

Justice Quarterly

ISSN: 0741-8825 (Print) 1745-9109 (Online) Journal homepage: www.tandfonline.com/journals/rjqy20

Ambient Population and Crime: The Fragility of Causal Links in Urban Environments

Ariadna Albors Zumel, Michele Tizzoni, Wilson Hernández & Gian Maria Campedelli

To cite this article: Ariadna Albors Zumel, Michele Tizzoni, Wilson Hernández & Gian Maria Campedelli (10 Nov 2026): Ambient Population and Crime: The Fragility of Causal Links in Urban Environments, Justice Quarterly, DOI: 10.1080/07418825.2026.2747582

To link to this article: https://doi.org/10.1080/07418825.2026.2747582

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group on behalf of Academy of Criminal Justice Sciences.

Published online: 10 Nov 2026.

Submit your article to this journal

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=rjqy20

JUSTICE QUARTERLY
https://doi.org/10.1080/07418825.2026.2747582

OPEN ACCESS Check for updates

Ambient Population and Crime: The Fragility of Causal Links in Urban Environments

Ariadna Albors Zumel[a], Michele Tizzoni[a], Wilson Hernández[b,c] and Gian Maria Campedelli[a,d]

[a]Department of Sociology and Social Research, University of Trento, Trento, Italy; [b]Department of Criminology, University of Pennsylvania, Philadelphia, PA, USA; [c]GRADE, Lima, Peru; [d]Mobile and Social Computing Lab, Fondazione Bruno Kessler, Trento, Italy

ABSTRACT

The relationship between ambient population and crime has long been studied, with competing theories suggesting that pedestrian activity may enhance guardianship and informal surveillance, but also increase suitable targets. Yet, empirical evidence remains mainly correlational, despite causal identification being essential for policy. We assess how robustly this link can be causally identified using high-resolution smartphone-derived footfall data from Baltimore, Chicago, and Philadelphia (2023–2025), estimating increasingly rigorous models—from naïve baselines to two-stage residual inclusion with distributed lags, fixed effects, and spatial spillovers. The estimated associations remain small and essentially unchanged after addressing endogeneity, while precision drops. The few significant effects are sensitive to crime type, temporal granularity, and context, and none survives multiple-testing correction. This fragility reflects operationalization, not theory: competing mechanisms move with headcounts, but in opposite directions, making null and offsetting effects indistinguishable. Ambient population should therefore be measured by who is present, where, and why, not how many.

ARTICLE HISTORY
Received 17 March 2026
Accepted 28 September 2026

KEYWORDS
Urban crime; ambient population; Poisson; fixed effects; causality; instrumental variables

Introduction

Urban environments are complex, dynamic systems where spatial, social, and economic factors interact to shape patterns of human activity and behavior. Crime is one such pattern, and one of particular concern in highly populated cities. Foundational theories offer competing mechanisms on how human presence affects its occurrence. On the one hand, the "eyes on the street" perspective, proposed by Jacobs (1961), posits that pedestrian activity enhances public safety: more people on the street across different times of day means more potential witnesses, which deters would-be offenders through continuous informal social surveillance. This deterrence effect is also central to routine activity theory (RAT) (Cohen & Felson, 1979), which formalizes the role of the capable guardian—a person whose mere presence can discourage a motivated

CONTACT Ariadna Albors Zumel ariadna.alborszumel@unitn.it
© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group on behalf of Academy of Criminal Justice Sciences. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

offender from committing a crime. On the other hand, the opportunity perspective suggests the opposite: increased foot traffic concentrates potential victims in predictable locations and times, elevating the risk of crime. This is also clearly reflected in RAT, where higher ambient population also increases the number of potential targets, hence creating more crime opportunities, and in crime pattern theory (CPT) (Brantingham & Brantingham, 1993), which states that crime clusters in locations that are part of people's daily routines, which are precisely locations where foot traffic is highest. Therefore, whether foot traffic increases or decreases crime depends on which mechanism is dominating. This tension, together with additional confounding factors that influence crime, such as police presence, residential density, and neighborhood socioeconomic conditions, makes it particularly challenging to make causal claims regarding the relationships between human mobility and crime. This challenge is further amplified by the fact that criminology has historically shied away from causal inference approaches for observational data, with most existing research focusing on correlation rather than causation (Sampson et al., 2013). Yet understanding causation is essential for theory testing and development as well as informing policy interventions, since knowing only that two phenomena co-occur does not tell us whether changing one will affect the other. Hence, without causal identification, research-informed policies risk implementing misguided interventions, potentially wasting resources or producing the opposite of the intended effect. This raises important questions: can the effect of ambient population on crime be causally identified at this level of aggregation and how difficult is it to causally identify the relationship between ambient population and crime? How important are, in this regard, our modeling choices? These questions also have theoretical ramifications, since the theories discussed above predict that ambient population activates several competing mechanisms pulling in opposite directions, so determining what an aggregate measure of human presence can and cannot reveal about them has direct implications for how these theories should be operationalized and tested. This matters for whether ambient population should be measured as a headcount or in terms of who is present, where, and why.

To address these questions, we leverage crime incidents data and high-resolution points-of-interest-centered *footfall* data, which captures the number of individuals with a cellphone present in specific points of interest (POIs) at specific times and serves as a measure of ambient population, from three U.S. cities collected between June 2023 and June 2025, at very fine temporal and spatial granularities (6 h and hexagons of 0.74 sq. km, respectively). We then integrate these data with relevant confounders and other covariates to examine the footfall–crime relationship. To this end, we take a cautionary approach, systematically testing whether footfall–crime links are robust to diverse factors. To achieve this, we begin by estimating progressively sophisticated models that address key sources of bias, with the goal of establishing whether the effect of ambient population on crime can be causally identified at this level of spatial aggregation. Afterwards, we test whether causal claims from our primary temporal granularity (6 h) remain stable when using three alternative temporal resolutions (1, 3, and 12 h), since temporal aggregation may shape which causal relationships are observable. Furthermore, we investigate the effect of the sample correction approach used and identify sources of heterogeneity in the

footfall–crime relationship by examining whether causal effects differ across hexagons with distinct socioeconomic and mobility characteristics. Therefore, the aim of this paper is not to establish a universal law regarding the effect of ambient population on crime, but to examine how robustly such a link can be identified using high-resolution POI-based footfall data, which, despite its coverage limitations, is one of the most spatiotemporally detailed measures of ambient population currently available for crime research, and to show how sensitive any such claim is to methodological and contextual choices.

Background

More People, Less Crime: Natural Surveillance and Guardianship

In the seminal work *The Death and Life of Great American Cities*, Jane Jacobs introduced the concept of "eyes on the street." This concept states that the almost unconscious and voluntary natural surveillance by residents and users of urban spaces serves as the primary mechanism for maintaining public safety, and that no amount of police can replace it when it has broken down. Furthermore, she states that there are three conditions necessary for effective natural surveillance. First, there must be a clear separation between public and private space so that the area requiring surveillance has defined boundaries. Second, there must be "eyes upon the street," belonging to the natural proprietors of the street (i.e. residents, shopkeepers, and regular users who have both the opportunity and motivation to observe their surroundings), so the buildings have to face the street. Third, sidewalks must have continuous users to have more effective eyes and to make the streets more interesting and entertaining to look at. This can be achieved through the presence of stores, bars, and restaurants, particularly through a mix of different businesses that give people reasons to use those streets and the ones connecting them. Once these three conditions are satisfied, having people on the street becomes an asset rather than a danger (Jacobs, 1961).

Jacobs' insights laid the groundwork for crime prevention through environmental design (CPTED), which emerged in the 1970s as an approach to reducing crime opportunities through architectural and urban design choices that maximize natural surveillance (Jeffery, 1971). CPTED also draws on rational choice theory in criminology, which posits that the decision to commit a crime is more strongly influenced by the perceived risk of being caught than by the expected reward (Cornish & Clarke, 1986). When more people are present and visible in a public space—creating more "eyes on the street"—potential offenders perceive heightened scrutiny and fewer escape routes, making criminal activity less attractive.

This deterrence logic is further formalized in RAT, which introduced the concept of the capable guardian, referring not only to formal agents such as police officers or security guards, but also ordinary citizens whose presence can deter crime. From this perspective, a higher ambient population increases the number of potential guardians, strengthening the deterrence effect described by Jacobs (Cohen & Felson, 1979).

However, despite its remarkable influence—or perhaps precisely for its remarkable influence—Jacobs' concept has faced significant academic scrutiny. Newman (1972)

argued that surveillance alone is not sufficient to deter crime, and that urban design must trigger a "natural human territoriality" by dividing public space into proprietary zones that motivate residents to intervene and self-police. Building on this, Linning and Eck (2021) introduced the "Neo-Jacobian" perspective, arguing that the most effective "eyes" actually belong to place managers (i.e. shopkeepers and property owners) rather than ordinary residents. Therefore, this perspective reframes social control not as an organic, emergent process, but as the deliberate, goal-driven actions of a small group that consciously shape or suppress criminal behavior at their specific properties. Besides this, researchers have found that a high density of "eyes" can lead individuals to be less likely to intervene because they assume others will—a phenomenon known as the "bystander effect"—posing a major limitation to the "eyes on the street" concept (Darley & Latané, 1968). Furthermore, there is a profound racial blind spot in the Jacobian framework since it has been shown that the "eyes" do not monitor everyone equally, with, for instance, Black men being subject to significantly higher levels of suspicion regardless of their behavior (Duneier, 1999).

These debates and critiques suggest that the protective effect of ambient population depends on who is watching, who is being watched, whether observers feel territorial responsibility, or whether diffusion of responsibility undermines collective vigilance. These fragilities are associated with structural disadvantages that some theories have conceptualized as limiting how communities are organized toward a common goal. In that line of thought, social disorganization theory (SDT), developed by the Chicago School (Shaw & McKay, 1942), argues that neighborhoods characterized by poverty, residential instability, and ethnic heterogeneity tend to develop weakened social networks and diminished capacity for informal social control. In such environments, the collective willingness to intervene—what Sampson et al. (1997) called "collective efficacy"—erodes precisely where it is most needed. Therefore, natural surveillance also depends on the socioeconomic and demographic composition of neighborhoods.

More People, More Crime: Targets, Offenders, and Opportunity

While RAT formalizes the protective role of guardians, it simultaneously highlights the opposite mechanism: higher ambient population also increases the pool of suitable targets and motivated offenders converging in space and time. More specifically, this theory posits that crime occurs when three elements converge in time and space: a motivated offender, a suitable target, and the absence of a capable guardian. This "crime triangle" shifted criminological attention away from offender pathology and toward the situational conditions that enable criminal events to occur (Cohen & Felson, 1979).

Integrating insights from RAT with environmental criminology gave rise to CPT, developed by Brantingham and Brantingham (1993), which is based on the observation that crime is not randomly distributed but concentrates at particular places and times. Specifically, this theory defines the locations that individuals routinely visit in their daily lives as "nodes" and the routes traveled between them as "paths." Through repeated travel along these paths and visitation of these nodes, individuals develop a mental map of their surroundings—called the "awareness space"—that includes

knowledge of potential targets, guardianship levels, and escape routes. CPT posits that offenders commit crimes primarily within their awareness space because they possess the local knowledge necessary to identify opportunities and assess risks. Within this framework, Brantingham and Brantingham (1995) further classified nodes by their relationship to crime, distinguishing crime generators (i.e. locations that draw large crowds for non-criminal reasons and produce crime as a by-product of the high footfall) from crime attractors (i.e. locations that attract motivated offenders specifically because of their known criminal opportunities), as well as crime detractors (i.e. locations which discourage offending). Although traditional formulations treated the criminogenic potential of these nodes as time-invariant, research has shown that it depends on the time of day, since offenders are more likely to commit crimes in locations they have regularly visited at the same time of day as the offense (van Sleeuwen et al., 2021).

Building directly on CPT, a substantial body of work has shown that the physical layout of the street network itself shapes how people move and, consequently, where crime concentrates (Davies & Bowers, 2018; Hillier, 2004; Kim & Hipp, 2020; Summers & Johnson, 2017). Therefore, from this perspective, what matters is not solely the count of people or their motivations, but how the underlying network morphology influences their movement.

Another opportunity-side mechanism is anonymity: large concentrations of people, particularly nonresidents, can lower offenders' perceived risk of recognition and apprehension. This lack of recognition weakens residents' sense of responsibility and their exercise of informal social control, leading to increased problems related to a lack of control (Taylor et al., 1981).

Therefore, from the opportunistic perspective, higher ambient population increases the pool of suitable targets and motivated offenders converging in space and time, as well as providing anonymity, which could result in an increase in crime. Combining this perspective with the deterrence mechanisms discussed above (namely, the increased presence of capable guardians who provide natural surveillance), we can conclude that ambient population activates four competing mechanisms (i.e. more targets, more offenders, more guardians, and more anonymity), which can offset one another. Therefore, the net effect of ambient population on crime depends not simply on the number of people present but on their composition (i.e. the relative proportions of the different elements), the specific locations where they converge, and the time of day at which this convergence occurs.

From Correlation to Causation: Challenges and Recent Advances

Despite the extensive empirical work done in crime research, this field has predominantly produced correlational rather than causal evidence (Owens & Ludwig, 2013). This is reflected in the fact that many of the studies published in leading criminology journals rely on cross-sectional data, meaning they capture a single snapshot in time and cannot be used to establish causality (Weisburd & Piquero, 2008). Therefore, a major critique of the field is that it has largely failed to accumulate verified knowledge or to falsify its major theoretical claims (Bernard, 1990; Weisburd & Piquero, 2008). Part of what underlies this stagnation is that the strongest empirical support for

established theories tends to emerge from the weakest studies (Pratt & Cullen, 2005), where empirical results are often too loosely connected to the theories they claim to test (Wikström & Kroneberg, 2022), and where causal conclusions are extracted from correlational results without adequately addressing identification problems (Nunes et al., 2025; Sampson et al., 2013). Moreover, evidence from a recent review suggests that more than half of all criminological studies lack sufficient statistical power to detect the observed effect sizes, with approximately a quarter exhibiting "dangerously low" levels of power, a condition that leads to inflated false-discovery rates and overestimated effect sizes (Barnes et al., 2020). However, it should be acknowledged that a growing subset of crime researchers has begun tackling these shortcomings by more explicitly incorporating approaches designed for causal inference for targeting a range of questions in recent years (e.g. Chalfin et al., 2022; Humphries et al., 2025; Lawrence et al., 2024; Tahamont, 2019), although this approach remains the exception rather than the rule in the field.

These shortcomings—the reliance on correlational designs and weak causal identification—are particularly pronounced in the study of ambient population and crime, where two issues stand out. The first is endogeneity, which manifests through both reverse causality (crime can also influence ambient population) and omitted variable bias (numerous factors, such as land use and socioeconomic characteristics, simultaneously affect both ambient population and crime). In the last decades, several methodologies have been developed and refined to address endogeneity, including difference-in-differences (Bertrand et al., 2004), regression discontinuity designs (Lee & Lemieux, 2010), instrumental variable (IV) approaches (Angrist & Krueger, 2001), and generalized method of moments (Hall, 2005). The applicability of these methods depends on the research design and the underlying generating process of a given dataset. Difference-in-differences requires discrete treatment events and parallel trends; regression discontinuity designs require clear threshold discontinuities; and generalized method of moments is best suited for panels with short time dimensions, as it suffers from instrument proliferation when the temporal dimension is large (Roodman, 2009). In contrast, IV approaches can handle panels with large temporal dimensions by leveraging exogenous sources of variation to isolate causal effects when a valid instrument is available. Therefore, they are more appropriate for this study given our panel structure with high temporal granularity. Within the IV approaches, two-stage least squares (2SLS) is the most common estimator for linear models (Angrist & Krueger, 2001). The alternative for nonlinear models is two-stage residual inclusion (2SRI), which uses a control function approach where first-stage residuals are included as additional regressors in the second stage to account for endogeneity (Terza et al., 2008). Hence, this method is particularly suited for count data models like Poisson regression, making it relevant for crime research where the outcome is the number of crime incidents. Particularly for this study, the use of this approach is crucial in order to be able to make causal claims, because part of the variation in footfall in a given place and time is generated by the same factors that also affect crime there (e.g. a transit disruption, or a police operation). Therefore, a regression model that uses all of the variation in footfall mixes the effect we want (i.e. how crime responds to a change in footfall that occurs for reasons unrelated to crime itself) with the effect of these unobserved disturbances. 2SRI allows us to separate the two. Specifically,

the intuition behind this approach is that in a first stage, footfall is predicted using only variation that is plausibly unrelated to these disturbances (known as the instruments), alongside the fixed effects and covariates. Whatever that regression fails to predict (known as the residual) is where the contaminated variation is concentrated. In the second stage, this residual is entered into the crime model as an additional explanatory variable. Holding it constant absorbs this contaminated variation, so the remaining coefficient on footfall is estimated from the clean variation alone (more technical details can be found in the Methodology section). However, IV approaches, and hence also 2SRI, are subject to important limitations, including bias toward ordinary least squares under weak instruments and high sensitivity to within-cluster error correlation and heteroskedasticity (Cameron & Miller, 2015; Keane & Neal, 2023). These issues should be addressed through instrument strength diagnostics, such as the Staiger–Stock F-statistic benchmark, and cluster-bootstrap inference (Cameron & Miller, 2015; Staiger & Stock, 1997).

The second challenge is measurement. Until recently, ambient population data lacked the spatiotemporal granularity needed to test the theories outlined above, forcing researchers to rely on crude proxies: census residential counts (Boggs, 1965), commuting-derived workday population (Stults & Hasbrouck, 2015), time-use surveys (Lemieux & Felson, 2012), and transportation surveys (Felson & Boivin, 2015). These sources offered valuable but inherently limited snapshots—annual or single-day, aggregated to census tracts or administrative boundaries—and could not capture how population fluctuates across hours and neighborhoods within a single day. However, the proliferation of smartphone-derived mobility data has begun to change this. Mobile phone traces, geotagged social media posts, and GPS data from apps now allow researchers to estimate population presence at fine spatiotemporal resolutions (Campedelli et al., 2026; Hipp et al., 2019; Song et al., 2023), opening the door to the kind of within-day, within-neighborhood analysis that the footfall–crime question demands. This new data granularity has fostered a stream of research aimed at constructing human mobility-based measures that predict crime, including networked exposure approaches such as the ecological network framework of Browning et al. (2017) and mobility-based measures of networked disadvantage such as the "triple disadvantage" framework of Levy et al. (2020) (see also Graif et al., 2021; Mahfoud et al., 2021; Wu et al., 2025). However, while these studies demonstrate the predictive value of mobility-derived measures, they pursue a different question from ours: they build structural predictors of crime, whereas our aim is to assess whether a causal effect of ambient population on crime can be identified using an aggregate measure of human presence.

To the best of our knowledge, there have only been two (very recent) attempts to address these challenges to investigate the causal relationship between ambient population and crime. Both studies use SafeGraph mobility data,[1] which has been employed in several prior crime studies with descriptive or predictive goals (e.g. Albors

[1]SafeGraph is a data company that provides anonymized location data from smartphone applications. See https://www.safegraph.com/.

Zumel et al., 2026; Situ, 2024; Wu et al., 2025), and use IV approaches, though with different identification strategies. On the one hand, Lafrogne-Joussier and Rollet (2025) leveraged a shift-share strategy based on the geographic origins of visitors and temporary closures of "L" rail stations in Chicago, finding that while an increase in ambient density leads to more recorded crimes in absolute terms, it significantly reduces victimization rates. On the other hand, Vachuska (2025) used street closures as exogenous shocks to local mobility in New York City, finding that while two-way fixed effects models show a significant positive correlation between visitors and crime, the IV estimates indicate no statistically significant causal effect. However, both studies rely on temporary disruptions to public infrastructure that affect a small portion of the population and induce non-routine behavioral responses. Hence, they capture exceptional situations rather than daily routines. The contradictory results and methodological limitations of these studies, together with the scarcity of empirical research in this area, underscore the need for further investigation using alternative approaches and diverse urban contexts.

Data

Study Area and Time Period

For this study, we selected three U.S. cities based on their high crime rates, their availability of high-resolution mobility data, and their diversity in urban shape and sociodemographic composition. Specifically, we chose Baltimore, MD (Bal); Chicago, IL (Chi); and Philadelphia, PA (Phi). Baltimore is a mid-sized port city with relatively low overall population density concentrated in specific urban cores, and it exhibits some of the highest crime rates in the nation. Chicago, by contrast, is a large, densely populated metropolitan area with a complex polycentric urban structure and moderate crime rates. Philadelphia represents an intermediate case: a dense mid-sized city with a more compact, monocentric form and moderate crime rates. Table 1 provides a summary of key characteristics for each city. The present analysis covers a two-year period, from June 30, 2023 to June 29, 2025,[2] thereby excluding the pandemic period to better capture usual crime patterns and urban mobility dynamics.

Crime Data

We utilized city-wide crime datasets publicly available from the open data portals of the three cities. The datasets comprised four main variables: the date and time of each crime incident, the type of crime, and the latitude and longitude coordinates of the location of the crime (with a certain degree of obfuscation).

Although the crime categories available varied between cities, we focused on categories that were consistently reported across all three locations and that are

[2]The study period is constrained by the mobility dataset, which includes data only up to the 26th week of 2025.

Table 1. Overview of the three U.S. cities included in this study, including demographic, spatial, and crime characteristics.

City	Region	Population	Area (km^2)	Population density	Crime rate
Baltimore (Bal)	South	565,239	210	2696	6997
Chicago (Chi)	Midwest	2,664,452	591	4510	4381
Philadelphia (Phi)	Northeast	1,550,542	347	4467	4037

Population density is expressed as residents per square kilometer, and crime rates (per 100,000 residents) are obtained from FBI UCR data.

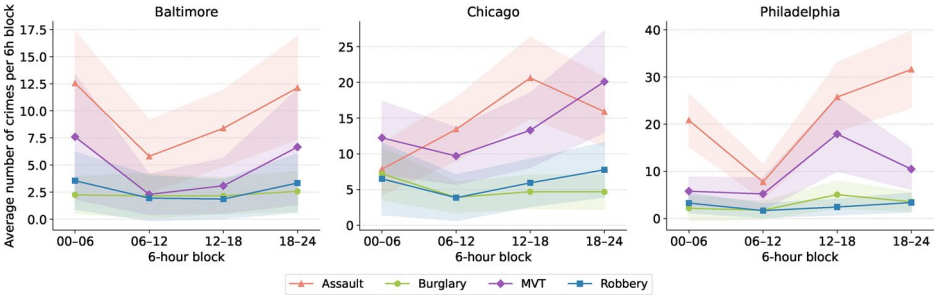


Figure 1. Average crime counts per crime type across the four 6-h blocks of the day for each city over the study period. Shaded areas indicate ±1 standard deviation (*SD*).

generally more reliably reported to law enforcement. This approach was critical to minimize the influence of underreporting biases (Xie & Baumer, 2019). Specifically, we selected burglary and motor vehicle theft (MVT) as property crimes, and assault and robbery as violent crimes.[3] Figure 1 illustrates the average crime count of each crime type across the day at our chosen temporal granularity, revealing that crime levels are generally lower during the morning hours [6:00–12:00) and peak in the evening [18:00–24:00).

Mobility Data

The second data source employed in this study is human mobility data obtained from Advan, a location intelligence company that provides micro-level mobility information derived from anonymized mobile device data from approximately 100 million cellphones of users aged 16 and older. Specifically, we utilized Advan's "Weekly Patterns" dataset[4] (Advan Research, 2022), which consists of data from over 11 million points of interest across the U.S. and Canada, starting from 2018 until the 26th week of 2025. From this dataset, we extracted the variable "visits_by_each_hour," which indicates the number of visits per hour at each POI (i.e. the footfall).

[3]Homicide was also consistently reported across the three cities and does not suffer from strong underreporting bias, but it was excluded from the analysis because of a data quality issue in one of the raw datasets.

[4]This dataset has recently been exploited in several studies examining urban environments from different perspectives. For some examples, see Amiri et al. (2025), Albors Zumel et al. (2026), and Wilk et al. (2026).

Table 2. Summary statistics of footfall over 6-h periods (after city-level sample correction) across cities during our study period, using two levels of spatial aggregation.

Aggregation	City	Mean	Median	Min	Max	SD	Var/Mean
City level	Bal	2,186,666	2,011,880	333,750	6,068,110	1,156,197	611,338
	Chi	14,246,660	13,031,425	2,049,600	40,470,060	8,343,621	4,886,478
	Phi	5,794,382	4,928,815	801,910	15,812,350	3,373,222	1,963,734
Hexagon level	Bal	24,448	21,382	1398	90,987	15,363	10,455
	Chi	78,953	68,005	3673	274,041	51,981	36,264
	Phi	55,769	44,997	2067	229,095	39,384	29,705

Hexagon-level statistics are computed per hexagon and then averaged across hexagons.

However, this dataset captures only a small sample of the total population present in these cities. Using the panel information provided by Advan,[5] we determined the sampling rate to be around 3% for our cities. Therefore, we corrected for sample bias following the approach used in Lafrogne-Joussier and Rollet (2025) (the specific procedure is explained in the methodology section). The summary statistics of the 6-h granularity footfall data after performing this city-level sample correction, aggregated at both the city and hexagon levels, can be found in Table 2. Summary statistics for the other granularities can be found in Appendix Table G2.

Confounders and Moderators

To ensure that we control for potential factors that may impact both the ambient population and the occurrence of crime, our dataset also considered three main variables that are known to have a confounding role in the relationship between ambient population and crime, which are holidays (Cohn & Rotton, 2003), major sport events (Copus & Laqueur, 2019), and other events with high attendance (Campedelli et al., 2024). The holidays for each city were obtained using the Python package "holidays,"[6] which provides federal holidays at the state level. We then complemented this list by manually adding city-specific holidays. The major sports events were constructed as a dummy variable indicating whether a major sports event occurred on a given day. This variable considered only home games (including playoffs) from Major League Baseball (MLB), National Basketball Association (NBA), National Football League (NFL, American football), and National Hockey League (NHL), and was manually generated based on schedules from Sports Reference.[7] The number of events for each city and each sports league can be found in Appendix A. Lastly, we compiled a dataset of other events that could disrupt ambient population and crime in each city. Specifically, we selected 22 unique events across the three cities, capturing a total of 43 event occurrences within the study period. These include large-scale music festivals, sporting events, cultural celebrations, and unexpected incidents. The complete list of events can be found in Appendix B.

As theoretical moderators for the heterogeneity analysis, we selected variables that are known to amplify or dampen the effect of ambient population on crime depending

[5]https://app.deweydata.io/data/advan/weekly-patterns/weekly-patterns-home-panel-summary
[6]https://pypi.org/project/holidays/
[7]https://www.sports-reference.com/

on the local context, and extracted a static measure for each of them at our spatial unit of analysis. Specifically, we selected two mobility-related variables and five sociodemographic-related variables.

As mobility-related moderators, we calculated the POI diversity and the percentage of night businesses, following Jacobs' argument that a mix of different establishments with diverse opening hours generates the continuous street activity necessary for effective natural surveillance (Jacobs, 1961). These moderators also help account for the fact that POIs are not equivalent in the amount of activity they draw or their relationship to crime, since they can be crime generators, crime attractors, crime detractors, or crime-neutral (Brantingham & Brantingham, 1995). The POI diversity index was computed by first mapping each POI into 11 categories, based on the NAICS system.[8] Then, for each spatial unit and each week, we counted the number of POIs in each category, computed the proportion of each category, and calculated the Shannon Diversity Index (SDI).[9] We then normalized it, resulting in a value between 0 (all POIs belong to one category) and 1 (perfectly even distribution across all categories). Lastly, we averaged each spatial unit's POI diversity value across all weeks in the study period to obtain a single static measure per spatial unit of analysis. To determine the percentage of night businesses, we classified each POI as a "night business" if its closing time exceeded 21:00 on any day of the week, based on the business's open hours. We then took the mode value per POI across all weeks to handle potential data inconsistencies. Finally, we calculated the proportion of POIs classified as night businesses in each spatial unit and multiplied it by 100.

As sociodemographic-related moderators, we selected median household income, percentage of unemployed, percentage of occupied housing units, percentage of single-parent households, and ethnic diversity based on empirical evidence indicating that economic disadvantage (Chamberlain & Hipp, 2015), residential stability (Boggess & Hipp, 2010), and social heterogeneity (Hipp, 2007) are fundamental dimensions of neighborhood context that shape crime dynamics and informal social control (Sampson & Groves, 1989). All five variables were derived from the American Community Survey (ACS) 5-year estimates for 2023 at the block group level, then aggregated to our spatial unit of analysis using area-weighted averages. The median household (HH) income was taken directly from the ACS. The percentage of unemployed was calculated as the number of unemployed persons divided by the labor force population, and the percentage of occupied housing units (HUs) as the number of occupied housing units divided by the total number of housing units, both multiplied by 100. The percentage of single-parent households was calculated as the number of single-parent families with own children under 18 (male or female householder with no spouse present) divided by the total number of families with own children under

[8]The specific categories used are: (1) Utilities and construction, (2) Manufacturing, (3) Retail and Wholesale Trade, (4) Transportation and Warehousing, (5) Business and Professional Services, (6) Educational Services, (7) Health care and Social Assistance, (8) Arts, entertainment, and recreation, (9) Accommodation and food services, (10) Public Administration, (11) Other Services.

[9]The SDI is defined as $H = -\sum_{i=1}^{N} p_i \ln(p_i)$, where p_i is the proportion of POIs belonging to category i and N is the total number of POI categories (11 in our case).

18, multiplied by 100. Lastly, the ethnic diversity was measured using a normalized SDI across six racial categories (i.e. White, Black, American Indian/Alaska Native, Asian, Native Hawaiian/Pacific Islander, and Other).

Methodology

Spatial Unit of Analysis

To conduct fine-grained, cross-city spatial analyses that estimate the area-level causal effect of ambient population on crime, our spatial unit of analysis must satisfy three requirements. First, it should have constant area, so that footfall and crime counts are comparable across units and not mechanically driven by differences in unit size. Second, it should exhibit uniform adjacency between neighboring cells, so that the spatial relationships between units are regular rather than an artifact of irregular boundaries. Third, it should be consistent across the three cities, in order to enable cross-city comparability. Traditional administrative units such as census blocks fail all three requirements since they vary significantly in size and shape across jurisdictions and within individual cities, so they lack constant area, exhibit irregular adjacency, and are not consistent across cities, introducing the modifiable-areal-unit-problem distortions our design is meant to avoid (Openshaw, 1984). Another common unit of analysis in crime research is street segments or street-network-defined communities (Grannis, 2009; Weisburd et al., 2012), which are designed to represent urban morphology. However, like administrative units, they have variable size, irregular adjacency, and city-specific structure, hence they fail the three requirements again. Moreover, they answer a different question since they are designed for micro-place crime concentration rather than for area-level causal estimation. Regular tessellations, by contrast, satisfy two of the requirements by construction, since they provide constant area and cross-city consistency. Among the three regular tessellations of the plane (i.e. triangles, squares, and hexagons), only hexagons provide unambiguous uniform adjacency, hence satisfying all three requirements, besides reducing edge effects and minimizing orientation bias (Birch et al., 2007; Ma et al., 2025). This makes hexagons particularly well-suited to modeling spatial processes such as crime and mobility, which often spread in all directions, and explains their growing presence in recent crime research (Harinam et al., 2022; Hoet, 2023; Khalfa & Hardyns, 2025; Kounadi et al., 2025; Pearcey et al., 2024). However, it should be noted that these properties of regular tessellations are obtained at the cost of abstracting away from the urban form that generates human mobility. Hexagons are regular precisely because they are arbitrary, since their boundaries carry no social or morphological meaning and cut across the street networks and land parcels that channel how people move through a city. We consider this an acceptable cost in this study because our estimand is an area-level causal effect that must be comparable across three cities, which is precisely what morphologically meaningful spatial units cannot provide. The implications of this abstraction are nonetheless discussed further in the limitations of this study. We consequently proceeded with hexagonal spatial units and implemented this tessellation using the H3 spatial indexing system (Uber Technologies, Inc., 2018), because its hierarchical structure provides efficient neighbor lookups and reproducible, consistent grids across

cities. However, any equal-area, uniformly adjacent grid (e.g. S2 or a custom hexagonal grid) would serve the same purpose. Specifically, we set the spatial resolution of H3 to 8, which corresponds to hexagons with an area of approximately 0.74 sq. km (0.3 sq. miles). It takes approximately 11 min to walk from one side of a hexagon to the other, which is a modest distance at the scale of everyday pedestrian activity. This resolution was chosen because it is fine enough to capture localized spatial variation while large enough to avoid excessive sparsity and is the most appropriate resolution for our analysis: resolution 7 (5.16 sq. km) would aggregate spatial units that are too large, whereas resolution 9 (0.11 sq. km) would generate overly small units, leading to severe sparsity in the data. Therefore, the chosen spatial resolution reflects a deliberate balance between spatial granularity and statistical reliability, which is a known methodological consideration in spatial crime research (Malleson et al., 2019).

From the obtained hexagons covering each city, we removed those mainly covered in water (>80% water based on OpenStreetMap), hexagons lacking sociodemographic data, and those containing no POIs, resulting in 255 hexagons for Baltimore, 657 for Chicago, and 413 for Philadelphia. A visualization of the resulting hexagonal grids can be found in Appendix E. However, when fitting each model, hexagons without any incidents for the focal crime over the entire study period are excluded from the analysis. The number of hexagons effectively used for each configuration can be found in Appendix Q.

Temporal Unit of Analysis

Regarding the temporal dimension, we aggregated our datasets into 6-h blocks, dividing each day into four intervals: night [0:00–6:00), morning [6:00–12:00), afternoon [12:00–18:00), and evening [18:00–24:00). This temporal resolution was chosen because 6-h blocks align with meaningful periods of daily activity that differ substantially in their social and commercial character and are each associated with distinct patterns of pedestrian activity and crime opportunities. For instance, residential burglaries tend to predominantly occur during morning and afternoon hours when residents are away at work or school (Coupe & Blake, 2006; Frith et al., 2022); and violent crimes such as street robbery tend to peak during evening and night hours, when darkness provides cover for offenders (Haberman & Ratcliffe, 2015; Tompson & Bowers, 2013). Similarly, pedestrian activity experiences commuting peaks on weekdays and more dispersed activity patterns on weekends (Dobler et al., 2021). Additionally, this granularity is coarse enough to reduce sparsity in crime data, yet fine enough to capture relevant patterns. However, to fully understand the sparsity–granularity tradeoff, we also estimate the model across alternative temporal resolutions: 1, 3, and 12 h.

Model Definition

Our model addresses several features of our data (i.e. the count nature of crime, the temporal dynamics in the footfall–crime relationship, and the spatial interdependence between neighboring areas), as well as two core empirical challenges (i.e. unobserved heterogeneity across locations and potential reverse causality). However, rather than presenting a single model that addresses all these issues, we take a cautionary approach by building a pipeline of increasingly complex models that gradually tackle

each feature and challenge of our setting. This structure is intended to mirror the sequence of modeling decisions that a criminologist would typically make when studying the relationship between ambient population and crime. Moreover, this pipeline illustrates how much methodological depth is needed before estimates can be given a causal interpretation and that reaching this point is not merely a technical refinement, but can fundamentally alter results.

We begin this gradual model construction by addressing the choice of the core regression model. Crime counts are discrete, non-negative integers that violate standard linear regression assumptions, requiring instead a generalized linear model designed for count data. Additionally, overdispersion (i.e. the conditional variance exceeding the conditional mean) is a well-recognized issue in criminological count data (Berk & MacDonald, 2008), which can bias standard errors. Therefore, to account for both the count nature of the data and any overdispersion present, we consider two models: Poisson and Negative Binomial. Both models focus on the expected value of the dependent variable (i.e. the expected crime counts) and use a log link function to relate the linear predictors to this expected outcome (Hilbe, 2011). However, although Negative Binomial relaxes Poisson's equal mean–variance assumption *via* a dispersion parameter, if its distributional assumption is itself misspecified (e.g. if the dispersion is not constant across observations), Negative Binomial estimates can be inconsistent (Cameron & Trivedi, 2013). Therefore, since our primary objective is causal inference rather than distributional fit, we choose Poisson, which is simpler, easier to interpret, and yields consistent estimates as long as the conditional mean is correctly specified, regardless of the true underlying distribution (Gourieroux et al., 1984; Silva & Tenreyro, 2006). However, while overdispersion does not threaten the consistency of its estimates, it can still downward bias the standard errors and distort inference. We therefore ensure valid inference under overdispersion by clustering standard errors at the hexagon level (specifically, using CRV1[10]), which is a well-known approach that is robust to overdispersion because it does not rely on the Poisson assumption of equal mean and variance (Cameron & Miller, 2015). Despite this, we also empirically assess overdispersion severity in the results section by comparing Poisson and Negative Binomial estimates for the baseline model. We thus define our baseline model (under both distributional assumptions to facilitate comparison, with Poisson as our preferred specification) to estimate the contemporaneous effect of footfall on crime counts, without any additional controls. Specifically, for each hexagon i and 6-h block t:

$$\log(\mu_{i,t}) = \alpha + \beta\,\text{Footfall}_{i,t}, \quad \text{with } \mu_{i,t} = \mathbb{E}[\text{Crime}_{i,t}\,|\bullet] \text{ and } \text{Crime}_{i,t} \sim \begin{cases} \text{P}\ (\mu_{i,t}) & \phi \to 0 \\ \text{NB}(\mu_{i,t}, \phi) & \phi > 0 \end{cases}, \quad (1)$$

where $\phi \geq 0$ is the dispersion parameter, with $Var[Crime_{i,t}] = \mu_{i,t} + \phi\mu_{i,t}^2$; α is the intercept, indicating the log of the expected baseline crime count when the footfall is zero;

[10]CRV1, also known as the standard CRVE, is a cluster-robust variance estimator that allows for arbitrary correlation of the error terms within clusters while assuming independence across clusters. For more details, see Cameron and Miller (2015).

and β is the slope coefficient, which indicates the expected change in the log of the crime count for a one-unit change in footfall. Hence, we will report our results as Incidence Rate Ratios (IRRs), $\exp(\beta)$, which give the multiplicative factor by which the expected crime count changes for a one-unit increase in footfall. An IRR greater than 1 indicates that the expected number of crimes increases with higher footfall, whereas an IRR below 1 indicates a decrease. The percentage change in the expected count is computed as $(IRR-1)\times 100$. Moreover, for interpretability, we report IRRs corresponding to a one average standard deviation (rounded to the nearest thousand) increase in footfall. Therefore, for example, an IRR of 1.012 means that approximately a one-standard-deviation increase in footfall is associated with a 1.2% increase in the expected number of crimes, while an IRR of 0.82 indicates that a one-standard-deviation increase in footfall is associated with an 18% decrease in the expected number of crimes.

Moreover, the effect of footfall on crime can occur over a certain time period, not only instantaneously during the same time block, meaning that crime may follow a build-up of ambient population rather than coinciding exactly with it. Therefore, in order to capture this prolonged exposure, as well as absorbing timestamp imprecision in crime reporting and controlling for short-run temporal autocorrelation in both footfall and crime, it is important to also include the hexagons' footfall during q previous time blocks. Adding this distributed lag structure to Equation 1, the model becomes:

$$\log(\mu_{i,t}) = \alpha + \sum_{k=0}^{q} \beta_k \text{Footfall}_{i,t-k}, \tag{2}$$

where the coefficient β_0 is interpreted as the immediate effect of footfall on the log of the expected crime count, while the sum $\sum_{k=0}^{q} \beta_k$ represents the long-run cumulative effect on the log of the expected crime count. We refer to this distributed lag specification as Model 2.

However, since Model 2 does not control for unobserved heterogeneity, which hinders any attempts to make causal claims, we introduce fixed effects (Wooldridge, 2010). This allows us to control for all unobserved characteristics that are constant over time within a hexagon (e.g. street layout and baseline socioeconomic status) and all common shocks that affect all hexagons at the same time (e.g. season, day of the week, and hour of the day). Modifying Equation 2 to include fixed effects, we obtain Model 3, which we define as:

$$\log(\mu_{i,t}) = \alpha_i + \delta_t + \sum_{k=0}^{q} \beta_k \text{Footfall}_{i,t-k} + \gamma X_{i,t}, \tag{3}$$

where α_i is the hexagon fixed effects—a unique intercept for each hexagon i controlling for all time-invariant, unobserved characteristics of that area. This makes the model more robust to omitted variable bias than a simple cross-sectional or pooled regression. The term δ_t represents the time fixed effects, which capture aggregate time-varying shocks that affect all hexagons simultaneously. Specifically, rather than including a full set of time-period dummies, it is implemented as a parametric decomposition consisting of four components: 6-h block of the day (4 categories), day of

the week (7 categories), month (12 categories), and year (3 categories). This accounts for predictable city-wide rhythms (e.g. rush hour and weekend activity), while still retaining sufficient within-period, cross-hexagon variation for identification. Lastly, the term $\gamma X_{i,t}$ is used to control for the time-varying covariates not captured by the fixed effects.

Yet crime patterns are not isolated to individual hexagons but are influenced by activity in neighboring areas. Failure to account for this spatial interdependence can lead to biased estimates. To address this, we extend Equation 3 to a modified Spatial Durbin model (SDM), which incorporates spillover effects by including both a spatio-temporally lagged dependent variable and spatially lagged independent variables (Anselin, 1988). This model has also been applied in criminology to capture spatial spillovers (Kim & Jeong, 2025; Maksuta et al., 2024; Zhang & Xiang, 2022). However, unlike the standard SDM, which includes contemporaneous spatial lags of the dependent variable inducing simultaneity, our modification additionally uses temporal lags to avoid endogeneity while maintaining the spatial spillover structure. Therefore, the full model incorporating the spatial spillover components, hereafter Model 4, is as follows:

$$\log(\mu_{i,t}) = \alpha_i + \delta_t + \rho(\mathbf{W}\text{Crime}_{t-1})_i + \sum_{k=0}^{q}(\beta_k \text{Footfall}_{i,t-k} + \theta_k(\mathbf{W}\text{Footfall}_{t-k})_i) + \gamma X_{i,t}, \quad (4)$$

where ρ is the spatial autoregressive coefficient, measuring how strongly neighboring hexagons' past expected log crime counts influence hexagon i's own expected log crime count. The parameter $\mathbf{W} \in \{0,1\}^{n \times n}$ is the spatial weights matrix, where n is the total number of hexagons. This matrix is defined such that for $i \neq j$, $w_{ij} = 1$ if hexagons i and j share a border and $w_{ij} = 0$ otherwise, and then row-standardized so that each row sums to one (see Appendix D). Hence, it is the row-standardized Queen contiguity matrix that describes the spatial relationship between all hexagons in our dataset (Cliff & Ord, 1981). Lastly, the coefficients θ_k measure the impact of neighboring footfall on the current hexagon's expected log crime count at lag k.

Despite Model 4 successfully controlling for unobserved heterogeneity, covariates, and spatial spillover effects, it may still suffer from several potential sources of endogeneity. The first source is reverse causality: while footfall affects crime, crime can also affect footfall (e.g. people may avoid areas after crimes). Second, there may be omitted time-varying confounders that affect both footfall and crime (e.g. transit disruptions and police activity), potentially biasing our coefficients. Therefore, footfall and the neighboring hexagons' footfall could be endogenous.

Since our most granular panel has a very large temporal dimension (hourly observations over two years) and the outcome is a nonlinear count, we address potential endogeneity of footfall using 2SRI, which is the recommended approach for nonlinear count models (Terza et al., 2008): despite its absence from the criminological literature, this approach is extensively employed in social science research, mostly in health economics, applied econometrics, development economics, and real estate economics (Benefield et al., 2020; Mery et al., 2016; Rahman et al., 2022; Yang & Zikos, 2023).

As the name indicates, 2SRI consists of two stages. In the first stage, every endogenous variable is modeled as a function of all the exogenous variables, fixed effects,

and instruments. In our case, building on Model 4 with $q = 2$, we identify three endogenous variables: the footfall at the current time block, since contemporaneous footfall and crime counts are likely subject to simultaneity bias or reverse causality; the footfall in neighboring hexagons at the current time block, since spatial spillovers imply that neighboring footfall may simultaneously affect and be affected by crime in the current hexagon; and the crime in the neighboring hexagons one time block earlier, since crime in nearby areas is often influenced by the same underlying factors (e.g. local events and policing activity) that affect crime in the current hexagon. Therefore, in the first stage, we run an ordinary least squares regression for each endogenous variable D with the following structure:

$$\widehat{D}_{i,t} = \alpha_i + \delta_t + \sum_{k=1}^{q}(\beta_k \text{Footfall}_{i,t-k} + \theta_k(\mathbf{W}\text{Footfall}_{t-k})_i) + \gamma X_{i,t} + \tau_1 Z_{1,i,t} + \tau_2 Z_{2,i,t}, \quad (5)$$

where the instruments Z_1 and Z_2 are lags 3 and 4 of the same variable for the footfall equations and lags 2 and 3 for the neighboring crime equation for the 6-h configuration. Regarding the other temporal granularities, we used the same lag structure, except for the footfall equations in the 1-h and 12-h configurations, where we used lags 4 and 5 and lags 2 and 3, respectively, in order to maintain a comparable temporal window across configurations (more details can be found in Appendix J). The validity of these instruments relies on the exclusion restriction, namely that these lagged variables used as instruments affect current crime only through their effect on their endogenous regressors and are otherwise uncorrelated with the structural error term. This assumption is plausible in our model because a violation of the exclusion restriction would require a shock that is hexagon-specific (otherwise absorbed by α_i); time-varying (otherwise absorbed by δ_t); not a holiday, sports event, or local disturbance (otherwise absorbed by $X_{i,t}$); and has no effect on the shorter lags included in the model but suddenly appears in the instrumented lags. This combination makes a violation highly unlikely. However, to further assess this assumption, we replicate our main results using deeper lags as instruments since consistent estimates would suggest that the exclusion restriction is not violated.

Afterwards, the residuals are calculated as the difference between the observed value and its prediction $\hat{u}_{i,t} = D_{i,t} - \widehat{D}_{i,t}$, and included in Model 4. This yields Model 5, specified as follows:

$$\log(\mu_{i,t}) = \alpha_i + \delta_t + \rho(\mathbf{W}\text{Crime}_{t-1})_i + \sum_{k=0}^{q}(\beta_k \text{Footfall}_{i,t-k} + \theta_k(\mathbf{W}\text{Footfall}_{t-k})_i) + \gamma X_{i,t} + \sum_{j=1}^{3}\lambda_j \hat{u}_{j,i,t}, \quad (6)$$

where λ_j is the coefficient associated with each of the three residual terms. If these coefficients are statistically significant, it indicates that the corresponding variable is indeed endogenous.

However, two additional steps are necessary to obtain reliable results from 2SRI. First, we need to verify instrument strength since weak instruments make 2SRI susceptible to finite-sample bias, which could result in unreliable estimates. We therefore conduct F-tests in each first-stage regression—following the conventional rule of

thumb, the F-statistics should exceed 10 to satisfy the weak instrument criterion (Staiger & Stock, 1997; Stock & Yogo, 2005). Second, 2SRI includes generated regressors in the second stage, and PPML standard errors ignore the estimation uncertainty this introduces. Therefore, in order to obtain asymptotically correct standard errors, it is necessary to use the non-parametric cluster bootstrap procedure (MacKinnon et al., 2023), resampling at the hexagon level and re-running the full 2SRI in each iteration (see Appendix I for details).

Sensitivity Analyses

Sample Correction

Mirroring the procedure followed by Lafrogne-Joussier and Rollet (2025), we corrected for sample bias by estimating each city's sampling rate as the average number of resident devices (obtained from Advan's Home Panel Summary) divided by the ACS population aged 16 and older. This resulted in sampling rates of 3.51% for Baltimore, 2.75% for Chicago, and 2.52% for Philadelphia. Afterwards, we multiplied the footfall data by the inverse of these sampling rates to approximate the true footfall.

However, using a uniform city-level scaling factor assumes spatially homogeneous device penetration, which is not perfectly satisfied in our dataset (see Figure H3). Therefore, we also performed a hexagon-level sample correction to evaluate the sensitivity of our results to this assumption. This was achieved by calculating the average number of resident devices per block group over 2023 and then their weighted contribution for each hexagon.[11] This resulted in one sampling rate per hexagon, so we multiplied the footfall data of each hexagon by the inverse of its sampling rate.

Despite the hexagon-level correction being more granular than the city-level correction, hexagon-level sampling rates can introduce bias by assuming that visitors have the same sampling rate as residents. Therefore, neither of these sample correction approaches perfectly corrects for sample bias, nor can either be considered better than the other, given the limited panel information available.

Spatial Heterogeneity

To investigate whether the footfall–crime relationship varies across different neighborhood contexts, we used our sociodemographic and POI-related static variables and performed a heterogeneity analysis using the interaction approach with a median split. Specifically, we added an interaction term to Model 5, defined as the interaction between contemporaneous footfall and all the hexagons whose heterogeneity variable was equal to or above the median value. Moreover, since the interaction with an endogenous variable is still endogenous, we included the interaction variable in the first stage and instrumented it using the interaction for lags 1 and 2. Therefore, in this analysis, the IRR for contemporaneous footfall represents the effect on crime in the low group (i.e. hexagons below the median), while the IRR for the interaction

[11]For a small number of hexagons whose overlapping block groups had no population estimates above 16 years of age—due to ACS data suppression in low-population areas—we imputed the average sampling rate.

term indicates the additional effect in the high group, with the total effect in the high group being the product of both IRRs. Since this approach does not provide the standard error, confidence interval, or *p*-value for the high group, we calculated them by estimating the total effect in each bootstrap iteration (i.e. the sum of the coefficients for the low group and the interaction) and deriving them from the bootstrap distribution.

Results

Overall Effects

In Appendix L we report the baseline model estimates under both Poisson and Negative Binomial for all crime types and cities. The results show that Poisson and Negative Binomial yield virtually identical results across all crime types and cities, suggesting that overdispersion is either negligible or adequately handled by the clustered standard errors. This validates our choice of Poisson as the preferred specification due to its consistency and interpretability.

Having validated the use of Poisson, the results from the baseline model and Models 2 to 4 show a systematic pattern across specifications for all crime types (see Figure 2): the baseline model yields the largest IRR, which decreases notably when including distributed lags (Model 2), decreases slightly again in Model 3, and remains relatively stable in Model 4. This was expected since, as we progressively account for different sources of bias, the inflated estimates are corrected downward. In particular, the decline from the baseline model to Model 2 suggests that contemporaneous and lagged footfall capture overlapping variation. Additionally, the slight decrease from Model 2 to Model 3 may reflect the fixed effects and covariates controlling for confounding or omitted variables. Finally, the stability from Model 3 to Model 4 suggests that spatial spillovers do not substantially alter the contemporaneous footfall–crime relationship. Moreover, from the significant results in Model 4, we can conclude that there is generally a positive, albeit extremely mild, association between footfall and crime, except for MVT in Baltimore and Philadelphia, which shows a negative association. In fact, the significant associations are small in magnitude across all four models, indicating that the relationship between ambient population and crime as captured by POI-centered footfall is modest.

It should be noted, however, that these models progressively account for confounding (through covariates, fixed effects, and spatial spillovers), but none address the endogeneity arising from reverse causality and simultaneity. Therefore, the estimates up to this point cannot yet be given a causal interpretation.

Correcting for Endogeneity

In the final model, we corrected for the remaining sources of endogeneity using 2SRI (Model 5) and performed cluster bootstrap with $B = 500$. Before interpreting the results, we assessed instrument strength using first-stage diagnostics. As shown in Appendix Table O25, 43 out of the 45 first-stage equations exceeded the conventional threshold of 10, with the exceptions being the instruments for the spatially lagged crime for

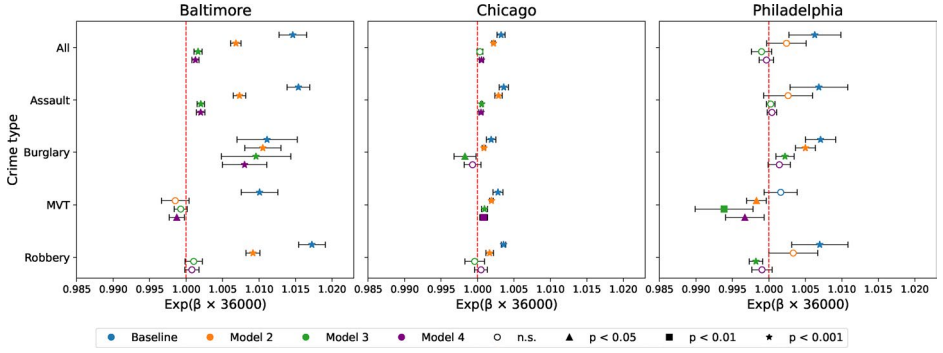


Figure 2. IRR estimates corresponding to a 36,000-person increase (one average standard deviation) in current footfall on different crime types across the three cities using our first four models. The error bars indicate 95% confidence intervals, and the red line indicates no effect.

assault in Baltimore, and robbery in Philadelphia. Therefore, the Model 5 estimates for these two configurations should be interpreted with caution since they could suffer from weak instrument bias. Afterwards, we evaluated instrument validity by replicating the estimates using deeper lags (lags 4–5) as alternative instruments for the footfall equations. As shown in Appendix Table M9, signs, magnitudes, and significance levels are virtually identical across instrument sets in almost every cell, providing empirical support for the exclusion restriction. With this support in place, we proceeded with the regression results. Throughout the remainder of the paper, we report significant estimates from this model as causal effects, since they are estimates from our causal specification; however, many of the effects that were significant in Model 4 became non-significant under Model 5 (see Figure 3) and, as shown in the following subsection, none of the effects in our main analysis survives correction for multiple testing. This underscores how substantially the picture can shift when moving from a correlational design to a causal one.

Additionally, from our full regression results (see Appendix N), we observe that in 7 out of 15 city–crime configurations at least one of the residual terms is statistically significant, confirming the presence of endogeneity and justifying the use of 2SRI. The footfall residuals are not significant in any of the configurations, but this should not be read as evidence of footfall being exogenous since these residuals suffer from the same reduced precision as the coefficients themselves.

Despite the overall reduction in significance, two causal effects remain: in Baltimore, a 36,000-person increase in footfall leads to a 0.9% increase in burglary; and in Philadelphia, the same increase leads to a 0.4% decrease in MVT. In absolute terms,[12] a 36,000-person increase in footfall within a single hexagon during a 6-h period leads to one additional burglary roughly every 8.5 years; and in Philadelphia, the same increase leads to one fewer MVT roughly every 7.2 years. These absolute magnitudes

[12]The effect in absolute terms is the percentage change multiplied by the average crime count per hexagon per 6-h period for the corresponding crime type (see Table K3), so it should be read as the effect for a hexagon at the mean count.

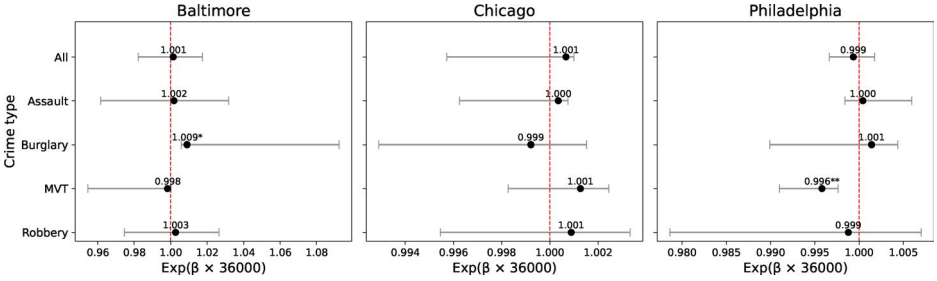


Figure 3. IRR estimates corresponding to a 36,000-person (one average standard deviation) increase in footfall across crime types and cities using Model 5. Error bars indicate 95% confidence intervals, and the red line denotes no effect. Stars indicate significance: *p<0.05, **p<0.01, ***p<0.001. Note: bootstrap confidence intervals use the percentile method, hence their asymmetry, and significance stars are based on two-sided bootstrap p-values.

are very small, reinforcing that the magnitude of the relationship between POI-centered footfall and crime is modest, even after the 2SRI correction. The remaining configurations show no significant causal effects due to widened standard errors rather than the point estimates shrinking toward zero (see Appendix N), which suggests that, under our causal specification, removing the remaining endogenous variation leaves too little identifying variation in an aggregate footfall measure to estimate these effects with sufficient precision.

Results with Alternative Temporal Granularities

We then repeated the analysis for Model 5 using 1-h, 3-h, and 12-h granularities. From the results in Figure 4, we see that which causal effects reach significance varies depending on the temporal granularity. Specifically, when using 1-h granularity, we find no significant effects with very wide CI for Burglary and MVT in Baltimore, likely due to the extreme data sparsity at this resolution. Similarly, the 12-h granularity yields mostly non-significant results, with only one exception: burglary in Chicago, where a 45,000-person increase in footfall causes a 0.2% decrease in burglary, hence one fewer burglary per hexagon roughly every 43.9 years. The limited significance at this coarser granularity is likely because this level of aggregation masks short-term temporal dynamics. Likewise, the 3-h granularity exhibits only one significant effect: in Chicago, a 19,000-person increase in footfall causes a 0.2% increase in overall crime (one additional crime per hexagon every 5.7 years). This variety of results suggests that different temporal granularities capture causal links from different perspectives, likely revealing temporal dynamics that qualitatively change when aggregating hours. This is because each temporal granularity defines a different exposure window and instrument configuration, and therefore these models estimate distinct causal contrasts rather than competing estimates of a single true effect. However, and importantly, after applying the Benjamini–Hochberg correction for multiple testing (Benjamini & Hochberg, 1995), across all configurations (including 6 h), none of the effects retain significance. In other words, the evidence, after accounting for the multiplicity of tests, is insufficient to confidently identify any specific causal relationship when studied at this level of aggregation.

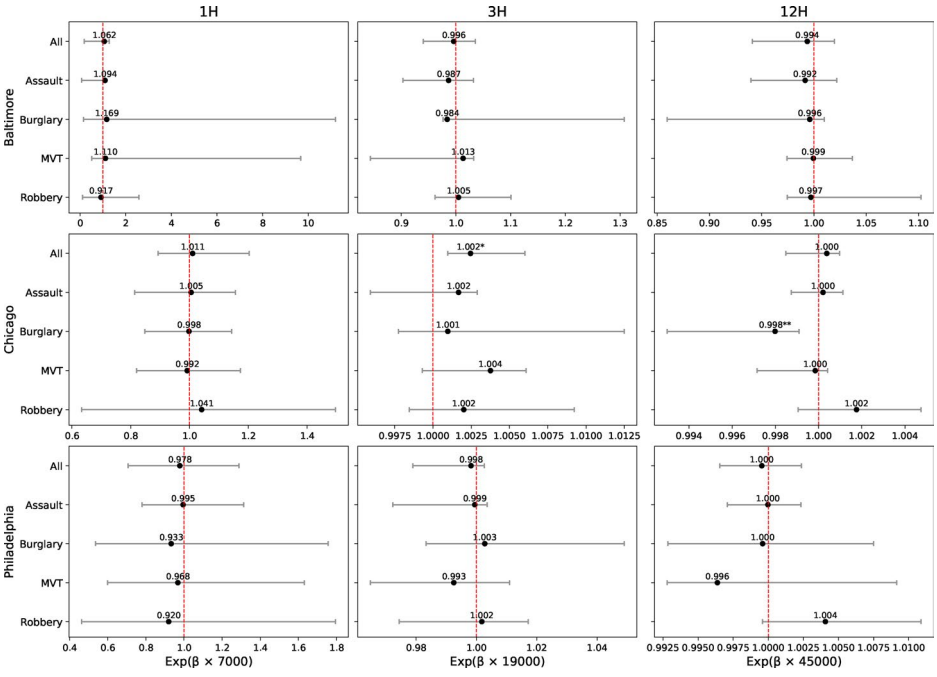


Figure 4. IRR estimates for three additional temporal granularities (1, 3, and 12 h), corresponding to a 1SD increase in current footfall (7000, 19,000, and 45,000 persons, respectively) across crime types and cities using Model 5. Error bars indicate 95% confidence intervals, and the red line denotes no effect. Stars indicate significance: *p<0.05, **p<0.01, ***p<0.001. Note: bootstrap confidence intervals use the percentile method, hence their asymmetry, and significance stars are based on two-sided bootstrap p-values.

Effects of Sample Correction

Repeating this analysis using a different sample correction (i.e. with hexagon-level instead of city-level sampling rates), we obtained very similar results, as shown in Figure 5. However, the hexagon-level correction results in generally wider confidence intervals, leading to the loss of one significant result obtained with the city-level correction: burglary in Baltimore. Despite this, the hexagon-level correction retains the significant causal effect for MVT in Philadelphia, although it has virtually no effect (IRR ≈1).

Heterogeneity Tests

Table 3 presents the results of our heterogeneity analysis for the configurations with significant interaction effects and at least one significant group effect, since these are the only cases that indicate meaningful heterogeneity in the footfall–crime relationship. The low-group effect represents the footfall–crime relationship in areas with low values of the heterogeneity variable (below median), and the high-group effect represents the relationship in areas with high values (above median).

The results vary considerably across cities. In Baltimore, we find no evidence that neighborhood characteristics moderate the footfall–crime relationship, since none of the interaction terms are significant. In Philadelphia, only two configurations show significant

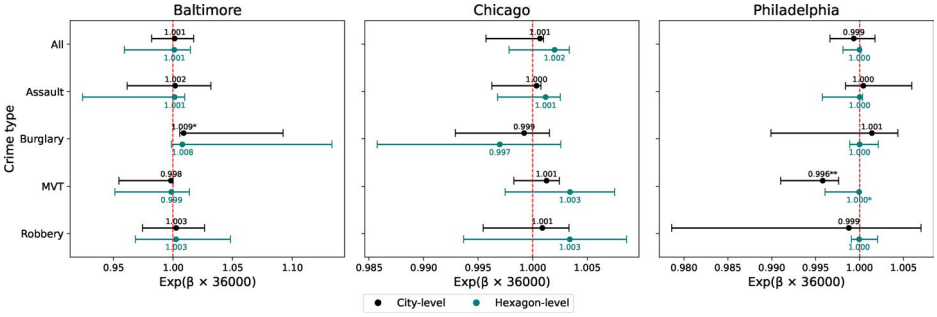


Figure 5. IRR estimates corresponding to a 36,000-person increase (one average standard deviation) in current footfall across crime types and cities using Model 5 with city-level and hexagon-level sample correction. Error bars indicate 95% confidence intervals, and the red line denotes no effect. Stars indicate significance: *p<0.05, **p<0.01, ***p<0.001. Note: bootstrap confidence intervals use the percentile method, hence their asymmetry, and significance stars are based on two-sided bootstrap p-values.

Table 3. IRR corresponding to a 36,000-person increase (one average standard deviation) in footfall for the configurations of Model 5 with significant interaction effects and at least one significant group effect.

	Crime type	City	Low	High
POI Diversity	Burglary	Chi	0.999	0.978**
	MVT	Phi	0.996*	0.980**
% Night Business	All	Phi	1.000	0.998**
	Burglary	Chi	0.999	0.965**
	MVT	Chi	1.001	1.010**
Median HH Income	All	Chi	1.084**	1.001
	Assault	Chi	1.081**	1.000
	MVT	Chi	1.077**	1.001
% Unemployed	All	Chi	1.001	1.015**
	Assault	Chi	1.000	1.018**
	Robbery	Chi	1.001	1.032**
% Occupied HUs	Burglary	Chi	0.999	0.953**
	MVT	Chi	1.001	1.010**
Ethnic Diversity	MVT	Chi	1.015**	1.001
% Single-parent HH	All	Chi	1.001	1.131**
	Assault	Chi	1.000	1.104*

Stars denote significance: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. Full results in Appendix P. *Note*. Significance stars are based on two-sided bootstrap p-values.

effects: a 36,000-person increase in footfall in a hexagon causes a 0.4% decrease in MVT in areas with low POI diversity, hence a 36,000-person increase in footfall within a single hexagon during a 6-h period leads to roughly one fewer MVT every 7.2 years (2.0% decrease in high POI diversity areas, or roughly one fewer MVT every 1.4 years), and has no significant effect on overall crime in areas with low percentage of night businesses but causes a 0.2% decrease in areas with high percentage of night businesses (about one fewer crime per hexagon every 3.8 years). In contrast, Chicago exhibits the strongest evidence of heterogeneity, particularly regarding median household income. In low-income areas, a 36,000-person increase in footfall causes an 8.4% increase in overall crime, an 8.1% increase in assault, and a 7.7% increase in MVT (so one additional crime per hexagon every 1.6 months, one additional assault every 4.6 months, and one additional MVT every 5.1 months). However, these effects are completely attenuated in high-income areas, where

the total effect is essentially null. A similar pattern emerges for ethnic diversity, where a 36,000-person increase in footfall causes a 1.5% increase in MVT (one additional MVT per hexagon every 2.2 years) in areas with low ethnic diversity, and this effect is again completely attenuated in high ethnic diversity areas. Conversely, in areas with a high proportion of unemployed residents, footfall increases overall crime by 1.5%, assault by 1.8%, and robbery by 3.2% (one additional crime per hexagon every 9.1 months, one additional assault every 1.7 years and one additional robbery every 2.3 years), with no significant effect in low-unemployment areas. Likewise, in areas with a high proportion of single-parent households, a 36,000-person increase in footfall increases overall crime by 13.1% and assault by 10.4%, with no significant effect in low single-parent-household areas (about one additional crime every month, and one additional assault every 3.6 months). Furthermore, in areas with a high proportion of occupied housing units, footfall decreases burglary by 4.7% but increases MVT by 1.0% (one fewer burglary per hexagon every 1.9 years and one additional MVT every 3.2 years), while in areas with a high percentage of night businesses, footfall decreases burglary by 3.5% but increases MVT by 1.0% (one fewer burglary every 2.5 years and one additional MVT every 3.2 years). Finally, in areas with high POI diversity, footfall decreases burglary by 2.2% (one fewer burglary every 4.0 years). Therefore, taken together, these results indicate that the causal effect of ambient population on crime at this level of aggregation is highly context-dependent for some cities.

Discussion and Conclusion

Given the longstanding theoretical and empirical controversy concerning the link between ambient population and crime, this study critically examined whether a causal relationship can be reliably established, and how sensitive any such claims are to methodological and contextual factors. Our results revealed the following key findings.

First, the gradual construction of our model and the change in the magnitude of the respective coefficients show that traditional regression models evaluating the effect of ambient population on crime are biased by the past footfall, confounders, omitted variables, and spillover effect from surrounding areas. Moreover, without controlling for endogeneity, these traditional modeling approaches give the impression that an increase in ambient population increases crime for most cities and crime types.

Second, once we correct for reverse causality and simultaneity, the point estimates remain essentially unchanged, but their bootstrap standard errors widen substantially, resulting in the loss of statistical significance for most results. Therefore, for most city–crime configurations, there is a modest association between ambient population and crime, but we cannot identify a causal effect of changes in ambient population on crime. However, it is important to note that this is not a story of large effects being shrunk to nothing. The original associations were already small, so the additional correction renders already-modest effects statistically nonsignificant, because their reduced precision makes them indistinguishable from zero. Moreover, these causal relationships (or their absence) vary substantially depending on the temporal granularity used. This reflects that the tradeoff between having enough resolution to capture meaningful crime patterns while simultaneously avoiding excessive crime data sparsity strongly determines which causal effects can be detected. This temporal scale dependence is itself a substantive finding because each temporal granularity defines

a different exposure window and instrument configuration, making causal claims about footfall and crime inherently scale-specific. The Benjamini–Hochberg correction for multiple testing performed on the main analysis (i.e. excluding the two sensitivity tests) strengthens this conclusion: given the volume of tests performed across cities, crime types, and temporal granularities, the handful of significant effects detected cannot be distinguished from what chance alone would produce.

Third, causal claims are sensitive to the sample correction approach used, since the hexagon-level correction generally produced wider confidence intervals than the city-level correction, reducing statistical significance in some configurations.

Fourth, the heterogeneity results reveal notable city-level differences in how neighborhood context shapes the footfall–crime relationship at 6-h granularity. In Baltimore, none of the interaction effects reach significance, so we find no evidence that neighborhood characteristics moderate this relationship, implying that city-level factors may dominate over local context. In contrast, several suggestive patterns emerged in Chicago and Philadelphia. First, in Chicago, areas with high occupancy and sustained nighttime activity tend to suppress burglary but facilitate MVT, while areas with diverse establishments only suppress burglary. This suggests that activity-generating environments enhance surveillance and guardianship against some crime types while increasing target availability for others, which would be consistent with routine activity theory. In Philadelphia, areas with diverse establishments and sustained nighttime activity tend instead to suppress MVT and overall crime, respectively. Second, in Chicago, an increase in ambient population increases crime in socioeconomically disadvantaged areas (i.e. low median household income, high unemployment, and a high proportion of single-parent households), with effects completely attenuated in more advantaged areas, suggesting that disadvantaged neighborhoods lack the social controls needed to offset the crime-generating potential of increased ambient population, which would be consistent with social disorganization theory. Lastly, the strong positive effect of ambient population on MVT in low ethnic diversity areas in Chicago may reflect reduced informal social control in homogeneous environments, where social familiarity could diminish vigilance. However, since this is a sensitivity analysis outside the multiple-testing correction, we interpret these patterns with caution and do not treat them as establishing which neighborhood characteristics moderate the effect of footfall. Instead, the substantial within- and between-city variation observed in the footfall–crime relationship indicates that a single, context-invariant parameter is unlikely to be the right measure for this effect and is itself a warning against making universal claims.

Our findings also carry theoretical implications linked to our operationalization of ambient population. As we described in the Background section, current theories and perspectives distinguish four competing mechanisms (i.e. more targets, more offenders, more guardians, and more anonymity) activated by ambient population which simultaneously affect crime in opposite directions. However, we operationalized ambient population using a headcount, which represents the net offsetting of these effects, and hence a null estimate on this aggregate measure cannot distinguish between no mechanisms operating, or several operating and canceling each other out. This is an identification problem inherent in the standard operationalization of the theory relating ambient population and crime, not merely a limitation of our data, and our results express the same problem in an empirical form: once the variation in footfall that is endogenous to crime is removed, an aggregate headcount is not sufficient to identify

a stable causal effect, with signs varying across cities and crime types. This is then further compounded by our POI-based measure, since it observes human presence anchored to establishments rather than on the street itself, and the four mechanisms are not evenly distributed across these settings (e.g. the anonymity provided by crowds is a property of street-level human presence, while the guardianship provided by place managers is mainly confined to their establishments). Unlike the aggregation problem described above, this second constraint is specific to our measurement. Therefore, the causal fragility that we report should not be read as evidence against any of these theories, but as evidence that aggregate measures of ambient population cannot differentiate the postulated competing mechanisms. This suggests that what should be reconsidered is not these theories' core claims, but the presumption that aggregate headcounts are a sufficient measure of ambient population. Hence, this measure should be reconceptualized in terms of who is present and why rather than how many. The elements to carry this out are already available in the literature reviewed above. The Neo-Jacobian perspective indicates that guardianship depends on whose eyes are on the street, with place managers exerting a stronger form of social control than regular pedestrians (Linning & Eck, 2021); social disorganization and collective efficacy indicate that whether human presence acts as informal social control depends on the setting's capacity to organize it (Sampson et al., 1997); the empirical work of Song et al. (2023) decomposes ambient population by role, showing that residents, workers, and visitors relate differently to crime; networked approaches show that visitors carry the disadvantages of their origin neighborhoods with them, so local crime depends not only on the characteristics of an area but also on those of the areas its visitors come from (Browning et al., 2017; Levy et al., 2020); and time-centered extensions of crime pattern theory indicate that the same volume of human presence carries different criminogenic meaning depending on when it occurs, since its composition varies (van Sleeuwen et al., 2021). A causal test of the relationship between changes in ambient population and crime should therefore decompose ambient population along these dimensions rather than using a single aggregate. Our heterogeneity results are consistent with this reading; however, since they are sensitivity analyses outside the multiple-testing correction carried out in our main results, we do not interpret them as identifying which moderators matter in which cities. Nevertheless, they do indicate that a single, context-invariant measure is the wrong target for this literature, since the same increase in human presence is associated with different responses across diverse neighborhood contexts, which is precisely what would be expected if distinct mechanisms dominate in distinct settings.

While the extensive sensitivity analyses partially address the robustness of these findings, the present study has several limitations. Arguably, the main limitation is the quality of our human mobility data. Although the high temporal and spatial resolution of our dataset improves upon prior available data, it is still not a perfect representation of ambient population. This is primarily because only visits to POIs are counted, so individuals remaining on the streets are excluded. Consequently, our measure of ambient population is suffering from selection bias toward POI-related activity, hence likely underestimating human presence in residential areas and in outdoor public spaces. As discussed in the previous paragraph, this also restricts which of the competing mechanisms our measure is able to observe. More critically, this dataset represents

only ~3% of the population, so sample correction approaches were necessary. However, sampling rates could only be calculated using the resident population, so the sample corrections applied could not account for visitors. Hence, while these sample corrections reduced sample bias, they also introduced their own bias and could not be used to recover the unobserved non-POI activity. These measurement limitations may be partially reflected in the consistently small magnitudes we observe across all model specifications, since a POI-based measure that captures only a portion of ambient population would tend to attenuate any underlying effect on crime. Additionally, this dataset only provided one mobility measure, footfall, lacking the more nuanced information that could be obtained from origin–destination flows. Future research should explore human mobility flow datasets that have already been sample-corrected and that give a measure of everyone who is in that area, not just POI visitors.

Another data-related limitation is that the ambient population associated with specific events (which we used as covariates) is not homogeneous. That is due to attendance at such events being non-random—biased, for instance, toward individuals with the economic means to attend—and shaped by factors such as intoxication and event policing that our data cannot capture. Future work could mitigate these biases using additional event-specific information, but doing so was beyond the scope of this paper.

Moreover, a limitation of our study period is that although it begins in mid-2023 to avoid the core pandemic period, we cannot rule out that lingering post-pandemic shifts in commuting, downtown activity, and crime reporting influenced the early part of the study window. Future work could test this by redoing the analysis with later, fully post-pandemic periods once the data becomes available.

Another important limitation concerns our spatial unit of analysis. First, at 0.74 sq. km, our hexagons are coarser than the micro-places at which routine activity theory is increasingly operationalized, such as street segments, addresses, and individual facilities (Weisburd et al., 2012). Therefore, if guardianship and the convergence of offenders and targets operate at this micro scale, our coefficients are spatial averages over many such micro-environments, within which crime-generating and crime-suppressing effects may offset one another. This resolution reflects the best available balance between granularity and sparseness given our crime and mobility data, but it is a tradeoff, and the possibility that it masks more micro-level effects is a genuine cost of our design. Our estimates should therefore be read as area-level effects, and the absence of a detectable effect at this resolution should not be taken as evidence of an absence at the micro-place level. Second, with the current set-up we were unable to test other spatial granularities since H3 does not offer enough granularity choices. Therefore, we could not assess whether causal links are also sensitive to the chosen unit of analysis. Future work should build its own self-consistent hexagonal grids, without using H3, to explore this dimension. Lastly, as already anticipated, hexagonal units are arbitrary with respect to urban form, since they capture spatial proximity but not the morphology of the street network that channels movement and crime within them. This matters for interpretation as well as for design, since the boundaries are drawn without regard to where one functional environment ends and another begins. Therefore, footfall and crime belonging to one underlying process can be divided between neighboring observations, while different underlying processes can fall within the same spatial unit despite generating very different dynamics. Future research could thus incorporate network-derived measures to assess how the street network conditions the effect.

Lastly, the heterogeneity analysis employed a limited set of variables, since it was only used as a sensitivity test in this study. Further research should expand this analysis using additional variables, as it is possible that Baltimore (or other cities) could exhibit significant heterogeneous effects when examined through additional dimensions not considered in the present study.

In conclusion, this study sought to make a crucial contribution by showing that the causal relationship between ambient population and crime is far more difficult to isolate than previously recognized. For decades, foundational criminological theories and perspectives have posited that ambient population influences crime through mechanisms such as natural surveillance, guardianship, target availability, anonymity, and collective efficacy. Critically, we show that while associations between aggregate ambient population and crime are widespread, they are also modest in absolute terms, and once reverse causality and simultaneity are addressed, the point estimates remain essentially unchanged, but their precision drops sharply and no effects survive correction for multiple testing. Our findings thus serve as a cautionary tale against drawing strong causal claims from correlational evidence, particularly when empirical results are intended to inform policy, and against using aggregate measures of human presence to test the causal relationship between changes in ambient population and crime. Despite the limitations noted, this work advances the field by highlighting the modeling complexity needed to address all biases encountered when studying this causal relationship, noting that even with rigorous bias correction, the detected causal effects remain highly sensitive to methodological and contextual factors. Future research should therefore prioritize nuanced causal analyses over universal claims and should use more accurate human mobility data that captures not only how many people are present, but who is present, where, and why.

Acknowledgment

The authors acknowledge support from the Trento Sociology Department Open Access Initiative.

Author contributions

CRediT: **Ariadna Albors Zumel**: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing; **Michele Tizzoni**: Conceptualization, Funding acquisition, Project administration, Supervision, Writing – review & editing; **Wilson Hernández**: Conceptualization, Resources, Writing – review & editing; **Gian Maria Campedelli**: Conceptualization, Funding acquisition, Methodology, Project administration, Supervision, Validation, Writing – review & editing.

Disclosure Statement

No potential conflict of interest was reported by the author(s).

Funding

Data acquisition has been possible due to the financial support from Fondazione Bruno Kessler.

ORCID

Ariadna Albors Zumel http://orcid.org/0009-0000-8193-1111
Gian Maria Campedelli http://orcid.org/0000-0002-7734-7956
Michele Tizzoni http://orcid.org/0000-0001-7246-2341
Wilson Hernández http://orcid.org/0000-0002-1280-5709

Data Availability Statement

All code, generated confounders & moderators datasets, and links to the publicly available crime data sources used in this study have been deposited in Zenodo and are accessible via DOI: 10.5281/zenodo.20698336.

References

Advan Research. (2022). *Foot traffic/weekly patterns* [dataset]. https://doi.org/10.82551/X1PP-1F65

Albors Zumel, A., Tizzoni, M., & Campedelli, G. M. (2026). Deep learning for crime forecasting: The role of mobility at fine-grained spatiotemporal scales. *Journal of Quantitative Criminology*, *42*(3), 511–568. https://doi.org/10.1007/s10940-025-09629-3

Amiri, H., Song, J., Xiong, L., & Züfle, A. (2025). Generating and understanding large datasets of social contact patterns from foot-traffic data. In *Proceedings of the 1st ACM SIGSPATIAL International Workshop on Geospatial Computing for Public Health* (pp. 14–17). ACM. https://doi.org/10.1145/3764918.3770162

Angrist, J. D., & Krueger, A. B. (2001). Instrumental variables and the search for identification: From supply and demand to natural experiments. *Journal of Economic Perspectives*, *15*(4), 69–85. https://doi.org/10.1257/jep.15.4.69

Anselin, L. (1988). *Spatial econometrics: Methods and models* (Vol. 4). Springer.

Barnes, J. C., TenEyck, M. F., Pratt, T. C., & Cullen, F. T. (2020). How powerful is the evidence in criminology? On whether we should fear a coming crisis of confidence. *Justice Quarterly*, *37*(3), 383–409. https://doi.org/10.1080/07418825.2018.1495252

Benefield, J. D., Cain, C. L., & Maynard, N. (2020). Solving old puzzles with new tricks: Addressing endogeneity and nonlinearity in time-on-market research. *Journal of Real Estate Research*, *42*(3), 420–438. https://doi.org/10.1080/08965803.2020.1844476

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Berk, R., & MacDonald, J. (2008). Overdispersion and Poisson regression. *Journal of Quantitative Criminology*, *24*(3), 269–284. https://doi.org/10.1007/s10940-008-9048-4

Bernard, T. J. (1990). Twenty years of testing theories: What have we learned and why? *Journal of Research in Crime and Delinquency*, *27*(4), 325–347. https://doi.org/10.1177/0022427890027004002

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics*, *119*(1), 249–275. https://doi.org/10.1162/003355304772839588

Birch, C. P., Oom, S. P., & Beecham, J. A. (2007). Rectangular and hexagonal grids used for observation, experiment and simulation in ecology. *Ecological Modelling*, *206*(3–4), 347–359. https://doi.org/10.1016/j.ecolmodel.2007.03.041

Boggess, L. N., & Hipp, J. R. (2010). Violent crime, residential instability and mobility: Does the relationship differ in minority neighborhoods? *Journal of Quantitative Criminology*, *26*(3), 351–370. https://doi.org/10.1007/s10940-010-9093-7

Boggs, S. L. (1965). Urban crime patterns. *American Sociological Review*, *30*(6), 899–908. https://doi.org/10.2307/2090968

Brantingham, P., & Brantingham, P. (1993). Environment, routine, and situation: Toward a pattern theory of crime. In R. V. Clarke & M. Felson (Eds.), *Routine activity and rational choice* (pp. 259–294, Vol. 5). Transaction Publishers.

Brantingham, P., & Brantingham, P. (1995). Criminality of place: Crime generators and crime attractors. *European Journal on Criminal Policy and Research*, *3*(3), 5–26. https://doi.org/10.1007/BF02242925

Browning, C. R., Calder, C. A., Boettner, B., & Smith, A. (2017). Ecological networks and urban crime: The structure of shared routine activity locations and neighborhood-level informal control capacity. *Criminology*, *55*(4), 754–778. https://doi.org/10.1111/1745-9125.12152

Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, *50*(2), 317–372. https://doi.org/10.3368/jhr.50.2.317

Cameron, A. C., & Trivedi, P. K. (2013). *Regression analysis of count data* (2nd ed.). Cambridge University Press.

Campedelli, G. M., Jelveh, Z., Chalfin, A., Semenza, D., Piza, E., Albors Zumel, A., Lepri, B., & Sharkey, P. (2026). Computational approaches and the future of urban crime research. *Nature*, *655*(8122), 315–326. https://doi.org/10.1038/s41586-026-10622-4

Campedelli, G. M., Piza, E. L., Piquero, A. R., & Kurland, J. (2024). Evidence on the impact of the Prudential Center on crime in downtown Newark. *Journal of Experimental Criminology*, *20*(4), 1225–1251. https://doi.org/10.1007/s11292-023-09576-8

Chalfin, A., Hansen, B., Lerner, J., & Parker, L. (2022). Reducing crime through environmental design: Evidence from a randomized experiment of street lighting in New York City. *Journal of Quantitative Criminology*, *38*(1), 127–157. https://doi.org/10.1007/s10940-020-09490-6

Chamberlain, A. W., & Hipp, J. R. (2015). It's all relative: Concentrated disadvantage within and across neighborhoods and communities, and the consequences for neighborhood crime. *Journal of Criminal Justice*, *43*(6), 431–443. https://doi.org/10.1016/j.jcrimjus.2015.08.004

Cliff, A., & Ord, J. (1981). *Spatial processes: Models & applications*. Pion.

Cohen, L. E., & Felson, M. (1979). Social change and crime rate trends: A routine activity approach. *American Sociological Review*, *44*(4), 588–608. https://doi.org/10.2307/2094589

Cohn, E. G., & Rotton, J. (2003). Even criminals take a holiday: Instrumental and expressive crimes on major and minor holidays. *Journal of Criminal Justice*, *31*(4), 351–360. https://doi.org/10.1016/S0047-2352(03)00029-1

Copus, R., & Laqueur, H. (2019). Entertainment as crime prevention: Evidence from Chicago sports games. *Journal of Sports Economics*, *20*(3), 344–370. https://doi.org/10.1177/1527002518762551

Cornish, D. B., & Clarke, R. V. (1986). *The reasoning criminal: Rational choice perspectives on offending*. Springer-Verlag.

Coupe, T., & Blake, L. (2006). Daylight and darkness targeting strategies and the risks of being seen at residential burglaries. *Criminology*, *44*(2), 431–464. https://doi.org/10.1111/j.1745-9125.2006.00054.x

Darley, J. M., & Latané, B. (1968). Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology*, *8*(4), 377–383. https://doi.org/10.1037/h0025589

Davies, T., & Bowers, K. J. (2018). Street networks and crime. In G. J. Bruinsma & S. D. Johnson (Eds.), *The Oxford handbook of environmental criminology*. Oxford University Press.

Dobler, G., Vani, J., & Dam, T. T. L. (2021). Patterns of urban foot traffic dynamics. *Computers, Environment and Urban Systems*, *89*, Article 101674. https://doi.org/10.1016/j.compenvurbsys.2021.101674

Duneier, M. (1999). *Sidewalk*. Farrar, Straus and Giroux.

Felson, M., & Boivin, R. (2015). Daily crime flows within a city. *Crime Science*, *4*(1), 31. https://doi.org/10.1186/s40163-015-0039-0

Frith, M., Bowers, K., & Johnson, S. (2022). Household occupancy and burglary: A case study using COVID-19 restrictions. *Journal of Criminal Justice*, *82*, Article 101996. https://doi.org/10.1016/j.jcrimjus.2022.101996

Gourieroux, C., Monfort, A., & Trognon, A. (1984). Pseudo maximum likelihood methods: Applications to Poisson models. *Econometrica*, *52*(3), 701–720. https://doi.org/10.2307/1913472

Graif, C., Freelin, B. N., Kuo, Y.-H., Wang, H., Li, Z., & Kifer, D. (2021). Network spillovers and neighborhood crime: A computational statistics analysis of employment-based networks of neighborhoods. *Justice Quarterly*, *38*(2), 344–374. https://doi.org/10.1080/07418825.2019.1602160

Grannis, R. (2009). *From the ground up: Translating geography into community through neighbor networks*. Princeton University Press.

Haberman, C. P., & Ratcliffe, J. H. (2015). Testing for temporally differentiated relationships among potentially criminogenic places and census block street robbery counts. *Criminology*, *53*(3), 457–483. https://doi.org/10.1111/1745-9125.12076

Hall, A. R. (2005). *Generalized method of moments*. Oxford University Press.

Harinam, V., Bavcevic, Z., & Ariel, B. (2022). Spatial distribution and developmental trajectories of crime versus crime severity: Do not abandon the count-based model just yet. *Crime Science*, *11*(1), 14. https://doi.org/10.1186/s40163-022-00176-x

Hilbe, J. M. (2011). *Negative binomial regression* (2nd ed.). Cambridge University Press.

Hillier, B. (2004). *Can streets be made safe*? *Urban Design International*, *9*(1), 31–45. https://doi.org/10.1057/palgrave.udi.9000079

Hipp, J. R. (2007). Income inequality, race, and place: Does the distribution of race and class within neighborhoods affect crime rates? *Criminology*, *45*(3), 665–697. https://doi.org/10.1111/j.1745-9125.2007.00088.x

Hipp, J. R., Bates, C., Lichman, M., & Smyth, P. (2019). Using social media to measure temporal ambient population: Does it help explain local crime rates? *Justice Quarterly*, *36*(4), 718–748. https://doi.org/10.1080/07418825.2018.1445276

Hoet, M. J. (2023). Crime concentration and hot spot dynamics: An examination of homicides in Santa Fe, Argentina. *International Criminology*, *3*(4), 313–327. https://doi.org/10.1007/s43576-023-00103-3

Humphries, J. E., Ouss, A., Stavreva, K., Stevenson, M. T., & van Dijk, W. (2025). Conviction, incarceration, and recidivism: Understanding the revolving door. *The Quarterly Journal of Economics*, *140*(4), 2907–2962. https://doi.org/10.1093/qje/qjaf040

Jacobs, J. (1961). *The death and life of great American cities*. Random House.

Jeffery, C. R. (1971). *Crime prevention through environmental design*. Sage Publications.

Keane, M., & Neal, T. (2023). Instrument strength in IV estimation and inference: A guide to theory and practice. *Journal of Econometrics*, *235*(2), 1625–1653. https://doi.org/10.1016/j.jeconom.2022.12.009

Khalfa, R., & Hardyns, W. (2025). Comparing machine learning-based crime predictions across micro-geographic units: Street segments, rectangular grids, and hexagonal grids. *Applied Spatial Analysis and Policy*, *18*(3), 79. https://doi.org/10.1007/s12061-025-09683-1

Kim, Y.-A., & Hipp, J. R. (2020). Pathways: Examining street network configurations, structural characteristics and spatial crime patterns in street segments. *Journal of Quantitative Criminology*, *36*(4), 725–752. https://doi.org/10.1007/s10940-019-09428-7

Kim, Y.-A., & Jeong, J. (2025). Making spatial more spatial: Decomposing direct, indirect, and total impacts in neighborhood and crime studies. *Journal of Quantitative Criminology*. https://doi.org/10.1007/s10940-025-09646-2

Kounadi, O., Vallejo-Velázquez, M., Huang, Q., & Pődör, A. (2025). Are places where people feel safe or unsafe similar to spatial crime patterns? Extracting the collective truth of the spatial crime perception gap using digital sketch maps. *Urban Science*, *9*(10), 397. https://doi.org/10.3390/urbansci9100397

Lafrogne-Joussier, R., & Rollet, V. (2025). Ambient density and urban crime: Evidence from smartphone data. *SSRN*. https://doi.org/10.2139/ssrn.4440554

Lawrence, D., Peterson, B., White, M., Cunningham, B., & Coldren, J. (2024). Effects of correctional body-worn cameras on responses to resistance: A randomized controlled trial in a jail setting. *Justice Quarterly*, *41*(6), 892–914. https://doi.org/10.1080/07418825.2024.2327301

Lee, D. S., & Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, *48*(2), 281–355. https://doi.org/10.1257/jel.48.2.281

Lemieux, A. M., & Felson, M. (2012). Risk of violent crime victimization during major daily activities. *Violence and Victims*, *27*(5), 635–655. https://doi.org/10.1891/0886-6708.27.5.635

Levy, B. L., Phillips, N. E., & Sampson, R. J. (2020). Triple disadvantage: Neighborhood networks of everyday urban mobility and violence in U.S. cities. *American Sociological Review*, *85*(6), 925–956. https://doi.org/10.1177/0003122420972323

Linning, S., & Eck, J. (2021). *Whose 'eyes on the street' control crime? Expanding place management into neighborhoods*. Cambridge University Press.

Ma, Y., Li, G., Zhao, L., & Yao, X. (2025). Accuracy evaluation method for vector data based on hexagonal discrete global grid. *ISPRS International Journal of Geo-Information*, *14*(1), 5. https://doi.org/10.3390/ijgi14010005

MacKinnon, J. G., Nielsen, M. Ø., & Webb, M. D. (2023). Cluster-robust inference: A guide to empirical practice. *Journal of Econometrics*, *232*(2), 272–299. https://doi.org/10.1016/j.jeconom.2022.04.001

Mahfoud, M., Bernasco, W., Bhulai, S., & van der Mei, R. (2021). Forecasting spatio-temporal variation in residential burglary with the integrated Laplace approximation framework: Effects of crime generators, street networks, and prior crimes. *Journal of Quantitative Criminology*, *37*(4), 835–862. https://doi.org/10.1007/s10940-020-09469-3

Maksuta, K. D., Zhao, Y., & Yang, T.-C. (2024). Race, disadvantage, and violence: A spatial exploration of macrolevel covariates of police-involved homicides within and between US counties. *Social Science Research*, *119*, 102985. https://doi.org/10.1016/j.ssresearch.2024.102985

Malleson, N., Steenbeek, W., & Andresen, M. A. (2019). Identifying the appropriate spatial resolution for the analysis of crime patterns. *PLoS ONE*, *14*(6), e0218324. https://doi.org/10.1371/journal.pone.0218324

Mery, G., Wodchis, W. P., & Laporte, A. (2016). The determinants of the propensity to receive publicly funded home care services for the elderly in Canada: A panel two-stage residual inclusion approach. *Health Economics Review*, *6*(1), 8. https://doi.org/10.1186/s13561-016-0086-6

Newman, O. (1972). *Defensible space: Crime prevention through urban design*. Macmillan.

Nunes, K. L., Hatton, C. E., Pham, A. T., Blank, C., & Maimone, S. (2025). Causal interpretations of correlational evidence regarding violence. *Journal of Interpersonal Violence*, *40*(15–16), 3712–3728. https://doi.org/10.1177/08862605241285996

Openshaw, S. (1984). *The modifiable areal unit problem. Concepts and Techniques in Modern Geography*, *38*, 1–41.

Owens, E. G., & Ludwig, J. (2013). Using regression discontinuity designs in crime research. In B. C. Welsh, A. A. Braga, & G. J. N. Bruinsma (Eds.), *Experimental criminology: Prospects for advancing science and public policy* (pp. 194–222). Cambridge University Press.

Pearcey, J., Ariel, B., Harinam, V., & Assaraf, N. (2024). Violence against women and girls in Dorset, United Kingdom: An epidemiological study of perpetrators and locations based on police records. *Humanities and Social Sciences Communications*, *11*(1), 1634. https://doi.org/10.1057/s41599-024-04021-1

Pratt, T. C., & Cullen, F. T. (2005). Assessing macro-level predictors and theories of crime: A meta-analysis. *Crime and Justice*, *32*, 373–450. https://doi.org/10.1086/655357

Rahman, M. S., Huang, W.-C., Toiba, H., & Efani, A. (2022). Does adaptation to climate change promote household food security? Insights from Indonesian fishermen. *International Journal of Sustainable Development & World Ecology*, *29*(7), 611–624. https://doi.org/10.1080/13504509.2022.2063433

Roodman, D. (2009). A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics*, *71*(1), 135–158. https://doi.org/10.1111/j.1468-0084.2008.00542.x

Sampson, R. J., & Groves, W. B. (1989). Community structure and crime: Testing social-disorganization theory. *American Journal of Sociology*, *94*(4), 774–802. https://doi.org/10.1086/229068

Sampson, R. J., Raudenbush, S. W., & Earls, F. (1997). Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science*, *277*(5328), 918–924. https://doi.org/10.1126/science.277.5328.918

Sampson, R. J., Winship, C., & Knight, C. (2013). Translating causal claims: Principles and strategies for policy-relevant criminology. *Criminology & Public Policy*, *12*(4), 587–616. https://doi.org/10.1111/1745-9133.12027

Shaw, C. R., & McKay, H. D. (1942). *Juvenile delinquency and urban areas*. University of Chicago Press.

Silva, J. M. C. S., & Tenreyro, S. (2006). The log of gravity. *The Review of Economics and Statistics*, *88*(4), 641–658. https://doi.org/10.1162/rest.88.4.641

Situ, X. (2024). From mobility to crime: Collective patterns of human mobility and gun violence in Baltimore City. *Journal of Criminal Justice*, *94*, Article 102256. https://doi.org/10.1016/j.jcrimjus.2024.102256

Song, G., Zhang, Y., Bernasco, W., Cai, L., Liu, L., Qin, B., & Chen, P. (2023). Residents, employees and visitors: Effects of three types of ambient population on theft on weekdays and weekends in Beijing, China. *Journal of Quantitative Criminology*, *39*(2), 385–423. https://doi.org/10.1007/s10940-021-09538-1

Staiger, D., & Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica*, *65*(3), 557–586. https://doi.org/10.2307/2171753

Stock, J. H., & Yogo, M. (2005). Testing for weak instruments in linear IV regression. In D. W. K. Andrews & J. H. Stock (Eds.), *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg* (pp. 80–108). Cambridge University Press.

Stults, B. J., & Hasbrouck, M. (2015). The effect of commuting on city-level crime rates. *Journal of Quantitative Criminology*, *31*(2), 331–350. https://doi.org/10.1007/s10940-015-9251-z

Summers, L., & Johnson, S. D. (2017). Does the configuration of the street network influence where outdoor serious violence takes place? Using space syntax to test crime pattern theory. *Journal of Quantitative Criminology*, *33*(2), 397–420. https://doi.org/10.1007/s10940-016-9306-9

Tahamont, S. (2019). The effect of facility security classification on serious rules violation reports in California prisons: A regression discontinuity design. *Journal of Quantitative Criminology*, *35*(4), 767–796. https://doi.org/10.1007/s10940-019-09405-0

Taylor, R. B., Gottfredson, S. D., & Brower, S. (1981). Territorial cognitions and social climate in urban neighborhoods. *Basic and Applied Social Psychology*, *2*(4), 289–303. https://doi.org/10.1207/s15324834basp0204_5

Terza, J. V., Basu, A., & Rathouz, P. J. (2008). Two-stage residual inclusion estimation: Addressing endogeneity in health econometric modeling. *Journal of Health Economics*, *27*(3), 531–543. https://doi.org/10.1016/j.jhealeco.2007.09.009

Tompson, L., & Bowers, K. (2013). A stab in the dark?: A research note on temporal patterns of street robbery. *The Journal of Research in Crime and Delinquency*, *50*(4), 616–631. https://doi.org/10.1177/0022427812469114

Uber Technologies, Inc. (2018). *H3: A hexagonal hierarchical spatial index*. Uber Technologies, Inc.

Vachuska, K. (2025). Urban mobility and crime: Causal inference using street closures as an instrumental variable. *Frontiers in Big Data*, *8*, 1579332. https://doi.org/10.3389/fdata.2025.1579332

van Sleeuwen, S., Ruiter, S., & Steenbeek, W. (2021). Right place, right time? Making crime pattern theory time-specific. *Crime Science*, *10*(1), 1–10. https://doi.org/10.1186/s40163-021-00139-8

Weisburd, D., Groff, E. R., & Yang, S.-M. (2012). *The criminology of place: Street segments and our understanding of the crime problem*. Oxford University Press.

Weisburd, D., & Piquero, A. (2008). How well do criminologists explain crime? Statistical modeling in published studies. *Crime and Justice*, *37*(1), 453–502. https://doi.org/10.1086/524284

Wikström, P.-O. H., & Kroneberg, C. (2022). Analytic criminology: Mechanisms and methods in the explanation of crime and its causes. *Annual Review of Criminology*, *5*, 179–203. https://doi.org/10.1146/annurev-criminol-030920-091320

Wilk, T., Deza, M., Hodge, T. R., & Danagoulian, S. (2026). Couch-locked with the munchies: Effects of recreational marijuana laws on exercise and nutrition. *Journal of Policy Analysis and Management*, *45*(2), e70080. https://doi.org/10.1002/pam.70080

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT Press.

Wu, L., Lee, C. S., & Zhu, C. (2025). Mobility-based neighborhood disadvantage and crime in Houston. *Journal of Criminal Justice*, *101*, 102508. https://doi.org/10.1016/j.jcrimjus.2025.102508

Xie, M., & Baumer, E. P. (2019). Crime victims' decisions to call the police: Past research and new directions. *Annual Review of Criminology*, *2*(1), 217–240. https://doi.org/10.1146/annurev-criminol-011518-024748

Yang, L., & Zikos, V. (2023). Mental health and smoking behavior. *Economic Modelling*, *126*, 106407. https://doi.org/10.1016/j.econmod.2023.106407

Zhang, J., & Xiang, G. (2022). Revisiting the relationship between tourism and crime based on a dynamic spatial durbin model. *Crime & Delinquency*, *68*(11), 2146–2171. https://doi.org/10.1177/00111287211064790

Appendix A. Number of Home Games per Sports League

Table A1. Number of major league sports events by city, league, and year.

	2023 (second half)				2024				2025 (first half)			
City	MLB	NBA	NFL	NHL	MLB	NBA	NFL	NHL	MLB	NBA	NFL	NHL
Bal	40	0	8	0	79	0	10	0	39	0	2	0
Chi	78	20	8	17	149	38	9	41	80	26	0	24
Phi	40	16	8	16	81	43	8	43	36	27	5	23

Appendix B. Extra Covariates Events

- **Baltimore:**
  - *Preakness Stakes* (May 20, 2023; May 18, 2024; May 17, 2025): Annual thoroughbred horse race held at Pimlico Race Course.
  - *Baltimore Pride* (June 19–25, 2023; June 10–16, 2024; June 8–15, 2025): Week-long celebration of LGBTQ+ pride.
  - *AFRAM Festival* (June 17–18, 2023; June 22–23, 2024; June 21–22, 2025): Major African American cultural festival held at Druid Hill Park.
  - *Artscape* (September 22–24, 2023; August 2–4, 2024; May 24–25, 2025): America's largest free outdoor arts festival. The 2023 edition was partially canceled due to the weather.
  - *Brooklyn Homes Mass Shooting* (July 2, 2023): Mass shooting during a neighborhood party that resulted in two fatalities and 28 people wounded. It is the largest shooting incident in the city's history.
  - *Maryland Cycling Classic* (September 3, 2023; September 6, 2025): Professional cycling race. The 2024 edition was canceled following the Key Bridge collapse.
  - *Baltimore Running Festival* (October 14, 2023; October 19, 2024; October 18, 2025): Annual running event including a marathon and half-marathon.

- **Chicago:**
  - *Sueños Music Festival* (May 27–28, 2023; May 25–26, 2024; May 24–25, 2025): Latin music festival held in Grant Park.
  - *Chicago Blues Festival* (June 8–11, 2023; June 6–9, 2024; June 5–8, 2025): World's largest free blues festival.
  - *Chicago Pride Fest* (June 17–18, 2023; June 22–23, 2024; June 21–22, 2025): Two-day street festival preceding the Pride Parade.
  - *Chicago Pride Parade* (June 25, 2023; June 30, 2024; June 29, 2025): One of the largest Pride parades in the U.S.
  - *NASCAR Chicago Street Race* (July 1–2, 2023; July 6–7, 2024): Street circuit race held downtown; the 2023 edition was partially canceled due to flooding.
  - *Lollapalooza* (August 3–6, 2023; August 1–4, 2024; July 31–August 3, 2025): Multi-day international music festival in Grant Park.
  - *Northalsted Market Days* (August 12–13, 2023; August 10–11, 2024; August 8–10, 2025): Annual summer street festival in the Boystown neighborhood.
  - *Air and Water Show* (August 19–20, 2023; August 10–11, 2024; August 16–17, 2025): Large-scale air and water performance along the lakefront.
  - *Taste of Chicago* (September 8–10, 2023; September 6–8, 2024; September 5–7, 2025): Iconic food festival featuring local vendors and live music.
  - *Chicago Marathon* (October 8, 2023; October 13, 2024; October 12, 2025): One of the six World Marathon Majors.

- **Philadelphia:**
  - *Philadelphia Auto Show* (January 28–February 5, 2023; January 13–21, 2024; January 11–20, 2025): Annual automotive exhibition at the Pennsylvania Convention Center.

- *Philadelphia Flower Show* (March 4–12, 2023; March 2–10, 2024; March 1–9, 2025): The nation's largest and longest-running horticultural event.
- *ODUNDE Festival* (June 11, 2023; June 9, 2024; June 8, 2025): African American cultural festival held in South Philadelphia.
- *Philadelphia Pride* (June 4, 2023; June 2, 2024; June 1, 2025): LGBTQ+ Pride parade and festival.
- *Philadelphia Marathon Weekend* (November 17–19, 2023; November 22–24, 2024; November 21–23, 2025): Includes the marathon, half marathon, and 8K races.

Appendix C. Data Aggregation at the Hexagon Level

To aggregate the datasets into our defined hexagonal grids, we had to use slightly different strategies for each dataset.

C.1. Crime Data Aggregation

For each crime incident, we extracted the latitude and longitude coordinates and performed a point-in-polygon test to assign it to the corresponding hexagon. This was done using an iterative search through all hexagon polygons until finding the one that contained the crime's point geometry. Each crime's timestamp was then converted into an hour-of-year index, representing the number of hours elapsed since midnight on January 1st of that year. For example, a crime occurring at 3:00 PM on January 2nd would receive index 39 (24 h for January 1st plus 15 h into January 2nd). Lastly, we built a sparse matrix where rows correspond to unique hexagons and columns correspond to hourly time bins (i.e. the hour-of-year indexes) and iterated through each crime record to extract its corresponding hexagon and hour-of-year index, incrementing the corresponding matrix cell by one. Annual matrices were then concatenated column-wise across years and trimmed to only contain data between the 30th of June 2023 and the 29th of June 2025, producing a dataset where rows are the hexagons and columns represent the number of crimes occurring in each hour of the study period.

C.2. Mobility Data Aggregation

The mobility data was provided as weekly files containing POIs with their hourly visit counts across 168 h (one week). Therefore, we first created a POI-to-hexagon mapping dictionary by assigning each unique POI to its containing hexagon *via* the same point-in-polygon procedure as done for the crime dataset. This mapping was computed once and stored. Then, for each weekly file, we applied this mapping, grouped the data by hexagon, and summed the hourly visit counts across all POIs within each hexagon. Each weekly dataset was then concatenated into annual matrices, with column names converted to hour-of-year indices, then combined across years to produce the final time series. Moreover, due to the weekly structure of the dataset, additional cleaning steps were applied to remove overlapping hours at year boundaries and to ensure the dataset was trimmed to only contain data between the 30th of June 2023 and the 29th of June 2025. Therefore, in the final mobility dataset, each row represents a hexagon and each column the aggregated hourly footfall throughout the study period.

Appendix D. Construction of the Contiguity Matrix

It has been observed that urban phenomena rarely respect strict geographic boundaries. For instance, high pedestrian activity in a commercial district may draw visitors to neighboring streets, and crime hotspots can displace human activity to nearby areas or generate fear that reduces mobility in the surrounding streets. Treating each area (i.e. each hexagon) independently would ignore these important spatial interdependencies that create spillover effects. To ad-

dress this, we built a spatial contiguity matrix **W** that formalizes the neighborhood structure of our hexagonal grids, hence allowing us later on to assess how the properties in one area are or are not affected by conditions in its neighboring areas.

More specifically, the contiguity matrix **W** is an $n \times n$ matrix where n is the total number of hexagons for that city. Each row i corresponds to a spatial unit (hexagon i) and each column j corresponds to another spatial unit (hexagon j). The matrix is then constructed as follows: w_{ij} is set to 1 if hexagons i and j share a border or a vertex, otherwise it is set to 0. This is known as Queen contiguity and it yields a square symmetric zero-diagonal matrix with high sparsity, where each row represents a focal location and the columns are its neighbors (Cliff & Ord, 1981). However, we then row-standardize the matrix by dividing each element by its row sum, so we have

$$w_{ij}^{R} = \frac{w_{ij}}{\sum_{j} w_{ij}}. \tag{7}$$

This results in each row summing to 1, but also means that our contiguity matrix is probably no longer symmetric. Row-standardization is necessary because without it, hexagons with more neighbors would mechanically have larger spillover values simply due to summing over more units, rather than reflecting meaningful differences in spatial influence. By ensuring each row sums to 1, we convert the raw neighbor counts into normalized weights that represent the average conditions in the surrounding area.

Therefore, the spatial spillover y of a variable x over our hexagonal grid can be obtained by doing

$$\mathbf{y} = \mathbf{W}\mathbf{x}, \tag{8}$$

where $\mathbf{x}$ is an $n \times 1$ vector containing the values of variable x for all hexagons and $\mathbf{y}$ is the resulting $n \times 1$ vector of spatially lagged values.

Appendix E. Hexagon Units Used

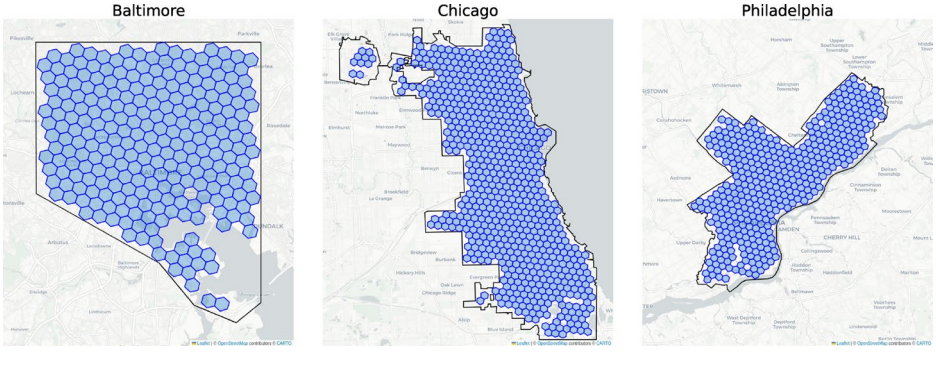


Figure E1. Map of the initial set of hexagons considered for each city. Each hexagon covers an area of 0.74 sq. km.

Appendix F. Crime Spatial Distribution

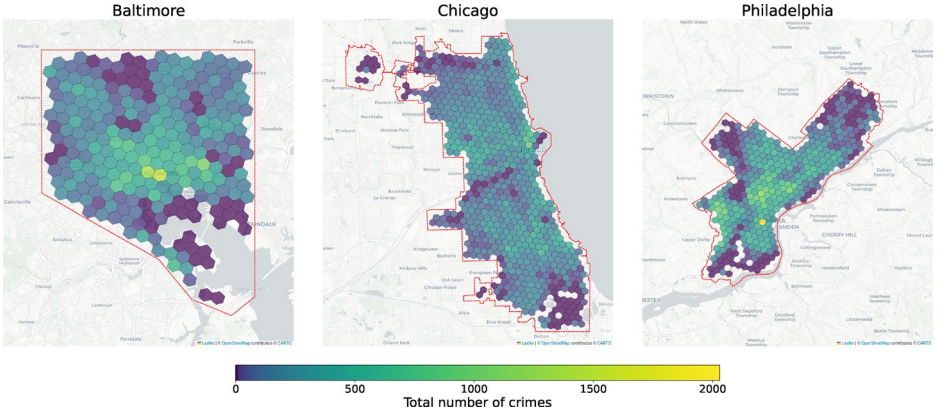


Figure F2. Heatmap of the total number of crimes included in the study per hexagon and city.

Appendix G. Footfall Descriptive Statistics per Temporal Granularity

Table G2. Hexagon-level footfall descriptive statistics by temporal granularity for each city during our study period, after performing the city-level sample correction.

Granularity	City	Mean	Median	Min	Max	SD	Var/Mean
1h	Bal	4074.63	3278.99	49.42	22,569.19	3081.74	2589.01
	Chi	13,158.83	10,806.87	123.42	64,309.31	9949.32	8011.55
	Phi	9294.86	7067.80	87.58	55,024.20	7770.84	7009.69
3h	Bal	12,223.89	10,498.95	337.31	52,541.69	8282.62	6148.02
	Chi	39,476.48	33,989.48	898.83	156,690.63	27,508.95	20,347.68
	Phi	27,884.59	22,255.96	517.42	134,869.92	21,224.03	17,325.90
6h	Bal	24,447.79	21,382.05	1398.33	90,986.93	15,362.72	10,454.88
	Chi	78,952.96	68,004.38	3672.59	274,041.11	51,980.65	36,264.37
	Phi	55,769.18	44,996.81	2066.71	229,094.55	39,384.49	29,704.90
12h	Bal	48,895.57	46,396.00	9624.42	133,272.51	19,472.68	8719.68
	Chi	157,905.92	151,064.37	28,155.18	399,791.33	65,446.32	29,472.99
	Phi	111,538.35	106,462.05	18,187.89	308,112.84	49,818.40	24,678.28

Hexagon-level statistics are computed per hexagon and then averaged across hexagons.

Appendix H. Penetration Rate

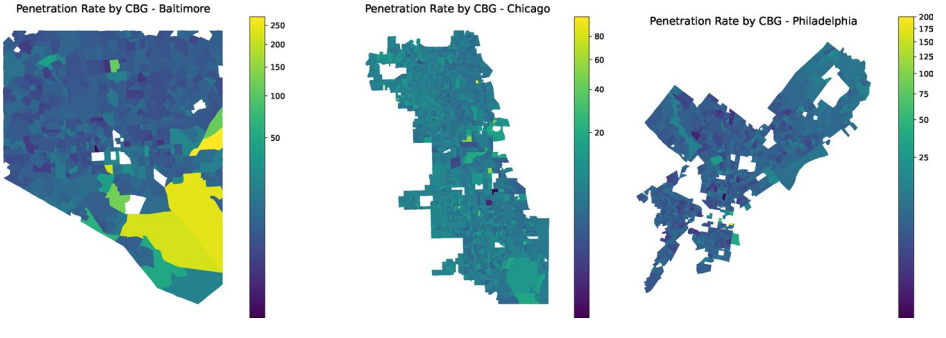


Figure H3. Choropleth map indicating the penetration rate for each census block group during 2023.

Appendix I. Bootstrap Correction

This bootstrapping method consists of resampling at the cluster level (i.e. hexagon id), by randomly selecting with replacement from the original set of hexagons. Moreover, when a hexagon is selected, the entire time series corresponding to that hexagon is included in the bootstrap sample. This allows us to preserve the strong temporal autocorrelation within each spatial unit, their inherent fixed effects, and also accounts for arbitrary spatial dependence that may exist across neighboring hexagons. For each bootstrap replication, the entire 2SRI is executed to obtain the replicated coefficient estimates. This procedure is repeated over B iterations, in order to obtain the final standard errors, p-values, and confidence intervals as follows.

The bootstrap standard error is calculated as the sample standard deviation of the collection of B bootstrap coefficients:

$$\mathrm{SE}_{boot}(\beta_{orig}) = \sqrt{\frac{1}{B-1}\sum_{b=1}^{B}(\hat{\beta}^{(b)} - \bar{\beta})^2},$$

where β_{orig} is the coefficient of interest from the original second-stage Poisson regression, $\hat{\beta}^{(b)}$ is the second-stage coefficient obtained during iteration b of bootstrap, and $\bar{\beta} = \frac{1}{B}\sum_{b=1}^{B}\hat{\beta}^{(b)}$ is the mean of the bootstrap coefficients.

The bootstrap p-value tests the null hypothesis that the true coefficient is zero ($H_0 : \beta_{orig} = 0$), which is calculated empirically by determining how often the bootstrap estimates are on the opposite side of zero. So the two-sided bootstrap p-value is given by

$$p\text{-value}(\beta_{orig}) = 2 \times \min\left\{\frac{1}{B}\sum_{b=1}^{B}\mathbb{1}\left(\hat{\beta}^{(b)} \le 0\right), \frac{1}{B}\sum_{b=1}^{B}\mathbb{1}\left(\hat{\beta}^{(b)} > 0\right)\right\}, \tag{9}$$

where $\mathbb{1}(\cdot)$ is the indicator function. Any empirical p-value of exactly zero is replaced by $2/(B+1)$, indicating the minimum detectable p-value given B resamples.

Finally, the bootstrap 95% confidence interval is calculated using the percentile method, which is preferred for nonlinear models like Poisson because it does not assume that the sampling distribution of the bootstrap coefficients is perfectly normal or symmetric. Hence, for a 95% confidence interval, we have

$$\mathrm{CI}_{95\%} = \left[Q_{0.025}\left(\hat{\beta}^{(1)}, \ldots, \hat{\beta}^{(B)}\right), Q_{0.975}\left(\hat{\beta}^{(1)}, \ldots, \hat{\beta}^{(B)}\right)\right], \tag{10}$$

where $Q_q(\cdot)$ denotes the q-th empirical quantile of the bootstrap distribution.

Appendix J. 2SRI Specification for Each Temporal Granularity

In order to keep approximately equivalent model definitions for each temporal granularity, we had to use different temporal lags and number of instruments. These are the specific configurations used for each temporal granularity:

- *For 1h granularity:*
  - We set $q = 3$.
  - *Endogenous variables:* ['footfall_lag0', 'footfall_lag1', 'Wfootfall_lag0', 'Wfootfall_lag1', 'WCrime_lag1']
  - *Instruments:* For 'footfall_lag0' and 'footfall_lag1', ['footfall_lag4', 'footfall_lag5']; for 'Wfootfall_lag0' and 'Wfootfall_lag1', ['Wfootfall_lag4', 'Wfootfall_lag5']; and for 'WCrime_lag1', ['WCrime_lag2', 'WCrime_lag3'].
  - *Exogenous variables:* ['footfall_lag2', 'footfall_lag3', 'Wfootfall_lag2', 'Wfootfall_lag3', 'sport_event', 'holiday', 'extra_event'] + fixed effects from ['hour_of_day', 'day_of_week', 'month', 'year']

- *For 3h and 6h granularity:*
  - We set $q = 2$.
  - *Endogenous variables:* ['footfall_lag0', 'Wfootfall_lag0', 'WCrime_lag1']
  - *Instruments:* For 'footfall_lag0', ['footfall_lag3', 'footfall_lag4']; for 'Wfootfall_lag0', ['Wfootfall_lag3', 'Wfootfall_lag4']; and for 'WCrime_lag1', ['WCrime_lag2', 'WCrime_lag3'].
  - *Exogenous variables:* ['footfall_lag1', 'footfall_lag2', 'Wfootfall_lag1', 'Wfootfall_lag2', 'sport_event', 'holiday', 'extra_event'] + fixed effects from ['three_hour_block' (or 'six_hour_block'), 'day_of_week', 'month', 'year']
- *For 12h granularity:*
  - We set $q = 1$.
  - *Endogenous variables:* ['footfall_lag0', 'Wfootfall_lag0', 'WCrime_lag1']
  - *Instruments:* For 'footfall_lag0', ['footfall_lag2', 'footfall_lag3']; for 'Wfootfall_lag0', ['Wfootfall_lag2', 'Wfootfall_lag3']; and for 'WCrime_lag1', ['WCrime_lag2', 'WCrime_lag3'].
  - *Exogenous variables:* ['footfall_lag1', 'Wfootfall_lag1', 'sport_event', 'holiday', 'extra_event'] + fixed effects from ['twelve_hour_block', 'day_of_week', 'month', 'year']

Appendix K. Average Number of Crimes per Hexagon per Temporal Granularity

Table K3. Average crime counts per hexagon per temporal granularity, by city and crime type.

City	Crime type	1h granularity	3h granularity	6h granularity	12h granularity
Bal	All	0.0128	0.0384	0.0769	0.1537
	Assault	0.0063	0.0190	0.0381	0.0762
	Burglary	0.0015	0.0045	0.0090	0.0180
	MVT	0.0032	0.0096	0.0193	0.0385
	Robbery	0.0018	0.0053	0.0105	0.0210
Chi	All	0.0100	0.0300	0.0600	0.1200
	Assault	0.0037	0.0110	0.0220	0.0440
	Burglary	0.0013	0.0039	0.0078	0.0156
	MVT	0.0035	0.0105	0.0211	0.0421
	Robbery	0.0015	0.0046	0.0092	0.0183
Phi	All	0.0150	0.0449	0.0898	0.1797
	Assault	0.0087	0.0260	0.0520	0.1039
	Burglary	0.0013	0.0038	0.0076	0.0151
	MVT	0.0040	0.0119	0.0238	0.0476
	Robbery	0.0011	0.0033	0.0065	0.0130

Appendix L. Comparison Baseline Model Results Between Negative Binomial and Poisson

Table L4. Comparison of Poisson and Negative Binomial regression results for all crimes together across cities.

	Bal		Chi		Phi	
	P	NB	P	NB	P	NB
Footfall	1.015***	1.017***	1.003***	1.004***	1.006***	1.007**
	(9.70×10^{-4})	(1.80×10^{-3})	(2.80×10^{-4})	(4.03×10^{-4})	(1.82×10^{-3})	(2.71×10^{-3})
Intercept	−2.566	−2.569	−2.803	−2.804	−2.419	−2.421
	(5.98×10^{-2})	(5.96×10^{-2})	(2.84×10^{-2})	(2.84×10^{-2})	(4.76×10^{-2})	(4.75×10^{-2})
Num. Obs.	731,000	731,000	1,874,284	1,874,284	1,198,840	1,198,840
Log-Likelihood	−212,061.71	−204,311.96	−440,836.04	−438,763.71	−376,967.19	−372,681.69
Pearson χ^2	1,102,641.18	1,021,810.69	2,025,191.03	1,908,317.32	1,385,607.32	1,270,387.50
Pseudo R^2	0.005	0.005	0.003	0.003	0.003	0.003

Footfall is reported as IRR for +36,000 persons, Exp(36000β), with its clustered standard error in parentheses. Stars denote significance: *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$.

Table L5. Comparison of Poisson and Negative Binomial regression results for assault across cities.

	Bal		Chi		Phi	
	P	NB	P	NB	P	NB
Footfall	1.015***	1.017***	1.004***	1.004***	1.007***	1.008**
	(7.98×10^{-4})	(1.16×10^{-3})	(3.23×10^{-4})	(4.09×10^{-4})	(2.02×10^{-3})	(2.97×10^{-3})
Intercept	−3.251	−3.253	−3.802	−3.802	−2.952	−2.955
	(6.64×10^{-2})	(6.63×10^{-2})	(3.19×10^{-2})	(3.19×10^{-2})	(5.30×10^{-2})	(5.28×10^{-2})
Num. Obs.	716,380	716,380	1,859,664	1,859,664	1,178,372	1,178,372
Log-Likelihood	−124,163.95	−119,937.73	−202,225.08	−201,861.90	−249,538.93	−247,423.88
Pearson χ^2	1,108,486.81	1,066,031.41	1,918,198.18	1,875,967.43	1,315,472.71	1,248,813.76
Pseudo R^2	0.006	0.006	0.004	0.004	0.004	0.004

Footfall is reported as IRR for +36,000 persons, Exp(36000β), with its clustered standard error in parentheses. Stars denote significance: * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

Table L6. Comparison of Poisson and Negative Binomial regression results for burglary across cities.

	Bal		Chi		Phi	
	P	NB	P	NB	P	NB
Footfall	1.011***	1.011***	1.002***	1.002***	1.007***	1.007***
	(2.11×10^{-3})	(2.23×10^{-3})	(3.29×10^{-4})	(3.36×10^{-4})	(1.06×10^{-3})	(1.20×10^{-3})
Intercept	−4.639	−4.639	−4.808	−4.808	−4.812	−4.813
	(7.00×10^{-2})	(7.00×10^{-2})	(3.19×10^{-2})	(3.19×10^{-2})	(5.74×10^{-2})	(5.74×10^{-2})
Num. Obs.	684,216	684,216	1,821,652	1,821,652	1,099,424	1,099,424
Log-Likelihood	−38,212.84	−37,686.44	−87,169.06	−86,832.44	−52,905.53	−52,745.53
Pearson χ^2	880,814.73	872,241.22	1,964,583.34	1,948,574.61	1,158,157.09	1,148,759.95
Pseudo R^2	0.001	0.001	0.000	0.000	0.003	0.003

Footfall is reported as IRR for +36,000 persons, Exp(36000β), with its clustered standard error in parentheses. Stars denote significance: * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

Table L7. Comparison of Poisson and Negative Binomial regression results for motor vehicle theft across cities.

	Bal		Chi		Phi	
	P	NB	P	NB	P	NB
Footfall	1.010***	1.010***	1.003***	1.003***	1.002	1.002
	(1.28×10^{-3})	(1.39×10^{-3})	(3.40×10^{-4})	(3.85×10^{-4})	(1.16×10^{-3})	(1.18×10^{-3})
Intercept	−3.920	−3.921	−3.834	−3.834	−3.706	−3.706
	(5.03×10^{-2})	(5.03×10^{-2})	(2.74×10^{-2})	(2.74×10^{-2})	(4.10×10^{-2})	(4.10×10^{-2})
Num. Obs.	716,380	716,380	1,850,892	1,850,892	1,166,676	1,166,676
Log-Likelihood	−71,162.45	−70,571.97	−195,812.67	−195,223.96	−136,233.57	−135,543.31
Pearson χ^2	811,409.92	795,368.39	1,947,075.99	1,905,480.24	1,264,688.72	1,234,279.97
Pseudo R^2	0.001	0.001	0.002	0.002	0.000	0.000

Footfall is reported as IRR for +36,000 persons, Exp(36000β), with its clustered standard error in parentheses. Stars denote significance: *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$.

Table L8. Comparison of Poisson and Negative Binomial regression results for robbery across cities.

	Bal		Chi		Phi	
	P	NB	P	NB	P	NB
Footfall	1.017***	1.018***	1.004***	1.004***	1.007***	1.007**
	(9.36×10^{-4})	(1.14×10^{-3})	(1.57×10^{-4})	(1.77×10^{-4})	(1.97×10^{-3})	(2.22×10^{-3})
Intercept	−4.469	−4.471	−4.624	−4.625	−4.895	−4.896
	(7.20×10^{-2})	(7.19×10^{-2})	(3.70×10^{-2})	(3.69×10^{-2})	(5.85×10^{-2})	(5.84×10^{-2})
Num. Obs.	663,748	663,748	1,760,248	1,760,248	1,026,324	1,026,324
Log-Likelihood	−43,931.35	−42,378.11	−98,769.41	−98,305.25	−46,070.65	−46,003.02
Pearson χ^2	1,076,605.43	1,063,967.62	1,902,064.62	1,883,358.97	1,048,762.77	1,040,875.67
Pseudo R^2	0.007	0.008	0.003	0.003	0.003	0.003

Footfall is reported as IRR for +36,000 persons, Exp(36000β), with its clustered standard error in parentheses. Stars denote significance: * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

Appendix M. Exclusion Restriction Results

Table M9. IRRs for a 36,000-person increase in footfall across cities and crime types at 6-h granularity, comparing results using lags 3–4 versus lags 4–5 as instruments for the footfall equations.

	Bal		Chi		Phi	
	Lags 3–4	Lags 4–5	Lags 3–4	Lags 4–5	Lags 3–4	Lags 4–5
All	1.001	1.002	1.001	1.001	0.999	0.999
	(0.008)	(0.008)	(0.001)	(0.001)	(0.003)	(0.003)
Assault	1.002	1.002	1.000	1.000	1.000	1.001
	(0.017)	(0.017)	(0.001)	(0.001)	(0.002)	(0.002)
Burglary	1.009*	1.008**	0.999	0.999	1.001	1.001
	(0.033)	(0.036)	(0.002)	(0.002)	(0.004)	(0.005)
MVT	0.998	1.001	1.001	1.001	0.996**	0.996**
	(0.017)	(0.013)	(0.001)	(0.001)	(0.002)	(0.002)
Robbery	1.003	1.002	1.001	1.000	0.999	0.999
	(0.011)	(0.012)	(0.002)	(0.002)	(0.007)	(0.006)

Stars denote significance: * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

Appendix N. Regression Results

Table N10. Regression results for all crimes together in Baltimore.

	Baseline	Model 2	Model 3	Model 4	Model 5
Footfall lag 0	1.015***	1.007***	1.002***	1.001***	1.001
	(9.70×10^{-4})	(3.65×10^{-4})	(2.79×10^{-4})	(2.53×10^{-4})	(8.10×10^{-3})
Footfall lag 1	–	0.999	0.994***	0.994***	0.995
		(9.96×10^{-4})	(3.40×10^{-4})	(4.26×10^{-4})	(9.39×10^{-3})
Footfall lag 2	–	1.014***	1.003***	1.002***	1.002
		(4.55×10^{-4})	(3.82×10^{-4})	(2.48×10^{-4})	(6.23×10^{-3})
W Footfall lag 0	–	–	–	1.146	1.185
				(1.00×10^{-1})	(1.49×10^{-1})
W Footfall lag 1	–	–	–	0.911	0.970
				(5.71×10^{-2})	(1.68×10^{-1})
W Footfall lag 2	–	–	–	1.411***	1.392**
				(1.41×10^{-1})	(1.90×10^{-1})
W Crime lag 1	–	–	–	0.148***	0.659
				(2.67×10^{-2})	(9.34×10^{-1})
Footfall resid lag 0	–	–	–	–	1.000
					(9.44×10^{-3})
W Footfall resid lag 0	–	–	–	–	0.929
					(1.01×10^{-1})
WCrime resid lag 1	–	–	–	–	−0.512
					(9.28×10^{-1})
Intercept	−2.566	−2.577	–	–	–
	(0.060)	(0.059)			
Num. Obs.	731,000	730,500	730,500	730,500	730,000
Log-Likelihood	−212,061.71	−210,982.39	−185,053.47	−185,013.49	−184,829.94
Pearson χ^2	1,102,641.18	1,096,207.87	1,002,870.53	1,001,268.75	1,000,761.44
Pseudo R^2	0.005	0.010	0.131	0.132	0.132
Fixed effects?	No	No	Yes	Yes	Yes
Covariates?	No	No	Yes	Yes	Yes
Spillover effects?	No	No	No	Yes	Yes
Endogeneity correction?	No	No	No	No	Yes

For footfall variables, we report the IRR for +36,000 persons, Exp(36000β), with its robust standard error in parentheses. Stars denote significance: * $p<0.05$, ** $p<0.01$, and *** $p<0.001$. For Model 5, standard errors are cluster-bootstrap standard deviations and are used for reference only since significance is based on two-sided bootstrap p-values, and may therefore not match Wald-type intervals computed from these standard errors.

Table N11. Regression results for Assault in Baltimore.

	Baseline	Model 2	Model 3	Model 4	Model 5
Footfall lag 0	1.015***	1.007***	1.002***	1.002***	1.002
	(7.98×10^{-4})	(4.32×10^{-4})	(2.53×10^{-4})	(3.02×10^{-4})	(1.73×10^{-2})
Footfall lag 1	–	0.998	0.993***	0.993***	0.993
		(9.64×10^{-4})	(3.60×10^{-4})	(5.63×10^{-4})	(1.63×10^{-2})
Footfall lag 2	–	1.015***	1.005***	1.004***	1.005
		(3.54×10^{-4})	(3.95×10^{-4})	(2.53×10^{-4})	(8.63×10^{-3})
W Footfall lag 0	–	–	–	1.020	1.108
				(1.20×10^{-1})	(1.68×10^{-1})
W Footfall lag 1	–	–	–	0.852	0.712
				(1.40×10^{-1})	(6.29×10^{-1})
W Footfall lag 2	–	–	–	1.406***	1.523
				(1.08×10^{-1})	(4.06×10^{-1})
W Crime lag 1	–	–	–	0.068	−1.703
				(5.65×10^{-2})	(9.45×10^{0})
Footfall resid lag 0	–	–	–	–	1.000
					(1.86×10^{-2})
W Footfall resid lag 0	–	–	–	–	0.832
					(2.49×10^{-1})
WCrime resid lag 1	–	–	–	–	1.770
					(9.45×10^{0})
Intercept	−3.251	−3.263	–	–	–
	(0.066)	(0.066)			
Num. Obs.	716,380	715,890	715,890	715,890	715,400
Log-Likelihood	−124,163.95	−123,499.32	−109,544.28	−109,532.50	−109,441.44
Pearson χ^2	1,108,486.81	1,103,486.64	1,006,567.72	1,005,530.33	1,004,820.85
Pseudo R^2	0.006	0.010	0.122	0.122	0.122
Fixed effects?	No	No	Yes	Yes	Yes
Covariates?	No	No	Yes	Yes	Yes
Spillover effects?	No	No	No	Yes	Yes
Endogeneity correction?	No	No	No	No	Yes

For footfall variables, we report the IRR for +36,000 persons, Exp(36000β), with its robust standard error in parentheses. Stars denote significance: * $p<0.05$, ** $p<0.01$, and *** $p<0.001$. For Model 5, standard errors are cluster-bootstrap standard deviations and are used for reference only since significance is based on two-sided bootstrap p-values, and may therefore not match Wald-type intervals computed from these standard errors.

Table N12. Regression results for Burglary in Baltimore.

	Baseline	Model 2	Model 3	Model 4	Model 5
Footfall lag 0	1.011***	1.011***	1.010***	1.008***	1.009*
	(2.11×10^{-3})	(1.27×10^{-3})	(2.43×10^{-3})	(1.55×10^{-3})	(3.27×10^{-2})
Footfall lag 1	–	0.988***	0.982***	0.984***	0.984*
		(2.49×10^{-3})	(1.95×10^{-3})	(1.20×10^{-3})	(3.95×10^{-2})
Footfall lag 2	–	1.014***	1.014***	1.012***	1.013
		(6.64×10^{-4})	(1.43×10^{-3})	(9.33×10^{-4})	(2.21×10^{-2})
W Footfall lag 0	–	–	–	1.782**	2.121
				(3.92×10^{-1})	(6.97×10^{-1})
W Footfall lag 1	–	–	–	0.526***	0.406**
				(9.18×10^{-2})	(1.19×10^{-1})
W Footfall lag 2	–	–	–	2.094**	2.954**
				(5.04×10^{-1})	(1.16×10^{0})
W Crime lag 1	–	–	–	0.437*	24.493
				(1.97×10^{-1})	(1.29×10^{1})
Footfall resid lag 0	–	–	–	–	0.998
					(2.50×10^{-2})
W Footfall resid lag 0	–	–	–	–	0.654
					(2.46×10^{-1})
WCrime resid lag 1	–	–	–	–	−24.070
					(1.29×10^{1})
Intercept	−4.639	−4.643	–	–	–
	(0.070)	(0.070)			
Num. Obs.	684,216	683,748	683,748	683,748	683,280
Log-Likelihood	−38,212.84	−38,147.67	−34,919.06	−34,904.45	−34,859.57
Pearson χ^2	880,814.73	879,802.16	822,673.78	822,275.17	822,174.60
Pseudo R^2	0.001	0.002	0.087	0.087	0.087
Fixed effects?	No	No	Yes	Yes	Yes
Covariates?	No	No	Yes	Yes	Yes
Spillover effects?	No	No	No	Yes	Yes
Endogeneity correction?	No	No	No	No	Yes

For footfall variables, we report the IRR for +36,000 persons, Exp(36000β), with its robust standard error in parentheses. Stars denote significance: * $p<0.05$, ** $p<0.01$, and *** $p<0.001$. For Model 5, standard errors are cluster-bootstrap standard deviations and are used for reference only since significance is based on two-sided bootstrap p-values, and may therefore not match Wald-type intervals computed from these standard errors.

Table N13. Regression results for MVT in Baltimore.

	Baseline	Model 2	Model 3	Model 4	Model 5
Footfall lag 0	1.010***	0.999	0.999	0.999*	0.998
	(1.28×10^{-3})	(9.61×10^{-4})	(4.51×10^{-4})	(5.25×10^{-4})	(1.73×10^{-2})
Footfall lag 1	–	1.007***	1.003**	1.002**	1.003*
		(1.19×10^{-3})	(1.23×10^{-3})	(7.48×10^{-4})	(2.32×10^{-2})
Footfall lag 2	–	1.008***	0.997***	0.996***	0.994**
		(1.43×10^{-3})	(4.30×10^{-4})	(4.59×10^{-4})	(1.46×10^{-2})
W Footfall lag 0	–	–	–	1.162	1.045
				(1.76×10^{-1})	(2.83×10^{-1})
W Footfall lag 1	–	–	–	1.822*	2.176*
				(4.26×10^{-1})	(6.52×10^{-1})
W Footfall lag 2	–	–	–	1.150	0.807
				(1.29×10^{-1})	(1.74×10^{-1})
W Crime lag 1	–	–	–	1.064***	6.900**
				(9.99×10^{-2})	(1.62×10^{0})
Footfall resid lag 0	–	–	–	–	1.001
					(2.49×10^{-2})
W Footfall resid lag 0	–	–	–	–	1.208
					(4.12×10^{-1})
WCrime resid lag 1	–	–	–	–	−5.852**
					(1.60×10^{0})
Intercept	−3.920	−3.927	–	–	–
	(0.050)	(0.050)			
Num. Obs.	716,380	715,890	715,890	715,890	715,400
Log-Likelihood	−71,162.45	−70,964.88	−63,135.22	−63,035.10	−62,936.54
Pearson χ^2	811,409.92	808,686.12	832,022.50	822,753.37	822,132.83
Pseudo R^2	0.001	0.003	0.113	0.114	0.114
Fixed effects?	No	No	Yes	Yes	Yes
Covariates?	No	No	Yes	Yes	Yes
Spillover effects?	No	No	No	Yes	Yes
Endogeneity correction?	No	No	No	No	Yes

For footfall variables, we report the IRR for +36,000 persons, Exp(36000β), with its robust standard error in parentheses. Stars denote significance: *$p<0.05$, **$p<0.01$, and ***$p<0.001$. For Model 5, standard errors are cluster-bootstrap standard deviations and are used for reference only since significance is based on two-sided bootstrap p-values, and may therefore not match Wald-type intervals computed from these standard errors.

Table N14. Regression results for Robbery in Baltimore.

	Baseline	Model 2	Model 3	Model 4	Model 5
Footfall lag 0	1.017***	1.009***	1.001	1.001	1.003
	(9.36×10^{-4})	(4.83×10^{-4})	(6.07×10^{-4})	(4.93×10^{-4})	(1.13×10^{-2})
Footfall lag 1	–	0.998*	0.995***	0.995***	0.998
		(9.30×10^{-4})	(6.62×10^{-4})	(7.19×10^{-4})	(2.12×10^{-2})
Footfall lag 2	–	1.017***	1.003***	1.002***	1.000
		(5.77×10^{-4})	(4.52×10^{-4})	(4.62×10^{-4})	(2.25×10^{-2})
W Footfall lag 0	–	–	–	1.131	1.159
				(2.11×10^{-1})	(3.74×10^{-1})
W Footfall lag 1	–	–	–	0.799	1.647
				(1.63×10^{-1})	(9.49×10^{-1})
W Footfall lag 2	–	–	–	1.617**	1.605*
				(3.01×10^{-1})	(4.76×10^{-1})
W Crime lag 1	–	–	–	0.325	29.014
				(1.89×10^{-1})	(2.10×10^{1})
Footfall resid lag 0	–	–	–	–	0.995
					(1.82×10^{-2})
W Footfall resid lag 0	–	–	–	–	0.941
					(4.12×10^{-1})
WCrime resid lag 1	–	–	–	–	−28.689
					(2.10×10^{1})
Intercept	−4.469	−4.487	–	–	–
	(0.072)	(0.070)			
Num. Obs.	663,748	663,294	663,294	663,294	662,840
Log-Likelihood	−43,931.35	−43,671.98	−39,850.84	−39,842.04	−39,810.04
Pearson χ^2	1,076,605.43	1,072,516.24	1,001,670.28	1,000,009.94	999,365.22
Pseudo R^2	0.007	0.013	0.099	0.099	0.100
Fixed effects?	No	No	Yes	Yes	Yes
Covariates?	No	No	Yes	Yes	Yes
Spillover effects?	No	No	No	Yes	Yes
Endogeneity correction?	No	No	No	No	Yes

For footfall variables, we report the IRR for +36,000 persons, Exp(36000β), with its robust standard error in parentheses. Stars denote significance: *$p<0.05$, **$p<0.01$, and ***$p<0.001$. For Model 5, standard errors are cluster-bootstrap standard deviations and are used for reference only since significance is based on two-sided bootstrap p-values, and may therefore not match Wald-type intervals computed from these standard errors.

Table N15. Regression results for all crimes together in Chicago.

	Baseline	Model 2	Model 3	Model 4	Model 5
Footfall lag 0	1.003***	1.002***	1.000	1.001***	1.001
	(2.80×10^{-4})	(1.18×10^{-4})	(2.11×10^{-4})	(1.49×10^{-4})	(1.15×10^{-3})
Footfall lag 1	–	1.001***	1.000	1.000	1.000
		(8.72×10^{-5})	(1.66×10^{-4})	(1.93×10^{-4})	(9.41×10^{-4})
Footfall lag 2	–	1.002***	1.000*	1.000***	1.000*
		(2.84×10^{-4})	(7.92×10^{-5})	(1.47×10^{-4})	(1.01×10^{-3})
W Footfall lag 0	–	–	–	0.972	0.965
				(1.95×10^{-2})	(3.01×10^{-2})
W Footfall lag 1	–	–	–	0.969**	0.941*
				(1.10×10^{-2})	(2.12×10^{-2})
W Footfall lag 2	–	–	–	1.040*	1.005
				(1.78×10^{-2})	(2.63×10^{-2})
W Crime lag 1	–	–	–	0.163***	5.129**
				(2.44×10^{-2})	(1.13×10^{0})
Footfall resid lag 0	–	–	–	–	1.000
					(9.54×10^{-4})
W Footfall resid lag 0	–	–	–	–	1.019
					(2.88×10^{-2})
WCrime resid lag 1	–	–	–	–	−4.970**
					(1.13×10^{0})
Intercept	−2.803	−2.808	–	–	–
	(0.028)	(0.028)			
Num. Obs.	1,874,284	1,873,002	1,873,002	1,873,002	1,871,720
Log-Likelihood	−440,836.04	−440,087.10	−407,591.80	−407,543.37	−407,174.10
Pearson χ^2	2,025,191.03	2,023,174.92	1,962,204.21	1,962,042.66	1,960,733.64
Pseudo R^2	0.003	0.004	0.078	0.078	0.078
Fixed effects?	No	No	Yes	Yes	Yes
Covariates?	No	No	Yes	Yes	Yes
Spillover effects?	No	No	No	Yes	Yes
Endogeneity correction?	No	No	No	No	Yes

For footfall variables, we report the IRR for +36,000 persons, Exp(36000β), with its robust standard error in parentheses. Stars denote significance: *$p<0.05$, **$p<0.01$, and ***$p<0.001$. For Model 5, standard errors are cluster-bootstrap standard deviations and are used for reference only since significance is based on two-sided bootstrap p-values, and may therefore not match Wald-type intervals computed from these standard errors.

Table N16. Regression results for Assault in Chicago.

	Baseline	Model 2	Model 3	Model 4	Model 5
Footfall lag 0	1.004***	1.003***	1.001***	1.000***	1.000
	(3.23×10^{-4})	(2.53×10^{-4})	(5.89×10^{-5})	(1.46×10^{-4})	(1.06×10^{-3})
Footfall lag 1	–	1.000***	1.000	1.000	1.000
		(7.19×10^{-5})	(1.64×10^{-4})	(2.20×10^{-4})	(1.14×10^{-3})
Footfall lag 2	–	1.001***	1.000	1.000	1.000
		(2.21×10^{-4})	(1.45×10^{-4})	(2.01×10^{-4})	(7.62×10^{-4})
W Footfall lag 0	–	–	–	1.014	1.016
				(2.42×10^{-2})	(3.46×10^{-2})
W Footfall lag 1	–	–	–	0.962*	0.935
				(1.79×10^{-2})	(7.07×10^{-2})
W Footfall lag 2	–	–	–	1.014	1.013
				(2.76×10^{-2})	(3.76×10^{-2})
W Crime lag 1	–	–	–	0.155**	3.582
				(5.97×10^{-2})	(8.72×10^{0})
Footfall resid lag 0	–	–	–	–	1.000
					(8.58×10^{-4})
W Footfall resid lag 0	–	–	–	–	1.002
					(5.92×10^{-2})
WCrime resid lag 1	–	–	–	–	−3.432
					(8.73×10^{0})
Intercept	−3.802	−3.805	–	–	–
	(0.032)	(0.032)			
Num. Obs.	1,859,664	1,858,392	1,858,392	1,858,392	1,857,120
Log-Likelihood	−202,225.08	−201,991.81	−186,951.44	−186,946.46	−186,775.79
Pearson χ^2	1,918,198.18	1,918,567.84	1,881,346.28	1,881,733.78	1,880,486.02
Pseudo R^2	0.004	0.004	0.078	0.078	0.078
Fixed effects?	No	No	Yes	Yes	Yes
Covariates?	No	No	Yes	Yes	Yes
Spillover effects?	No	No	No	Yes	Yes
Endogeneity correction?	No	No	No	No	Yes

For footfall variables, we report the IRR for +36,000 persons, Exp(36000β), with its robust standard error in parentheses. Stars denote significance: *$p<0.05$, **$p<0.01$, and ***$p<0.001$. For Model 5, standard errors are cluster-bootstrap standard deviations and are used for reference only since significance is based on two-sided bootstrap p-values, and may therefore not match Wald-type intervals computed from these standard errors.

Table N17. Regression results for Burglary in Chicago.

	Baseline	Model 2	Model 3	Model 4	Model 5
Footfall lag 0	1.002***	1.001***	0.998*	0.999	0.999
	(3.29×10^{-4})	(1.24×10^{-4})	(7.64×10^{-4})	(5.86×10^{-4})	(2.08×10^{-3})
Footfall lag 1	–	0.999	0.999*	0.999*	0.999
		(3.30×10^{-4})	(4.10×10^{-4})	(5.18×10^{-4})	(1.58×10^{-3})
Footfall lag 2	–	1.003***	0.999	1.000	0.999
		(3.87×10^{-4})	(6.82×10^{-4})	(7.40×10^{-4})	(1.90×10^{-3})
W Footfall lag 0	–	–	–	0.890*	0.882
				(4.69×10^{-2})	(7.46×10^{-2})
W Footfall lag 1	–	–	–	1.026	1.117*
				(4.19×10^{-2})	(7.82×10^{-2})
W Footfall lag 2	–	–	–	0.988	1.004
				(4.43×10^{-2})	(7.07×10^{-2})
W Crime lag 1	–	–	–	0.862***	35.217**
				(1.72×10^{-1})	(1.17×10^{1})
Footfall resid lag 0	–	–	–	–	1.000
					(2.11×10^{-3})
W Footfall resid lag 0	–	–	–	–	1.039
					(1.13×10^{-1})
WCrime resid lag 1	–	–	–	–	−34.366**
					(1.17×10^{1})
Intercept	−4.808	−4.813	–	–	–
	(0.032)	(0.032)			
Num. Obs.	1,821,652	1,820,406	1,820,406	1,820,406	1,819,160
Log-Likelihood	−87,169.06	−87,041.43	−82,224.66	−82,203.45	−82,151.21
Pearson χ^2	1,964,583.34	1,961,162.14	1,923,330.10	1,921,859.97	1,920,312.80
Pseudo R^2	3.84×10^{-4}	0.001	0.056	0.057	0.057
Fixed effects?	No	No	Yes	Yes	Yes
Covariates?	No	No	Yes	Yes	Yes
Spillover effects?	No	No	No	Yes	Yes
Endogeneity correction?	No	No	No	No	Yes

For footfall variables, we report the IRR for +36,000 persons, Exp(36000β), with its robust standard error in parentheses. Stars denote significance: *$p<0.05$, **$p<0.01$, and ***$p<0.001$. For Model 5, standard errors are cluster-bootstrap standard deviations and are used for reference only since significance is based on two-sided bootstrap p-values, and may therefore not match Wald-type intervals computed from these standard errors.

Table N18. Regression results for MVT in Chicago.

	Baseline	Model 2	Model 3	Model 4	Model 5
Footfall lag 0	1.003***	1.002***	1.001***	1.001**	1.001
	(3.40×10^{-4})	(1.19×10^{-4})	(2.09×10^{-4})	(2.74×10^{-4})	(1.22×10^{-3})
Footfall lag 1	–	1.001***	0.999*	0.999	0.999
		(1.59×10^{-4})	(3.06×10^{-4})	(3.19×10^{-4})	(9.59×10^{-4})
Footfall lag 2	–	1.001**	0.999***	0.999*	1.000
		(3.78×10^{-4})	(1.73×10^{-4})	(2.58×10^{-4})	(6.48×10^{-4})
W Footfall lag 0	–	–	–	1.016	1.015
				(1.80×10^{-2})	(3.99×10^{-2})
W Footfall lag 1	–	–	–	0.985	0.929*
				(1.81×10^{-2})	(2.62×10^{-2})
W Footfall lag 2	–	–	–	0.996	0.966
				(2.88×10^{-2})	(4.36×10^{-2})
W Crime lag 1	–	–	–	0.726***	19.197**
				(6.44×10^{-2})	(2.39×10^{0})
Footfall resid lag 0	–	–	–	–	0.999
					(8.77×10^{-4})
W Footfall resid lag 0	–	–	–	–	0.995
					(5.79×10^{-2})
WCrime resid lag 1	–	–	–	–	−18.491**
					(2.39×10^{0})
Intercept	−3.834	−3.838	–	–	–
	(0.027)	(0.027)			
Num. Obs.	1,850,892	1,849,626	1,849,626	1,849,626	1,848,360
Log-Likelihood	−195,812.67	−195,567.34	−183,554.14	−183,481.41	−183,235.00
Pearson χ^2	1,947,075.99	1,945,461.66	1,921,841.01	1,920,199.14	1,918,876.03
Pseudo R^2	0.002	0.002	0.063	0.064	0.064
Fixed effects?	No	No	Yes	Yes	Yes
Covariates?	No	No	Yes	Yes	Yes
Spillover effects?	No	No	No	Yes	Yes
Endogeneity correction?	No	No	No	No	Yes

For footfall variables, we report the IRR for +36,000 persons, Exp(36000β), with its robust standard error in parentheses. Stars denote significance: * $p<0.05$, ** $p<0.01$, and *** $p<0.001$. For Model 5, standard errors are cluster-bootstrap standard deviations and are used for reference only since significance is based on two-sided bootstrap p-values, and may therefore not match Wald-type intervals computed from these standard errors.

Table N19. Regression results for Robbery in Chicago.

	Baseline	Model 2	Model 3	Model 4	Model 5
Footfall lag 0	1.004***	1.002***	1.000	1.000	1.001
	(1.57×10^{-4})	(2.63×10^{-4})	(6.85×10^{-4})	(4.49×10^{-4})	(2.16×10^{-3})
Footfall lag 1	–	1.001***	1.001**	1.001**	1.001**
		(1.63×10^{-4})	(1.82×10^{-4})	(3.31×10^{-4})	(2.32×10^{-3})
Footfall lag 2	–	1.002***	1.000	0.999**	0.999
		(3.17×10^{-4})	(1.47×10^{-4})	(2.73×10^{-4})	(2.05×10^{-3})
W Footfall lag 0	–	–	–	0.882*	0.850
				(5.29×10^{-2})	(7.28×10^{-2})
W Footfall lag 1	–	–	–	0.959	1.021
				(2.64×10^{-2})	(6.10×10^{-2})
W Footfall lag 2	–	–	–	1.121**	0.995
				(4.16×10^{-2})	(6.50×10^{-2})
W Crime lag 1	–	–	–	0.846***	21.695**
				(1.27×10^{-1})	(6.55×10^{0})
Footfall resid lag 0	–	–	–	–	0.999
					(1.68×10^{-3})
W Footfall resid lag 0	–	–	–	–	1.104
					(1.02×10^{-1})
WCrime resid lag 1	–	–	–	–	−20.860**
					(6.55×10^{0})
Intercept	−4.624	−4.636	–	–	–
	(0.037)	(0.036)			
Num. Obs.	1,760,248	1,759,044	1,759,044	1,759,044	1,757,840
Log-Likelihood	−98,769.41	−98,431.52	−90,783.14	−90,715.19	−90,653.61
Pearson χ^2	1,902,064.62	1,897,946.91	1,838,535.01	1,835,247.98	1,834,098.23
Pseudo R^2	0.003	0.006	0.083	0.084	0.084
Fixed effects?	No	No	Yes	Yes	Yes
Covariates?	No	No	Yes	Yes	Yes
Spillover effects?	No	No	No	Yes	Yes
Endogeneity correction?	No	No	No	No	Yes

For footfall variables, we report the IRR for +36,000 persons, Exp(36000β), with its robust standard error in parentheses. Stars denote significance: * $p<0.05$, ** $p<0.01$, and *** $p<0.001$. For Model 5, standard errors are cluster-bootstrap standard deviations and are used for reference only since significance is based on two-sided bootstrap p-values, and may therefore not match Wald-type intervals computed from these standard errors.

Table N20. Regression results for all crimes together in Philadelphia.

	Baseline	Model 2	Model 3	Model 4	Model 5
Footfall lag 0	1.006***	1.002	0.999	1.000	0.999
	(1.82×10^{-3})	(1.38×10^{-3})	(7.14×10^{-4})	(5.06×10^{-4})	(2.82×10^{-3})
Footfall lag 1	–	1.004***	1.001**	1.000	1.001
		(6.23×10^{-4})	(3.80×10^{-4})	(2.46×10^{-4})	(3.45×10^{-3})
Footfall lag 2	–	1.002*	1.000	1.000	1.000
		(1.12×10^{-3})	(5.05×10^{-4})	(2.72×10^{-4})	(1.34×10^{-3})
W Footfall lag 0	–	–	–	0.880***	0.907
				(2.81×10^{-2})	(5.95×10^{-2})
W Footfall lag 1	–	–	–	1.117**	1.091
				(4.06×10^{-2})	(6.08×10^{-2})
W Footfall lag 2	–	–	–	0.928	0.987
				(4.36×10^{-2})	(7.85×10^{-2})
W Crime lag 1	–	–	–	0.140***	−0.597
				(1.81×10^{-2})	(7.36×10^{-1})
Footfall resid lag 0	–	–	–	–	1.001
					(1.97×10^{-3})
W Footfall resid lag 0	–	–	–	–	0.934
					(1.09×10^{-1})
WCrime resid lag 1	–	–	–	–	0.739
					(7.33×10^{-1})
Intercept	−2.419	−2.426	–	–	–
	(0.048)	(0.048)			
Num. Obs.	1,198,840	1,198,020	1,198,020	1,198,020	1,197,200
Log-Likelihood	−376,967.19	−376,007.26	−318,294.88	−318,253.59	−317,978.43
Pearson χ^2	1,385,607.32	1,381,437.53	1,242,734.83	1,241,748.50	1,240,802.00
Pseudo R^2	0.003	0.005	0.158	0.158	0.158
Fixed effects?	No	No	Yes	Yes	Yes
Covariates?	No	No	Yes	Yes	Yes
Spillover effects?	No	No	No	Yes	Yes
Endogeneity correction?	No	No	No	No	Yes

For footfall variables, we report the IRR for +36,000 persons, Exp(36000β), with its robust standard error in parentheses. Stars denote significance: * $p<0.05$, ** $p<0.01$, and *** $p<0.001$. For Model 5, standard errors are cluster-bootstrap standard deviations and are used for reference only since significance is based on two-sided bootstrap p-values, and may therefore not match Wald-type intervals computed from these standard errors.

Table N21. Regression results for Assault in Philadelphia.

	Baseline	Model 2	Model 3	Model 4	Model 5
Footfall lag 0	1.007***	1.003	1.000	1.000	1.000
	(2.02×10^{-3})	(1.71×10^{-3})	(2.93×10^{-4})	(3.38×10^{-4})	(2.49×10^{-3})
Footfall lag 1	–	1.004***	1.001***	1.001	1.000
		(6.33×10^{-4})	(2.03×10^{-4})	(4.48×10^{-4})	(3.77×10^{-3})
Footfall lag 2	–	1.003*	1.000	1.000	1.001
		(1.39×10^{-3})	(3.18×10^{-4})	(2.82×10^{-4})	(1.43×10^{-3})
W Footfall lag 0	–	–	–	0.951	0.986
				(4.07×10^{-2})	(6.28×10^{-2})
W Footfall lag 1	–	–	–	1.125*	1.100
				(5.86×10^{-2})	(7.44×10^{-2})
W Footfall lag 2	–	–	–	0.899*	0.962
				(4.56×10^{-2})	(8.47×10^{-2})
W Crime lag 1	–	–	–	0.098**	−0.959
				(3.29×10^{-2})	(1.29×10^{0})
Footfall resid lag 0	–	–	–	–	1.000
					(2.83×10^{-3})
W Footfall resid lag 0	–	–	–	–	0.901
					(1.20×10^{-1})
WCrime resid lag 1	–	–	–	–	1.057
					(1.29×10^{0})
Intercept	−2.952	−2.961	–	–	–
	(0.053)	(0.053)			
Num. Obs.	1,178,372	1,177,566	1,177,566	1,177,566	1,176,760
Log-Likelihood	−249,538.93	−248,752.64	−211,122.51	−211,111.00	−210,938.51
Pearson χ^2	1,315,472.71	1,310,896.41	1,203,284.70	1,202,901.80	1,201,890.25
Pseudo R^2	0.004	0.006	0.157	0.157	0.157
Fixed effects?	No	No	Yes	Yes	Yes
Covariates?	No	No	Yes	Yes	Yes
Spillover effects?	No	No	No	Yes	Yes
Endogeneity correction?	No	No	No	No	Yes

For footfall variables, we report the IRR for +36,000 persons, Exp(36000β), with its robust standard error in parentheses. Stars denote significance: * $p<0.05$, ** $p<0.01$, and *** $p<0.001$. For Model 5, standard errors are cluster-bootstrap standard deviations and are used for reference only since significance is based on two-sided bootstrap p-values, and may therefore not match Wald-type intervals computed from these standard errors.

Table N22. Regression results for Burglary in Philadelphia.

	Baseline	Model 2	Model 3	Model 4	Model 5
Footfall lag 0	1.007***	1.005***	1.002***	1.001	1.001
	(1.06×10^{-3})	(6.95×10^{-4})	(6.36×10^{-4})	(7.77×10^{-4})	(4.16×10^{-3})
Footfall lag 1	–	1.002**	0.999*	0.999	0.998
		(7.57×10^{-4})	(5.76×10^{-4})	(8.07×10^{-4})	(4.00×10^{-3})
Footfall lag 2	–	1.001	1.000	1.000	0.999
		(8.02×10^{-4})	(8.34×10^{-4})	(7.44×10^{-4})	(3.05×10^{-3})
W Footfall lag 0	–	–	–	1.184	1.245
				(1.28×10^{-1})	(2.17×10^{-1})
W Footfall lag 1	–	–	–	0.924	0.745**
				(6.63×10^{-2})	(1.07×10^{-1})
W Footfall lag 2	–	–	–	1.045	0.974
				(1.60×10^{-1})	(1.85×10^{-1})
W Crime lag 1	–	–	–	0.832***	35.488**
				(1.95×10^{-1})	(1.25×10^{1})
Footfall resid lag 0	–	–	–	–	1.000
					(4.40×10^{-3})
W Footfall resid lag 0	–	–	–	–	0.879
					(2.06×10^{-1})
WCrime resid lag 1	–	–	–	–	−34.681**
					(1.24×10^{1})
Intercept	−4.812	−4.817	–	–	–
	(0.057)	(0.058)			
Num. Obs.	1,099,424	1,098,672	1,098,672	1,098,672	1,097,920
Log-Likelihood	−52,905.53	−52,855.81	−47,297.10	−47,284.72	−47,235.75
Pearson χ^2	1,158,157.09	1,157,954.62	1,142,995.10	1,142,118.62	1,142,648.44
Pseudo R^2	0.003	0.004	0.108	0.109	0.109
Fixed effects?	No	No	Yes	Yes	Yes
Covariates?	No	No	Yes	Yes	Yes
Spillover effects?	No	No	No	Yes	Yes
Endogeneity correction?	No	No	No	No	Yes

For footfall variables, we report the IRR for +36,000 persons, Exp(36000β), with its robust standard error in parentheses. Stars denote significance: * $p<0.05$, ** $p<0.01$, and *** $p<0.001$. For Model 5, standard errors are cluster-bootstrap standard deviations and are used for reference only since significance is based on two-sided bootstrap p-values, and may therefore not match Wald-type intervals computed from these standard errors.

Table N23. Regression results for MVT in Philadelphia.

	Baseline	Model 2	Model 3	Model 4	Model 5
Footfall lag 0	1.002	0.998*	0.994**	0.997*	0.996**
	(1.16×10^{-3})	(6.81×10^{-4})	(2.02×10^{-3})	(1.34×10^{-3})	(2.29×10^{-3})
Footfall lag 1	–	1.005***	1.000	0.999	1.000
		(1.29×10^{-3})	(1.11×10^{-3})	(1.13×10^{-3})	(3.57×10^{-3})
Footfall lag 2	–	0.999	1.001	1.001	1.000
		(6.96×10^{-4})	(1.26×10^{-3})	(1.37×10^{-3})	(2.32×10^{-3})
W Footfall lag 0	–	–	–	0.617***	0.663**
				(4.64×10^{-2})	(9.79×10^{-2})
W Footfall lag 1	–	–	–	1.295**	1.379*
				(1.22×10^{-1})	(1.86×10^{-1})
W Footfall lag 2	–	–	–	0.978	0.899
				(8.55×10^{-2})	(1.23×10^{-1})
W Crime lag 1	–	–	–	0.743***	14.217**
				(6.40×10^{-2})	(2.05×10^{0})
Footfall resid lag 0	–	–	–	–	1.002
					(2.31×10^{-3})
W Footfall resid lag 0	–	–	–	–	0.849
					(1.71×10^{-1})
WCrime resid lag 1	–	–	–	–	−13.495**
					(2.05×10^{0})
Intercept	−3.706	−3.707	–	–	–
	(0.041)	(0.041)			
Num. Obs.	1,166,676	1,165,878	1,165,878	1,165,878	1,165,080
Log-Likelihood	−136,233.57	−136,109.39	−121,408.91	−121,314.39	−121,167.60
Pearson χ^2	1,264,688.72	1,263,362.39	1,235,766.45	1,234,098.50	1,234,158.49
Pseudo R^2	6.43×10^{-5}	2.51×10^{-4}	0.108	0.109	0.109
Fixed effects?	No	No	Yes	Yes	Yes
Covariates?	No	No	Yes	Yes	Yes
Spillover effects?	No	No	No	Yes	Yes
Endogeneity correction?	No	No	No	No	Yes

For footfall variables, we report the IRR for +36,000 persons, Exp(36000β), with its robust standard error in parentheses. Stars denote significance: * $p<0.05$, ** $p<0.01$, and *** $p<0.001$. For Model 5, standard errors are cluster-bootstrap standard deviations and are used for reference only since significance is based on two-sided bootstrap p-values, and may therefore not match Wald-type intervals computed from these standard errors.

Table N24. Regression results for Robbery in Philadelphia.

	Baseline	Model 2	Model 3	Model 4	Model 5
Footfall lag 0	1.007***	1.003	0.998***	0.999	0.999
	(1.97×10^{-3})	(1.71×10^{-3})	(4.74×10^{-4})	(7.10×10^{-4})	(6.68×10^{-3})
Footfall lag 1	–	1.003***	1.002***	1.002***	1.002
		(3.32×10^{-4})	(3.81×10^{-4})	(4.98×10^{-4})	(4.91×10^{-3})
Footfall lag 2	–	1.004***	0.997*	0.999	0.999
		(1.03×10^{-3})	(1.16×10^{-3})	(8.38×10^{-4})	(4.91×10^{-3})
W Footfall lag 0	–	–	–	0.843	0.806**
				(8.17×10^{-2})	(1.08×10^{-1})
W Footfall lag 1	–	–	–	1.063	0.915
				(6.95×10^{-2})	(1.99×10^{-1})
W Footfall lag 2	–	–	–	0.813***	1.030
				(4.82×10^{-2})	(3.29×10^{-1})
W Crime lag 1	–	–	–	0.201	−52.839
				(2.24×10^{-1})	(5.96×10^{1})
Footfall resid lag 0	–	–	–	–	1.001
					(4.65×10^{-3})
W Footfall resid lag 0	–	–	–	–	1.147
					(3.18×10^{-1})
WCrime resid lag 1	–	–	–	–	53.034
					(5.96×10^{1})
Intercept	−4.895	−4.909	–	–	–
	(0.059)	(0.058)			
Num. Obs.	1,026,324	1,025,622	1,025,622	1,025,622	1,024,920
Log-Likelihood	−46,070.65	−45,932.96	−41,564.90	−41,561.97	−41,522.39
Pearson χ^2	1,048,762.77	1,047,718.05	1,021,989.49	1,022,156.91	1,022,074.66
Pseudo R^2	0.003	0.006	0.100	0.100	0.100
Fixed effects?	No	No	Yes	Yes	Yes
Covariates?	No	No	Yes	Yes	Yes
Spillover effects?	No	No	No	Yes	Yes
Endogeneity correction?	No	No	No	No	Yes

For footfall variables, we report the IRR for +36,000 persons, Exp(36000β), with its robust standard error in parentheses. Stars denote significance: * $p<0.05$, ** $p<0.01$, and *** $p<0.001$. For Model 5, standard errors are cluster-bootstrap standard deviations and are used for reference only since significance is based on two-sided bootstrap p-values, and may therefore not match Wald-type intervals computed from these standard errors

Appendix O. Instruments Strength

Table O25. First-stage F-statistics for the endogenous variables from Model 5 with 6-h granularity for each city and crime type.

City	Instrumented variable	All	Assault	Burglary	MVT	Robbery
Bal	Footfall lag 0	19191.11	19175.91	19495.86	19176.67	19675.93
	W Footfall lag 0	62199.80	61781.97	62518.56	61944.81	60747.32
	W Crime lag 1	115.10	**6.83**	46.53	406.22	15.97
Chi	Footfall lag 0	30501.37	30492.66	30551.09	30487.24	30541.01
	W Footfall lag 0	43561.39	43528.17	43087.39	43515.43	43214.13
	W Crime lag 1	281.07	36.42	68.89	274.48	85.00
Phi	Footfall lag 0	417.30	417.14	416.73	417.48	415.64
	W Footfall lag 0	5581.43	5643.13	5675.60	5655.77	6539.99
	W Crime lag 1	201.55	112.87	61.86	343.31	**9.57**

Values below the conventional threshold of 10 are shown in bold.

Appendix P. Heterogeneity Analysis Results

Table P26. IRR per 36,000 additional pedestrians for Baltimore, where "Low" is the effect in the low group, "High" the effect in the high group, and "Int." the interaction effect.

Measure	Effect	All	Assault	Burglary	MVT	Robbery
POI Diversity	Low	1.002	1.002	1.009	0.998	1.003
	High	1.028	1.038	1.022	0.993	1.051
	Int.	1.026	1.036	1.013	0.995	1.047
% Night Business	Low	0.978	0.973	1.017	0.976	1.007
	High	1.001	1.002	1.009*	0.998*	1.003
	Int.	1.024	1.030	0.992	1.022	0.995
Median HH Income	Low	1.002	1.025	1.000	0.948	1.015
	High	1.001	1.002	1.009*	0.998	1.003
	Int.	0.999	0.977	1.009	1.052	0.988
% Unemployed	Low	1.011	1.024	1.046**	0.977*	0.991
	High	1.001	1.002	1.008	0.998	1.003
	Int.	0.991	0.978	0.964	1.022	1.012
% Occupied HUs	Low	1.001	1.002	1.009	0.998	1.003
	High	1.014	1.026	1.039	1.000	0.982
	Int.	1.012	1.024	1.030	1.001	0.979
Ethnic Diversity	Low	0.989	0.979	1.105	0.988	1.012
	High	1.001	1.002	1.009*	0.998	1.003
	Int.	1.012	1.023	0.913	1.011	0.991
% Single-parent HH	Low	1.002	1.002	1.009	0.998*	1.003
	High	1.037	1.052	1.030	1.001	1.055
	Int.	1.036	1.050	1.020	1.003	1.051

Stars indicate significance and are based on two-sided bootstrap *p*-values: * p <0.05, ** p <0.01, *** p <0.001.

Table P27. IRR per 36,000 additional pedestrians for Chicago, where "Low" is the effect in the low group, "High" the effect in the high group, and "Int." the interaction effect.

Measure	Effect	All	Assault	Burglary	MVT	Robbery
POI Diversity	Low	1.001	1.000	0.999	1.001	1.001
	High	1.003	1.009	0.978**	1.002	1.003
	Int.	1.002	1.008	0.979*	1.001	1.002
% Night Business	Low	1.001	1.000	0.999	1.001	1.001
	High	1.001	1.000	0.965**	1.010**	0.995
	Int.	1.000	1.000	0.965**	1.008*	0.994
Median HH Income	Low	1.084**	1.081**	1.069	1.077**	1.086
	High	1.001	1.000	0.999	1.001	1.001
	Int.	0.923**	0.925**	0.935	0.930**	0.922
% Unemployed	Low	1.001	1.000	0.999	1.001	1.001
	High	1.015**	1.018**	1.009	1.010	1.032**
	Int.	1.014**	1.018**	1.010	1.009	1.031**
% Occupied HUs	Low	1.001	1.000	0.999	1.001	1.001
	High	1.000	1.002	0.953**	1.010**	0.993
	Int.	0.999	1.002	0.953**	1.008**	0.992
Ethnic Diversity	Low	1.000	0.994	0.978	1.015**	0.980
	High	1.001	1.000	0.999	1.001	1.001
	Int.	1.001	1.007*	1.022	0.986**	1.021
% Single-parent HH	Low	1.001	1.000	0.999	1.001	1.001
	High	1.131**	1.104*	1.057	1.122	1.085
	Int.	1.131**	1.104*	1.058	1.120	1.084

Stars indicate significance and are based on two-sided bootstrap *p*-values: * p <0.05, ** p <0.01, *** p <0.001.

Table P28. IRR per 36,000 additional pedestrians for Philadelphia, where "Low" is the effect in the low group, "High" the effect in the high group, and "Int." the interaction effect.

Measure	Effect	All	Assault	Burglary	MVT	Robbery
POI Diversity	Low	0.999	1.001	1.001	0.996*	0.999
	High	1.003	1.014	0.956	0.980**	0.968
	Int.	1.003	1.014	0.955	0.984*	0.969
% Night Business	Low	1.000	1.001	1.002	0.997	0.998
	High	0.998**	1.000	0.998	0.994*	1.001
	Int.	0.998**	0.998*	0.996	0.997	1.003
Median HH Income	Low	1.001	1.009	1.015	1.006	0.989
	High	0.999	1.000	1.001	0.996**	0.999
	Int.	0.998	0.992	0.987	0.990	1.010
% Unemployed	Low	0.999	1.000	1.001	0.997	0.999
	High	0.998	1.004	0.988	0.983**	0.990*
	Int.	0.998	1.004	0.986	0.985	0.992
% Occupied HUs	Low	0.999	1.000	1.001	0.997	0.999
	High	0.998	1.003	0.996	0.985	0.991*
	Int.	0.999	1.003	0.995	0.988	0.992
Ethnic Diversity	Low	1.000	1.003	1.006	0.996	0.995
	High	0.999	1.000	1.001	0.996**	0.999
	Int.	0.999	0.997	0.995	1.000	1.004
% Single-parent HH	Low	0.999	1.000	1.001	0.996**	0.999
	High	0.989	1.009	1.010	0.979	0.986
	Int.	0.990	1.009	1.008	0.983	0.988

Stars indicate significance and are based on two-sided bootstrap *p*-values: * p <0.05, ** p <0.01, *** p <0.001.

Appendix Q. Number of Effective Hexagons by Configuration

Table Q29. Number of hexagons effectively used per city and crime type.

City	All	Assault	Burglary	MVT	Robbery
Bal	250	245	234	245	227
Chi	641	636	623	633	602
Phi	410	403	376	399	351

Counts are identical across all five models and four granularities.