of violent crime and theft in both 2015 and 2020 than those that were eligible to gentrify but did not. However, theft rates do fall between 2015 and 2020 in tracts undergoing gentrification while rise in other tracts. Tracts that undergo gentrification also see a much higher percentage of residents with bachelor's degrees (19%) in 2010 than tracts that do not undergo gentrification (11%). This may say something about why particular areas gentrified. It would be very interesting to see what crime patterns were across tracts in 2010 and prior.

8.2.3 Violent Crime Modeling

Table 6 presents three models of 2020 violent crime counts in Saint Paul census tracts, with Model 1 being our original Poisson model, Model 2 the Quasi-Poisson model, and Model 3 the spatial Poisson.

	Poisson	Quasi-Poisson	Spatial-Poisson
	(1)	(2)	(3)
Cannot Gentrify	0.721***	0.721*	0.737*
	(0.655, 0.791)	(0.517, 0.998)	(0.5505, 0.972)
Gentrified	1.037	1.037	1.025
	(0.934, 1.145)	(0.71, 1.477)	(0.726, 1.43)
Violent Crime 2015	1.019***	1.019***	1.019***
	(1.017, 1.02)	(1.014, 1.024)	(1.014, 1.025)
Note:		*p<0.05; **p<0.01; ***p<0.001	

Table 6: Our three Saint Paul violent crime models, each having the dependent variable being the 2020 violent crime count in each of 82 census tracts. Coefficients and respective 95% confidence intervals are exponentiated.

Conclusions are broadly the same as those from the Minneapolis models. *Cannot Gentrify* is significant across all three models, implying that accounting for 2015 crime rate, tracts that were ineligible to gentrify in 2010 see significantly lower violent crime rates in 2020 than tracts that were eligible to gentrify but did not. Our significant *Violent Crime 2015* confirms that 2015 violent crime is a strong indicator of 2020 violent crime, holding gentrification status constant. Our coefficient of interest, *Gentrified*, is insignificant across all three models, implying that there is not evidence that tracts undergoing gentrification see meaningfully lower rates of violent crime than "eligible-but-didn't" tracts, accounting for 2015 violent crime rate.

Aside our missing data likely not being missing at random, another limitation to both the Saint Paul violent crime and theft models is that the first full year of recorded incidents that we have access to is 2015. It is reasonable to assume that using a year that is closer to our outcome year of 2020 makes this predictor more highly correlated with the outcome, and may result in decreased significance of our gentrification coefficients compared to if we use 2010 violent crime as a predictor.

8.2.4 Theft Modeling

Table 7 presents three models of 2020 theft counts in Saint Paul census tracts, with Model 1 being our original Poisson model, Model 2 the Quasi-Poisson model, and Model 3 the spatial Poisson.

	Poisson	Quasi-Poisson	Spatial-Poisson
	(1)	(2)	(3)
Cannot Gentrify	0.774***	0.774*	0.823
	(0.749, 0.799)	(0.613, 0.975)	(0.64, 1.051)
Gentrified	1.097***	1.097	0.919
	(1.046, 1.149)	(0.766,1.532)	(0.673, 1.235)
Theft 2015	1.001***	1.001***	1.0015***
	(1.001, 1.001)	(1.0008, 1.0013)	(1.0011, 1.0021)
Note:			*p<0.05; **p<0.01; ***p<0.001

Table 7: Our three Saint Paul theft models, each having the dependent variable being the 2020 theft count in each of 82 tracts. Coefficients and respective 95% confidence intervals are exponentiated.

Table 7 displays similar results to our Minneapolis theft models (Table 3). Our *Cannot Gentrify* coefficient is significant and negative in both the Poisson and Quasi-Poisson model, though becomes insignificant in the spatial-Poisson model. While in both Model 1 and 2 the *Gentrified* coefficient is positive, it becomes negative once spatial autocorrelation is accounted for (though remains insignificant). There are two broad takeaways from this set of models, both of which are very similar to the conclusions we draw from the Minneapolis analysis. First, as we saw in Minneapolis, prior theft in a census tract is a strong indicator of future theft, holding gentrification status constant. Second, there is not evidence that census tracts undergoing gentrification see meaningful reductions in theft rate compared to "eligible but didn't" tracts, at least over the period of this study.

References

[Barton, 2016] Barton, M. S. (2016). Gentrification and Violent Crime in New York City. *Crime amp; Delinquency*, 62(9):1180–1202.

[Bates, 2013] Bates, L. (2013). Gentrification and Displacement Study: Implementing an Equitable Inclusive Development Strategy in the Context of Gentrification.

[Bostic and Martin, 2003] Bostic, R. W. and Martin, R. W. (2003). Black Home-owners as a Gentrifying Force? Neighbourhood Dynamics in the Context of Minority Home-ownership. *Urban Studies*, 40(12):2427–2449.

[Bureau, 2022] Bureau, U. C. (2022). American Community Survey (ACS).

[Cambon et al., 2021] Cambon, J., Hernangómez, D., Belanger, C., and Possenriede, D. (2021). tidygeocoder: An R package for geocoding. *Journal of Open Source Software*, 6(65):3544. R package version 1.0.5.

[Covington and Taylor, 1989] Covington, J. and Taylor, R. B. (1989). Gentrification and Crime. *Urban Affairs Quarterly*, 25(1):142–172.

[Ding et al., 2016] Ding, L., Hwang, J., and Divringi, E. (2016). Gentrification and Residential Mobility in Philadelphia. *Regional Science and Urban Economics*, 61:38–51.

[Feder, 2020] Feder, S. (2020). Gentrification disproportionately affects minorities.

[Freeman, 2005] Freeman, L. (2005). Displacement or Succession?: Residential Mobility in Gentrifying Neighborhoods. *Urban Affairs Review*, 40(4):463–491.

[Gelman and Hill, 2009] Gelman, A. and Hill, J. (2009). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.

[Goetz et al., 2019] Goetz, E. G., Lewis, B., Damiano, A., and Calhoun, M. (2019). The Diversity of Gentrification: Multiple Forms of Gentrification in Minneapolis and St. Paul.

[Heggeseth, 2022] Heggeseth, B. (2022). Correlated Data Notes.

[Hyra, 2017] Hyra, D. S. (2017). Race, Class, and Politics in the Cappuccino City. *Chicago: University of Chicago Press*.

[Kirk and Laub, 2010] Kirk, D. and Laub, J. (2010). Neighborhood Change and Crime in the Modern Metropolis. *Crime and Justice*, 39(1):441–502.

[Kreager et al., 2011] Kreager, D. A., Lyons, C. J., and Hays, Z. R. (2011). Urban Revitalization and Seattle Crime, 19822000. *Social Problems*, 58(4):615–639.

[Lee, 2013] Lee, D. (2013). CARBayes: An R Package for Bayesian Spatial Modeling with Conditional Autoregressive Priors. *Journal of Statistical Software*, 55(13):1–24.

[Lees et al., 2008] Lees, L., Slater, T., and Wyly, E. (2008). Gentrification.

[Lüdecke et al., 2021] Lüdecke, D., Ben-Shachar, M. S., Patil, I., Waggoner, P., and Makowski, D. (2021). performance: An R Package for Assessment, Comparison and Testing of Statistical Models. *Journal of Open Source Software*, 6(60):3139.

[Maciag, 2015] Maciag, M. (2015). Gentrification in America Report.

[McDonald, 1986] McDonald, S. C. (1986). Does Gentrification Affect Crime Rates? *Crime and Justice*, 8:163–201.

[Mendez, 2022] Mendez, C. (2022). Baltimore officials walk through communities looking to upgrade areas with blighted, vacant homes.

[Metropolitan, 2021] Metropolitan (2021). Twin Cities Population is Growing and Diversifying.

[Papachristos et al., 2011] Papachristos, A. V., Smith, C. M., Scherer, M. L., and Fugiero, M. A. (2011). More Coffee, Less Crime? The Relationship between Gentrification and Neighborhood Crime Rates in Chicago, 1991 to 2005. *City amp; Community*, 10(3):215–240.

[Roback and Legler, 2021] Roback, P. and Legler, J. (2021). *Poisson Regression*. CRC Press.

[Sganga, 2023] Sganga, N. (2023). Auto thefts, carjackings in major U.S. cities spike, new report finds.

[Smith, 2012] Smith, C. M. (2012). The Influence of Gentrification on Gang Homicides in Chicago Neighborhoods, 1994 to 2005. *Crime amp; Delinquency*, 60(4):569–591.

[States Department of Housing and Urban Development, 2016] States Department of Housing, U. and Urban Development (2016). Insights into Housing and Community Development Policy.

[Walker and Herman, 2022] Walker, K. and Herman, M. (2022). *tidycensus: Load US Census Boundary and Attribute Data as 'tidyverse' and 'sf'-Ready Data Frames*. R package version 1.1.2.9000.

[Wyly and Hammel, 1999] Wyly, E. K. and Hammel, D. J. (1999). Islands of Decay in Seas of Renewal: Housing Policy and the Resurgence of Gentrification. *Housing Policy Debate*, 10(4):711–771.