Gentrification, Gun Violence, and Coordination Failure

Zachary Porreca*

West Virginia University

September 2021

Abstract

In this study, I demonstrate the causal linkage between gentrification and gun violence. I develop a theoretical model of competition in the unregulated illegal drug market, and draw the conclusion that violence in the market is, in part, caused by the officially unenforceable nature of territorial claims. Exogenous shocks, such as gentrification, keep viable territory in a state of constant flux, preventing sustained cooperation between these illegal actors. I then specify a two-way fixed effects differences-in-differences estimator to empirically test the model's prediction that the gentrification of one block will lead to increases in violence across the surrounding neighborhood. I find a robust result, that some 5,800 (21%) of Philadelphia's shootings over the decade of this study's window can be attributed to spillover effects from gentrification. This effect is nearly ten times stronger, when it is a high drug crime block that gentrifies. This study further contributes a new easily replicable empirical measurement of gentrification drawn primarily from property sales, along with building, zoning, and alteration permit issuance. This new measurement is able to capture gentrification at its finest and most realistic resolution: the individual block level.

Keywords: Gentrification, Illegal Drugs, Violence, Crime, Guns

JEL Codes: K42, J10, R11

*Department of Economics, John Chambers College of Business and Economics, 1601 University Avenue, Morgantown WV 26506, zjp00003@mix.wvu.edu

1 Introduction

On the night of February 21, 2019 a shootout erupted on the corner of 4th and West Huntington Streets in a section of North Philadelphia infamous for its open-air heroin markets: a neighborhood known locally as the "Badlands". Some 25 shots rang out from at least two shooters. As those shooters fled, a teenager was left dead on the sidewalk, having suffered multiple gunshot wounds[1].

A few blocks over, a large construction project was underway. Nine brand new three and four story triplexes were being built by an out-of-town developer, completely changing the face and character of the neighborhood surrounding the corner of Front and West Huntington Streets. The gentrification of this run-down residential neighborhood was well underway.

In this study, I demonstrate the causal linkage between gentrification and gun violence. I develop a theoretical model of competition in the unregulated illegal drug market, and draw the conclusion that violence in the market is, in part, caused by the officially unenforceable nature of territorial claims. Exogenous shocks, such as gentrification, keep viable territory in a state of constant flux, preventing sustained cooperation between these illegal actors.

I then specify a two-way fixed effects differences-in-differences estimator to empirically test the model's prediction that the gentrification of one block will lead to increases in violence across the surrounding neighborhood. I find a robust result showing that on average gentrification increases levels of gun violence on neighbor blocks. This effect is more pronounced when the gentrified block has a history of drug crime, with an average increase of nearly nine shootings in the surrounding neighborhood. Back of the envelope calculations suggest that roughly 5,800 shootings, or 21% of the city's shootings across the ten year study

[1] From 6ABC Philadelphia. https://6abc.com/philadelphia-kensington-shooting-teen-killed-teenager/5150764/

window can be attributed to spillover effects from gentrification.

This study further contributes a new easily replicable empirical measurement of gentrification drawn primarily from property sales along with building, zoning, and alteration permit issuance. This new measurement is able to capture gentrification at its finest and most realistic resolution: the individual block level.

That a relationship between gentrification and crime exists is not a new revelation. O'Sullivan (2005) provided a theoretical model positing a causal relationship between exogenously falling crime rates and an influx of higher income residents. The author went further, to assert that gentrification is self-reinforcing in that the influx of higher income residents will further decrease crime rates, which further incentives an even greater displacement of low income residents in favor of those with a higher income. Similarly, Ellen et al. (2019) found that reductions in crime rates invite increases in rates of wealthy and educated residents moving into the central city.

Numerous authors have attempted to look the other direction at the effect that gentrification has on crime. Papachristos et al. (2011) regressed homicide and robbery rates on Chicago neighborhood demographic factors drawn from census data, along with a novel proxy for gentrification: the number of coffee shops in a neighborhood. The authors found that their markers of gentrification were associated with declining homicide rates, but increasing robbery rates. Smith (2012) utilized a similar approach, regressing gang related homicides on census provided neighborhood demographic factors, coffee shops, and an indicator for the demolition of public housing. The author finds decreases in homicides as markers of gentrification grow more prevalent. More recently, Autor et al. (2017) found that rapid gentrification precipitated by the end of rent control in Cambridge, Massachusetts led to significant decreases in the overall crime rate.

This study employs a different approach. Neighborhood, census tract, or other artificially drawn borders and boundaries do not necessarily capture the reality of a city. A city is a collection of streets, the lattice-work of their intersections creating the individual blocks that serve as the base level points from which change across a city can be observed. Violence, gentrification, and crime begin as extremely localized phenomena. As such, this study examines block level variation in relevant markers and adopts a quasi-experimental approach: taking block level gentrification as an exogenous treatment so as to allow causal inferences to be made regarding the effect that gentrification has on gun violence.

This study seeks to explain a counter-intuitive result that runs against the major current of gentrification and crime related literature. Gentrification does not impose a positive externality on the surrounding community in regards to gun violence. Instead, these spikes in new real estate interest push violence outward and into the surrounding neighborhood Much of this increase in violence is the result of competition in the illegal drug markets that exist within pockets of the urban environment. The peculiarities of this sort of competition have been largely unexamined in the literature.

Skaperdas (2001), Varese (2010), and Skarbek (2016) have all posited similarly that organized criminal governance can only arise in the vacuum of official forms of governance. However, the presence of such a vacuum does not guarantee that such an organized outcome will arise. Skaperdas (2001) posits a reason for the failure to coordinate: that criminal competition more closely resembles an "arms race", where market share can be taken through violent means that decrease overall market welfare. It this proposition that this study provides theoretical and empirical backing to.

This study seeks to together explain both the reasons why what is a seeming improvement to an area is able to cause increases in violence and why an unregulated market so violently fails to give rise to coordination among competitors. This study posits that the failure to coordinate strategies in this competitive environment is a product of the ability of a firm to gain market share through aggressive and militaristic tactics. The likelihood that all firms across a market should elect not to engage in such strategies decreases as the size of the market to be competed for decreases. Exogenous shocks such as gentrification are responsible for this market concentration and give rise to the increases in gun violence observed.

Several other recent studies have taken a similar approach in examining spatial dimensions of competition among criminal organizations. DeAngelo (2012) builds a theoretical model of spatial competition, and with this model predicts that increases in law enforcement can deter lower productivity criminals' entrance into the market. Sobrino (2019) examines Mexican drug cartel violence as a competition for control over lucrative territory. She finds empirically that violence coincided with the entrance of additional cartels into a municipality. Castillo et al. (2020) also examines Mexican cartel violence, modeling municipal level competition as a contest for revenue. They find that scarcity caused by cocaine seizures led to significant increases in levels of violence. Bruhn (2021) examines gang presence across the city of Chicago. He finds that as a gang enters new territory, levels of crime (including violent crime) in that area increase. While these studies focus on the entrance and exit of competitors and the impact of supply shocks on the illegal market, this study documents the impact that the lack of officially enforceable territorial claims has on the dynamics of competition.

The model developed in this study builds upon the more common models of tacit collusion in price competition by the inclusion of a contest function to determine each organization's latent market share. These market shares can be directly influenced by an organization's choice of strategy, as criminal organizations have the additional option to utilize violent force to expand a share of the market. The inclusion of violence or military capacity in a contest function is not a new contribution. Hirshleifer (1989), Skaperdas (1992), Polo (1995), and Skaperdas (1996) have well established these foundations. Collusion on a strategy space that allows for violent conflict and this market dynamic's implications for urban gentrification are novel contributions.

This study offers both theoretical and empirical evidence for the cause of criminal firms' failure to collude, which is in turn posited to be the mechanism by which gentrification causes increases in violence.

2 Theory and Model

The markets for illegal goods are notably absent formal regulation. There is no protection of property rights and no legal enforcement of contracts or territorial claims. This unregulated market allows criminal organizations to compete for profit both through the traditional vehicles of price and quantity as well as through competitive strategies that are unparalleled in traditional market dynamics. The forcible seizure of a rival's assets or territory is a very real possibility.

To set the framework for this model, the illegal drug market is modeled by an infinitely repeated price competition among N organizations. For simplicity, I will assume a homogeneous product. Some exogenous history has determined the current division of of the market, which is represented by a contest function:

$$S_i = \frac{\eta_i}{\eta_i + \sum_{j \neq i}^{N} \eta_j} = \frac{k_i}{k}$$

Where η_i can be thought of as the defensive effort or reputational advantage needed to maintain an organization's existing position. This market is overlaid over a structure of k city blocks, which for simplicity I will assume are all equally productive. As such, an organization's market share is equal to the fraction of total blocks that are under that organization's control.

Demand is linear and negatively sloping in price. All organization's face the same constant marginal cost and no fixed cost.

Thus, the profit function of firm i is equal to:

$$\pi_i = (p - c)(a - bp)(\frac{\eta_i}{\eta_i + \sum_{j \neq i}^{N} \eta_j})$$

When these organizations fail to coordinate on prices, competition drives prices down to marginal cost. Coordination, however, can allow overall market revenue to be maximized by colluding on the monopoly price, $p^m = \frac{a-Q}{b}$. This allows each organization to earn a higher profit, π_i^m .

By deviating from this cooperative strategy and offering a price marginally lower than p^m, namely $p^* = p^m - \epsilon$, the non-cooperative organization is able to serve the entirety of the market demand, earning all of the market profit. The profit of the deviating firm will be:

$$\pi_i^* = (p^* - c)(a - bp^*)$$

The profit of all other firms will be zero. In this simple case, the other organizations can retaliate against the deviating organization by dropping prices down to marginal cost and in turn driving market profits to zero.

Thus, collusion at this monopoly price is only sustainable as long as the following condition holds:

$$\pi_i^m(\frac{1}{1-\delta_i}) \geq \pi_i^* \quad \forall i \in [1, ..., N]$$

Where: δ_i is the discount factor for organization i; a measure of that organization's "impatience" in discounting future earnings

As such, collusion at that price is only possible when the following condition holds:

$$\delta_i \geq 1 - (\frac{\eta_i}{\eta_i + \sum_{j \neq i}^{N} \eta_j}) \quad \forall i \in [1, ..., N]$$

This is the basic model of collusion in price competing markets. However, the illicit market is unique from those of regulated industries in that there is no enforcement of property rights and organizations are able to compete for market share through violence. This necessitates the inclusion of an additional input into the model. Now, each organization is also able to select a level of military capacity to invest in for use in the next period. This will be represented by γ_{it}, and can be thought of as encompassing all dimensions of violent conflict. It is in essence the level of effort put towards violent competition in a given period. An organization selecting a γ_{it} higher than that of its competitors is able to increase its share of the market at the expense of others in the next period.

Now, a given organization's market share is determined by:

$$S_i^* = \frac{\eta_i + \gamma_{it}}{(\eta_i + \gamma_{it}) + \sum_{j \neq i}^{N} (\eta_j + \gamma_{jt})}$$

This military expenditure is not cost-less. Without loss of generality, I will assume that each unit of military expenditure is normalized at a cost of one and that these expenditures cannot be negative. Further, since this investment must be made in the period before its effect is realized, it is subject to the further constraint that it must be less than the previous stage's profit (for the sake of the model's simplicity, I will exclude the possibility of credit or inter-temporal savings).

The organization's profit function can now be represented as:

$$\pi_{it} = (p - c)(a - bp)(\frac{\eta_i + \gamma_{it}}{(\eta_i + \gamma_{it}) + \sum_{j \neq i}^{N} (\eta_j + \gamma_{it})}) - \gamma_{it+1}$$

If this market is to avoid both welfare-destroying price competition and violent conflict, all organizations must coordinate both on price being set at the monopoly level as well as on the sum of all military effort being zero. This scenario will maintain market shares at their original levels.

Given the nature of Bertrand price competition, there is no incentive for an organization to deviate in both price and military expenditure. Deviation in price negates the impact of a given organization's market share as the deviating organization will now have the lowest price and will command the entirety of market profits. As such, I will assume collusion on the price dimension at the monopoly price level and instead focus on an organization attempting to increase its market share through military spending.

Baring in mind that the maximum a deviating firm is able to expend on military investment is: $\gamma_{it} = \pi_{it-1}$, and assuming collusion in the preceding period and that all other organizations intend to continue colluding, the largest possible market share that can be obtained by the deviating organization in the current period is:

$$S_{it}^{*} = \frac{\eta_i + \pi_{it-1}}{(\eta_i + \pi_{it-1}) + \sum_{j \neq i}^{N} \eta_j}$$

Before investment in military capacity for the next period, this organization will now have a profit of $\Pi_t^m(S_{it}^*)$. This leaves a profit of $\Pi_t^m(S_{jt}^*) \quad \forall j \neq i$ to be distributed among all other organizations. As in the preceding price collusion scenario, when an organization deviates on price, it is expected that the other organizations will retaliate in equal measure. Here, this would mean expending the entirety of each organization's profits from period t on military capacity for period $t+1$. Thus, in this grim trigger strategy:

$$\sum_{j \neq i}^{N} \gamma_{jt+1} = \Pi_t^m(1 - S_{it}^*)$$

And the resulting market share of organization i in period $t+1$ would be equal to:

$$S_{it+1}^{*} = \frac{\eta_i + \gamma_{it} + \gamma_{it+1}}{(\eta_i + \gamma_{it} + \gamma_{it+1}) + \sum_{j \neq i}^{N} (\eta_j) + \sum_{j \neq i}^{N} \gamma_{jt+1}}$$

Where: $\gamma_{it} = \pi_{it-1}$ and $\gamma_{it+1} = \pi_{it}$

Collusion (avoiding the arms race of military expenditure) can only occur when the following condition holds:

$$\Pi_{t-1}^{m}(S_{it-1})(\frac{1}{1-\delta_i}) \geq \Pi_t^m(S_i^*) + \delta_i \Pi_{t+1}^m(S_{it+1}^*) \quad \forall i \in [1, ..., N]$$

This expression is cleaned up by substituting $\frac{k_{it}}{k}$ for S_{it}, leaving the following simplified condition:

$$\delta_i^2(\Pi_{t+1}^m k_{it+1}) - \delta_i(\Pi_{t+1}^m k_{it+1} - \Pi_t^m k_{it}) - k(\Pi_{t-1}^m k_{it-1}) \geq \Pi_t^m k_{it}$$

The expression allows for δ_i to be isolated, providing an expression for the minimum threshold level of δ_i at which an organization will choose to collude in this environment. Taking the derivative of this threshold with respect to k provides a comparative static that serves the crux of this study's empirical basis [2]. In the context of the world that this model approximates, changes in k can be caused by exogenous shocks that remove or add viable territory.

$$\frac{\partial \delta_i^*}{\partial k} = -\frac{k_{it-1}\Pi_{t-1}^m}{\sqrt{k_{it}^2 \Pi_t^{m2} - 2k_{it+1}k_{it}\Pi_{t+1}^m \Pi_t^m - 4k_{it-1}k_{it+1}k\Pi_{t+1}^m \Pi_{t-1}^m + 5k_{it+1}^2 \Pi_{t+1}^{m2}}}$$

Where δ_i^* represents the lower threshold of organization i's discount factor at which collusion can occur.

This expression states that partial derivative of the threshold level of discount factor at which collusion on zero military expenditure is possible with respect to the number of divisions of the market to be contested is negative. Thus, as the number of blocks to be contested shrinks, the threshold discount factor at which violence can be avoided increases.

We will assume that the discount rate, δ_i, is a random variable uniformly distributed $\sim U[0, 1]$ among the N organizations. Let, $F(\delta_i)$ be the cumulative distribution function of these discount rates. As such, as the threshold level δ_i^* increases, the probability that a given organization's δ_i is below that threshold increases.

[2] The entirety of these derivations are provided in the appendix

Proposition: As the number of divisions in the market to competed for decreases, collusion becomes less likely.

Proof: Through the model's derivations above, δ_i^* is decreasing in k. Thus, for $k* > k$, $F(\delta_i^*|k^*)^N > F(\delta_i^*|k)^N$.

We are able to conclude, that the higher the threshold level of δ_i^*, the less likely we are to observe the δ_i values for all organizations being above that threshold. Collusion grows less likely. In the case where violence is a means of competition, this means that violence within the market grows more likely. Making the assumption (as will be argued below) that the gentrification of a viable drug block reduces k allows us an identification strategy by which to test this model's proposition.

3 Data

3.1 Description of Data

The majority of the data utilized in this study was made publicly available by the City of Philadelphia through the OpenDataPhilly portal. Shapefiles representing the city's Census Block Groups, Neighborhoods, and street corner intersections were downloaded independently and mapped to one another. Further, the Philadelphia Police Department's (PPD) crime incidents data set, the city's Department of Licensing and Inspections' (DLI) building and zoning permits data set, and the city's Office of Property Assessment's (OPA) property data set were all utilized. Data from the years 2010-2020 were made use[3].

The OPA property assessment data includes geocoded locations of each property and the date of that property's last sale. Each property was mapped to its

[3] Only the years 2011-2020 were actually made use of in the study, so as to limit the study period to an even ten years. 2010 data was used solely for determining lags and deltas of variables

nearest street corner and is assigned as belonging to that block.

The DLI building and zoning permit data contains records of each permit issued within the city. the permits consist of structural alterations, demolitions, new construction, and change of zoning status permits. Each record has the date of the permit's issuance, it's permit type, and the geocoded location of the property that the permit pertains to. Each permit was mapped and assigned to its closest matching city block.

The PPD crime incidents data contains incident dates, locations, and a brief descriptor of the type of incident. Here, only those incidents described as "aggravated assault-firearm", "homicide- criminal", and "narcotic/ drug law violations" were kept. The section on heterogeneous treatment effects found later in this paper will go into greater detail regarding the drug incidents. Each crime was mapped to its nearest block.

Summary statistics for these series are displayed in the table below.

Statistic	N	Mean	St. Dev.	Min	Max
Number of Shootings	6027120	0.109	0.461	0	69
Number of Drug Crimes	6027120	0.386	2.244	0	163
Property Sales	6027120	0.174	0.653	0	32
Building, Zoning, and Alteration Permits	6027120	1.967	5.269	0	461

Table 1: Summary statistics for block level variables. Note that each observation in the table is for an individual block.

From this block level data, I have built a directional spatial network. Each block is linked to its 24 closest neighbors. This is roughly equivalent to a two block radius surrounding each block. Block $a's$ link to block b is one observation, while block $b's$ link to block a is a separate observation. In total, this provides 602,712 unique linkages. The first node is referred to herein as the "signal block". The node it is linked to is referred to as the "response block". Treatments will occur to signal block, while outcomes will be measured at the response block.

In total, I have indexed the data in linkage by year panels. There are 25,113 city blocks each linked to their 24 closest neighbors, charted over the course of ten years, for a total of 6,027,120 observations. As such, the primary specifications of the models specified will have over 600,000 fixed effects. These fixed effects will be able to capture any variation that could have been picked up by the inclusion of covariates into the models. Further, at the fine level of resolution at which data is made use of in this study, reliable covariate data is not available. Census data (ACS American Community Survey) to provide demographic controls is not available at this block level. Data for larger groupings are unreliable and offered at lengthy intervals. Further, most of the event level data provided by the City of Philadelphia that is able to be geocoded to specific blocks is likely to be collinear to the data already made use of here.

It is of note that each year roughly fifty homicides occur in the City of Philadelphia that are not firearm related. As such, as much as 3% of what are here identified as shootings may be mislabeled. However, it is likely that more shootings than this go unreported each year. The extent of this underreporting is unclear, as shootings without wounded victims may not attract police attention when on average five people are shot in the city each day [4].

Below is a charting on the total number of shootings in each year of this study.

[4] Per the Philadelphia Inquirer

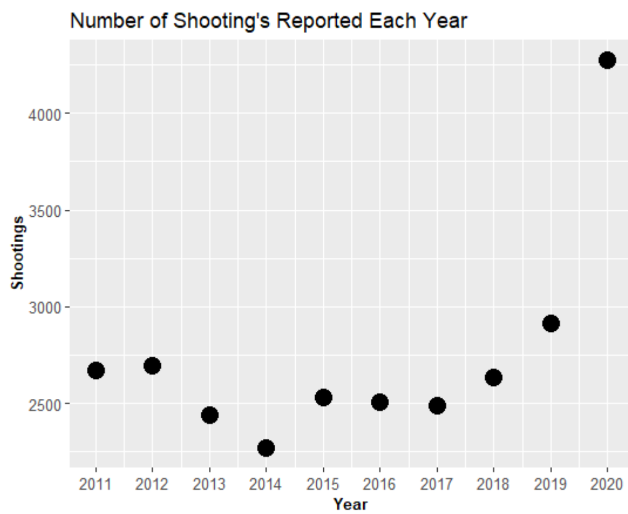

Figure 1: "Total Shooting Incidents by Year"

It is evident that there was a large-scale increase in gun violence within the city in 2020. While the results of this study do not seem to be unduly influenced by 2020's unprecedented high shooting total, the primary specification and its disaggregation are replicated in the supplemental appendix with the year 2020 removed.

The crux of this study's empirical strategy rests upon network link analysis, estimating average changes that occur in one node when the link's opposite node is treated. To meet this purpose, a k-nearest neighbor adjacency matrix was first constructed, tying each block to its 24 closest neighbors. This square matrix has a length and a width equal to the total number of city blocks, and each row and column indexes a particular block. Every cell in the matrix has a value of zero, aside from those blocks that are among the set of k-nearest neighbors to one another. Their intersections will have a value of one (while the main diagonal, representing a block's intersection with itself will have a value of zero as well). For example, if block three and block four are neighbors, cell (4,3) and cell (3,4) will have values of one while cells (4,4) and (3,3) will have values of zero.

To generate the directional links, each coordinate pairing with a value of one is taken as its own observation for each year. With this strategy, the average treatment effect on linked blocks is able to be efficiently estimated.

The map below shows the locations of Philadelphia's highest gun violence areas. These are those blocks experiencing a k-neighbor shooting total greater than three standard deviations above the city-wide mean. The locations of these blocks are overlaid on a map of the city's officially designated neighborhoods, the boundaries of which are primarily the city's main surface streets.

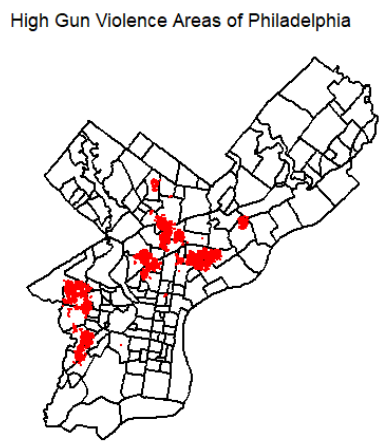

Figure 2: "Map depicting the city's highest gun violence clusters: blocks with neighbor shooting totals more than three standard deviations above the city-wide mean"

3.2 Measuring Gentrification

First and foremost, there exists no overarching consensus regarding the method by which gentrification should be identified quantitatively (Barton, 2014). Glaeser et al. (2020) notes that there are nearly as many different measures of gentrification as there are papers written on the subject. Some of these different definitions and measurements are discussed in Appendix A.

Data limitations are the primary factor preventing the adoption of some wide-spread empirical standard. Due to the unprecedented fine resolution at which I am attempting to measure gentrification and its effects, none of the established gentrification measures are feasible. Instead, I adopt an approach to measurement adapted from that of Holms and Schulz (2017) and Glaeser et al. (2020). The "GentriMap" model of Holms and Schulz (2017) is a two pronged approach to identifying gentrification, utilizing both real estate factors (upgrades to properties and value changes) and social demographic factors as identifying criteria. The approach of Glaeser et al. (2020) is similar. The authors first identify neighborhoods that are capable of gentrifying, based on poverty rates from the first five-year American Community Survey (ACS) to occur in the window of their study. Next, to identify gentrification, the authors track growth in home rental prices; with those that have above median rental price growth rates being labeled as gentrifying.

The approach I utilize in this study identifies gentrification at its finest resolution, at the level of an individual city block. As in the above mentioned approaches, the feasible set of blocks that are capable of gentrification are identified by demographic characteristics. Here, I make use of household income levels from the 2011-2015 ACS for this purpose. The finest resolution that this data is reliably available for is the Census Block Group [5]. As such, each city block is mapped to the Census Block Group that encompasses it. Each block that is located in a block group having an average household income level below the city-wide median is labeled as having the potential to gentrify.

Of those blocks identified as having the potential to gentrify, the gentrification treatment is defined as beginning based upon real estate trends. Since home sale and rental prices (or home valuations) are not easily available for

[5] Census Blocks mentioned here are not the same as the city blocks utilized as the basic level of observation in this study. Philadelphia has 384 Census Tracts, 1336 Census Block Groups, 18872 Census Blocks, and 25,113 actual city blocks

the area of study, I made use of home sales and building/renovation permits to approximate the real estate component of gentrification. Spikes in home sales are likely to be accompanied by rising costs. Helms (2003) provides justification for the use of permitting data[6], finding that trends in building permit issuance do in fact match the types and locations of properties that are predicted to experience real estate interest brought on by gentrification.

To be labeled as gentrifying, the block in question must either have a change from the previous year in the number of new construction or building alteration permits issued that is more than three standard deviations above the city-wide mean, have a change in the number of home sales that is greater than three standard deviations above the city-wide mean, or have both a change in the number of permits issued and a change in the number of home sales that each exceed two standard deviations above the city-wide mean. The numerical values of these classification criteria are summarized in the table below.

Year	Conditions to be met in addition to below median household income condition	# newly meeting criteria
2011	ΔSales>1 or ΔPermits>11 or (ΔSales>0 & ΔPermits>7)	265
2012	ΔSales>1 or ΔPermits>11 or (ΔSales>1 & ΔPermits>7)	292
2013	ΔSales>1 or ΔPermits>12 or (ΔSales>1 & ΔPermits>8)	268
2014	ΔSales>1 or ΔPermits>13 or (ΔSales>1 & ΔPermits>8)	241
2015	ΔSales>2 or ΔPermits>14 or (ΔSales>1 & ΔPermits>9)	131
2016	ΔSales>2 or ΔPermits>15 or (ΔSales>1 & ΔPermits>10)	152
2017	ΔSales>2 or ΔPermits>14 or (ΔSales>1 & ΔPermits>9)	189
2018	ΔSales>2 or ΔPermits>15 or (ΔSales>1 & ΔPermits>10)	152
2019	ΔSales>2 or ΔPermits>18 or (ΔSales>1 & ΔPermits>12)	115
2020	ΔSales>2 or ΔPermits>17 or (ΔSales>1 & ΔPermits>11)	35
Total		1840

Table 2: Criteria for Gentrification Classification

Utilizing changes in permitting and sales, rather than absolute numbers, ensures that areas of longer term sustained real estate interest do not get mislabeled as newly gentrifying. The primary results of the study are replicated with varying definitions of gentrification in the supplemental appendix. Once a

[6]Using permitting data as a justification for gentrification implies a lagged effect on the treatment, as there is undoubtedly a delay between a permit being issued and any sort of residential demographic shifts. This is addressed in the treatment dynamics subsection of the robustness checks portion of this paper. Further, the primary results are replicated with a lagged treatment definition in the supplemental appendix.

block is labeled as having begun the gentrification treatment, it maintains the gentrification label for the remainder of the study. The map below displays the locations of gentrifying blocks in the city.

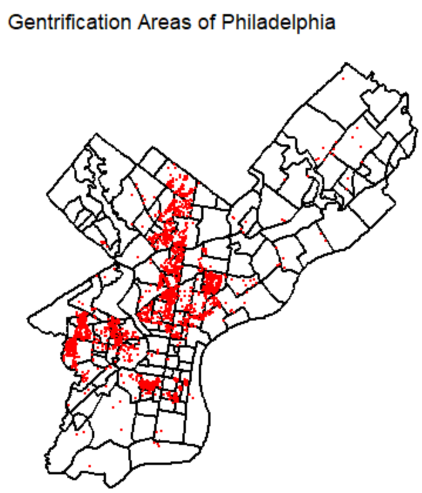

Figure 3: "Map depicting blocks that gentrify during the period of the study"

Appendix C details an alternative gentrification measure where gentrification is treated as a continuous variable.

4 Estimation Strategy

4.1 Primary Specification

To test the validity of the theoretical model's prediction, a two-way fixed effect differences-in-differences model is specified to identify the causal relationship between gentrification and gun violence. City blocks gentrify independently of one another across time. This provides an ideals setting from which to estimate gentrification's impact on violence.

This primary differences-in-differences specification is reported below.

$$shootings_{it} = \beta_i + \psi_t + \alpha_1 Gentrified_{it} + \epsilon_{it} \quad (1)$$

Where, i indexes block links, t indexes years, β_i is a vector of block link fixed effects, and ψ_t is a vector of year fixed effects.

4.2 Justification for Identification

The driving assumption behind this paper's empirical strategy is that a block's gentrification makes that block no longer suitable for drug competition. There are multiple potential rationales that would support this assumption. For one, gentrification includes a turnover of residents and as such likely reduces the original resident base that supported the illicit activity. Ellen, Horn & O'Regan (2012) found that gentrified neighborhoods grow wealthier, more educated, exhibit higher rates of home ownership, and experience significant racial demographic changes. This can be inferred to represent a replacement of many of the neighborhood's original residents; thus making it more difficult for a criminal organization to find the level of community support necessary to operate openly. Further, empirical evidence has demonstrated that gentrification leads to increased policing and the adoption of more punitive policing practices (Laniyonu, 2017). It is reasonable to believe that increases in policing would deter criminal activity, as famously predicted by Becker (1974). Regardless of the actual mechanism at play, it is reasonable to assume that gentrification will make a block less suitable for drug competition.

Further, to empirically test the validity of this identifying assumption, the following difference-in-differences estimation model is specified and tested. This specification seeks the causal impact of gentrification on the levels of drug crime observed on a block following that block's gentrification [7]. For this study's iden-

[7]This specification mirrors the primary estimation strategy, for gentrification's effect on

tifying assumption to hold, the post-gentrification treatment coefficient should be negative and significant.

$$Drugs_{it} = \beta_i + \psi_t + \alpha_1 Gentrified_{it} + \epsilon_{it} \tag{2}$$

As is typical, i indexes individual block links, t indexes time periods, β_i is a vector of block link fixed effects, ψ_t is a vector of year fixed effects, ϵ_{it} is an error term, and α_1 is the parameter of interest. Since these blocks are solely linked to themselves in this specification, these block link fixed effects essentially reduce to a vector of individual block fixed effects and the ensuing regression is identical to a panel of 25,113 individual blocks across the ten year window. The results of this specification are reported below. As the result demonstrates, there is sufficient evidence to support the identifying assumption. Newly gentrified blocks do not support the same level of drug trade post-gentrification as they did pre-gentrification.

	Dependent variable:
	Immediate Block Drug Crimes
Gentrification Indicator	-0.218***
	(0.065)
Block Level Fixed Effects	✓
Year Fixed Effects	✓
Mean	0.39
Observations	6,027,120
Adjusted R^2	0.56
AIC	22504829
Standard Errors	Clustered at Census Block Group Level
Note: Standard errors in parentheses	*p<0.1; **p<0.05; ***p<0.01

Table 3: Output from differences-in-differences estimation of gentrification's negative causal effect on immediate block drug crime.

The theoretical model predicts that the removal of viable territory from an

shootings, which will be analyzed in much greater detail.

illegal market will increase the relative value of the remaining territory to be competed for. I have provided evidence that gentrification reduces drug crime on the immediate gentrified block, implying that this block has been removed from the set of viable territory to be competed for. This is not necessarily enough to completely tie the theoretical model to this empirical setting. It needs to be seen that in the short run the gentrification of one block leads to increases in drug crime on linked neighbor blocks. As such, I will limit the treatment variable to it's effect in the immediate period of treatment[8] and replicate the same specification with linked response block drug crime as the response variable.

	Dependent variable:
	Neighbor Block Drug Crimes
Immediate Year Gentrification Indicator	0.06**
	(0.027)
Block Level Fixed Effects	✓
Year Fixed Effects	✓
Mean	0.38
Observations	6,027,120
Adjusted R^2	0.56
AIC	22408438
Standard Errors	Clustered at Census Block Group Level
Note: Standard errors in parentheses	*p<0.1; **p<0.05; ***p<0.01

Table 4: Output from differences-in-differences estimation of gentrification's short run positive causal effect on neighboring block drug crime.

In the short run there is a clear increase in neighbor block drug crime following a block's gentrification. When taken in conjunction with gentrification's effect on drug crime on the immediate block there is a clear suggestion that gentrification pushes drug crime away from the immediate block and into the surrounding neighborhood. This provides a sufficient justification for tying the theoretical model to this empirical setting. Gentrification does reduce the k

[8]This is done by replacing the Gentrification Indicator with a new indicator that is equal to one if a given block is newly gentrified in that particular period.

parameter, and as such should be associated with a decreased likelihood of co-operation across the illicit market and a measurable increase in gun violence.

5 Results

5.1 Primary Result

The results of fitting this primary empirical model to the data are reported below.[9] Throughout this remainder of this study, I will focus most of my attention to the OLS specification.

	Dependent variable:				
	Neighbor Block Shootings				
	OLS	Poisson	Negative Binomial	Log-Linear	Log IHS
Gentrification Indicator	0.02***	0.04	0.02	0.009**	0.012**
	(0.006)	(0.024)	(0.023)	(0.003)	(.004)
Block Level Fixed Effects	✓	✓	✓	✓	✓
Year Fixed Effects	✓	✓	✓	✓	✓
Mean	0.11	0.11	0.11	0.066	0.085
Observations	6,027,120	2,391,250	2,391,250	6,027,120	6,027,120
Adjusted R^2	0.13	0.15	0.15	0.16	0.16
AIC	7512901	3390800	3338845	-822808	2248264
All Standard Errors Clustered at Census Block Group Level					

Note: R^2 is squared correlation in Poisson and NegBin *p<0.1; **p<0.05; ***p<0.01
Note: For log-linear, log(y)=log(y+1). Log IHS is log inverse hyperbolic sine transformation
Note: Standard errors are in parentheses

Table 5: Output from primary differences-in-differences specification from various estimation methods

Gentrification has a positive effect on levels of gun violence observed in the neighborhood surrounding the gentrified block. This result will be analyzed in detail in the robustness section of this paper. First, though, I will examine the heterogeneous treatment effect; disaggregating this observed effect into its

[9] Estimation methods aimed at addressing the large amount of zeros in the data are reported in the appendix.

different component effects on high drug crime blocks and non-high drug crimes blocks.

5.2 Heterogeneous Treatment Effects

5.2.1 When High Drug Crime Blocks Gentrify

The theoretical model derived in Section 2 is specific in its setting. It provides a prediction of the effect that changes in the total quantity of territory to be competed over will have on aggregate levels of violence across an illegal market. The empirical results, as reported thus far, are generalized to the entirety of the city. Here, I analyze the heterogeneous treatment effects, providing evidence that the treatment has a significantly stronger effect on those blocks that house the illegal drug trade.

To begin, as mentioned previously, drug incidents from the same Philadelphia Police Department crime incidents database were mapped to their nearest city block. These incidents were listed as "narcotic/ drug law violations". Those blocks that had a number of incidents in the previous year that exceeded three standard deviations above the city-wide mean were labeled as high drug crime areas. Once labeled a drug block that block will maintain that label for the remainder of the study. Allowing for blocks to gain this label at later time periods is meant to capture some of the dynamic nature of the movement of the drug trade across the city. It is of note that the drug incident count that is being made use of here is not a neighbor-based metric metric, as the shooting count was. These are drug incidents occurring directly on that block. A map of the location of these blocks is shown below.

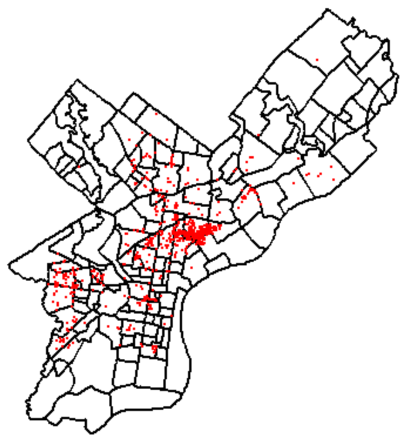

Figure 4: “Map depicting blocks designated as high drug crime areas during the period of study”

The table below charts the number of drug blocks both initially identified and gentrified in the city in each year. It is evident that drug blocks are both a relative rarity in the city and that their gentrification is an even rarer occurrence. Across the ten year window examined, the threshold for this classification is on average at least seven drug incidents occurring on that block in a given year.

Year	Blocks Gentrified	New Blocks Labeled High Drug crime	Drug Blocks Gentrified
2011	265	282	23
2012	292	0	24
2013	268	71	21
2014	241	62	23
2015	131	34	5
2016	152	35	13
2017	189	42	7
2018	152	49	13
2019	115	29	10
2020	35	50	4
Total	1,840	654	143

Table 6: Counts of unique blocks labeled as gentrified or as drug blocks by year

With this drug block indicator, the primary difference in difference specification is able to be disaggregated into the treatment effect on non-drug blocks and the effect on drug blocks.

$$shootings_{it} = \beta_i + \psi_t + \alpha_1 NonDrugBlockGentrified_{it} + \alpha_2 DrugBlockGentrified_{it} + \epsilon_{it} \tag{3}$$

	Dependent variable:				
	Neighbor Block Shootings				
	OLS	Poisson	Negative Binomial	Log-Linear	Log IHS
Drug Block Gentrification Indicator	0.12***	0.18**	0.16***	0.052***	0.068***
	(0.029)	(0.062)	(0.056)	(0.013)	(0.17)
Non-Drug Block Gentrification Indicator	0.01**	0.02	0.004	0.006**	0.008**
	(0.005)	(0.022)	(0.022)	(0.002)	(0.003)
Block Level Fixed Effects	✓	✓	✓	✓	✓
Year Fixed Effects	✓	✓	✓	✓	✓
Mean	0.11	0.11	0.11	0.066	0.085
Observations	6,027,120	2,391,250	2,391,250	6,027,120	6,027,120
Adjusted R^2	0.13	0.15	0.15	0.16	0.16
AIC	7512571	3390748	3338813	-823076	2247992
All Standard Errors Clustered at Census Block Group Level					

Note: R^2 is squared correlation in Poisson and NegBin *p<0.1; **p<0.05; ***p<0.01
Note: For log-linear, log(y)=log(y+1). Log IHS is log inverse hyperbolic sine transformation
Note: Standard errors are in parentheses

Table 7: Output from disaggregating difference and difference coefficient with various estimation methods

As is evident in all specifications, the gentrification of a drug block has a highly significant positive effect on the level of gun violence exhibited in the surrounding neighborhood as predicted by the theoretical model. Further, it is of note that the effect on drug blocks is consistently of a higher magnitude than that on non-drug blocks. Focusing on the OLS results, the treatment effect is more than ten times as large when it is a drug block that gentrifies. Testing the equality of these two (OLS) coefficients returns a Wald Chi Square of 12.38, with a corresponding p-value below 0.01. These two coefficients are significantly different than one another.

To further confirm this heterogeneous treatment effect, a differences-in-differences-in-differences (triple-difference) model is specified. A positive coefficient on the triple-difference estimator signifies that the effect on drug blocks is of a greater magnitude than the overall treatment effect.

$$shootings_{it} = \beta_i + \psi_t + \alpha_1 Gentrified_{it} + \alpha_2 (Gentrified_{it} \cdot DrugBlock_i) + \epsilon_{it} \tag{4}$$

Where $DrugBlock_i$ is an indicator equal to one if the signal block is designated as a drug block at any point during the study. α_2 is the parameter of interest.

	Dependent variable:
	Neighbor Blocks Shootings
Gentrification Indicator	0.014**
	(0.005)
Drug Block Gentrification Indicator	0.10***
	(0.026)
Block Level Fixed Effects	✓
Year Fixed Effects	✓
Mean	0.11
Observations	6,027,120
Adjusted R^2	0.13
AIC	7512621
Standard Errors	Clustered at Census Block Group Level
Note: Standard errors in parentheses	*p<0.1; **p<0.05; ***p<0.01

Table 8: Staggered triple differences-in-differences estimator. Drug block gentrification is the parameter of interest

As is evident from the results table above, the triple-difference coefficient is positive and significant. Gentrification has a statistically significantly different effect on high drug crime blocks. This effect is of a greater magnitude, and is able to explain some of the higher levels of violence observed on drug blocks.

5.2.2 Gentrification's Impact on Other Gentrified Blocks

Gentrification does not occur disparately across the city. As evidenced in the map displayed earlier, gentrification tends to occur in clusters. The table below shows, for the final year of this study (2020), the degree of clustering among gentrified blocks. At the end of the study, there were 1,840 blocks recorded as having gentrified. The table displays the frequency of a gentrified block having a number of gentrified neighbors in its 2 block radius (k=24).

Range	Count
0 to 5	246
6 to 10	378
11 to 15	535
16 to 20	518
21 to 24	163

Table 9: Frequency of gentrified observations with a number of gentrified neighbors falling into a specific range

To examine the heterogeneous impact that gentrification has on other gentrified blocks versus non-gentrified blocks, the data is subset into linkages to gentrified blocks and linkages to non-gentrified blocks. The primary model specification was replicated on these two subsets.

	Dependent variable:	
	Neighbor Blocks Shootings	
	(Gentrified Response Block Subset)	(Non-Gentrified Response Block Subset)
Gentrification Indicator	0.011	0.016^{**}
	(0.010)	(0.005)
Block Level Fixed Effects	✓	✓
Year Fixed Effects	✓	✓
Mean	0.245	0.103
Observations	292,231	5,734,889
Adjusted R^2	0.13	0.13
AIC	601560	6799962
Standard Errors	Clustered at Census Block Group Level	
Note: Standard errors in parentheses		$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 10: Comparison

There is no effect observed on the gentrified response blocks. To verify this, the primary regression specification is altered to allow the inclusion of an indicator for whether the response block is gentrified and an interaction term to capture the different effect that gentrification has on other gentrified blocks. This new specification is displayed below.

$$shootings_{it} = \beta_i + \psi_t + \alpha_1 Gentrified_{it} + \alpha_2 Response\ Block\ Gentrified_{it}$$
$$+\alpha_3 (Gentrified \cdot \ Response\ Block\ Gentrified)_{it} + \epsilon_{it} \tag{5}$$

	Dependent variable:
	Neighbor Block Shootings
Gentrification Indicator	0.013**
	(0.005)
Response Block Gentrified Indicator	0.031***
	(0.007)
Interaction	0.011
	(0.008)
Block Level Fixed Effects	✓
Year Fixed Effects	✓
Mean	0.11
Observations	6,027,120
Adjusted R^2	0.13
AIC	7512447
Standard Errors	Clustered at Census Block Group Level
Note: Standard errors in parentheses	*p<0.1; **p<0.05; ***p<0.01

Table 11: Regression to test whether the treatment effect is different for response blocks that are themselves treated.

The insignificance of the interaction terms suggests that there is no difference in treatment effect between gentrified and non-gentrified response blocks. This, taken in conjunction, with the previous regressions showing no significant treatment effect at all on gentrified response blocks, provides evidence that the effect being captured in this paper's primary regression specification is largely occurring on non-gentrified blocks.

6 Robustness

6.1 Common Trends Assumption and Treatment Dynamics

Given that the empirical model makes use of a staggered treatment timing difference and difference estimator, this estimator's unbiasedness rests upon the assumption of parallel trends in outcome between both control and treatment groups (prior to those treatment blocks receiving their gentrification treatment). Following the bulk of recent literature, an event-study plot is used to visually demonstrate the average treatment effect at a number of periods before or after treatment.

This plot is made my specifying the following regression model, adapted from He and Wang (2017).

$$shootings_{it} = \sum_{k=-9, k\neq -1}^{k=9} D_{it}^{k} \cdot \delta_k + \gamma_t + \psi_i + \epsilon_{it} \tag{6}$$

Here, the coefficients of interest are δ_k. D_{it}^{k} represents a vector of dummy variables equal to one, if block i in period t is k periods away from gentrification. $k = 0$ in the initial treatment period. Further, as in He and Wang (2017) omitted $k = -1$ so that post-treatment event study estimators are relative to the period immediately before treatment.

The results are reported and plotted in the table and graphic below.

	Dependent variable:	
	Neighbor Block Shootings	
9 years before	-0.03	(0.019)
8 years before	-0.02	(0.015)
7 years before	-0.02	(0.011)
6 years before	-0.01	(0.009)
5 years before	-0.004	(0.008)
4 years before	-0.001	(0.008)
3 years before	-0.01**	(0.005)
2 years before	-0.002	(0.005)
Treatment Period	0.007 *	(0.004)
1 year after	0.01**	(0.005)
2 years after	0.01**	(0.006)
3 years after	0.02***	(0.006)
4 years after	0.02***	(0.007)
5 years after	0.03***	(0.008)
6 years after	0.05***	(0.012)
7 years after	0.05***	(0.010)
8 years after	0.09***	(0.015)
9 years after	0.11***	(0.017)
Block Level Fixed Effects	✓	
Year Fixed Effects	✓	
Observations	6,027,120	
Adjusted R^2	0.13	
AIC	7512078	
Standard Errors	Clustered at Census Block Group Level	
Note: Standard errors in parentheses	*p<0.1; **p<0.05; ***p<0.01	

Table 12: Event study estimators of leads and lags of treatment effect

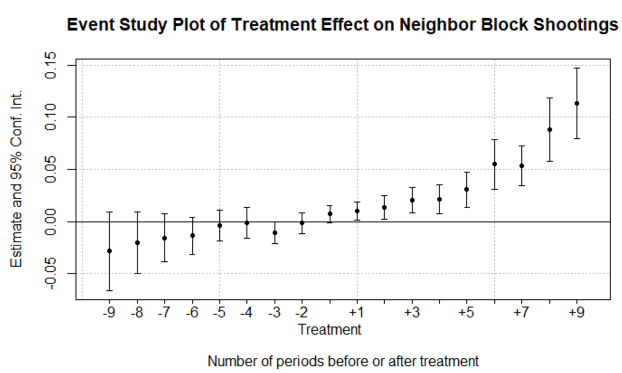

Figure 5: "Observed treatment effect by number of years pre or post treatment"

Of the eight pre-treatment coefficients, only one is significant at the 95% confidence level. This result is not dissimilar from those surveyed by Roth (2020). However, the primary criteria by which the common trends assumption will be tested is on is a joint-significance χ^2 test. All of the pre-period coefficients will be tested together against a null hypothesis of joint insignificance.

This test returns a Wald statistic of 1.41, which has a p-value of 0.18. As such, the null hypothesis cannot be rejected. There is not evidence that the assumption of common trends is violated. Further evidence of this will be provided from the results of falsification/ placebo tests later in this paper.

6.2 Differences-in-Differences Timing and Cohort Effects

Recent literature (Callaway and Sant'anna, 2020; Goodman-Bacon, 2021) on differences-in-differences methodology has shown us that the differences-in-differences estimator is not always an ideal measure of the average treatment effect for treated population (ATT) that it is aimed at estimating. As such, it is important to examine the cohort dynamics of treated units, when treatment occurs at staggered intervals. Below are the cohort specific estimators, and the overall ATT obtained using the Callaway and Sant'anna (2020) doubly robust estimator.

Overall ATT:				
	ATT	Std Error	95% CI Lower Bound	95% CI Upper Bound
	0.0218*	0.0046	0.0127	0.0309
Group Effects:				
Group	ATT	Std Error	95% CI Lower Bound	95% CI Upper Bound
2012	0.0380*	0.0097	0.0109	0.0650
2013	0.0109	0.0106	-0.0188	0.0406
2014	0.0208	0.0077	-0.0007	0.0424
2015	0.0356*	0.0101	0.0075	0.0638
2016	-0.0066	0.0115	-0.0389	0.0257
2017	-0.0070	0.0082	-0.0300	0.0160
2018	0.0400*	0.0108	0.0098	0.0702
2019	0.0447	0.0160	0.0000	0.0894
2020	0.0507	0.0439	-0.0721	0.1735
	Signif. codes: '*' confidence band does not cover 0			
Control Group:	Never Treated		Anticipation Periods:	0
Estimation Method:	Doubly Robust			

Table 13: cohort level results from doubly robust Callaway Sant'anna estimator

The graphic below, shows simple OLS cohort specific estimators obtained using this study's primary differences-in-differences specification. It is of note that these estimators are largely consistent in magnitude. This balance implies robustness of the overall estimator.

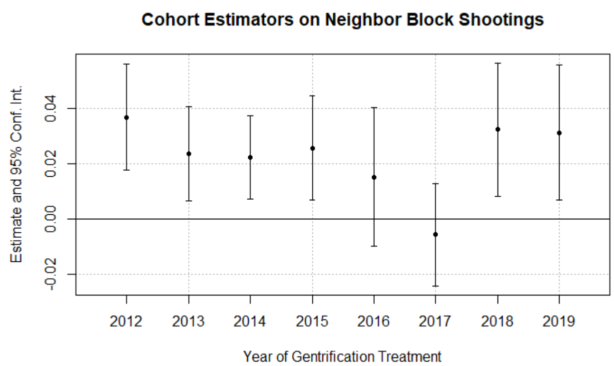

Figure 6: "Relatively consistent and positive estimators for each cohort "

Goodman-Bacon (2021) explored the mechanics behind two-way fixed effects difference-in-difference estimators such as the one utilized in the present study. He demonstrates that the estimator can be decomposed into a weighted average of all two-way difference-in-difference estimators that are able to be constructed from the data. This determines the impact of variations in treatment timing on the estimator. Decomposition of the estimator is able to show which units are driving the estimator's perceived effect, and whether the estimated effect dissipates within treated units over time.

Here, Goodman-Bacon's method, commonly referred to as "Bacon Decomposition", will be used to explore the robustness of the primary result displayed in the previous section. Decomposing that estimator yields the following result:

Table 14: Bacon Decomposition of the primary DID estimator

Type	Weight	Average Estimate
Earlier vs Later Treated	0.01047	0.00502
Later vs Always Treated	0.01095	-0.01393
Later vs Earlier Treated	0.01734	-0.01342
Treated vs Untreated	0.96124	0.02214

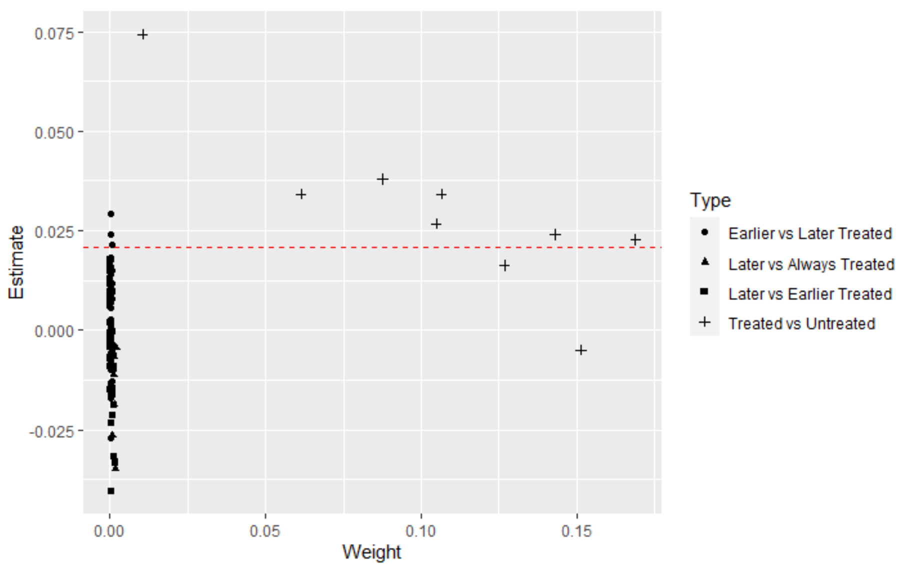

Figure 7: "Graphic depicting result from Bacon decomposition of primary DID estimator, with DID estimator value (0.0209) as dashed line"

As is evident from these decomposition results, the effect is driven nearly entirely by the difference between treated-vs-untreated units.

6.3 Falsification Tests

To test the validity of the treatment, several falsification tests are identified. these tests consist of pushing the treatment date of treated units forward and replicating the primary specification with this new placebo treatment indicator. For the purpose of these tests, observations for the year 2020 were dropped as the high outcome values for that year could unduly skew these falsification tests' results. These tests were replicated moving the gentrification treatment year forward one, two, three, and four years forward each. Insignificant coefficients imply that the placebo treatment is invalid, and give credence to the validity of the actual treatment. These results are reported in the table below.

	Dependent variable:			
	Number of Shootings			
	Minus One Year	Minus Two Years	Minus Three Years	Minus Four Years
Gentrification Indicator Placebo	0.007	0.006	0.002	0.004
	(0.006)	(0.006)	(0.007)	(0.006)
Block Level Fixed Effects	✓	✓	✓	✓
Year Fixed Effects	✓	✓	✓	✓
Observations	5,359,608	5,296,536	5,238,648	5,186,592
Adjusted R^2	0.13	0.13	0.13	0.13
AIC	6187331	6073616	5844546	5716283
	All Standard Errors Clustered at Census Block Group Level			
Note:	*p<0.1; **p<0.05; ***p<0.01			

Table 15: Output from falsification tests of primary differences-in-differences specification

These falsification tests show no significance for the placebo dummy. As such, they provide significant evidence that the gentrification treatment dummy is properly located temporally. This provides more evidence of the validity of the common trends assumption, and to the unbiasedness of the differences-in-differences estimator.

7 Discussion

This study has demonstrated a robust positive causal relationship between gentrification and gun violence. Following a block's gentrification, there is a substantial increase in shootings in the surrounding neighborhood. The theoretical model presented provides a potential mechanism behind some of these increases. I have argued that due to a lack of enforceable territorial claims, competition in illegal markets grows more violent with shocks to the total viable territory available and that gentrification serves as such a shock. However, this is not the only mechanism behind gentrification's role in spurring gun violence.

It is likely that gentrification forces intra-city migration. Residents are displaced from their long term homes, and forced into the remaining viable tracts of affordable housing. These sorts of situations, where disaffected low-income residents are forced to live in unfamiliar neighborhoods surrounded by similarly disaffected and displaced neighbors, have the potential to cause excessive tension. That this can give rise to explosions in gun violence is not surprising.

Unfortunately, as is the major limitation in this study, data availability makes empirical analysis of that potential mechanism near impossible. Reliable data documenting the movement of urban residents across the city on an individual level is not readily available. Similarly, regularly updated demographic data at the requisite granular block level is not easily accessible. This has prevented a more in-depth and formal analysis of what constitutes gentrification. Further, computational limitations have prevented the examination of wider spread impacts beyond the immediate neighborhood.

Despite these limitations, this study has presented a readily replicable measure of gentrification at an extremely fine resolution. Further, through leveraging a massive networked data set, I have been able to generate robust significant estimates of the causal effect that gentrification has on levels of gun violence.

All of the coefficient's estimated are of admittedly small magnitudes. However, when considering that each of these estimates is of the average causal effect on the each single linkage between a newly gentrified block and its neighbor, these numbers add up quickly to a startling total. Over the ten year window of the study, back of the envelope calculations suggest that the city experienced some 5,800 shootings that can be attributed to gentrification. This means, that of the 27,000 shootings that occurred across the city during this time period, around 21% were spillover effects of gentrification.

Exploring the mechanism posited by this study's theoretical model, the estimator was significantly higher. The gentrification of drug blocks accounted for roughy 2,400 additional shootings during the ten year time span. This attributes some 8% of the city's gun violence across the decade to gentrification giving rise to instability in the city's illicit drug markets.

These estimates are striking. Each episode of gun violence has the potential to forever impact the lives of victims, perpetrators, families, and community members at large. As such, in light of this unintended impact, it is crucial that urban development occur responsibly and intentionally. Forced displacement, through pricing out residents, has very real effects on the surrounding neighborhood.

Future research would do well to examine the natural limitations to where illegal drug markets can feasibly exist. Models that are able to predict the spatial movements of these markets across time can help to prevent these high levels of violence from continuing. Further, the theoretical model developed in this study should be expanded upon and tested empirically in other settings.

8 Appendix

8.1 Appendix A: Discussion of Other Gentrification Definitions

First and foremost, there exists no overarching consensus regarding the method by which gentrification should be identified quantitatively (Barton, 2014). Authors have employed novel identification strategies, such as utilizing the number of coffee shops in a given neighborhood [10] (Papachristos et al., 2011). Easton

[10] Such a strategy would not be useful for a city such as Philadelphia which embodies a striking divide between residential and commercial neighborhoods. As such, this method is

et al. (2019) have stated that the difficulties in developing a quantitative identification strategy for gentrification is primarily a product of insufficient data. They posit that gentrification is the *displacement* of lower income residents, and that neighborhoods undergoing gentrification can be identified by demographic changes, real estate trends, and survey data. However, the authors concede that this sort of data is limited in its usefulness and may not necessarily capture all of this sort of resident displacement.

Helms (2003) provides a thorough model of real estate dynamics as a product of gentrification. His model's framework provides much of foundation on which this study's quantitative definition of gentrification rests. Using data from Chicago to test his theoretical model, Helms (2003) finds that trends in building permit issuing do in fact match the types and locations of properties that are predicted to experience real estate interest brought on by gentrification. This finding supports this study reliance on building/ renovation permit and home sale data to serve as a marker of gentrification.

Holm and Schulz (2017) provide a robust statistical model for identifying gentrification. As is typical of this literature, the authors concede that data availability is the largest limitation to the quantitative identification of gentrification. Their "GentriMap" model includes both real estate factors (upgrades to properties and value changes) and social demographic factors as identifying criteria.

This study embodies a similar approach to that of the "GentriMap" model, subject to the data limitations. Sales counts are used in place of value changes. This proxy is a reasonable substitution, given data limitations. An increase in sales reflects an increase in demand which in turn should imply an increase in property valuations. Permit data is used in the same manner as in other gentrification models. Median household incomes are utilized as a demographic

not replicated in this study

factor indicative of a block's ability to become gentrified -as most definitions agree that low income household displacement is a necessary condition.

Several other recent papers have put forth measurements of gentrification. Dragan et al. (2019) made use of 5-year ACS data at the Census Tract level (the third smallest level of Census data aggregation) to identify potentially gentrifying blocks based in part on changes in demographic characteristics captured by these surveys. This data was augmented by administrative data identifying the addresses of individuals, to track their movement across the city over the study period. Ellen et al. (2019) confidential census data to identify household demographic characteristics and residency, to track high-income movements into the city that are associated with gentrification. This analysis takes place at the Census Tract level.

Autor et al. (2017) use spatial exposure to the buildings that end their rent control status as a measure of gentrification. This is done on the basis of some literature the authors cite, that establish the abatement of rent control as being associated with numerous changes in neighbor housing stock quality and pricing. Glaeser et al. (2020) puts forth both a continuous and a binary gentrification measure, built on 5 year ACS survey data and rental price data. That analysis is conducted at the zip-code level.

As is evident, this study is not alone in offering forth an empirical measure of gentrification. Nor is the measurement strategy utilized herein disparate from that utilized in the bulk of the literature. However, this study differs in that it measures gentrification at the finest resolution availble, while using solely open access freely available data.

8.2 Appendix B: Continuous Definition of Gentrification

It is worth briefly examining a potential continuous measure of gentrification, in which gentrification exists on a spectrum. More real estate interest on more impoverished blocks leads to higher values on this spectrum, while less interest of wealthier blocks leads to lower values.

The real estate component is represented by the following index.

$$Real\ Estate\ Index = \frac{(\frac{\Delta sales_{it}}{\Delta \bar{sales}_t} + \frac{\Delta permits_{it}}{\Delta \bar{permits}_t})}{\frac{1}{N} \cdot \sum_{i=1}^{N}(\frac{\Delta sales_{it}}{\Delta \bar{sales}_t} + \frac{\Delta permits_{it}}{\Delta \bar{permits}_t})} \tag{7}$$

This index value is block specific. It is then multiplied by its constituent Census Block Group's poverty rate to provide a simple continuous measure of gentrification.

$$Gentrification\ Index = Real\ Estate\ Index \cdot Poverty\ Rate \tag{8}$$

Replacing the binary gentrification treatment indicator from the initial specification with this continuous measure yields the following results.

$$shootings_{it} = \beta_i + \psi_t + \alpha_1 Gentrification\ Index_{it} + \epsilon_{it} \tag{9}$$

This measure has a mean of 0.032, a standard deviation of 1.05, and ranges from -10.07 to 9.3. Higher values are associated with higher degrees of gentrification.

	Dependent variable:
	Neighbor Block Shootings
Continuous Gentrification Measure	0.0007**
	(0.0003)
Block Level Fixed Effects	✓
Year Fixed Effects	✓
Mean	0.11
Observations	6,027,120
Adjusted R^2	0.13
AIC	7512447
Standard Errors	Clustered at Census Block Group Level
Note: Standard errors in parentheses	*p<0.1; **p<0.05; ***p<0.01

Table 16: Continuous model of gentrification's effect

8.3 Appendix C: Addressing Non-standard Distribution of Shootings

As is to be expected, the distribution of shootings is heavily clustered at zero. The vast majority of blocks and linkages do not experience any shootings. Typically, this is addressed with with a Tobit model. However, due to the large number of fixed effects used in the empirical model, this is extremely computationally costly and not feasible. This sort of estimation would rely on specialty atypical maximization routines that would still provide inconsistent estimations of the variance (Henningsen, 2020).

As such, a random effects Tobit model is specified here, the results of which are reported in the table below

	Dependent variable:
	Number of Shootings
Gentrification Indicator	0.11***
	(0.0008)
Log Sigma Mu	-0.62***
	(0.0007)
Log Sigma Nu	-0.78***
	(0.00002)
Block Level Random Effects	✓
Year Random Effects	✓
Mean	0.11
Observations	6,027,120
Log Liklihood	9419719
Standard Errors	Clustered at Census Block Group Level
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 17: Random Effects Tobit Model

The disaggregation of this coefficient, as shown earlier with OLS estimators in the heterogeneous treatment effects section, is displayed below.

	Dependent variable:
	Number of Shootings
Drug Block Gentrification	0.40***
	(0.003)
Non-Drug Block Gentrification	0.08***
	(0.0009)
Log Sigma Mu	-0.62***
	(0.0007)
Log Sigma Nu	-0.78***
	(0.00002)
Block Level Random Effects	✓
Year Random Effects	✓
Mean	0.11
Observations	6,027,120
Log Liklihood	9419083
Standard Errors	Clustered at Census Block Group Level
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 18: Random Effects Tobit Model- disaggregation of primary treatment estimator

Now, addressing the fixed effects issue, a hurdle model is approximated. This will provide two separate estimators: one, a logit estimator of the probability of the shooting outcome variable being equal to zero and the other a differences-in-differences OLS estimator ran solely on the subset of the data in which the number of shootings is greater than zero.

As above, the disaggregated coefficient is reported as well.

	Dependent variable:	
	Pr(Number of Shootings>0)	Number Shootings When Number of Shootings>0
	Logit	OLS
Gentrification Indicator	0.017	0.06***
	(0.028)	(0.022)
Block Level Fixed Effects	✓	✓
Year Fixed Effects	✓	✓
Mean	0.08	1.37
Observations	2,391,250	483,363
Adjusted R^2	0.12	0.14
AIC	2627291	1548094
All Standard Errors Clustered at Census Block Group Level		
Note: Logit Adjusted R^2 is Squared Correlation	*p<0.1; **p<0.05; ***p<0.01	

Table 19: Output from replicating primary difference and difference specification with hurdle approximation

	Dependent variable:	
	Pr(Number of Shootings>0)	Number Shootings When Number of Shootings>0
	Logit	OLS
Drug Block Gentrification Indicator	0.18***	0.19**
	(0.068)	(0.060)
Non-Drug Block Gentrification Indicator	-0.001	0.04*
	(0.027)	(0.022)
Block Level Fixed Effects	✓	✓
Year Fixed Effects	✓	✓
Mean	0.08	1.37
Observations	2,391,250	483,363
Adjusted R^2	0.12	0.14
AIC	2627269	1548063
All Standard Errors Clustered at Census Block Group Level		
Note: Logit Adjusted R^2 is Squared Correlation	*p<0.1; **p<0.05; ***p<0.01	

Table 20: Output from disaggregating difference and difference coefficient with hurdle approximation

9 References

References

[1] Autor, D., Palmer, C., & Pathak, P. (2017). Gentrification and the amenity value of Crime reductions: Evidence from rent deregulation. https://doi.org/10.3386/w23914

[2] Baker, A., Larcker, D. F., & Wang, C. C. (2021). How much should we trust staggered difference-in-differences estimates? SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3794018

[3] Barnum, J. D., Campbell, W. L., Trocchio, S., Caplan, J. M., & Kennedy, L. W. (2016). Examining the environmental characteristics of drug Dealing Locations. Crime & Delinquency, 63(13), 1731–1756. https://doi.org/10.1177/0011128716649735

[4] Barton, M. (2014). An exploration of the importance of the strategy used to identify gentrification. Urban Studies, 53(1), 92–111. https://doi.org/10.1177/0042098014561723

[5] Becker, G. S. (1974). Crime and Punishment: An Economic Approach. In Essays in the economics of crime and punishment (pp. 1–54). essay, National Bureau of Economic Research.

[6] Bruhn, J. (2021). Competition in the Black Market: Estimating the Causal Effect of Gangs in Chicago. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3837695

[7] Callaway, B., & Sant'Anna, P. H. C. (2020). Difference-in-differences with multiple time periods. Journal of Econometrics. https://doi.org/10.1016/j.jeconom.2020.12.001

[8] Castillo, J. C., Mejía, D., & Restrepo, P. (2020). Scarcity without Leviathan: The Violent Effects of Cocaine Supply Shortages in the Mexican Drug War. The Review of Economics and Statistics, 102(2), 269–286. https://doi.org/10.1162/rest$_{a0}$0801

[9] Chalfin, A., & McCrary, J. (2017). Criminal Deterrence: A Review of the Literature. Journal of Economic Literature, 55(1), 5–48. https://doi.org/10.1257/jel.20141147

[10] Clogg, C. C., Petkova, E., & Haritou, A. (1995). Statistical methods for comparing regression coefficients between models. American Journal of Sociology, 100(5), 1261–1293. https://doi.org/10.1086/230638

[11] DeAngelo, G. (2012). Making space for crime: A spatial analysis of criminal competition. Regional Science and Urban Economics, 42(1-2), 42–51. https://doi.org/10.1016/j.regsciurbeco.2011.04.008

[12] de Chaisemartin, C., & d'Haultfoeuille, X. (2020). Difference-in-differences estimators of intertemporal treatment effects. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3731856

[13] Dewey, M., Míguez, D. P., & Saín, M. F. (2016). The strength of collusion: A conceptual framework for interpreting hybrid social orders. Current Sociology, 65(3), 395–410. https://doi.org/10.1177/0011392116661226

[14] Dragan, K., Ellen, I. G., & Glied, S. (2020). Does gentrification displace poor children and their Families? New evidence from Medicaid data in New York City. Regional Science and Urban Economics, 83, 103481. https://doi.org/10.1016/j.regsciurbeco.2019.103481

[15] Easton, S., Lees, L., Hubbard, P., & Tate, N. (2019). Measuring and mapping displacement: The problem of quantification in the battle against gentrification. Urban Studies, 57(2), 286–306. https://doi.org/10.1177/0042098019851953

[16] Ellen, I. G., Horn, K. M., & Reed, D. (2019). Has falling crime invited gentrification? Journal of Housing Economics, 46, 101636. https://doi.org/10.1016/j.jhe.2019.101636

[17] Glaeser, E. L., Luca, M., & Moszkowski, E. (2020). Gentrification and Neighborhood Change: Evidence from Yelp. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3751621

[18] Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. Journal of Econometrics. https://doi.org/10.1016/j.jeconom.2021.03.014

[19] Grossman, H. I., & Kim, M. (1995). Swords or Plowshares? A Theory of the Security of Claims to Property. Journal of Political Economy, 103(6), 1275–1288. https://doi.org/10.1086/601453

[20] Harbring, C. (2006). The effect of communication in incentive systems—an experimental study. Managerial and Decision Economics, 27(5), 333–353. https://doi.org/10.1002/mde.1266

[21] He, G., & Wang, S. (2017). Do college graduates serving as village officials help rural china? American Economic Journal: Applied Economics, 9(4), 186–215. https://doi.org/10.1257/app.20160079

[22] Helms, A. C. (2003). Understanding gentrification: an empirical analysis of the determinants of urban housing renovation. Journal of Urban Economics, 54(3), 474–498. https://doi.org/10.1016/s0094-1190(03)00081-0

[23] Henningsen, A. (2020, August 4). Estimating Censored Regression Models in R using the censReg Package. Retrieved September 23, 2021, from https://cran.r-project.org/web/packages/censReg/vignettes/censReg.pdf.

[24] Hirshleifer, J. (1989). Conflict and rent-seeking success functions: Ratio vs. difference models of relative success. Public Choice, 63(2), 101–112.

[25] Holm, A., & Schulz, G. (2017). GentriMap: A Model for Measuring Gentrification and Displacement. Gentrification and Resistance, 251–277. https://doi.org/10.1007/978-3-658-20388-7$_1$0

[26] Laniyonu, A. (2017). Coffee Shops and Street Stops: Policing Practices in Gentrifying Neighborhoods. Urban Affairs Review, 54(5), 898–930. https://doi.org/10.1177/1078087416689728

[27] Manson, A., Schroeder, J., Van Riper, D., Kugler, T., & Ruggles, S. IPUMS National Historical Geographic Information System: Version 16.0 [dataset]. Minneapolis, MN: IPUMS. 2021. http://doi.org/10.18128/D050.V16.0

[28] O'Sullivan, A. (2005). Gentrification and crime. Journal of Urban Economics, 57(1), 73–85. https://doi.org/10.1016/j.jue.2004.08.004

[29] Papachristos, A. V., Smith, C. M., Scherer, M. L., & Fugiero, M. A. (2011). More Coffee, Less Crime? The Relationship between Gentrification and Neighborhood Crime Rates in Chicago, 1991 to 2005. City & Community, 10(3), 215–240. https://doi.org/10.1111/j.1540-6040.2011.01371.x

[30] Polo, M. (2005). Internal cohesion and competition among criminal organizations. In The economics of organized crime (pp. 87–109). essay, Cambridge University Press.

[31] Roth, J. (2019). Pre-test with Caution: Event-study Estimates After Testing for Parallel Trends. Department of Economics, Harvard University, Unpublished Manuscript.

[32] Skaperdas, S. (1992). Cooperation, Conflict, and Power in the Absence of Property Rights. The American Economic Review, 82, 720–739.

[33] Skaperdas, S. (1996). Contest success functions. Economic Theory, 7(2), 283–290. https://doi.org/10.1007/bf01213906

[34] Skaperdas, S. (2001). The political economy of organized crime: providing protection when the state does not. Economics of Governance, 2(3), 173–202. https://doi.org/10.1007/pl00011026

[35] Skarbek, D. (2016). Covenants without the Sword? Comparing Prison Self-Governance Globally. American Political Science Review, 110(4), 845–862. https://doi.org/10.1017/s0003055416000563

[36] Smith, C. M. (2012). The Influence of Gentrification on Gang Homicides in Chicago Neighborhoods, 1994 to 2005. Crime & Delinquency, 60(4), 569–591. https://doi.org/10.1177/0011128712446052

[37] Sobrino, F. (2019). Mexican Cartel Wars: Fighting for the U.S. Opioid Market.

[38] Strumpf, E. (2011). Medicaid's effect on single women's labor supply: Evidence from the introduction of Medicaid. Journal of health economics, 30(3), 531-548.

[39] Varese, F. (2010). What is organized crime?, In Organised Crime: Critical Concepts in Criminology (Vol. 1, pp. 11–33). introduction, Routledge.

10 Declarations

Funding: Not applicable

Conflicts of interest/Competing interests: Not applicable

Availability of data and material: All data is freely available from OpenData Philly and the National Historical Geographic Information System. Code and cleaned data is available upon request.

Code Availability: All econometrics were conducted in R. Code utilized by the author is available upon request.

Authors' Contributions: Not applicable

Acknowledgements: I would like to acknowledge the guidance offered by my advisor, Dr. Bryan McCannon, the inside insight into Philadelphia's illegal opiate markets provided by John Izzi and Dillon Kinsey, and the helpful feedback and suggestions provided at West Virginia University's Brown Bag internal seminar series.

11 Supplemental Appendix

11.1 Alternate Specifications

11.1.1 Drug Block Definition

Here, the definition of a high drug crime block is altered. Now, blocks with a number of drug incidents that is greater than two standard deviations above the city wide mean are labeled as high drug crime blocks.

	Dependent variable:
	Neighbor Block Shootings
Drug Block Gentrification Indicator	0.10***
	(0.023)
Non-Drug Block Gentrification Indicator	0.01*
	(0.005)
Block Level Fixed Effects	✓
Year Fixed Effects	✓
Mean	0.11
Observations	6,027,120
Adjusted R^2	0.13
AIC	7512461
Standard Errors	Clustered at Census Block Group Level
Note: Standard errors in parenthese	*p<0.1; **p<0.05; ***p<0.01

Table 21: Output from disaggregation of the primary specification differences-in-differences estimator with drug block classification standard lowered

11.1.2 Gentrification Treatment Definition

Here, the definition of the gentrification treatment is altered. The demographic classification criteria remains the same, but the real estate component is changed. Now, to be labeled as gentrifying, the block in question must either have a change from the previous year in the number of new construction or building alteration permits issued that is more than two standard deviations above the city-wide mean, have a change in the number of home sales that is greater than two standard deviations above the city-wide mean, or have both a change in the number of permits issued and a change in the number of home sales that each exceed one standard deviation above the city-wide mean.

	Dependent variable:
	Neighbor Block Shootings
Gentrification Indicator	0.02***
	(0.005)
Block Level Fixed Effects	✓
Year Fixed Effects	✓
Mean	0.11
Observations	6,027,120
Adjusted R^2	0.13
AIC	7512830
Standard Errors	Clustered at Census Block Group Level
Note: Standard error in parentheses	*p<0.1; **p<0.05; ***p<0.01

Table 22: Main differences-in-differences specification with classification standard lowered

	Dependent variable:
	Neighbor Block Shootings
Drug Block Gentrification Indicator	0.12***
	(0.028)
Non-Drug Block Gentrification Indicator	0.09*
	(0.004)
Block Level Fixed Effects	✓
Year Fixed Effects	✓
Mean	0.11
Observations	6,027,120
Adjusted R^2	0.13
AIC	7512359
Standard Errors	Clustered at Census Block Group Level
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 23: Output from disaggregation of the primary specification differences-in-differences estimator with gentrification classification standard lowered

11.1.3 Drug Block and Gentrification Definitions

Here, both gentrification and drug block classification standards are lowered in the manner described previously.

	Dependent variable:
	Neighbor Block Shootings
Drug Block Gentrification Indicator	0.10***
	(0.022)
Non-Drug Block Gentrification Indicator	0.01***
	(0.004)
Block Level Fixed Effects	✓
Year Fixed Effects	✓
Mean	0.11
Observations	6,027,120
Adjusted R^2	0.13
AIC	7512211
Standard Errors	Clustered at Census Block Group Level
Note: Standard errors in parentheses	*p<0.1; **p<0.05; ***p<0.01

Table 24: Output from disaggregation of the primary specification differences-in-differences estimator with both gentrification and drug block classification standards lowered

11.1.4 Removing Year 2020 Observations

	Dependent variable:
	Neighbor Block Shootings
Gentrification Indicator	0.009*
	(0.005)
Block Level Fixed Effects	✓
Year Fixed Effects	✓
Mean	0.10
Observations	5,424,408
Adjusted R^2	0.13
AIC	6372697
Standard Errors	Clustered at Census Block Group Level
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 25: Main differences-in-differences specification with year 2020 observations removed.

	Dependent variable:
	Neighbor Block Shootings
Drug Block Gentrification Indicator	0.07***
	(0.02)
Non-Drug Block Gentrification Indicator	0.004
	(0.005)
Block Level Fixed Effects	✓
Year Fixed Effects	✓
Mean	0.10
Observations	5,424,408
Adjusted R^2	0.13
AIC	6372557
Standard Errors	Clustered at Census Block Group Level
Note: Standard errors in parentheses	*p<0.1; **p<0.05; ***p<0.01

Table 26: Output from disaggregation of the primary specification differences-in-differences estimator with year 2020 observations removed.

11.2 Alternate Clustering of Standard Errors

	Dependent variable:			
	Neighbor Blocks Shootings			
Gentrification Indicator	0.02*	0.02***	0.02***	0.02
	(0.01)	(0.003)	(0.002)	(0.013)
Block Level Fixed Effects	✓	✓	✓	✓
Year Fixed Effects	✓	✓	✓	✓
Mean	0.11	0.11	0.11	0.11
Observations	6,027,120	6,027,120	6,027,120	6,027,120
Adjusted R^2	0.13	0.13	0.13	0.13
AIC	7512901	7512901	7512901	7512901
Standard Error Clustering	Police District	Block	Block Link	None
Note:	*p<0.1; **p<0.05; ***p<0.01			

Table 27: Output from primary difference and difference specification with various levels of standard error clustering

	Dependent variable:			
	Number of Neighbor Blocks Shootings			
Drug Block Gentrification Indicator	0.12**	0.12**	0.12***	0.12**
	(0.050)	(0.003)	(0.011)	(0.047)
Non-Drug Block Gentrification Indicator	0.01	0.01***	0.01***	0.01
	(0.009)	(0.016)	(0.002)	(0.011)
Block Level Fixed Effects	✓	✓	✓	✓
Year Fixed Effects	✓	✓	✓	✓
Observations	6,027,120	6,027,120	6,027,120	6,027,120
Adjusted R^2	0.13	0.13	0.13	0.13
AIC	7512571	7512571	7512571	7512571
Standard Error Clustering	Police District	Block	Block Link	None
Note:	*p<0.1; **p<0.05; ***p<0.01			

Table 28: Output from disaggregating difference and difference coefficient with various levels of standard error clustering