Re-Examining the Law of Crime Concentration: Between- and Within-City Evidence*

Aaron Chalfin
University of Pennsylvania

Jacob Kaplan
University of Pennsylvania

Maria Cuellar
University of Pennsylvania

February 7, 2020

Abstract

Objectives: In his 2014 Sutherland address to the American Society of Criminology, David Weisburd demonstrated empirically that the share of crime that is accounted for by the most crime-ridden street segments is notably high and strikingly similar across cities, an empirical regularity that Weisburd refers to as the "law of crime concentration." We build upon recent work in this area that points out that the concentration of a large share of crime amongst a small number of street segments can, in some cases, be a mechanical artifact of the degree of crime density in a city rather than the signature of an empirical law.
Methods: Using data from three of the largest cities in the United States, we identify crime concentration by comparing observed crime concentration to a counterfactual distribution of crimes generated by randomizing crimes to street segments. We show that this method avoids a key pitfall that causes existing methods of measuring crime concentration to overstate the degree of crime concentration in a city.
Results: Most (but not all) crimes are, in fact, concentrated amongst a small number of hot spots but the precise relationship is weaker — sometimes considerably so — than has been documented in the empirical literature. We further show that, within a city, crime is least concentrated in the neighborhoods that experience the largest number of crimes. Accordingly, the law of crime concentration sometimes holds only tenuously in the communities in which crimes are most prevalent.
Conclusions: We conclude that Weisburd's original intuition remains largely intact though the law of crime concentration requires some qualification.

Keywords: Criminology of place, hot spots, microgeography

*We thank John MacDonald for helpful comments on a previous version of this manuscript. Of course, all remaining errors are our own. Corresponding author: Aaron Chalfin, Department of Criminology, 558 McNeil Building, University of Pennsylvania, E-Mail: `achalfin [at] sas.upenn.edu`.

1 Introduction

A large and growing literature in criminology documents the importance of place — in particular, microgeographic places like street segments — one of the two faces of a standard city block — in explaining crime. Across a large number of places and in a variety of contexts, crime is found to be highly concentrated (Sherman et al., 1989; Eck et al., 2007; Weisburd, 2015; Andresen et al., 2017) and persistent over time (Weisburd et al., 2009; Gorr and Lee, 2015). Taken as a whole, the substantial geographic concentration of crime, particularly violent crime, suggests that the social and physical features of the urban landscape might potentially play an important role in the crime production function and therefore that crime hot spots are an appropriate target over which a social planner can focus resources and ultimately intervene.

The empirical regularity that crime is highly spatially concentrated has been central to the study of criminal justice policy and has promulgated a number of important research literatures that have become a mainstay of empirical criminology including a large literature on hot spots policing (Weisburd and Green, 1995; Sherman and Weisburd, 1995; Braga, 2001; Braga and Bond, 2008; Weisburd and Telep, 2014; Braga et al., 2014) and the equally important literature on the importance of environmental design including research on restoring vacant lots (Branas et al., 2011; Garvin et al., 2013; Bogar and Beyer, 2016; Kondo et al., 2016; Branas et al., 2018; South et al., 2018; Moyer et al., 2019), reducing physical disorder (Kelling et al., 1982; Keizer et al., 2008; Skogan, 2012; Braga et al., 2015) and improving ambient lighting (Farrington and Welsh, 2002; Welsh and Farrington, 2008; Doleac and Sanders, 2015; Chalfin et al., 2019).

In his 2014 Edmund H. Sutherland address to the American Society of Criminology, David Weisburd summarized the research on the importance of place and noted that places have been studied far less by criminologists than other natural units of analysis. Weisburd further notes the extent to which crime is concentrated among the most crime-ridden street segments is remarkably consistent across cities and proposes that this empirical regularity is sufficiently strong to be characterized as a "law of crime concentration."[1] Incredibly, across eight cities of varying sizes, the top one percent of street segments, ranked by crime incidence, accounted for approximately 25 percent of crimes in that city and the top 5 percent of street segments accounted for half of the

[1]In Weisburd's own words, "for a defined measure of crime at a specific microgeographic unit, the concentration of crime will fall within a narrow bandwidth of percentages."

crimes. The stability of these estimates is noteworthy and forms the basis for the claim that this pattern can be characterized as a law.[2]

Despite the abundance of research inspired by the law of crime concentration, recent scholarship has raised a number of key measurement issues in how crime concentration should be measured (Bernasco and Steenbeek, 2017; Hipp and Kim, 2017; Levin et al., 2017). In particular, past research notes that the fact that a small share of street segments accounts for a large share of the crime over a given time period does not necessarily mean that crime must be concentrated. To see this, consider that for an uncommon crime such as homicide, the empirical relationship that is documented in Weisburd (2015) can be substantively replicated through a process in which crimes are randomly assigned to street segments. The linchpin of this claim is that even in the cities with most the challenging crime problems, the number of street segments far exceeds the number of homicides known to law enforcement over any reasonable time window. For instance, consider a city like New York in which there are approximately 120,000 street segments and 300 homicides annually. In this case, it is trivial to see that, even if each homicide occurs on a different street segment (thus, by definition, there would be no concentration of crime), 0.25 percent of the street segments would account for 100 percent of the homicides.[3]

Thus, using the standard metric of crime concentration, the extent to which at least some types of crimes are concentrated will be biased upward. Similarly, the standard metric does now allow for a comparative analysis of concentration among different types of crimes since rarer crimes will, for mechanical reasons, appear to be more concentrated than more common crimes. Recent scholarship has proposed modifications to the law of crime concentration that address these concerns (Hipp and Kim, 2017; Levin et al., 2017). A particularly clever and ubiquitous approach proposed by Levin et al. (2017) is to measure crime concentration *only among street segments that experienced at least one crime*. The idea behind this approach is that crimes can only be concentrated where they, in fact, occur. This modification to the measurement of crime concentration does tend to reduce the degree of the bias in the standard measure of crime concentration but, as we show, this approach does not address the issue completely. Indeed, in most empirical applications, conditioning on

[2]As Weisburd notes, there is a great deal of evidence that remains to be documented – for instance, whether the law of crime concentration holds over a larger number of cities and, critically, whether there are circumstances in which the law of crime concentration does not apply.

[3]Even over a period of ten such years, if every homicide occurred on a different street segment, we would observe that just 2.5 percent of the street segments account for 100 percent of the homicides.

crime-free street segments will continue to lead to a substantial overestimate of the extent to which crimes are concentrated.

In this article, we propose a different way to measure crime concentration that fully addresses the concerns outlined above. Building upon insights in Levin et al. (2017) and Hipp and Kim (2017), we motivate a metric that compares the actual spatial distribution of crimes to a counterfactual condition in which crimes are not concentrated by construction. Specifically, our metric compares actual concentration — for instance, the share of street segments accounting for 25 percent or 50 percent of the crimes — to a counterfactual level of crime concentration that is constructed by randomly assigning crimes to street segments, with replacement. This randomization process generates a spatial distribution of crime across the street segments of a city that is not the result of concentration.

Using our proposed metric and data from New York City, Chicago and Philadelphia, three of the five largest cities in the United States, we show that while most types of crimes exhibit considerable concentration, the degree to which crimes are actually concentrated is smaller than has been supposed in prior literature. For rare crimes such as auto theft and robbery, we find that the law of crime concentration holds only to a limited degree.[4] While the stability of crime concentration across our cities is variable for individual crime types, across all crime types, there is an incredible degree of stability suggesting that the law of crime concentration holds — only at a lower intensity than has been previously documented.

We further extend empirical testing of the law of crime concentration to neighborhoods within cities, asking whether the law of crime concentration holds at the sub-city level. This analysis takes the perspective of a district commander in a police department who is charged with deciding how to allocate resources at the community level. We present novel evidence that the law of crime concentration does not hold equally among neighborhoods. While crime is highly concentrated in the safest neighborhoods, the highest crime communities — those which receive the greatest amount of police resources and which are in the greatest need of effective intervention strategies — are precisely the neighborhoods in which crimes are *least* concentrated.

[4] The law of crime concentration for auto thefts does hold in a statistical sense. Specifically, actual concentration for these crimes falls outside the range that is observed in Monte Carlo simulation. However, as noted by Hipp and Kim (2017), given the stability of estimates arising from Monte Carlo simulation, this is a particularly weak test of crime concentration.

The remainder of the article is organized as follows. In Section 2, we provide further context for our contribution by considering the recent literature that has proliferated since Weisburd (2015), paying particular attention to related contributions which propose modifications to the measurement of crime concentration. In Section 3, we lay out a framework for developing a corrected metric of crime concentration. Section 4 provides a description of our data. Section 5 provides an empirical assessment of the extent to which the law of crime concentration holds using the framework we have proposed. Section 6 summarizes our contributions, namely the novel method we propose for measuring crime concentration and the intuition behind our resulting calculations of crime concentration.

2 Prior Literature

2.1 Empirical Evidence on Crime Concentration

As noted by Weisburd (2015), the term "criminology of place" can be traced back to a 1989 article in *Criminology* by Sherman et al. (1989) which was among the first endeavors to systematically measure the concentration of crime among microgeographic areas. However, the recognition that a large share of crime is clustered in a small share of places is an observation that is nearly as old as modern cities (Quetelet, 1831; Weisburd et al., 2009). Over the last few decades, a literature has proliferated to establish that crime is both highly concentrated among a small number of crime hot spots (Eck et al., 2007; Weisburd, 2015) and that these hot spots, at least to an extent, persist over time (Weisburd et al., 2004, 2009; Gorr and Lee, 2015). Research has found that this pattern is not limited to low-impact crimes and applies equally, if not more forcefully, to some of the most costly criminal activity including gun crimes (Braga et al., 2010) and robbery (Braga et al., 2011).[5] Since Weisburd's influential 2015 article, a small but growing literature, buoyed by a 2017 special issue on the criminology of place in the *Journal of Quantitative Criminology*, has developed to further test and clarify the law of crime concentration and the extent to which it holds across time, place, and types of criminal activity. In this section, we summarize the recent literature and lay the groundwork for understanding how our paper helps to reconcile several remaining disconnects

[5] These findings are subject to important criticisms regarding the measurement of crime concentration by Hipp and Kim (2017), Levin et al. (2017) and Curiel et al. (2018) which we discuss in Section 2.2.

that remain in the literature.

In his 2014 Sutherland address, Weisburd notes that he uses a convenience sample of cities and reminds us of the importance of ascertaining whether the law of crime concentration applies more broadly across a large sample of places. In response, recent scholarship documents robust evidence that the law of crime concentration substantively holds in other U.S. cities including Chicago (Schnell et al., 2017), Seattle (Hibdon et al., 2017), St. Louis (Levin et al., 2017) and a large number of cities in California (Hipp and Kim, 2017), in a number of non-U.S. cities including Vancouver, Canada (Andresen et al., 2017) Milan, Italy (Favarin, 2018) and among various cities in the United Kingdom (Oliveira et al., 2017) and Latin America (Ajzenman and Jaitman, 2016) as well as in a suburban setting — Brooklyn Park, Minnesota (Gill et al., 2017). In every setting in which the law of crime concentration has been tested, the law, as proposed, holds up substantively.

A second set of papers consider whether the law of crime concentration holds equally for all crime types which is critical to understand given that different crimes will often require different policy responses.[6] This research concludes that, while some crime types are more concentrated than others, most crimes are concentrated to a reasonable degree among a small number of microgeographic units (Andresen et al., 2017). Building upon prior work by Braga et al. (2010) and Braga et al. (2011), Haberman et al. (2017) shows that a high degree of spatial concentration can be found among street crimes — in particular outdoor robberies. Similarly, Hibdon et al. (2017) show that the law of crime concentration is substantively replicated when an additional data source — 911 calls for emergency service — are used to explore the concentration of illegal drug activity.

A third set of papers consider whether the law of crime concentration holds among different geographic units and finds evidence that the majority of the variation in crime at the city-level is within-neighborhood rather than between-neighborhood variation, further bolstering the importance of micro- rather than macrogeography in explaining spatial variation in urban crime, (Steenbeek and Weisburd, 2016; Schnell et al., 2017).

Taken as a whole, the prevailing literature suggests that the following propositions are true:

1. Broadly speaking, the law of crime concentration holds, to a reasonable degree, in every city or non-urban setting in which crime concentration has been studied.

[6] Andresen and Linning (2012), among others, note that, for several reasons, it may not be desirable to aggregate individual crime types into aggregate crime.

2. Consistent with the instincts of a number of scholars in this area, crime concentration at the city level is not simply a function of the clustering of crime among high-crime neighborhoods — most of the variation in crime can be found at the street segment level.

3. The degree to which crimes are concentrated varies by crime type and, if anything, is too conservative for some of the most socially costly crimes such as robberies and assaults.

4. Trajectory analyses tend to find that the majority of street segments exhibit stable crime concentrations. However, there is sometimes considerable variability in which street segments are at the top of the distribution, depending upon the crime type.

2.2 Empirical Challenges to Measuring Crime Concentration

The four conclusions laid out in Section 2.1 are based the standard metric for computing crime concentration that is a mainstay of the extant literature — the share of street segments that account for 25 or, alternatively, 50 percent of the crimes. However, we note that several related contributions have sought to clarify the conditions under which the standard metric proposed to measure crime concentration in Weisburd (2015) will accurately characterize the degree to which crimes are substantively concentrated. We summarize these critiques here.

As noted by Levin et al. (2017), Hipp and Kim (2017) and Curiel et al. (2018) among others, the chief challenge to interpreting the standard concentration metric is that this metric does not account for the fact that, unless crime data are very dense, crimes will tend to be disproportionately concentrated in a small number of places even if they are randomly distributed. That is, in the types of crime samples used in this research, even if crimes are randomized to places, this does not guarantee uniformity in the spatial distribution of crimes — i.e., that the top k percent of microgeographic places, ranked in descending order by crime density, will account for exactly k percent of crimes.[7]

The intuition behind this result, which we formalize in Section 3, is multi-faceted. But it is, in part, based on the idea that for uncommon crime types, the number of crimes relative to places

[7] A corresponding literature has tackled in this issue in the context of measuring concentration among criminal offending among a cohort of individuals. See e.g., Tseloni and Pease (2005).

is small and, accordingly, many places do not experience any crimes over a given time period.[8] As a result, it is often true, by construction, that a small number of the places will account for many — and even all — of the crimes in a given city. This issue is not merely academic since, in many cities, a large number of street segments do not experience crime over a given time period (Curman et al., 2015) and, as it turns out, this issue has enormous implications for the conclusions that are drawn about which crimes are concentrated and the extent to which they are. For instance, as noted in Hipp and Kim (2017), while the standard crime concentration metric suggests that violent crimes are more concentrated than property crimes, after correcting the problem identified above, there is clear evidence that the degrees of concentration among violent and property crimes are, in fact, similar.

The literature has proposed two ways of dealing with this problem each of which depends upon a corrected measure of crime concentration that is robust to the problem of crime-free places. First, Levin et al. (2017) and Andresen et al. (2017) propose a simple but important tweak to the standard measure of crime concentration. In particular, they call upon researchers to measure the share of crimes that occur among the top k percent of street segments, limiting the data to *the street segments that experienced at least one crime*. The intuition behind such a correction is straightforward: since many street segments do not experience any crime at all, these zero crime blocks will tend to make crime appear more concentrated than it actually is at the top of the distribution. Accordingly, they propose to focus on the segments in which crimes do occur. By addressing bias in measures of crime concentration that is an artifact of crime-free places, their proposed metric moves us closer to correctly estimating the extent to which crime is substantively concentrated. To see how this works, consider a city in which half of blocks do not receive crime. If 2 percent of all street segments account for one quarter of the crimes, then it will be the case that 4 percent of street segments *which experience non-zero crime counts* account for one quarter of the crimes. Thus, the standard concentration metric will be two times too small.

A second proposal comes from Hipp and Kim (2017) who note the same issue with the standard measure of crime concentration and propose a clever adjustment to the standard metric that

[8]This discussion is central to the exposition in Levin et al. (2017) who argue for a crime concentration metric that uses only crime-free street segments but, as we note in Section 3, crime-free segments are not the only threat to measuring crime concentration using the standard metric.

leverages temporal variation in the data to build a counterfactual expectation.[9] In particular, they suggest that we sort the street segments in a city based on the number of crimes of a specific type (from highest to lowest) in some base year, t-1, and then compute the share of crimes that occurred on the top 5 percent of these street segments, identified during the base year, in the current year, t. That is, among the top 5 percent of street segments in the prior street, what share of crimes occurred on these segments in the current year? Consistent with the intuition behind shrinkage estimators such as ridge regression (Hoerl and Kennard, 1970) and the LASSO (Tibshirani, 1996) in statistics and computer science, this method has the virtue of using cross-validation to train and test crime data on different samples, thus shrinking the resulting estimate of crime concentration towards its true value. However, as we note in the next section and as the authors themselves note, their proposed metric measures a concept that is substantively different from crime concentration.

2.3 Substantive Issues with Prior Corrections

In this section, we discuss the virtues and drawbacks of methods of measuring crime concentration that have been proposed by Levin et al. (2017) and Hipp and Kim (2017) and propose a simple alternative to each of their conceptions that we argue is ideally suited to the specific task of measuring crime concentration. Notably, our proposed metric avoids the problem of upward bias that is present in prior concentration metrics.

2.3.1 Hipp and Kim (2017)

Recall that Hipp and Kim (2017) propose to measure crime concentration in a given time period, t, by measuring the share of crimes that occurred on the top 5 percent of street segments, ranked according to crime density in year t-1. Here, we consider how this estimator works in practice and describe the virtues as well as the limitations of this measure for studying crime concentration.

We begin with a simple example. Consider a rare crime such as homicide, which, as Hipp and Kim (2017) note, is precisely the type of crime for which measuring concentration in the standard way is so problematic. Suppose that there are 10 homicides in a city with 1,000 street segments

[9] Hipp and Kim (2017) note that any law of crime concentration will require a definition that includes both (1) a true concentration component as well as (2) a statistical chance component. However, they are skeptical of the theoretical utility of pursuing this direction and thus do not pursue such a computation directly.

and that, in each year, these crimes occur on completely different street segments. In this world, using the standard (unadjusted) measure of crime concentration, we would compute that $\frac{10}{1{,}000}$ = 1 percent of the street segments account for 100 percent of the homicides, an estimate which clearly misrepresents the degree to which homicides are concentrated. Using the metric proposed by Hipp and Kim (2017), we would compute an adjusted concentration metric of 0 since all of the homicides occurred on different street segments in year t and, as such, the top 5 percent of street segments in year t-1 accounted for none of the homicides in year t. This appears to be a win for their proposed metric since, in fact, in this example, homicides are not concentrated.

However, next consider the unlikely but nevertheless instructive scenario in which the homicides in year t occurred on precisely the same 10 street segments as in year t-1. In other words, there is perfect persistence among the high-crime blocks. Using the metric of Hipp and Kim (2017) we would compute that the top 5 percent of street segments, ranked in year t-1 account for 100 percent of the homicides in year t. The implication is that homicides are highly concentrated. But homicides are, in fact, not concentrated — indeed if homicides were randomly assigned to street segments, we would see the same degree of crime concentration — that a very small share of the street segments account for all of the homicides.

What is going on here? The method proposed by Hipp and Kim (2017) jointly addresses two critically important but distinct issues: the extent to which crimes are concentrated in hot spots and the extent to which crime hot spots persist over time. This, we believe, is a critical distinction. Crime concentration merely describes the extent to which a small share of places substantively account for a large share of crime. Persistence, on the other hand, is about the *predictability* of crime and folds in the extent to which relatively fixed features of the built or social environment lead to stable crime hot spots. To see why this is a critical distinction, consider, for instance, a scenario in which a big box store was to move to a new location, thus shifting crimes like retail theft. We would probably expect to see the crime hot spots change as a result of the store's move, but we might not expect to see a change in the degree to which crimes are concentrated. Thus, a critical difference between concentration and persistence is that the former concept allows for human activity to change over time in response to exogenous shocks.

We thus conclude that while Hipp and Kim (2017)'s proposed metric shrinks measures of crime concentration in practice, this will tend to be an artifact of a lack of spatial persistence in crime rather than due to the non-uniformity of spatial crime distributions. This issue — that their metric does not explicitly account for the non-uniformity problem — is, in fact, noted by Hipp and Kim (2017) themselves (p. 624). However, it is worth noting that their metric has other considerable virtues that extend beyond measuring crime concentration. By conflating crime concentration and persistence, Hipp and Kim (2017)'s metric is, in a number of ways, ideally suited as a means of assessing the practical usefulness of identifying crime hotspots.[10]

2.3.2 Levin, Rosenfeld and Deckard (2017)

A second proposed metric is laid out in Levin et al. (2017) (hereafter "LRD") who propose a simple fix to the problem of non-uniformity in the spatial distribution of crime: re-estimate the share of street segments that account for k percent of crimes, *using only segments that experienced non-zero crime*. There is great virtue to this proposal — it is simple, easy to compute and understand and it does directly implicate the non-uniformity problem. However, as we demonstrate, their method will yield a metric of crime concentration that is biased upward – in some cases, considerably so. As we explain, the limitation to their approach is that street segments which experience zero crimes is not the sole reason why uniformity does not hold when crimes are assigned, at random, to street segments.

We begin by characterizing what LRD's proposed metric means for the counterfactual level of crime concentration that is expected to be found in the absence of concentration. The implication of removing zero crime street segments to correct the non-uniformity problem is that, in the absence of any crime concentration, their measure of crime concentration should be 1. That is, the

[10] We further note that the metric proposed by Hipp and Kim (2017) can be tweaked to instead measure the share of crimes that took place on the street segments that were among the top k percent of street segments in the prior year. For instance, we might wish to understand what share of crimes are accounted for by the street segments that accounted for one quarter of the crimes in the prior year. The extent to which this share is smaller than one quarter provides a sense for the lack of persistence among hot spots. This variant of their proposed measure offers several advantages. First, it is arguably closer in spirit to the unadjusted crime concentration metric. Second, it is more appropriate for rare crimes types such as homicide. Consider that, for a rare crime, there are typically very few crimes relative to street segments. Thus, it might be the case that fewer than 5 percent of street segments will account for all of the homicides. The analyst must then choose which zero-crime street segments to include among the top 5 percent of street segments. In cases where the software program used determines which zero-crime street segments are included, if it does not include them randomly, substantial bias could be introduced. For example, if the software uses the original order the data was in as a tie-breaker for zero-crime streets, any bias in that order, such as street segment shapefiles that are arranged geographically as some city's data are (e.g. The furthest west street is the first segment in the data and each consecutive street is further east), would be present in the top 5 percent of street segments.

top k percent of street segments that have non-zero crime should account for exactly k percent of the crimes i.e., 25 percent of the street segments will account for 25 percent of the crimes, 50 percent of the street segments will account for 50 percent of the crimes, etc.

As it turns out, this standard — that of uniformity — is an overly stringent standard that ultimately leads to an overestimate of the degree to which there is crime concentration. To see this consider again a simple example, this time involving a city in which there are 1,000 street segments and 100 crimes. Using the standard measure of crime concentration, we would compute that $\frac{100}{1,000}$ = 10 percent of blocks account for 100 percent of the crimes. Using LRD's proposed metric, what would zero concentration look like? Zero concentration would hold if each crime occurred on a different street segment, as would be required under uniformity. Accordingly, using the metric suggested by Levin et al. (2017), we would compute that $\frac{100}{100}$ = 100 percent of the blocks with crime account for 100 percent of the crimes and therefore that crime is not concentrated.

However, we point out that a scheme in which crimes are randomized to street segments is unlikely — in fact, very unlikely — to produce the result that all 100 crimes occurred on different street segments. As a result, when the LRD metric is applied to a dataset in which there is zero crime concentration by construction, it will indicate a positive amount of crime concentration. We show this using a simple simulation and later, in Section 5, we present evidence on the degree to which LRD's metric yields an overestimate of crime concentration in empirical data.

We simulate the random assignment of M crimes to n street segments with replacement for a hypothetical city, re-sampling 1,000 times.[11] In each trial, we store up the number of unique street segments in which at least one crime occurs, beginning with the example proposed above in which there are n = 1,000 street segments and M = 100 crimes. In this example, uniformity would require that each of the crimes occurs on a different street segment so that 100 street segments experience crimes and 900 street segments do not. However, in the simulated data, there is a 99.3 percent chance that at least two of the 100 crimes that occur in this city will occur on the same street segment. This is an artifact of the fact that we are sampling with replacement, as we must.

In the above example, it is very likely that at least one block will experience multiple crimes. There is likewise a 1 in 3 chance that at least six segments will experience multiple crimes.[12] What

[11]Re-sampling 10,000 times does not substantively change the results.

[12]Replication code, written using `Stata 15.0` is available upon request to the authors.

this means in practice is that data which were generated at random will yield evidence of crime concentration 99.7 percent of the time and substantial evidence of crime concentration perhaps as often as one third of the time. This problem is more severe when crimes are more common. For example, if the number of crimes is 300 among 1,000 blocks, then, in 95 percent of randomized simulations, the number of unique blocks experiencing crime will be between 251 and 268. A value of 300 — indicating that all 300 crimes occurred on different blocks (which is the measure of zero concentration that is required under Levin et al. (2017)'s test) is extremely rare and occurs at a rate of less than one in a million trials.

Using the method proposed by Levin et al. (2017), in 95 percent of random simulations, we would compute crime concentration equal to between $\frac{300}{251}$ and $\frac{300}{268}$ or between 1.07 and 1.2. Thus, While the method proposed by Levin et al. (2017) is reasonable and produces results that are directionally consistent, their proposed metric will nevertheless result in an *overestimate* of crime concentration. As we show in Section 5, the extent of the upward bias is not trivial.[13]

3 Perfecting The Measurement of Concentration

In this section, we use randomization to propose a simple but, we argue, optimal way to identify the extent to which crimes are spatially concentrated. [14] We begin with a simple example and lay out our proposed framework. Consider a city that has n street segments and experiences j crimes where $n >> j$. Let's say that we are interested in understanding the extent to which homicides are concentrated amongst a city's street segments. For instance, in New York City, $n = 119{,}000$ and, as of 2018, $j = 295$. Even if homicides are not concentrated at all — meaning that each of the city's 295 homicides occurred on a different block — it will be the case that $\frac{295}{119{,}000}$ or 0.25 percent of the street segments account for 100 percent of the crimes.

Clearly the standard crime concentration metric — the share of street segments that account for one quarter of the homicides — is not useful in this scenario. The question then is: In the absence

[13]A second implication of computing crime concentration using the metric proposed by LRD is that, in most applications, crime concentration will be greater in datasets that contain more years of data and thus have greater density of crimes per street segment. We show this empirically for our three cities in **Appendix Figure 1**.

[14]We note that we are not the first to propose that simulating the random assignment of crimes to street segments is of value in this domain. Indeed both Levin et al. (2017) and Hipp and Kim (2017) utilize simulation to elucidate the importance of a counterfactual in interpreting crime concentration statistics. However, neither paper utilizes randomization to generate a measure of marginal crime concentration.

of concentration, what share of street segments *should* account for one quarter of homicides? The extant literature suggests that we ought to expect uniformity — that is, k percent of street segments account for k percent of crimes. The contribution of Levin et al. (2017) is simply to note that we should expect to see uniformity only when we filter out crime-free segments.

But is uniformity the correct counterfactual? We consider the conditions under which this will be the case by running a simple simulation exercise. Consider a fictional city which has 1,000 street segments and a thought experiment in which the following number of crimes are assigned, at random with replacement, to these 1,000 blocks: 50, 100, 500, 1,000, 5,000, 10,000, 50,000, 100,000 and 1,000,000. What share of crimes would we expect to see represented among the top 25 percent of street segments, ranked according to the number of crimes experienced? Of course, under uniformity, we would expect that 25 percent and 50 percent of street segments to account for 25 percent and 50 percent of the crimes, respectively.

We present the results of the simulation exercise in **Figure 1**. In Figure 1, Panel A plots the share of all street segments and corresponds to the unadjusted measure of crime concentration that is the mainstay of the empirical literature. Panel B plots the share of crime concentration among blocks that actually experience crime, as suggested by Levin et al. (2017) among others. In each panel, we plot the share of street segments accounting for 25 percent of the crimes using the dashed gray line and the share of street segments accounting for 50 percent of the crimes using the dashed black line. Horizontal reference lines are drawn at both 25 and 50 percent along the y-axis and represent the levels of crime concentration at which uniformity is achieved.

We begin our discussion with Panel A. Here, we see that when crime density is low relative to street segments (e.g. j = 50 crimes amongst 1,000 segments), a very small share, approximately 1.2 percent of street segments, ranked by crime density, account for one quarter of the crimes. Likewise, just 2.4 percent of the street segments account for half of the crimes. As crimes become more common, each measure of crime concentration increases. When the number of crimes is 1,000 equaling the number of street segments, we see that 7.5 percent and 20 percent of street segments account for one quarter and one half of the crimes, respectively. At 10,000 crimes — or 10 crimes per street segment, approximately 15 percent of segments account for one quarter of the crimes and approximately 37 percent of segments account for one half of the crimes. At 1,000,000 crimes — 1,000 per street segment — uniformity is roughly met. As the number of crimes becomes infinite, uniformity will be achieved asymptotically. However, for relatively uncommon crimes or common crimes that are measured over a reasonably short window (e.g., one or two years), such an asymptotic result is unlikely to hold and, as such, an unadjusted measure of crime concentration will overstate the extent to which crimes are concentrated. Next, we turn to Panel B which considers the performance of LRD's suggested solution to the non-uniformity problem in crime data. At very low crime densities, LRD's metric performs admirably. Conditioning on non-zero crime street segments leads to near-uniformity at 50 crimes for 1,000 street segments — here, 24.4 percent of the street segments account for one quarter of the crimes and 49 percent of the street segments account for 50 percent of the crimes. Likewise, their metric performs well asymptotically — though, of course, so does the unadjusted metric. However, LRD's metric performs far less well in the middle of the crime density distribution where we have between 1 and 100 crimes per street segment. For instance, at 1,000 crimes or 1 crime per street segment, we see that 12 percent of the segments account for one quarter of the crimes and 34 percent of the segments account for one half of the crimes. These figures are between one third and one half smaller than uniformity and the result is that crime concentration will be overestimated by between one third and one half. Likewise, at 10,000 crimes or 10 crimes per street segment, we see that 16 percent of segments account for one quarter of the crimes and approximately 39 percent of segments account for one half of the crimes. Incredibly, at these intermediate densities LRDs proposed solution, removing crime-free blocks, performs only marginally better than the unadjusted metric. Since this window (between 1 and 10 crimes per street segment) is an extremely common density among the data that has been studied in the extant literature, the scope for LRD's method to overstate crime density is unfortunately quite high.[15]

The thought experiment presented in Figure 1 makes clear that uniformity is an asymptotic result and does not hold in most applications. We further see that removing the zero crime street segments does not substantively correct this issue at most crime densities. We thus propose a "corrected" metric of crime concentration that allows us to quantify the *marginal* degree of crime concentration above and beyond that which would be expected as an artifact of the density of the

[15] In our data which spans between 10 and 15 years in three of the largest cities in the United States, the number of overall crimes per street segment varies between 35 and 115. Individual crime types are far less dense and vary between 0.5 and 10.

crime data:

$$mcc^k_{ij} = cc^k_{ij} - cc^k_{ij} \tag{1}$$

In (1), mcc^k_{ij} represents the marginal crime concentration in city i for crime type j and crime share k, where, for our purposes, k = 25 or 50 percent. cc^k_{ij} is the crime concentration that is actually experienced in city i (i.e., the measure proposed by Weisburd) for crime type j and cc^k_{ij} is the crime concentration obtained under randomization with replacement. Note that every randomization will lead to different results. Thus, cc^k_{ij} will, in practice, be the mean crime concentration across a large number of trials.[16] For a given value of mcc^k_{ij}, the larger the value of mcc^k_{ij}, the greater the degree of true crime concentration. Consider, for instance, a crime for which cc^{25}_{ij} = 4 percent and cc^{25}_{ij} = 10 percent. What this means is that, under the randomization of crimes to street segments, we would expect 10 percent of street segments to account for one quarter of the crimes. In reality, only 4 percent of street segments accounted for one quarter of the crimes. Hence, mcc^{25}_{ij} = 10 percent - 4 percent = 6 percent. Accordingly, the marginal share of blocks needed to account for one quarter of the crimes under randomization is 6 percent. Critically, unlike the standard crime concentration metric, higher *marginal* crime concentration indicates that crime is more concentrated. In Section 5, we estimate mcc^{25}_{ij} and mcc^{50}_{ij} for a variety of different crime types for each of our three cities: New York City, Chicago and Philadelphia.

4 Data

We derive corrected estimates of the degree of marginal crime concentration using public crime microdata from three of the five largest cities in the United States: New York City (January 1st 2006 - December 31st, 2018), Chicago (January 1st, 2001 - May 4th, 2019) and Philadelphia (January 1st, 2006 - May 11th, 2019). We focus on these three cities because crimes from a fourth large city — Los Angeles — are coded primarily to intersections rather than street segments.[17]

The data correspond to all crimes known to the city's municipal law enforcement agency and are

[16] As we document later, in practice, there is little meaningful variation among trials in cc^{k*}_{ij}.

[17] The city of Houston does not provide a shapefile of the city's street segments, excluding that city from the analysis.

geolocated allowing us to extract the street on which the crime occurred.[18] The data also provide details on the type of offense, which we use to examine five categories on crime in addition to total crimes: murder, robbery, assault (simple and aggravated), motor vehicle theft, and larceny/theft.

In keeping with prior literature, we assign crimes to street segments by creating a 50-foot buffer around each street segment in the city and checking the location of each crime against these buffers to determine the street segment on which a given crime took place. Following Weisburd (2015), we drop any crime that occurs in an intersection (i.e. matches with two or more street segments) or does not match to any street segments. There are substantial differences in the number of crimes geocoded to a single street segment rather than an intersection between each city. For both New York City (71%) and Chicago (94.4%), the majority of crime incidents are located within 50 feet of only one street segment, significantly larger than Philadelphia's 41%.[19]

We continue our discussion of the data by presenting descriptive statistics on crime in our three cities. **Table 1** presents, for each of our cities, the number of street segments as well as the number of crimes in the complete data set and in 2018, the last full year of data available. The cities included in this study have a wide range in the number of street segments in the city, though in rough accordance to the population of each city. Philadelphia has slightly over 41,000 street segments, Chicago has about 56,000, and New York City has nearly 120,000.Chicago contains the largest number of crimes in our data, approximately 6.4 million, a function of the high crime rate in the city, the near complete matching of crime to a single street segment, and the fact that the available data extends as far back as 2001. Philadelphia and New York City have fewer crimes with 1 million and 4.6 million total crimes, respectively. While the total number of crimes differ between cities, the makeup of each city's crime is similar. Larceny in the most common crime in each city, consisting of between 21% (Chicago) and 27% (New York City) of crimes. In each city, murder is rare relative to other crimes, comprising just 0.1% of crimes reported in the city. These trends are roughly similar when examining crime that occurred in 2018, the last full year of data available.

Next, we consider crime concentration in each of our three cities, replicating the canonical figure from Weisburd (2015) which presents cc_{ij}^{25} and cc_{ij}^{50} for each of five large cities: Cincinnati

[18]Fewer than 1 percent of crimes in each city have missing coordinates.

[19]In Los Angeles, only 21% of crimes are located within 50 feet of a single street segment.

OH, New York, NY, Sacramento, CA, Seattle, WA and Tel Aviv-Yafo (Israel). These data are presented in **Figure 2**, Panel A (crime concentration = 25 percent) and B (crime concentration = 50 percent). The gray bars represent the original cities in Weisburd's convenience sample. The black bars represent the three large cities for which we have data. Note that NYC is in both samples — the estimates differ slightly insofar as the sample years are slightly different. In Weisburd's convenience sample, in general, between 1-2 percent of street segments account for 25 percent of the crimes and between 4 and 6 percent of the street segments account for 50 percent of the crime, depending on the city. In our very large cities, crime is a little bit less concentrated but not dramatically so.

Crime is most concentrated in NYC which is relatively safe — 1.2 percent of street segments account for one quarter of the crimes and 4.2 percent of street segments account for one half of the crimes. Crimes are less concentrated in Chicago and Philadelphia which have higher levels of crime. In Chicago 2.8 percent of segments account for one quarter of the crimes and 9.4 percent of the segments account for one half of the crimes. In Philadelphia, those numbers are 2.1 percent and 8.2 percent respectively. Hence, the empirical regularity documented in Weisburd (2015) appears to roughly hold in our sample of cities too. In the next section, we characterize the extent to which crimes are concentrated, relative to what we argue is the ideal counterfactual — that which is generated using randomization with replacement rather than uniformity. We also compare our preferred metric to those that have been proposed in the extant literature.

5 Results

5.1 Citywide Crime Concentration

We begin discussion of our findings by presenting an accounting of crime concentration in our three cities both using the standard (unadjusted) measure of crime concentration as well as what concentration would look like under the randomization of crimes to street segments. In **Figures 3A, 3B, and 3C**, we present these results for each of our cities using overall crimes as well as disaggregated crimes of following types: murder, robbery, assault, motor vehicle theft and larceny. Each figure has two panels: Panel A presents results for 25 percent concentration; Panel B presents results for 50 percent concentration. We begin with Figure 3A which presents the data for New York City. Consistent with computations presented in Weisburd (2015), in NYC, just over 1 percent of street segments account for one quarter of the crimes and just under 5 percent of street segments account for half of the crimes. As has been noted by many others, the law of crime concentration holds broadly for each type of crime.

Next, we consider the degree of crime concentration that is generated via randomization with replacement — this is shown by the gray bars. For overall crimes, we see that, under randomization, 15 percent of street segments account for one quarter of the crimes and approximately 35 percent of the street segments account for half of crimes. Thus, while crime is substantively concentrated, empirical crime concentration is between 7x (concentration at 50 percent of crime) and 14x (concentration at 25 percent of crime) greater than random chance rather than between 10-22x greater than random chance which is implied by the standard measure.

Next, we turn to robberies, a common street crime. Here, we see that the share of street segments that account for one quarter of the robberies in the empirical data — approximately 1 percent — is not dramatically different from the share of street segments in the simulated data — approximately 3 percent. Referring to our suggested computation of marginal crime concentration (mcc_{ij}^{k}), we would subtract cc_{ij}^{k} (1 percent) from cc_{ij}^{k} (3 percent) to obtain $mcc_{ij}^{k} = 2$ percentage points. In other words, actual crime concentration is just two percentage points smaller than would be expected under random assignment of crimes to street segments. Auto theft, like robbery, is concentrated to a small degree — by approximately 1 percentage point. Assaults and larcenies, on the other hand, are considerably more concentrated — by approximately 8-9 percentage points. Referring to Figures 3B and 3C, the data are substantively similar for both Chicago and Philadelphia.

We next compare our measure of marginal concentration to the marginal measure of crime concentration that is obtained using LRD's preferred method of computing the concentration of crimes, excluding crime-free street segments. Recall that, in Section 2.3 we claimed that LRD's method of removing zero crime segments would yield a measure of crime concentration that is biased upward, a prediction that is supported by the simulations summarized in Figure 1. Here, we present empirical evidence for this claim using data from New York City, Chicago and Philadelphia. Results are summarized in **Tables 2A and 2B**. These two tables describe concentration at 25 and 50 percent of crimes, respectively and have a parallel structure. We pause here to describe the structure of the tables. The tables report several key crime concentration metrics for each five crime types (murder, robbery, assault, motor vehicle theft and larceny) and aggregate crime in each of our three cities. Each table has five columns. The first column reports the proportion of street segments, when ranked in descending order of crime incidence, that account for k percent of each type of crime. This is the standard (unadjusted) measure of crime concentration referenced by Weisburd (2015). The second column reports the same quantity, conditioning on non-zero crime segments as proposed by Levin et al. (2017). The third column reports the same quantity in simulated data in which crimes are randomized to street segments, with replacement. The final two columns use the information in columns (1)-(3), to compute *marginal* crime concentration. Column (4) reports the measure of marginal concentration that we lay out in Section 3, equation (1). Column (5) reports the measure of crime concentration that is implied by the approach suggested by Levin et al. (2017).

As a reminder, our measure of marginal crime concentration compares the share of street segments that account for k percent of crime in the empirical data (column 1) to the share of street segments that account for k percent of crime in simulated data (column 3), subtracting the former from the latter. The resulting measure of marginal crime concentration is the share of street segments (in percentage points), relative to the random distribution that are needed to account for k percent of crimes. The measure of marginal crime concentration implied by LRD is given by:

$$k - nzcc_{ij}^{k} \tag{2}$$

where $nzcc_{ij}^{k}$ is the share of crimes that are accounted for by the top k percent of street segments with at least one crime for city i and crime type j. That is, by focusing on the crime-free blocks, LRD's measure of zero crime concentration corresponds to a situation in which there is perfect uniformity — that is, that k percent of street segments account for k percent of crimes.

We begin our discussion of Table 2A which corresponds with the share of segments that account for one quarter of crimes. Referring to Panel A which uses data from New York City, we see that 1.2 percent of street segments account for one quarter of the crimes in the raw data. Conditioning on segments which experienced at least one crime, one quarter of crimes accrued to the top 2.6 percent of street segments. In simulated data, one quarter of the crimes would accrue to the top 14.8 percent of segments. What does this imply for our measure of marginal crime concentration and for that of LRD? Using LRD's metric, we see evidence of appreciable crime concentration. Even removing crime-free street segments, a very small share of segments accounted for a disproportionate share of crime. This finding is reflected in their measure of marginal crime concentration of 22.4, indicating that crime concentration is 22.4 percentage points greater than is seen under uniformity —their implicit counterfactual. On the other hand, our measure of crime concentration is 13.7, indicating that crime concentration is 13.7 percentage points smaller than under the counterfactual of random assignment. Our measure of crime concentration is thus 50 percent smaller, a ratio that is consistent with the prediction that we might have made using the simulation exercise summarized in Figure 1.

Next, we turn to murder. In the raw data, 0.1 percent of street segments accounted for 25 percent of the murders in New York City. Conditioning out the crime-free blocks, just over one fifth of the street segments which experienced a murder explain one quarter of the murders. This yields a marginal crime concentration of 3.8 percentage points, using LRD's proposed metric. The implication is that, while murder is not very concentrated, it is to some degree. On the other hand, the share of street segments accounting for one quarter of murders is equal in the empirical and the simulated data implying that murders are not concentrated. The implication is that murders in NYC are not concentrated at all. However, due to the sparsity of murders in the data, caution should be exercised in drawing such a conclusion. The reason for our caution is that when crime data are very sparse, we will be underpowered to detect concentration.[20]

A similar story can be seen for auto theft. LRD find evidence of appreciable concentration for auto thefts when they condition on the segments that experienced at least one auto theft. However, once we account for the simulated distribution of crimes, we fail to see evidence of appreciable concentration — auto thefts are concentrated by only one percentage point more than what would occur via randomization. A similar story can be told for robbery which is concentrated but to a considerably smaller degree than is implied by LRD's proposed concentration metric.

[20] To see this, consider a society in which there are 1,000 street segments and 2 crimes. Even if those 2 crimes have an elevated probability of occurring on the same dangerous street segment, in most realizations of the data, those crimes will likely end up on different street segments thus implying that crimes are not concentrated.

Turning to assault and larceny, we see considerable evidence of crime concentration albeit less than has been measured in the prior literature. The measures are quite similar by city, especially for overall crime, thus providing support for the idea that crime concentration may well be highly stable across cities. The same relationship can be found in Table 2A which reports the share of street segments that account for half of the crimes.

We can also characterize the measure of crime concentration that is obtained using the method proposed by Hipp and Kim (2017). We present this information in **Table 3**. Column (1) presents Hipp and Kim's preferred metric, the current year's share of crimes among the top 5 percent of street segments, ranked using crime data from the prior year. Note that we computed their metric for each year and present averages across all of the years in our data. In column (2), we present a different variant of Hipp and Kim's metric: the share of crimes in the current year that are accounted for by the top 25 percent of street segments, ranked using crime data from the previous year. Column (3) presents our marginal crime concentration metric which first appeared in column (4) of Table 2B. Columns (4) and (5) present the same information for 50 percent crime concentration.

Across our three cities, we see that the top 5 percent of street segments, ranked in the prior year, account for an outsize share of crimes in the current year — those street segments account for 51 percent in New York City and just over 35 percent in Chicago and Philadelphia. The implication is that crime hot spots are, on the whole, persistent. However, there is considerable variation by crime type. For instance, in New York City, the Hipp-Kim metric for murder is just 4.4 percent; for auto theft it is 18.7 percent, implying that these crimes are less persistent. On the other hand, the Hipp-Kim metric for assaults, larcenies and robberies implies considerably more persistence. Patterns are similar in Chicago and Philadelphia.

Referring to column (2) of Table 3, we see yet more evidence for the persistence of overall crime hot spots — incredibly the street segments that accounted for one quarter of the crimes in the prior year continue to account for nearly one quarter of the crimes in the current year. The share is also very high for assault and, in two of our three cities, for larceny. On the other hand, there is little persistence among the murder hot spots; the street segments that accounted for 25 percent of the murders in a given year, on average, account for just 0.3-0.7 percent of murders in the following year.

Even though this variant of Hipp and Kim's metric measures something other than crime concentration, the statistic nevertheless compares favorably to our preferred approach of measuring marginal crime concentration. Using both approaches we draw the following conclusions. First, overall crime is concentrated to a great degree though to a lesser degree than has been documented in the extant literature. Second, crime hot spots are highly persistent over time. Third, across crime types, concentration and persistence appear to be correlated; those crime types that are highly spatially concentrated also tend to be persistent. This is intuitive — to the extent that there is an underlying physical or social feature of a given place that is criminogenic, we would expect that feature to cause crime to persist which, in turn, would tend to lead to concentration. On the other hand, crimes that are less concentrated also tend to be less persistent.

5.2 Within-City Crime Concentration

We next address the extent to which crime concentration varies *within a city*, a topic which has received little attention in the literature. We begin by characterizing the extent to which there is variation in crime concentration among precincts in the same city. This information is presented in **Tables 4A and 4B** which address crime concentration among 25 percent and 50 percent of crimes, respectively. Crime concentration measures are reported separately for each of our three cities. In Table 4A, column (1) reports the average share of street segments in a precinct that cumulatively account for 25 percent of crimes of each type.[21] Column (2) reports the share of street segments that account for 25 percent of crimes under simulation and Column (3) reports marginal crime concentration which is simply equal to Column (2) minus Column (1). In columns (4)-(6) we report standard deviations for each of the three measures.[22]

In drawing inferences from Table 4A, we focus primarily on the variation around the mean level of crime concentration at the precinct-level. We begin by focusing on the standard measure of crime concentration which does not adjust for the random distribution problem in low density data. In our three cities, in an average precinct, between 2.1 percent (Philadelphia) and 3.8 percent

[21] Note that these computations differ from citywide average since we are not weighting each precinct by its total crime count.

[22] We note that, in each city, one outlier police district is excluded. This includes New York City's Central Park precinct and the police districts covering O'Hare International Airport in Chicago and the Philadelphia International Airport. Each of these precincts contains very few street segments ($<$ 18) and, accordingly, concentration metrics do not behave normally.

(Chicago) of street segments account for one quarter of the overall crimes. The standard deviations around these means are 1.4 percent, 3.6 percent and 1 percent in New York City, Chicago and Philadelphia, respectively. While these standard deviations might seem small at first glance, we note that they are large relative to their respective means and, as such, are consistent with a large amount of variability in crime concentration within a city. For example, in New York City where 2.6 percent of the street segments account for one quarter of the crimes, the standard deviation in crime concentration is 1.4 which is more than 50 percent of the mean. Crime concentration ranges from 0.6 in the 100th precinct in the Far Rockaway section of Queens to 7.7 in Brooklyn's Bushwick neighborhood. In Chicago, where the mean and standard deviation are nearly identical, variability is even greater. Incredibly, the share of street segments accounting for one quarter of the crimes is as low as 0.4 in the Albany Park neighborhood on the city's north side and as high as 14.2 in Austin on the city's west side. Variability is correspondingly large for each of the five crime sub-types we consider.[23]

Next, we consider our measure of marginal crime concentration which is robust to the measurement issues discussed in Sections 2.2 and 2.3. Marginal crime concentration is high for overall crime as well as assault and larceny, is intermediate for robbery and auto theft and is nearly zero for murder — though we urge caution in interpreting estimated crime concentration for murder. There is likewise considerable variability around this measure within a city. In New York City, marginal crime concentration varies from a low of 13.7 in the 111th precinct in the Bayside section of Queens, one of the safest neighborhoods in the city, to a high of 22.5 in 14th precinct which covers the southern portion of Midtown Manhattan which includes Times Square. This is perhaps intuitive. In both neighborhoods, crimes are highly concentrated — the share of street segments that account for one quarter of the crimes is 1.3 in Midtown Manhattan and 2.1 in Bayside. However, in Bayside, a quiet residential neighborhood, there are relatively few crimes per street segment (28,924 over 3,850 blocks) whereas in Midtown South there are many more crimes (108,402) than streets (235), reflecting the fact that this small area experiences a very large amount of human activity. While crime appears to be concentrated in both areas, once differing crime densities are accounted for, we can see that crime is, in fact, far more concentrated in Midtown Manhattan than it is in Bayside.

[23]Precinct-level descriptive data for our three cities are available in Appendix tables 1A, 1B and 1C.

For readers who are interested, Table 4B presents the same information for the share of street segments that account for 50 percent of the crimes. The lessons from this table are similar. There is a great deal of heterogeneity in crime concentration throughout our three cities and, as such, while crime concentration is remarkably across our three cities as a whole, it does not appear as though there is a law of crime concentration that holds at lower levels of aggregation.

The anecdote presented above in which crime concentration in Bayside, Queens is contrasted with that of Midtown South raises a broader question: whether marginal crime concentration varies according to the underlying density of crime in a neighborhood. We have already seen that, for the raw measure of crime concentration this will be true. But is it also true for our corrected measure? We explore this relationship in **Figures 4A and 4B**, where marginal crime concentration is plotted on the y-axis and the number of crimes per street segment is plotted on the x-axis. In these figures we pool all 120 precincts across our three cities though we residualize out the city fixed effects in order to guard against confounding due to between-city differences in crime concentration.

In Figure 4A where we plot marginal crime concentration for 25 percent of crimes, each data point represents a precinct and a quadratic best fit line is drawn through the data. The small number of observations means that the highest-crime communities are highly leveraged and therefore that the slope of the curves drawn through the data are sensitive to outliers. As a result, we urge caution in interpretation. We begin by considering the relationship between marginal crime concentration and crime density for overall crime. At lower crime densities, the slope of the best fit curve is relatively flat, but the slope becomes negative at higher densities indicating that the highest-crime communities experiences lower levels of crime concentration. This relationship is different, however, for several serious street crimes. Robbery becomes *more* concentrated in higher-crime neighborhoods whereas assault is most concentrated towards the middle of the density distribution.

We next consider crime concentration for 50 percent of the crimes in Figure 4B. Here, the relationship between crime concentration and crime density is far more uniform. With the partial exception of robbery and murder (which is not concentrated at all), crime is concentrated to a far lesser degree in the highest crime communities. The contrast between the relationships presented in Figures 4A and 4B is interesting and merits further consideration. When we consider hot sports writ large — the street segments that account for half of the crimes — we see strong evidence that it takes considerably more of these hot spots to account for half the crimes in high-crime communities than in low-crime communities. However, when we focus on only the "hottest" of the hot spots, this is less true. One conclusion that rationalizes the data is that every community has a small number of areas in which crimes are strongly concentrated but that high-crime communities have a longer tail of less pervasive hot spots. The result is that it is, in general, more difficult to target resources in the highest crime parts of a city.

6 Conclusion

In this paper, we build upon recent methodological advances in the measurement of crime concentration and propose a novel method of measuring crime concentration that is robust to a key limitation of the dominant approaches in the extant literature. It has been noted that when crime concentration is defined using the share of street segments that account for some share of the crime in a city, this measure will tend to overstate the degree to which crimes are concentrated due to the presence of crime-free blocks. The intuition behind this result is that crimes can only be concentrated in the places in which crimes actually occur. The primary solution to this problem which has been proposed by Andresen et al. (2017) and Levin et al. (2017) among others is to measure crime concentration *among blocks that actually experience crime*. We note that while this approach will correct some of the upward bias in the measurement of crime concentration, appreciable bias will remain in many if not most empirical applications.

The reason why this approach does not completely solve what we refer to as the "non-uniformity problem" in the spatial distribution of crimes, is that, even among street segments that experience crime, by random chance some street segments will experience more crimes than others. We address the non-uniformity problem completely by comparing the actual distribution of crimes to the counterfactual spatial distribution of crimes that we generate by randomizing crimes to street segments with replacement. Our proposed solution — comparing the actual distribution of crimes to a distribution of crimes under the null hypothesis generated by randomization — allows us to generate an unbiased measure of crime concentration. However, we note that this methodology also has broad applicability to other domains in criminological research — for example to cohort studies of young people which invariably show that a small share of the population is responsible for an outsize share of the crimes and to "early warning systems" that police departments use to identify potentially problematic police officers and which are based on the premise that a small share of police officers are responsible for a disproportionate share of misconduct.

Like our predecessors we find considerable evidence that crimes are concentrated among cities and, accordingly, we provide additional support for the law of crime concentration. However, the extent to which the law of crime concentration applies requires some qualification. In this research, we note that in three of the largest cities in the United States, while crime is highly concentrated in the aggregate, murders are effectively unconcentrated and the robberies and auto thefts are only concentrated to a very small degree. On the other hand, assaults and larcenies exhibit fairly substantial concentration at the city level. These results are qualitatively different than many of those in the extant literature. We note that these discrepancies are, in large part, a function of how crime concentration has been measured and that previous research is likely to overestimate the degree of crime concentration.

We also extend this analysis to the study of crime concentration *within communities*. While the majority of the variation in crime within a city is explained by the within-community rather than between-community variation, we nevertheless note that many resource allocation problems are experienced at the community level. We find that while city-level concentration is remarkably stable across our three cities, the extent to which crimes are concentrated within our cities varies considerably. Within the same city, some communities experience crime concentration that is an order of magnitude larger than others. Accordingly, the law of crime concentration cannot be said to hold at lower levels of aggregation. We furthermore document evidence that crime tends to be less concentrated in higher-crime communities, indicating that crimes are less concentrated in precisely the communities in which efficient resource allocation is needed the most.

Since Weisburd (2015) noted the incredible explanatory power of place and the comparative inattention to place-based scholarship in the criminology literature, a large literature has proliferated to better describe and explain the extent to which and the conditions under which crimes are substantively concentrated in cities around the world. These are exciting developments and indeed there is much left to learn. We note that a number of papers already suggest fruitful substantive directions in which to take this research and instead suggest several ways in which the measurement of crime concentration can be improved. First, while the majority of research focuses on crime concentration at two important moments of the spatial distribution of crimes — 25 percent and 50 percent — crime concentration at other moments of the distribution might also be critical to explore. For instance, it would be useful to understand the extent to which crimes are concentrated at the very top of the distribution.[24] Second, given the variability of crime concentration among communities, it will be instructive to better understand which neighborhood characteristics predict crime concentration among communities with a city. Finally, greater research is needed to show the sensitivity of crime concentration measures to the length of the time period over which crime data are collected. In particular, to the extent that crime concentration is a function of the underlying density of crime, we might expect crime concentration to vary mechanically according to the sample size.

[24] Bernasco and Steenbeek (2017) explore this, to an extent, characterizing the distribution of crime concentration using the Gini coefficient.

References

Ajzenman, N. and L. Jaitman (2016). Crime concentration and hot spot dynamics in latin america. Technical report, Inter-American Development Bank.

Andresen, M. A., A. S. Curman, and S. J. Linning (2017). The trajectories of crime at places: understanding the patterns of disaggregated crime types. *Journal of Quantitative Criminology 33*(3), 427–449.

Andresen, M. A. and S. J. Linning (2012). The (in) appropriateness of aggregating across crime types. *Applied Geography 35*(1-2), 275–282.

Bernasco, W. and W. Steenbeek (2017). More places than crimes: Implications for evaluating the law of crime concentration at place. *Journal of Quantitative Criminology 33*(3), 451–467.

Bogar, S. and K. M. Beyer (2016). Green space, violence, and crime: A systematic review. *Trauma, Violence, & Abuse 17*(2), 160–171.

Braga, A. A. (2001). The effects of hot spots policing on crime. *The ANNALS of the American Academy of Political and Social Science 578*(1), 104–125.

Braga, A. A. and B. J. Bond (2008). Policing crime and disorder hot spots: A randomized controlled trial. *Criminology 46*(3), 577–607.

Braga, A. A., D. M. Hureau, and A. V. Papachristos (2011). The relevance of micro places to citywide robbery trends: a longitudinal analysis of robbery incidents at street corners and block faces in boston. *Journal of Research in Crime and Delinquency 48*(1), 7–32.

Braga, A. A., A. V. Papachristos, and D. M. Hureau (2010). The concentration and stability of gun violence at micro places in boston, 1980–2008. *Journal of Quantitative Criminology 26*(1), 33–53.

Braga, A. A., A. V. Papachristos, and D. M. Hureau (2014). The effects of hot spots policing on crime: An updated systematic review and meta-analysis. *Justice Quarterly 31*(4), 633–663.

Braga, A. A., B. C. Welsh, and C. Schnell (2015). Can policing disorder reduce crime? a systematic review and meta-analysis. *Journal of Research in Crime and Delinquency 52*(4), 567–588.

Branas, C. C., R. A. Cheney, J. M. MacDonald, V. W. Tam, T. D. Jackson, and T. R. Ten Have (2011). A difference-in-differences analysis of health, safety, and greening vacant urban space. *American Journal of Epidemiology 174*(11), 1296–1306.

Branas, C. C., E. South, M. C. Kondo, B. C. Hohl, P. Bourgois, D. J. Wiebe, and J. M. MacDonald (2018). Citywide cluster randomized trial to restore blighted vacant land and its effects on violence, crime, and fear. *Proceedings of the National Academy of Sciences 115*(12), 2946–2951.

Chalfin, A., B. Hansen, J. Lerner, and L. Parker (2019). Reducing crime through environmental design: Evidence from a randomized experiment of street lighting in new york city. Technical report, National Bureau of Economic Research.

Curiel, R. P., S. C. Delmar, and S. R. Bishop (2018). Measuring the distribution of crime and its concentration. *Journal of Quantitative Criminology 34*(3), 775–803.

Curman, A. S., M. A. Andresen, and P. J. Brantingham (2015). Crime and place: a longitudinal examination of street segment patterns in vancouver, bc. *Journal of Quantitative Criminology 31*(1), 127–147.

Doleac, J. L. and N. J. Sanders (2015). Under the cover of darkness: How ambient light influences criminal activity. *The Review of Economics and Statistics 97*(5), 1093–1103.

Eck, J. E., R. V. Clarke, R. T. Guerette, et al. (2007). Risky facilities: Crime concentration in homogeneous sets of establishments and facilities. *Crime Prevention Studies 21*, 225.

Farrington, D. P. and B. C. Welsh (2002). Improved street lighting and crime prevention. *Justice Quarterly 19*(2), 313–342.

Favarin, S. (2018). This must be the place (to commit a crime). testing the law of crime concentration in milan, italy. *European Journal of Criminology 15*(6), 702–729.

Garvin, E. C., C. C. Cannuscio, and C. C. Branas (2013). Greening vacant lots to reduce violent crime: a randomised controlled trial. *Injury Prevention 19*(3), 198–203.

Gill, C., A. Wooditch, and D. Weisburd (2017). Testing the law of crime concentration at place in a suburban setting: Implications for research and practice. *Journal of Quantitative Criminology 33*(3), 519–545.

Gorr, W. L. and Y. Lee (2015). Early warning system for temporary crime hot spots. *Journal of Quantitative Criminology 31*(1), 25–47.

Haberman, C. P., E. T. Sorg, and J. H. Ratcliffe (2017). Assessing the validity of the law of crime concentration across different temporal scales. *Journal of Quantitative Criminology 33*(3), 547–567.

Hibdon, J., C. W. Telep, and E. R. Groff (2017). The concentration and stability of drug activity in seattle, washington using police and emergency medical services data. *Journal of Quantitative Criminology 33*(3), 497–517.

Hipp, J. R. and Y.-A. Kim (2017). Measuring crime concentration across cities of varying sizes: Complications based on the spatial and temporal scale employed. *Journal of Quantitative Criminology 33*(3), 595–632.

Hoerl, A. E. and R. W. Kennard (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics 12*(1), 55–67.

Keizer, K., S. Lindenberg, and L. Steg (2008). The spreading of disorder. *Science 322*(5908), 1681–1685.

Kelling, G. L., J. Q. Wilson, et al. (1982). Broken windows. *The Atlantic Monthly 249*(3), 29–38.

Kondo, M., B. Hohl, S. Han, and C. Branas (2016). Effects of greening and community reuse of vacant lots on crime. *Urban Studies 53*(15), 3279–3295.

Levin, A., R. Rosenfeld, and M. Deckard (2017). The law of crime concentration: An application and recommendations for future research. *Journal of Quantitative Criminology 33*(3), 635–647.

Moyer, R., J. M. MacDonald, G. Ridgeway, and C. C. Branas (2019). Effect of remediating blighted vacant land on shootings: a citywide cluster randomized trial. *American Journal of Public Health 109*(1), 140–144.

Oliveira, M., C. Bastos-Filho, and R. Menezes (2017). The scaling of crime concentration in cities. *PloS One 12*(8), e0183110.

Quetelet, A. (1831). Research on the propensity to crime of different ages brussels: Hayez.

Schnell, C., A. A. Braga, and E. L. Piza (2017). The influence of community areas, neighborhood clusters, and street segments on the spatial variability of violent crime in chicago. *Journal of Quantitative Criminology 33*(3), 469–496.

Sherman, L. W., P. R. Gartin, and M. E. Buerger (1989). Hot spots of predatory crime: Routine activities and the criminology of place. *Criminology 27*(1), 27–56.

Sherman, L. W. and D. Weisburd (1995). General deterrent effects of police patrol in crime hot spots: A randomized, controlled trial. *Justice Quarterly 12*(4), 625–648.

Skogan, W. G. (2012). Disorder and crime. *The Oxford Handbook of Crime Prevention*, 173–188.

South, E. C., B. C. Hohl, M. C. Kondo, J. M. MacDonald, and C. C. Branas (2018). Effect of greening vacant land on mental health of community-dwelling adults: a cluster randomized trial. *JAMA Network Open 1*(3), e180298–e180298.

Steenbeek, W. and D. Weisburd (2016). Where the action is in crime? an examination of variability of crime across different spatial units in the hague, 2001–2009. *Journal of Quantitative Criminology 32*(3), 449–469.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological) 58*(1), 267–288.

Tseloni, A. and K. Pease (2005). Population inequality: the case of repeat crime victimization. *International Review of Victimology 12*(1), 75–90.

Weisburd, D. (2015). The law of crime concentration and the criminology of place. *Criminology 53*(2), 133–157.

Weisburd, D., G. J. Bruinsma, and W. Bernasco (2009). Units of analysis in geographic criminology: historical development, critical issues, and open questions. In *Putting Crime in its Place*, pp. 3–31. Springer.

Weisburd, D., S. Bushway, C. Lum, and S.-M. Yang (2004). Trajectories of crime at places: A longitudinal study of street segments in the city of seattle. *Criminology 42*(2), 283–322.

Weisburd, D. and L. Green (1995). Policing drug hot spots: The jersey city drug market analysis experiment. *Justice Quarterly 12*(4), 711–735.

Weisburd, D., N. A. Morris, and E. R. Groff (2009). Hot spots of juvenile crime: A longitudinal study of arrest incidents at street segments in seattle, washington. *Journal of Quantitative Criminology 25*(4), 443.

Weisburd, D. and C. W. Telep (2014). Hot spots policing: What we know and what we need to know. *Journal of Contemporary Criminal Justice 30*(2), 200–220.

Welsh, B. C. and D. P. Farrington (2008). Effects of improved street lighting on crime. *Campbell Systematic Reviews 13*, 1–51.

Table 1: Summary Statistics

	New York City	Chicago	Philadelphia
Time period	Jan 2006 - Dec 2018	Jan 2001 - May, 2019	Jan 2006 - May, 2019
Number of Street Segments	119,467	56,338	41,009
Percentage of Crime at Intersections	25.8%	5.4%	59%
Crime During Entire Studied Period			
All crimes	4,585,280	6,419,106	1,051,840
Murder	4,207	9,386	1,762
Robbery	131,492	237,517	29,037
Assault	684,394	1,593,989	222,651
Auto theft	82,196	296,561	21,156
Larceny	1,225,234	1,322,269	232,866
Crime During 2018			
All crimes	334,534	251,530	71,078
Murder	243	568	132
Robbery	8,412	9,163	1,684
Assault	55,667	67,385	16,499
Auto theft	3,953	9,559	1,082
Larceny	100,534	60,069	17,885

Note: Table presents descriptive data on the number of street segments and crimes in each of our three cities: New York City, Chicago and Philadelphia.

Table 2A: Marginal Crime Concentration at 25 Percent of Street Segments

	Share of Segments, Unadjusted (Weisburd)	Share of Segments, Non-Zero Crime Segments	Share of Segments, Simulated	Marginal Crime Concentration: Our proposed method	Marginal Crime Concentration: Levin-Rosenfeld-Deckard
A. New York City					
All crimes	1.2	2.6	14.8	13.7	22.4
Murder	.1	21.2	.1	0	3.8
Robbery	.8	8.1	2.9	2.1	16.9
Assault	.9	4	7.3	6.4	21
Auto theft	1.1	12.8	2.1	1	12.2
Larceny	.4	1.4	9.7	9.2	23.6
B. Chicago					
All crimes	2.7	3.8	18.3	15.6	21.2
Murder	.4	18.3	.6	.2	6.7
Robbery	1.8	6.2	6.3	4.5	18.8
Assault	2.2	4.1	13.5	11.3	20.9
Auto theft	3.1	8.5	7	3.9	16.5
Larceny	1.2	2.2	12.7	11.5	22.8
C. Philadelphia					
All crimes	2.1	3.2	16.8	14.6	21.8
Murder	.3	18	.4	.1	7
Robbery	.9	6	3.4	2.5	19
Assault	1.9	4.6	10.9	9	20.4
Auto theft	1.4	9.9	3.1	1.7	15.1
Larceny	.6	1.3	11.1	10.5	23.8

Note: This table reports the share of street segments that account for 25 and 50 percent of each of six crime types: total crime, murder, robbery, assault, auto theft and larceny and three cities: NYC (Panel A), Chicago (Panel B) and Philadelphia (Panel C). Column (1) reports crime concentration for all street segments, Column (2) reports crime concentration for street segments with non-zero crime and Column (3) reports simulated crime concentration arising generated by randomizing crimes to street segments, with replacement. The final two columns report *marginal* crime concentration using both our proposed method and by the method proposed by Levin-Rosenfeld-Deckard.

Table 2B: Marginal Crime Concentration at 50 Percent of Street Segments

	Share of Segments, Unadjusted (Weisburd)	Share of Segments, Non-Zero Crime Segments	Share of Segments, Simulated	Marginal Crime Concentration	
				Our proposed method	Levin-Rosenfeld-Deckard
A. New York City					
All crimes	4.5	9.9	34.7	30.1	40.1
Murder	.3	47.5	.3	0	2.5
Robbery	2.5	25.4	7.2	4.6	24.6
Assault	3.2	13.7	18.4	15.2	36.3
Auto theft	3.3	37.9	4.8	1.5	12.1
Larceny	2.8	8.7	23.7	20.9	41.3
B. Chicago					
All crimes	9.2	12.6	40.5	31.3	37.4
Murder	1	45.6	1.2	.2	4.4
Robbery	5.9	20.8	15.2	9.4	29.2
Assault	6.9	13	32.2	25.3	37
Auto theft	8.9	24.1	16.7	7.9	25.9
Larceny	6.6	11.8	30.6	24	38.2
C. Philadelphia					
All crimes	8.2	12.6	37.9	29.6	37.3
Murder	.7	45.4	.8	.1	4.6
Robbery	3.6	22.8	10.3	6.7	27.2
Assault	6.5	15.5	26.7	20.2	34.5
Auto theft	4.2	29.1	8.2	4	20.9
Larceny	4.8	10.2	27.1	22.4	39.8

Note: This table reports the share of street segments that account for 25 and 50 percent of each of six crime types: total crime, murder, robbery, assault, auto theft and larceny and three cities: NYC (Panel A), Chicago (Panel B) and Philadelphia (Panel C). Column (1) reports crime concentration for all street segments, Column (2) reports crime concentration for street segments with non-zero crime and Column (3) reports simulated crime concentration arising generated by randomizing crimes to street segments, with replacement. The final two columns report ***marginal*** crime concentration using both our proposed method and the method proposed by Levin-Rosenfeld-Deckard.

Table 3: Crime Concentration: Comparison between Chalfin-Kaplan-Cuellar and Hipp-Kim Metrics

		25 Percent		50 Percent	
	Share of Crimes Top 5 percent of segments, t-1	Share of Segments, 25 percent of crimes, t-1	Marginal Crime Concentration (CKC)	Share of Segments, 50 percent of crimes, t-1	Marginal Crime Concentration (CKC)
			A. New York City		
All crimes	51.0	23.7	13.7	47.1	30.1
Murder	4.4	0.3	0.0	0.6	0.0
Robbery	34.1	11.5	2.1	22.1	4.6
Assault	52.2	20.2	6.4	39.2	15.2
Auto Theft	18.7	4.5	1.0	7.7	1.5
Larceny	56.4	23.5	9.2	45.0	20.9
			B. Chicago		
All crimes	35.6	23.0	15.6	46.4	31.3
Murder	5.7	0.7	0.2	1.3	0.2
Robbery	28.7	13.8	4.5	25.4	9.4
Assault	38.9	21.1	11.3	42.3	25.3
Auto Theft	17.6	9.7	3.9	19.8	7.9
Larceny	41.6	23.0	11.5	42.4	24.0
			C. Philadelphia		
All crimes	36.0	21.5	14.6	41.7	29.6
Murder	3.6	0.4	0.1	0.6	0.1
Robbery	28.1	8.8	2.5	14.0	6.7
Assault	33.6	16.1	9.0	30.6	20.2
Auto Theft	14.4	4.5	1.7	6.9	4.0
Larceny	42.0	22.5	10.5	38.0	22.4

Note: This table reports a computation suggested by Hipp and Kim (2017): the share of crimes occurring in a given year that is accounted for by the top 5 percent of street segments in the prior year. We also report a modified version of their suggested computation: the share of crimes that occur in a given year on the top 25 or 50 percent of street segments, ranked according to crime incidence in the prior year. We also report our preferred measure: marginal crime concentration.

Table 3A: Descriptive Statistics: Marginal Crime Concentration by Precinct
25 percent concentration

	Mean			Standard Deviation		
	Share of Segments, Unadjusted	Share of Segments, Simulated	Marginal Crime Concentration	Share of Segments, Unadjusted	Share of Segments, Simulated	Marginal Crime Concentration
			A. New York City			
All crimes	2.6	20.6	18	1.4	1.7	1.7
Murder	.6	1.2	.6	.5	.9	.5
Robbery	2.2	9	6.9	1.3	3.7	2.7
Assault	2	15	13	1.1	3.3	2.8
Auto theft	2.9	7	4.1	1.3	2.1	1.2
Larceny	1.7	17	15.3	1.3	2.9	2.8
			B. Chicago			
All crimes	3.8	22.2	18.5	3.6	.7	3.1
Murder	1.5	2.8	1.4	1.7	1.6	.6
Robbery	2.6	13.5	11	2.3	2.5	1.5
Assault	3.1	19.6	16.6	3.1	1.4	2.1
Auto theft	5	14.5	9.5	4.3	1.7	3
Larceny	1.5	19.2	17.7	1.9	1.2	1.3
			C. Philadelphia			
All crimes	2.1	19.3	17.2	1	.8	.6
Murder	.5	1	.5	.4	.5	.2
Robbery	1	6.3	5.2	.5	1.5	1.1
Assault	2	14.1	12.2	1.2	1.6	.9
Auto theft	1.8	5.2	3.3	1.4	1.5	.7
Larceny	.6	14.5	13.9	.5	1.1	.8

Note: This table reports the precinct-level mean and standard deviation of street segments that account for 25 and 50 percent of each of six crime types: total crime, murder, robbery, assault, auto theft and larceny and three cities: NYC (Panel A), Chicago (Panel B) and Philadelphia (Panel C). Columns (1)-(3) report means; Columns (4)-(6) report standard deviations. Column (1) reports the mean share of street segments that account for 50 percent of the crimes in a precinct, Column (2) reports the mean simulated share of street segments that account for 50 percent of crimes in a precinct, generated by randomizing crimes to street segments with replacement and Column (3) reports marginal crime concentration which is equal to Column (2) minus Column (1). Columns (4)-(6) report the standard deviation for each of these three statistics. Note that, in each city, one outlier police district is excluded. This includes the Central Park precinct in NYC, and the police districts covering O'Hare International Airport in Chicago and the Philadelphia International Airport.

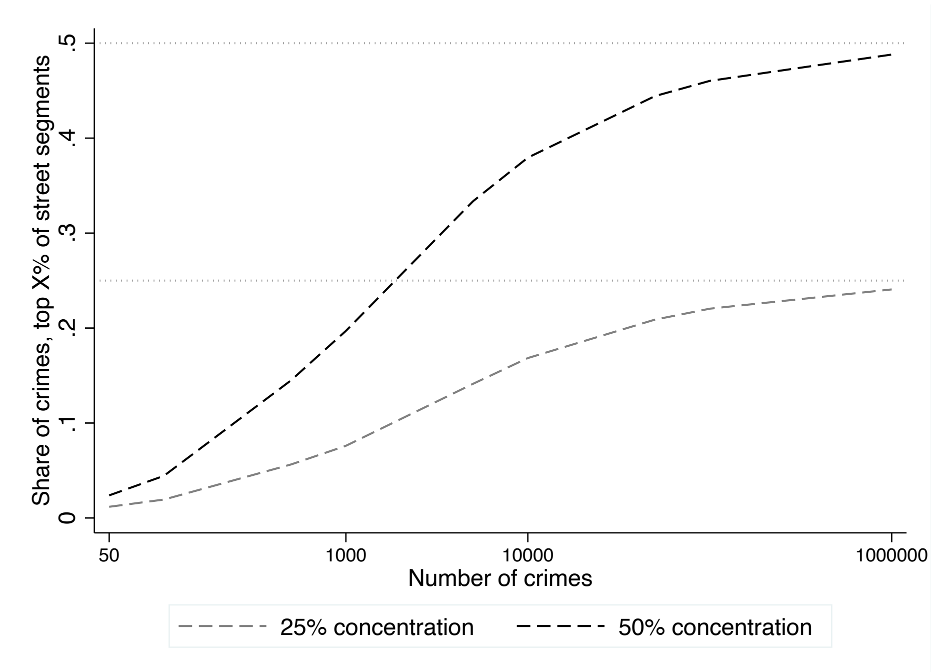
Table 3B: Descriptive Statistics: Marginal Crime Concentration by Precinct
50 percent concentration

	Mean			Standard Deviation		
	Share of Segments, Unadjusted	Share of Segments, Simulated	Marginal Crime Concentration	Share of Segments, Unadjusted	Share of Segments, Simulated	Marginal Crime Concentration
A. New York City						
All crimes	9.3	43.8	34.5	9.6	2.6	7.7
Murder	1.3	2.6	1.2	1.3	2.3	1.2
Robbery	7	22.1	15.1	7.3	8.2	4
Assault	6.3	34.3	28	6.6	6.3	4.1
Auto theft	8.7	17.5	8.8	6	4.9	2.3
Larceny	6.6	38	31.4	10.2	5.1	7.8
B. Chicago						
All crimes	11.9	46.3	34.5	8.6	1	7.8
Murder	4	7.2	3.2	3.9	4.4	1.2
Robbery	9.1	31.9	22.7	6.9	5.1	4
Assault	9.3	42.4	33.1	8.3	2.2	6.6
Auto theft	13.7	33.8	20.1	8.7	3.2	6
Larceny	8.7	41.9	33.1	6.2	2	4.5
C. Philadelphia						
All crimes	8.1	42	33.8	3.8	1.2	2.7
Murder	1.5	2	.5	.9	1.2	.4
Robbery	4.4	16	11.6	2.2	3.6	1.7
Assault	6.7	33.1	26.3	3.7	3.1	1.3
Auto theft	5.6	13.6	7.9	3.4	3	1.3
Larceny	5.2	33.9	28.7	2.5	2	1.4

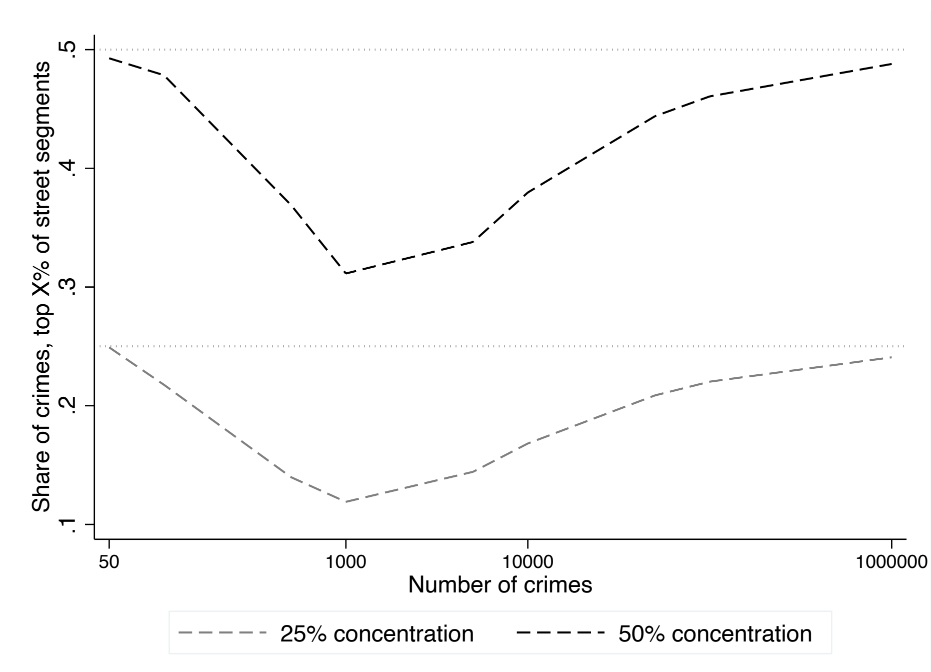
Note: This table reports the precinct-level mean and standard deviation of street segments that account for 25 and 50 percent of each of six crime types: total crime, murder, robbery, assault, auto theft and larceny and three cities: NYC (Panel A), Chicago (Panel B) and Philadelphia (Panel C). Columns (1)-(3) report means; Columns (4)-(6) report standard deviations. Column (1) reports the mean share of street segments that account for 50 percent of the crimes in a precinct, Column (2) reports the mean simulated share of street segments that account for 50 percent of crimes in a precinct, generated by randomizing crimes to street segments with replacement and Column (3) reports marginal crime concentration which is equal to Column (2) minus Column (1). Columns (4)-(6) report the standard deviation for each of these three statistics. Note that, in each city, one outlier police district is excluded. This includes the Central Park precinct in NYC, and the police districts covering O'Hare International Airport in Chicago and the Philadelphia International Airport.

Figure 1: Crime Concentration, Simulated Data for n = 1,000 blocks

Panel A: Unadjusted crime concentration, all street segments



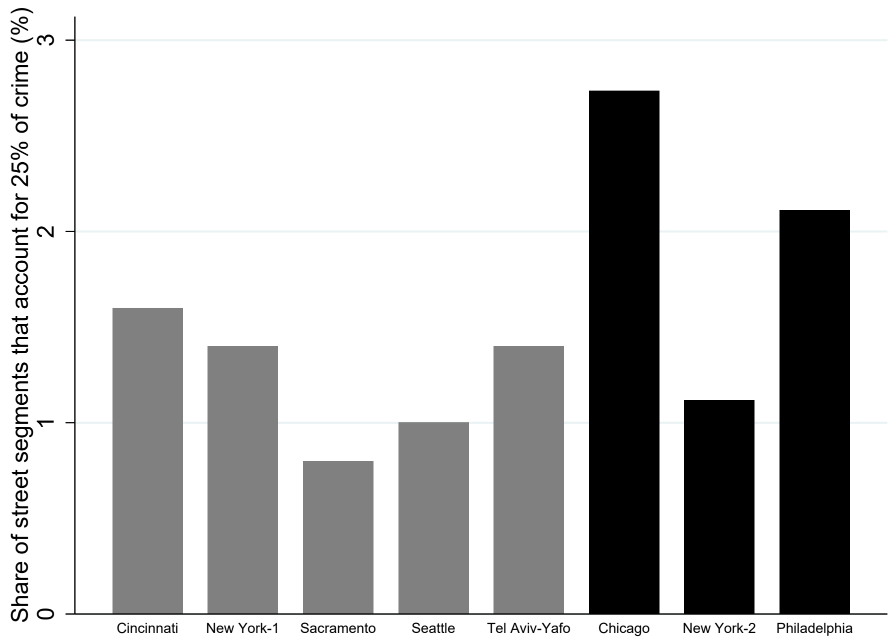
Panel B: Unadjusted crime concentration, non-zero crime street segments



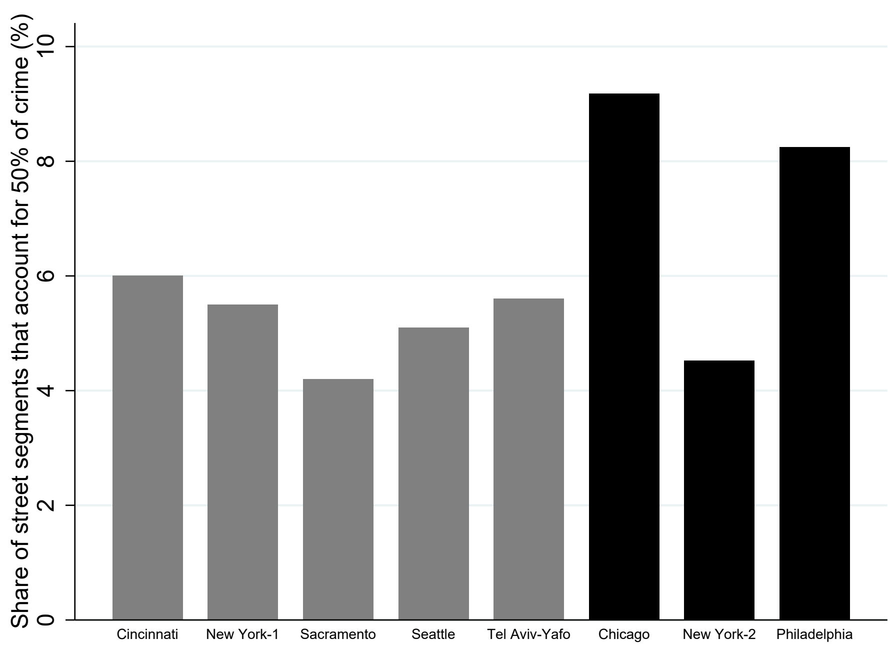
Note: Figures plot the share of street segments that account for 25 percent (Panel A) and 50 percent (Panel B) of crimes, in simulated data in which crimes are randomly assigned to street segments, with replacement. The number of street segments is fixed at 1,000 while the number of crimes is allowed to very along the x-axis. The x-axis has been transformed using a logarithmic scale.

Figure 2: Share of Crimes Among the Top 25 and 50 Percent of Street Segments, by City

Panel A: Unadjusted crime concentration, all street segments



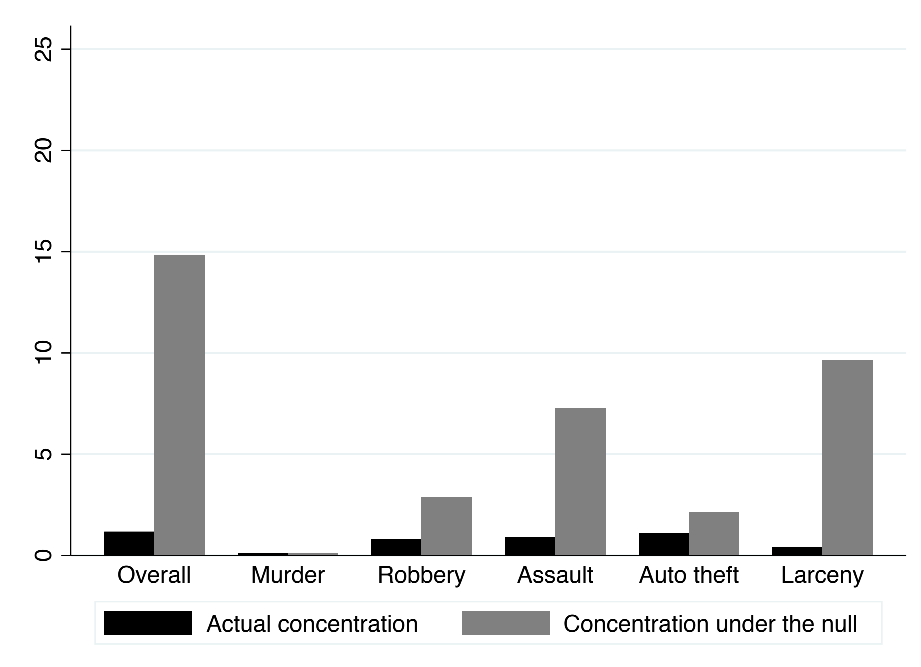
Panel B: Unadjusted crime concentration, non-zero crime street segments



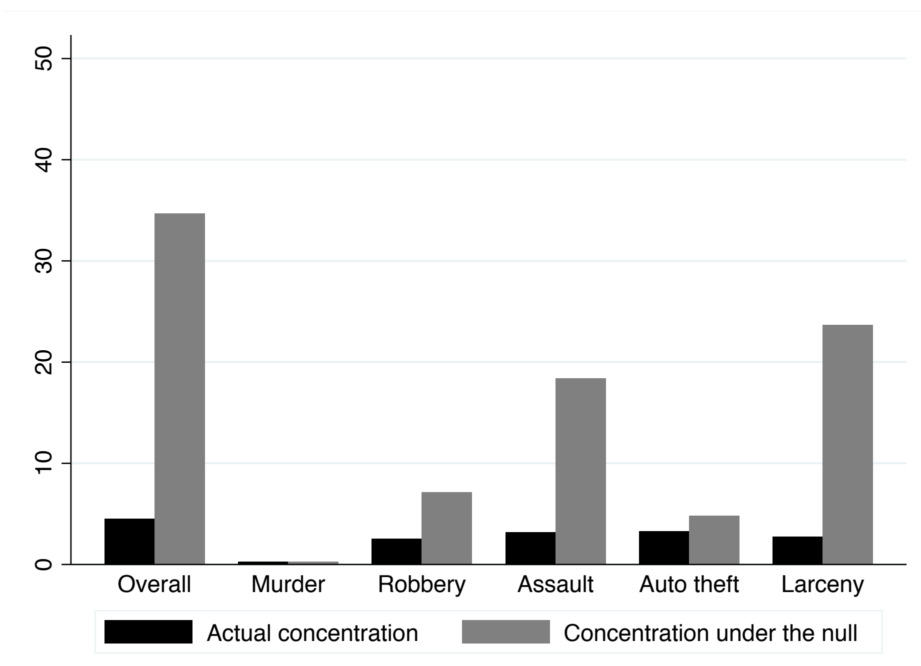
Note: Figures plot the share of street segments that account for 25 percent (Panel A) and 50 percent (Panel B) of crimes. The gray bars are replicated from Table 3 in Weisburd (2015). The black bars correspond to data from New York City, Chicago and Philadelphia (our sample).

Figure 3A: Actual vs. Simulated Share of Street Segments Accounting for 25 Percent and 50 Percent of Crimes, New York City

Panel A: 25 Percent of Crimes



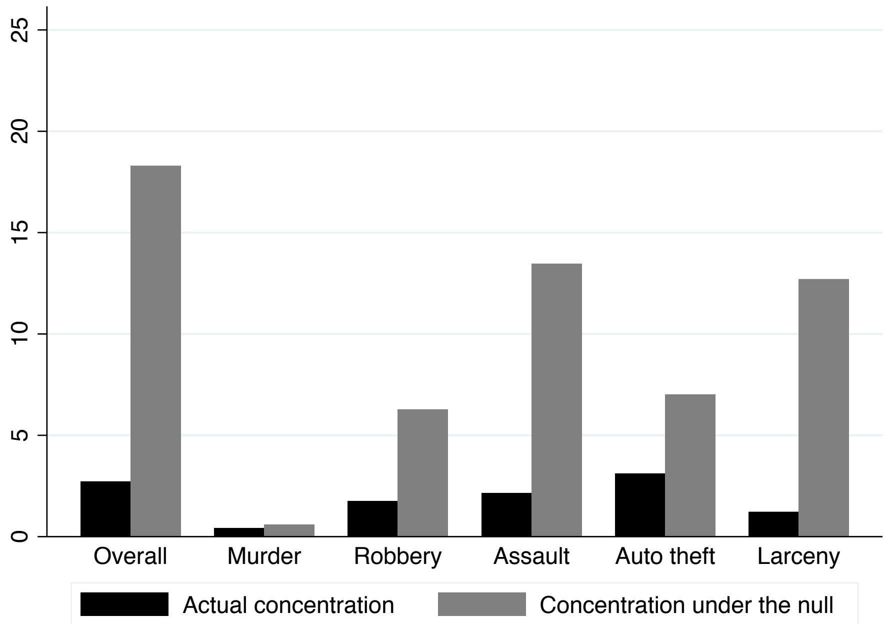
Panel B: 50 Percent of Crimes



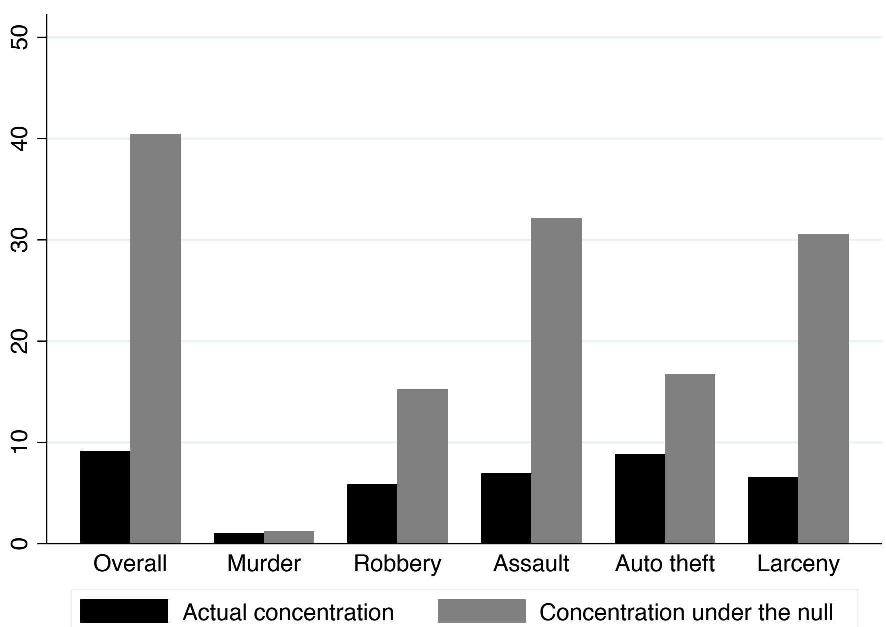
Note: Figures plot the share of street segments that account for 25 percent and 50 percent of crimes. The black bars represent the actual data the gray bars represent simulated data in which crimes are randomized to street segments, with replacement. Crimes are concentrated when the heights of the two bars are different.

Figure 3B: Actual vs. Simulated Share of Street Segments Accounting for 25 Percent and 50 Percent of Crimes, Chicago

Panel A: 25 Percent of Crimes



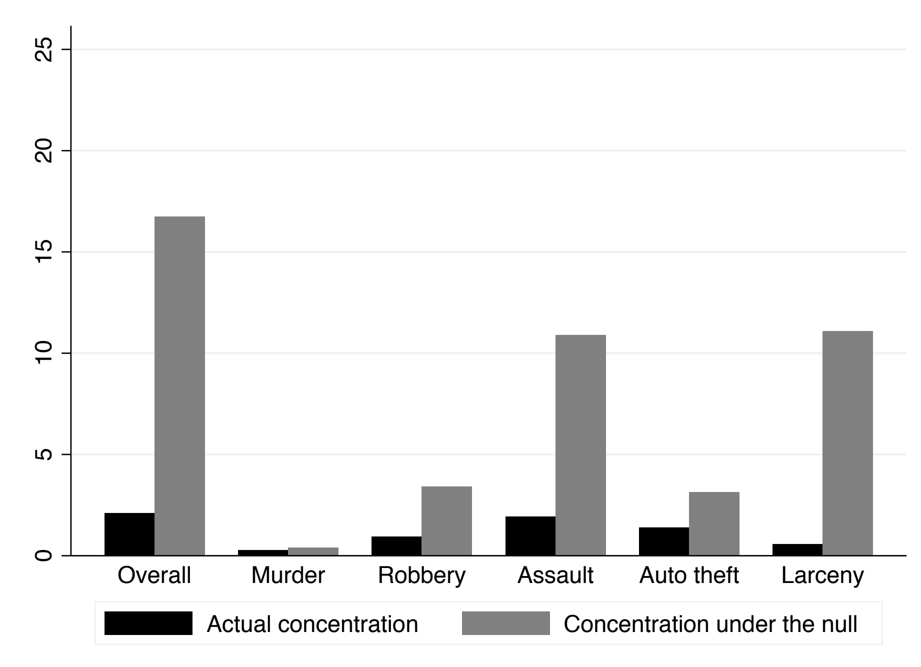
Panel B: 50 Percent of Crimes



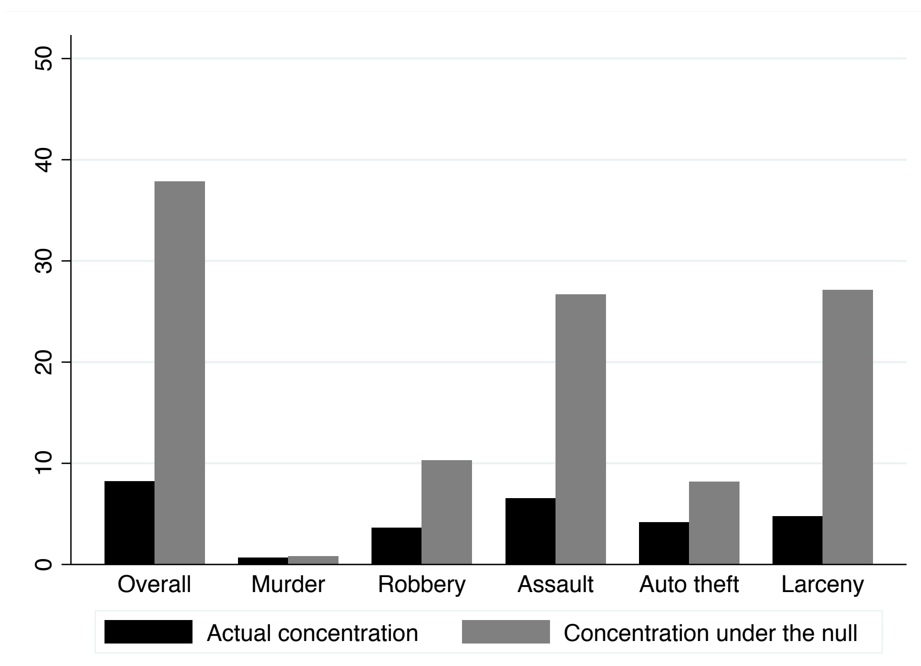
Note: Figures plot the share of street segments that account for 25 percent and 50 percent of crimes. The black bars represent the actual data the gray bars represent simulated data in which crimes are randomized to street segments, with replacement.

Figure 3C: Actual vs. Simulated Share of Street Segments Accounting for 25 Percent and 50 Percent of Crimes, Philadelphia

Panel A: 25 Percent of Crimes

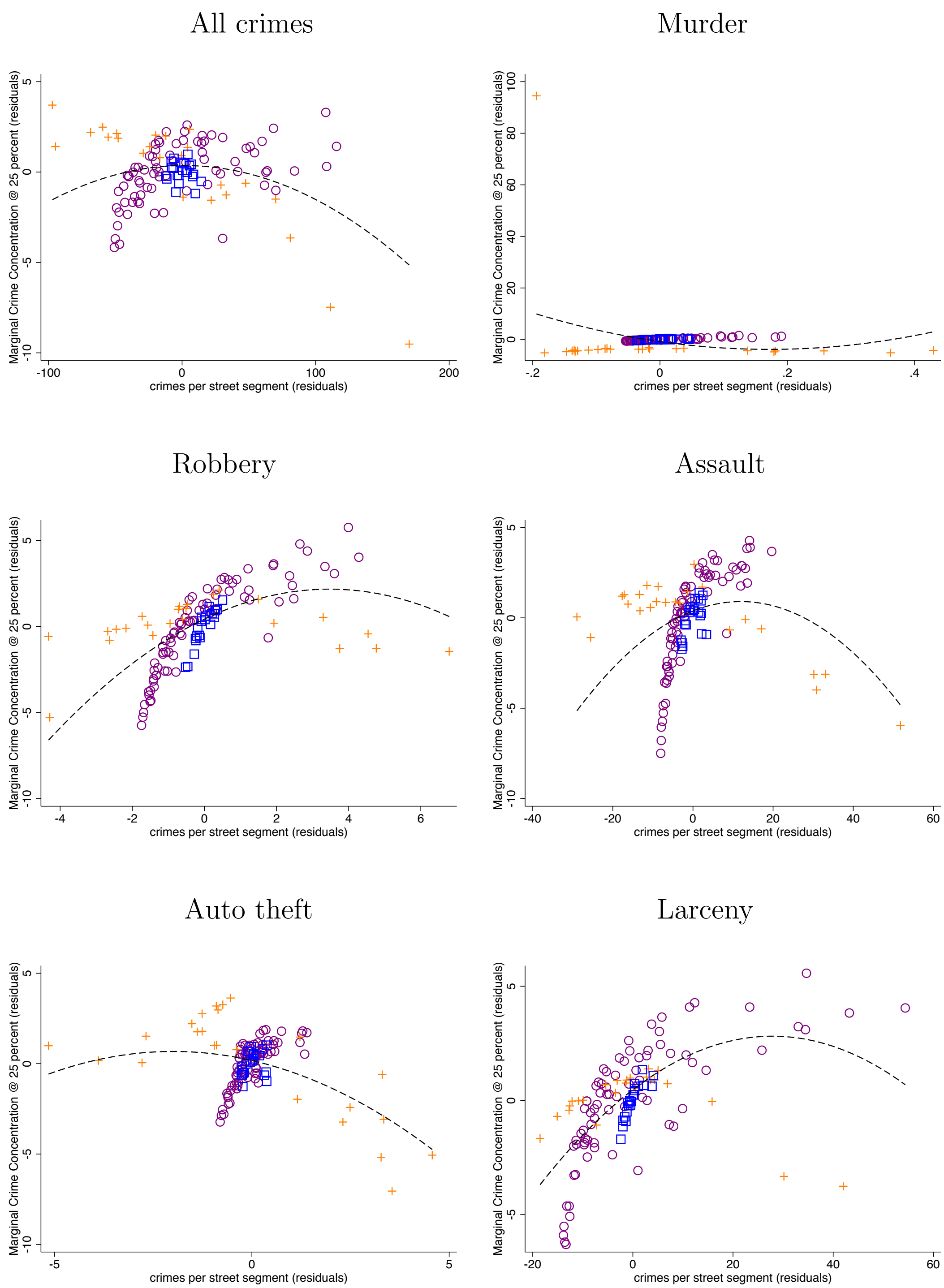


Panel B: 50 Percent of Crimes



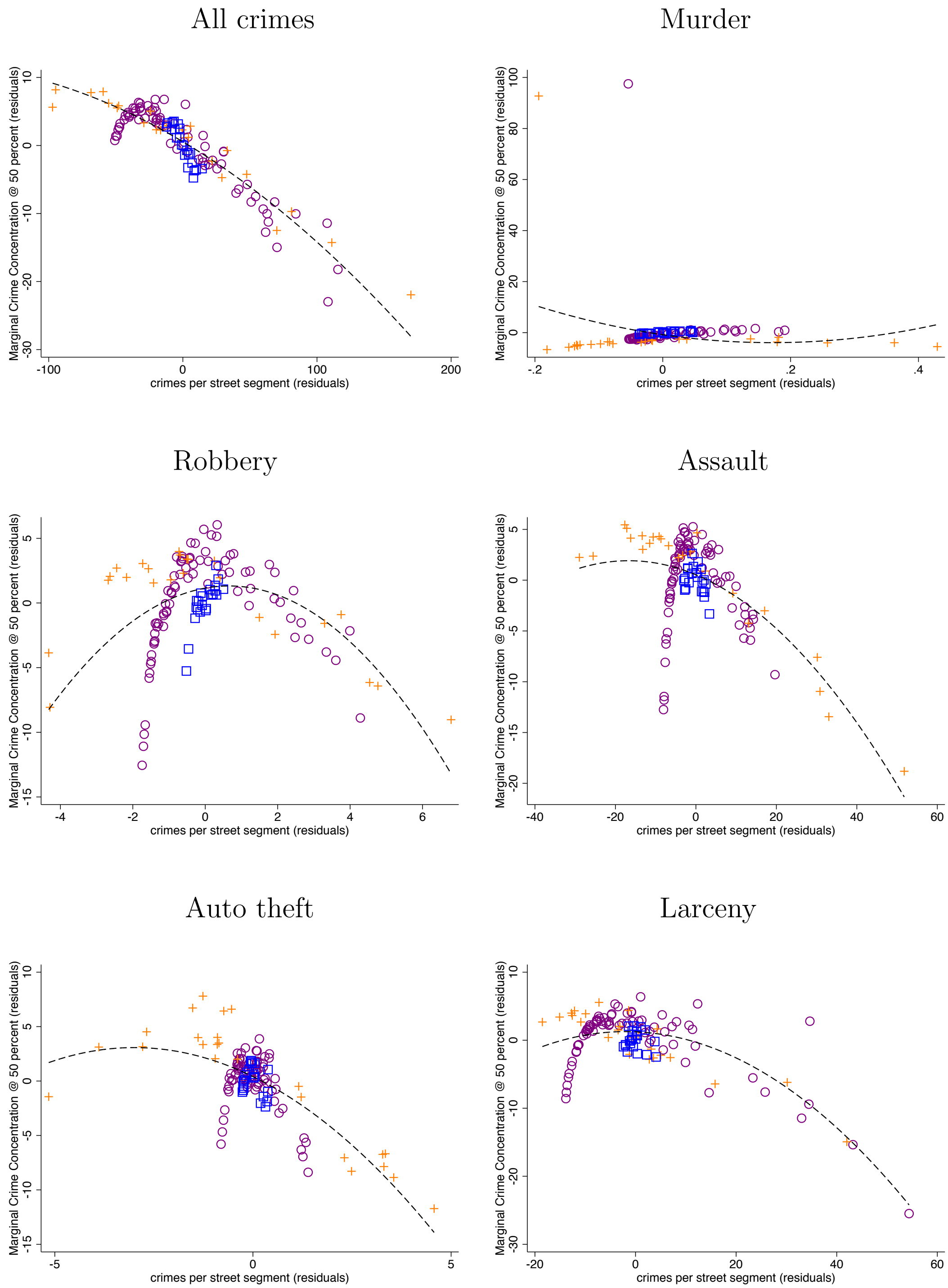
Note: Figures plot the share of street segments that account for 25 percent and 50 percent of crimes. The black bars represent the actual data the gray bars represent simulated data in which crimes are randomized to street segments, with replacement.

Figure 4A: Marginal Crime Concentration by Precinct, Street Segments Accounting for 25 Percent of Crimes



Note: Figures plot marginal crime concentration for 25 percent of crimes on the y-axis against the number of crimes per street segment in each police precinct on the x-axis. Both marginal crime concentration and crimes per street segment have been residualized, removing the city fixed effects. The circular markers plot data for precincts in New York City, the plus markers plot data for precincts in Chicago and the square markers plot data for precincts in Philadelphia. A quadratic best fit curve is drawn through the data points. A positive (negative) relationship in the data means that police precincts with more crimes per street segment experience greater (less) crime concentration.

Figure 4B: Marginal Crime Concentration by Precinct, Street Segments Accounting for 50 Percent of Crimes



Note: Figures plot marginal crime concentration for 50 percent of crimes on the y-axis against the number of crimes per street segment in each police precinct on the x-axis. Both marginal crime concentration and crimes per street segment have been residualized, removing the city fixed effects. The circular markers plot data for precincts in New York City, the plus markers plot data for precincts in Chicago and the square markers plot data for precincts in Philadelphia. A quadratic best fit curve is drawn through the data points to approximate the data generating process.

Appendix Table 1A: Precinct-Level Descriptive Statistics, New York City

				Crime, Concentration, 25 percent		Crime Concentration, 50 percent	
District	Number of segments	Crimes	Crimes per segment	Share of segments, Unadjusted	Marginal Crime Concentration	Share of segments Unadjusted	Marginal Crime Concentration
District 14	235	108402	461.3	1.3	22.5	47.2	1.2
District 18	440	76537	173.9	3.6	19.2	30.5	16.7
District 23	369	64156	173.9	3.5	19.3	30.1	17
District 9	371	61763	166.5	4.6	18.2	34.8	12.3
District 28	240	39793	165.8	1.7	21.2	23.3	23.8
District 24	324	46150	142.4	4.6	18	21.6	25.2
District 81	429	55094	128.4	5.6	16.9	26.3	20.3
District 32	427	54094	126.7	2.1	20.3	19.7	26.9
District 79	570	69520	122	4.4	18	22.5	24
District 77	518	62684	121	4.4	17.9	21.2	25.2
District 71	505	60575	120	5.2	17.2	24	22.5
District 13	551	65094	118.1	2.7	19.6	20.5	25.9
District 30	327	36781	112.5	3.4	19	18.6	27.8
District 46	875	95536	109.2	2.9	19.3	19.3	26.9
District 73	816	86712	106.3	2.9	19.2	16.7	29.5
District 20	381	38067	99.9	4.2	17.9	17.3	28.8
District 42	675	66090	97.9	3.6	18.5	17.8	28.3
District 26	320	28369	88.7	2.2	19.8	11.6	34.4
District 83	666	59034	88.6	7.7	14.2	11.6	34.3
District 6	488	42436	87	4.1	17.8	13.3	32.5
District 48	870	72824	83.7	3.8	18	13.9	31.8
District 40	1307	105097	80.4	1.8	20	12.6	33.1
District 34	625	48330	77.3	4.5	17.2	13.1	32.5
District 44	1318	98616	74.8	2	19.6	13.1	32.3
District 19	897	67060	74.8	3	18.6	10.4	35.1
District 70	967	70921	73.3	2.6	19	8.7	36.7
District 41	839	61490	73.3	1.7	19.9	12.6	32.8
District 52	1122	81710	72.8	2	19.5	11.9	33.5
District 7	561	39269	70	2	19.6	12.1	33.2
District 5	736	45636	62	.8	20.5	8.6	36.5
District 17	513	31609	61.6	4.5	16.9	10.7	34.3
District 67	1380	84378	61.1	2.8	18.4	7.4	37.6
District 78	626	37643	60.1	1.1	20.2	3.7	41.3
District 43	2023	119736	59.2	1.5	19.7	9.1	35.8
District 10	552	29741	53.9	1.6	19.5	10	34.7
District 75	2488	128292	51.6	2.4	18.5	6	38.5
District 33	777	38180	49.1	2.1	18.8	8.9	35.6
District 25	1130	52431	46.4	.6	20.1	6.3	38
District 90	1287	59288	46.1	3	17.7	6	38.3
District 62	1301	57640	44.3	5	15.7	2.1	42
District 60	1225	50715	41.4	1	19.6	4.9	39
District 47	1986	81834	41.2	2.9	17.7	4.5	39.4
District 88	725	29667	40.9	2.6	17.9	5.4	38.5
District 84	1027	41688	40.6	.9	19.7	5.7	38.2
District 49	1585	61988	39.1	2.3	18.1	4.5	39.2
District 103	1724	66101	38.3	1.9	18.5	3.4	40.3
District 1	1304	49910	38.3	.9	19.5	5.1	38.6
District 66	1188	44712	37.6	4.7	15.6	1.6	42
District 69	1200	44565	37.1	2.5	17.8	2.8	40.8
District 101	1125	40625	36.1	1.5	18.8	4.2	39.3
District 115	1604	56844	35.4	3.2	17	2.9	40.5
District 114	2524	85444	33.9	1.3	18.8	2.9	40.3
District 94	978	31470	32.2	3	17.1	2	41.1
District 76	829	24851	30	1.8	18	3.6	39.3
District 72	1418	38656	27.3	3	16.6	2.2	40.3
District 68	1705	45362	26.6	2.2	17.3	1.5	40.9
District 102	2050	54042	26.4	3.4	16.2	.9	41.4
District 61	2145	55772	26	2	17.4	1.6	40.7
District 50	1443	36672	25.4	1.3	18.2	3.2	39.1
District 106	2359	59493	25.2	3.1	16.3	.6	41.6
District 110	2358	54430	23.1	1	18.2	1.5	40.4
District 63	1891	42806	22.6	2.6	16.5	1.1	40.7
District 120	3443	74371	21.6	1.1	17.9	1.6	40
District 104	3011	63739	21.2	2.7	16.2	.8	40.8
District 109	4235	78190	18.5	.9	17.7	1.3	39.7
District 45	2889	52610	18.2	.9	17.7	1.4	39.4
District 113	3523	61188	17.4	2.9	15.6	.5	40.1
District 121	3677	63450	17.3	.7	17.7	.8	39.8
District 108	2639	40593	15.4	1.9	16.2	.5	39.6
District 112	1821	26466	14.5	.8	17.1	.8	39
District 105	5610	64020	11.4	3.2	13.9	.1	38.5
District 107	3160	34875	11	1.4	15.7	.5	38
District 122	5165	51475	10	1.8	14.9	.3	37.5
District 100	2232	20468	9.2	.6	15.9	.8	36.6
District 123	3958	32501	8.2	1.9	14.2	.2	36.6
District 111	3850	28924	7.5	2	13.7	.2	36

Appendix Table 1B: Precinct-Level Descriptive Statistics, Chicago

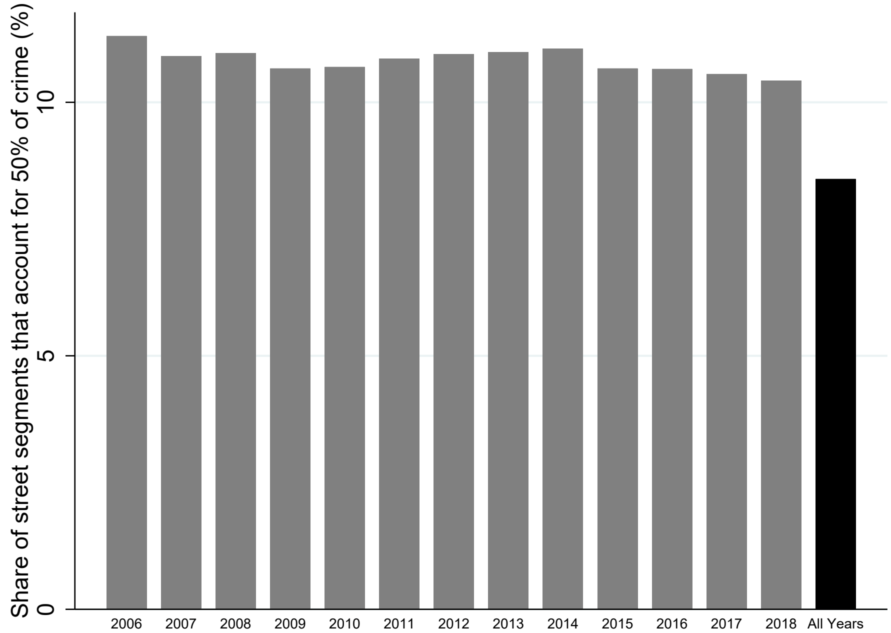
District	Number of segments	Crimes	Crimes per segment	Crime, Concentration, 25 percent: Share of segments, Unadjusted	Crime, Concentration, 25 percent: Marginal Crime Concentration	Crime Concentration, 50 percent: Share of segments Unadjusted	Crime Concentration, 50 percent: Marginal Crime Concentration
District 15	874	262690	300.6	14.2	9.1	35	12.8
District 11	1587	383392	241.6	11.9	11.1	27	20.4
District 3	1457	308214	211.5	8	15	22.3	25
District 7	1864	373899	200.6	5.7	17.1	25	22.2
District 6	2003	356789	178.1	4.7	18	16.6	30.5
District 18	1645	269004	163.5	5.3	17.4	12.9	34
District 10	1699	271246	159.7	4.7	17.9	16.9	30
District 2	1890	287900	152.3	5.5	17.1	14.4	32.4
District 25	2573	350481	136.2	1.4	21	9.1	37.5
District 24	1345	181181	134.7	2.5	20	10.7	35.9
District 1	1784	234183	131.3	5.2	17.2	10.7	35.9
District 19	2151	279734	130	2.8	19.5	9.5	37
District 14	2013	238218	118.3	1.6	20.6	8.7	37.6
District 5	2534	280563	110.7	1.5	20.7	9.2	37
District 20	946	100591	106.3	2.1	20	6.4	39.8
District 4	3626	367792	101.4	2.4	19.7	8.1	38
District 8	5021	415050	82.7	1.2	20.5	5.1	40.5
District 9	3763	307101	81.6	1	20.8	5.4	40.2
District 12	3726	280139	75.2	1	20.6	4.5	40.9
District 17	2481	176429	71.1	.4	21.1	2.7	42.6
District 22	3279	203417	62	.5	20.8	2.5	42.5
District 16	5250	187423	35.7	.2	20	.5	42.9

Appendix Table 1C: Precinct-Level Descriptive Statistics, Philadelphia

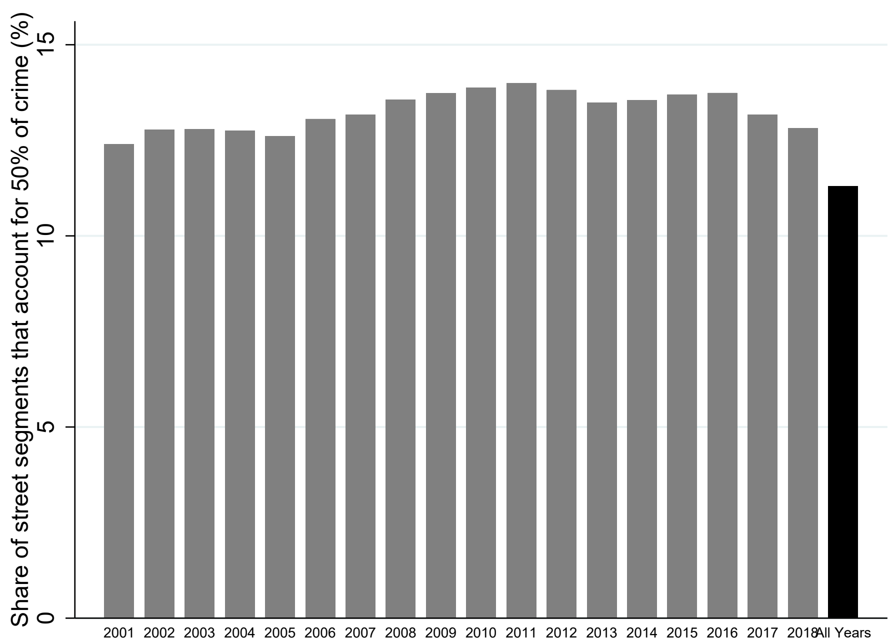
				Crime, Concentration, 25 percent		Crime Concentration, 50 percent	
District	Number of segments	Crimes	Crimes per segment	Share of segments, Unadjusted	Marginal Crime Concentration	Share of segments Unadjusted	Marginal Crime Concentration
District 18	1380	54664	39.6	3.8	16.7	13.3	30.4
District 24	1817	63957	35.2	4.2	16	13.1	30.3
District 12	1952	66080	33.9	3	17.1	13.1	30.1
District 25	1782	59082	33.2	3.1	17	14.1	29.1
District 19	1753	56428	32.2	2.3	17.6	11.8	31.3
District 2	2152	66472	30.9	2.3	17.6	10.2	32.8
District 35	2103	62306	29.6	1.6	18.2	10.3	32.5
District 22	2164	62968	29.1	2.5	17.2	12.1	30.6
District 15	3384	95825	28.3	2.4	17.3	9.5	33.1
District 39	1694	45297	26.7	1.9	17.7	10	32.4
District 14	2696	65731	24.4	1.6	17.8	8.1	33.9
District 9	1294	29664	22.9	2.6	16.6	5.5	36.4
District 16	1190	27209	22.9	2.2	17	6.9	34.9
District 8	2201	47964	21.8	2	17	4.6	37
District 6	1603	33111	20.7	2.8	16.1	6.4	35
District 17	1413	27207	19.3	.8	18	4.9	36.2
District 3	2745	51235	18.7	1.2	17.4	3.6	37.4
District 26	2122	37181	17.5	1	17.5	3.4	37.2
District 1	1189	20365	17.1	.6	17.8	4.4	36.2
District 5	1261	17473	13.9	.9	16.8	2.5	37.1
District 7	1994	26213	13.1	.6	17	2.8	36.6

Appendix Figure 1: Annual and Total Share of Non-Zero Crime Street Segments Accounting for 50 Percent of Crimes

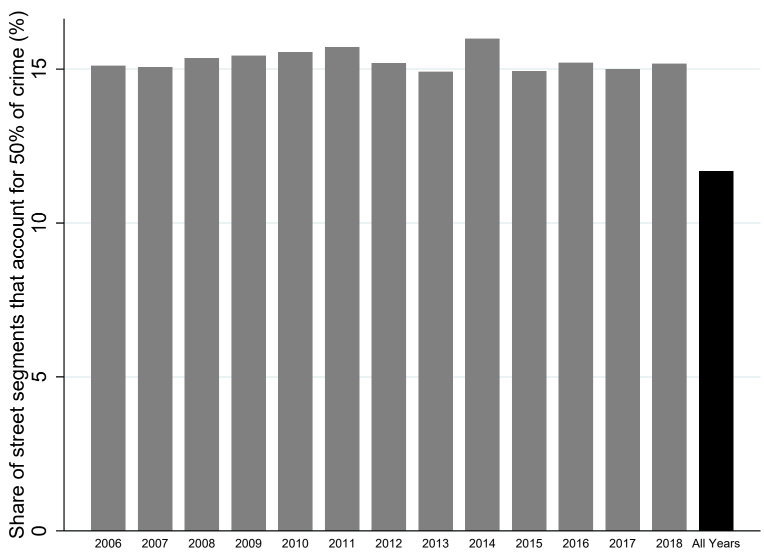
Panel A: New York



Panel B: Chicago



Panel C: Philadelphia



Note: Figures plot the share of non-zero crime street segments that account for 50 percent of total crimes in each city. The gray bars show the share of street segments for each year of data, the black bar shows the share when all years are combined.