



ARTICLE



<https://doi.org/10.1057/s41599-024-02610-8>

OPEN

Unraveling the influence of income-based ambient population heterogeneity on theft spatial patterns: insights from mobile phone big data analysis

Chong Xu¹, Zhenhao He ¹, Guangwen Song ¹✉ & Debao Chen²

While previous research has underscored the profound influence of the ambient population distribution on the spatial dynamics of crime, the exploration regarding the impact of heterogeneity within the ambient population, such as different income groups, on crime is still in its infancy. With the support of mobile phone big data, this study constructs an index of ambient population heterogeneity to represent the complexity of the social environment. After controlling for the effects of total ambient population, nonlocal rate, transportation accessibility, crime attractors, and crime generators, this study employs a negative binomial regression model to examine the influence of ambient population heterogeneity and different income groups on the spatial manifestations of thefts. The findings indicate that ambient population heterogeneity significantly escalates the incidence of thefts, with middle and upper-middle-income groups acting as more attractive targets, whereas the higher-income group exerts a deterrent effect. The interaction analysis shows that increased population heterogeneity contributes to social disorder, thereby amplifying the attractiveness of the ambient population to perpetrators. These conclusions highlight the crucial role of ambient population heterogeneity in explaining crime dynamics and therefore enrich the routine activity theory.

¹Center of GeoInformatics for Public Security, School of Geography and Remote Sensing, Guangzhou University, Guangzhou, China. ²Department of Geography, University of Cincinnati, Cincinnati, OH, USA. ✉email: songgwen@126.com

Introduction

In recent studies, researchers have recognized the significance of ambient population, which refers to individuals present at specific times and locations, in explaining and predicting crimes (Andresen, 2007). To explore this, scholars have turned to alternative sources of data, such as the LandScan global population database (Andresen, 2011), WiFi Sensor data (Kontokosta and Johnson, 2017), Geotagged Tweets data (Lan et al., 2019), and mobile phone signal data (Hanaoka, 2016; Malleson and Andresen, 2016), instead of relying solely on traditional census data. These big data have proven to be more precise in measuring the distribution of ambient population at finer spatio-temporal scales, thus providing insights into potential victimization.

Furthermore, scholars have become increasingly aware of how ambient populations with different social backgrounds can have varying impacts on crime (Haleem et al., 2021; He et al., 2020). For example, people from different income backgrounds have different spaces and patterns of activity and encounter different risks of crime (Song et al., 2023). Nevertheless, existing research has not yet reached a consensus regarding which income group experiences higher victimization risks. Additionally, the willingness of different income groups to engage in informal social control within their surroundings is also related to their socioeconomic status, including income (Reynald, 2010). Consequently, the effects of ambient population activities on crime display inconsistency across various income groups.

Currently, research indicates that ethnic/racial heterogeneity often exerts a promoting effect on both violent and property crimes (Taylor et al., 2015; Avison and Loring, 1986). Such communities tend to exhibit higher levels of anonymity and face challenges in establishing shared values, resulting in a decreased willingness among residents to maintain community order through informal social control (Bernasco and Luykx, 2003). Early studies on the impact of ethnic/racial heterogeneity on crime were primarily based on census data, considering only interactions among individuals within communities, while neglecting the effects of population mobility and heterogeneity changes (Gu et al., 2023). Moreover, these studies often measure heterogeneity based on different ethnic/racial groups, which may have limited effectiveness in some regions. For instance, in China, the predominant ethnic group is the Han Chinese, and there is no significant ethnic/racial heterogeneity in society. As a dimension of social diversity, income has been less explored as a measure of ambient population heterogeneity, leaving room for research focusing on income-based perspectives to understand its influence on crime.

With the support of mobile phone big data, this study focuses on the ZG city in China as the research area, distinguishing ambient populations based on their different income levels. Furthermore, a novel income-based ambient population heterogeneity index is constructed to investigate the following questions: (1) What is the impact of ambient population with varying incomes on thefts? (2) How does income-based ambient population heterogeneity influence thefts? (3) What role does the heterogeneity play in moderating the relationship between ambient population and thefts?

Literature review

The compositions of ambient population and crimes. Literature has long acknowledged the impact of population compositions with diverse socioeconomic attributes on crime. Social disorganization theory posits that individuals of various backgrounds may experience disparities in establishing common values and upholding community order, which may lead to an increase in criminal cases (Shaw and McKay, 1942). For instance,

certain marginalized groups, those residing in concentrated disadvantaged communities with lower wages and higher unemployment rates, are more susceptible to the influence of non-mainstream values and street culture (Wilson, 2011). Consequently, they may resort to illicit means, such as theft or violence, as a means of problem-solving, further exacerbating the occurrence of criminal incidents within the community. These studies take into account the emergence of criminal motives, considering them as products of the social and environmental structure (Rice and Csmith, 2002). They explain the influence of population composition within a community on crime from a macro-environmental perspective but simultaneously overlook the conditions required for the commission of crimes, namely, criminal opportunities. In contrast, the Routine activity theory and Situational Crime Prevention theory posit that crime is influenced not solely by individual traits or social structural factors but also by the specific characteristics of the crime situation and its environment (Eck and Clarke, 2019; Freilich and Newman, 2017; Cohen and Felson, 1979). People carry out activities for work, leisure, education, etc., and visit different places throughout the day, indoors or outdoors, in private or public spaces, all of which present different crime risks (Song et al., 2018a). Since criminal risks are constantly changing, it is essential to consider the impact of different individuals' activities within the environment on crime (Andresen, 2011).

Yet, most such studies rely solely on static surveys or census data for analysis, potentially leading to biased findings. This is because urban populations are characterized by persistent fluctuations, which is exemplified in residents' daily mobility patterns as they traverse diverse locations and are consequently exposed to varying degrees of crime risk (Kwan, 2012). A significant body of research indicates that a larger ambient population includes more potential targets and provides greater opportunities for crime (Song et al., 2021; Lan et al., 2019). And some studies turn their focus toward the impact of different groups' activities on crime.

For instance, Boivin and Felson (2018) probed into the impact of visitors, categorized by four different activity purposes—work, shopping, recreation, and education—on local crime. The study uncovered that visitors engaged in entertainment, shopping, and work activities could potentially act as offenders or become targets, thereby contributing to an increase in the local crime rates. Conversely, visitors engaged in educational activities often acted as capable guardians and exhibited a notable negative correlation with crime. In line with this trend, He et al. (2020) conducted a study utilizing mobile phone data to classify the ambient population into two distinct social groups: nonlocal residents and individuals who engaged in regular activities within the community, and found that the presence of regular social activity groups locally exerted an inhibitory effect on crime.

Likewise, Song et al. (2021) utilized Baidu LBS data and Wo et al. (2022) utilized geo-tagged tweet data to categorize the ambient population into residents (insiders) and visitors (outsiders). Both of these studies propose that individuals who are more familiar with the environment are more likely to become potential guardians. When they have a stronger sense of responsibility towards the people and places in the community, they are more inclined to provide more effective supervision (Felson, 1995). On the other hand, the influx of external visitors introduces heightened complexity to the environment and diminishes familiarity, thereby posing challenges for local residents in identifying potential offenders and detecting suspicious activities. Consequently, the effectiveness of guardianship behavior among insiders is reduced.

The empirical studies above underscore the significance of considering the composition of the ambient population when analyzing the spatial patterns of crimes. However, most of these studies have primarily categorized ambient populations into residents and other groups (such as migrants or employees), neglecting socioeconomic status, such as income level, which significantly influences criminals' decision-making processes.

Relationship between different income groups and crimes.

Literature has proved that different income groups would have varied impacts on the spatial distribution of crimes (Hipp, 2007; Pare and Felson, 2014). Many of them have investigated the correlation between low-income levels and higher crime rates in a particular area, which typically emphasizes how poverty, unstable housing, or family conditions influence informal social control within communities, consequently affecting crime (Pratt and Cullen, 2005). Meanwhile, in Fleisher's early research on 101 American cities, the results showed that higher family income was correlated with lower crime rates among young men (Fleisher, 1963). According to previous studies, both low-income and high-income individuals can serve as potentially motivated offenders, potential victims, or guardians in a given context.

In terms of motivated offenders, traditional theories, such as conflict theory (Taylor, Walton, and Young, 2013), subcultural theory (Cloward and Ohlin, 2013), and social disorganization theory (Shaw and McKay, 1942), share a common perspective that there exists a relationship between the economic status and criminal behavior. As Vold (1958) pointed out, "the lack of basic material needs for survival and minimum welfare will lead to a flourishing of criminal activities in a particular region." However, the assumption that individuals from lower socioeconomic backgrounds are more prone to engaging in criminal behavior than those with better economic conditions remains a subject of controversy (Patterson, 1991). All human behaviors, including crime, are the contextual outcomes of a perception-choice process (Wikström, 2006). The inclination of potential offenders to engage in criminal activities may be influenced by a myriad of factors, including individual cognitive abilities, access to information, moral standards, and risk tolerance, rather than solely focusing on the benefits (Bernasco and Nieuwebeerta, 2005).

From the perspective of potential victims, populations with different incomes exhibit varying degrees of attractiveness to motivated criminals. Residential areas often display visible features that reveal their value, implying the wealth of inhabitants. Studies by Bernasco and Luykx (2003) found that areas with higher housing values, proximity to CBD areas, and better accessibility are more susceptible to being targeted by burglars for residential theft. Similarly, research by Xiao et al. (2017) discovered that commercial buildings are more likely to be preferred targets for criminals compared to other types of residences. This is because affluent communities offer higher potential rewards, making them more appealing to criminals. In their research investigating street-level violent crime and property crime in Kentucky, Jones and Pridemore (2019) discovered that in relatively disadvantaged community streets, the diminished presence of customers drawn by commercial activities results in fewer valuable targets and reduced crime opportunities.

From the view of guardianship, the intervention of residents in the informal regulation of theft offenses may also be related to their income. Felson (2006) pointed out that the willingness of a capable guardian to intervene in crime depends not only on their knowledge and experience in dealing with crimes, and self-protection but also on their understanding of the role they play in preventing crime. This also means that their willingness to

intervene when necessary is influenced by the social context in which they are situated (Hipp, 2016).

More stable and higher-income residents, in order to protect community interests and property values, are more willing to provide informal surveillance (Bursik Jr., 1988). On the other hand, low-income groups residing in concentrated disadvantaged communities experience higher residential mobility and complex community compositions, which hinder the establishment of shared values among residents and the maintenance of effective social control (Sampson and Groves, 2017). In the study conducted by Reynald (2010), it was found that residents in communities with relatively higher crime rates and lower income exhibited a diminished willingness to participate in supervision and intervention when necessary.

In summary, different income groups exhibit distinct social and behavioral characteristics, leading to variations in their capacity to provide informal control, their likelihood of engaging in criminal behavior, and the risks they encounter with crime. These factors, in turn, have complex implications for theft-related offenses.

The importance of heterogeneity of ambient population with different incomes in China.

Many studies have examined the impact of racial/ethnic heterogeneity on crime based on the racial composition of residents within communities (Sun et al., 2004; Roncek and Maier, 1991). For instance, Sampson and Groves (2017) discovered that communities with greater racial differences had a higher likelihood of experiencing violent events. Hipp (2011) found that in cities with higher levels of racial/ethnic heterogeneity, increasing levels of racial segregation over time led to an overall rise in serious crimes. Such research has predominantly been conducted in Western countries, where there are more racial/ethnic groups with different characteristics. While China comprises numerous ethnic groups, the Han population constitutes the vast majority, and there are relatively few physical and lifestyle differences among these various groups (Xu et al., 2022).

Instead, income can be an important dimension in measuring demographic characteristics. Different income groups exhibit significant differences in spatial ranges and activity patterns in their daily lives. Existing literature has long highlighted the inequalities in residence and daily activities among various population groups from the perspective of socioeconomic status (Wang and Li, 2016). Zhou and Deng (2010) found that the activity spaces of high-income groups in Guangzhou generally exceed those of low-income groups. Similarly, Tao et al. (2020) discovered significant disparities in activity spaces between low-income groups and other income groups in urban areas and new towns. As for daily activity characteristics, Zhang and Chai (2011) analyzed the daily activity patterns of middle and low-income individuals in Beijing and found that low-income individuals experience fragmented activities, longer commuting times, and relatively dispersed activity spaces compared to middle-income groups. These findings indicate evident differences in transportation accessibility, available time, and spatial mobility among individuals with varying socioeconomic statuses and resources.

Meanwhile, scholars have extensively explored the relationship between income inequality and crime. Tsushima (1996) revealed a positive correlation between income inequality and theft, while poverty levels showed a significant positive correlation only with murder. Hipp (2011) argued that income inequality and poverty can lead to social disorder, strained social relationships, and cultural conflicts, which subsequently trigger social exclusion and a sense of relative deprivation. This, in turn, weakens informal social control and promotes crime. Metz and Burdina (2018)

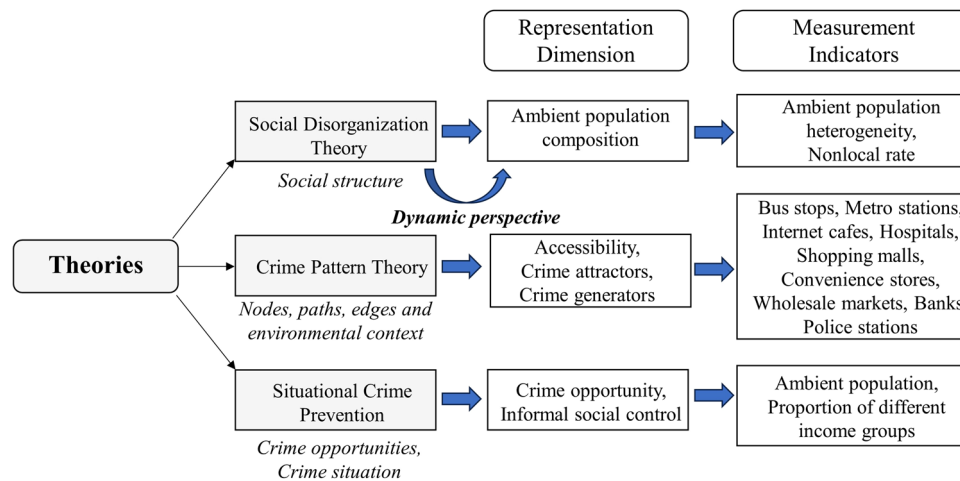


Fig. 1 Conceptual framework. Crime is influenced by both the social environment and the dynamics of criminal opportunities, which can be regarded as a behavioral outcome in the specific context.

found that as the income gap widens between more prosperous neighborhoods and the poorest adjacent blocks, property crime levels increase in the wealthier areas. The relationship between ambient population and crime across different income groups becomes more intricate due to population mobility. However, these studies have exclusively relied on static questionnaires or census data, neglecting to capture the true dynamics of activity spaces among different income groups and their compositions within specific areas. Song et al. (2023) utilized mobile phone user data to differentiate the impact of various ambient income groups on theft and found that no single group could consistently predict theft, as the influence of each income group fluctuated over time.

There have been few studies using ambient population data to measure population heterogeneity. Recently, Gu et al. (2023) innovatively proposed a racial heterogeneity index using ambient population data, which outperforms traditional indices in explaining street robbery crimes, and demonstrates a significant positive impact on street robbery, while the traditional indices do not show significance. This suggests that using ambient population data to measure heterogeneity can better explain crimes.

Summary. Compared to traditional census data, employing big data analytics to assess the data of the ambient population can offer a more comprehensive understanding of potential victims (Andresen, 2011). However, it is essential to recognize that the ambient population is not a homogeneous entity, and social groups with different education or income levels, may exert diverse influences on criminal behavior. Currently, there is limited research investigating the impact of ambient populations with different income levels on crime. Various income groups experience noticeable differences in their residential and activity spaces, leading to divergent levels of exposure to crime risks and potential victimization. Additionally, an individual's likelihood of involvement in theft-related crimes and their inclination to provide informal supervision for criminal activities are also closely correlated with their socioeconomic background.

The heterogeneity of populations has indeed been confirmed in previous research to have a promoting effect on criminal incidents. However, existing studies have predominantly relied on static data for investigation, overlooking the changes in population composition brought about by population mobility. This might not fully account for the dynamic social and environmental shifts prompted by the spatio-temporal mobility of various population groups. These limitations might lead to

incomplete understanding, especially in areas with smaller residential populations but larger transient populations, such as commercial districts or recreational parks.

Therefore, this study assesses ambient population levels based on community-level mobile phone user data, differentiating various income groups and examining their connections to thefts. Additionally, we introduce an innovative ambient population heterogeneity index based on income levels to explore its influence on crime. Lastly, we analyze the moderating effect of heterogeneity on the relationship between the ambient population and thefts by multiplying the ambient population with the heterogeneity index. This study focuses on the following hypotheses: (1) higher ambient population heterogeneity promotes thefts by increasing the anonymity of the social environment; (2) Ambient population heterogeneity has a facilitative moderating effect on the relationship between ambient population and thefts; and (3) High-income groups, residing in communities with better surveillance and being more inclined to offer informal supervision, exert a negative effect on theft.

Data and methods

Study area. ZG City, located in the southeastern coastal region of China, plays an instrumental role as a prominent hub for economic, political, and cultural activities within the South China region. The region accommodates an extensive population characterized by a vibrant interchange of diverse economic activities and a complex demographic composition.

This study adopts a community-based analysis unit. Communities are preferred for crime analysis because they exhibit relatively inherently homogeneous population structures. In ZG City, there are a total of 2,055 communities, with an average area of 1.81 square kilometers per community.

Theoretical framework and variables

Theoretical framework. This article establishes a theoretical framework based on Social Disorganization Theory, Crime Pattern Theory, Routine Activity Theory, and Situational Crime Prevention Theory (Fig. 1). Social Disorganization Theory is primarily employed to explain the impact of the social environment on crime. This study mainly discusses how the ambient population heterogeneity and nonlocal residents increase the anonymity of the social environment, weakening informal social control and potentially fostering criminal activities.

Crime pattern theory is closely related to the effects of aggregated elements of the physical environment (nodes, paths,

edges, and environmental contexts) as perceived by the offender on crime distribution. The selected indicators in this study include bus stops, subway stations, internet cafes, shopping malls, supermarkets, and banks, serving as covariates that reflect the influence of the built environment on crime.

Routine activity theory (RAT)/situational crime prevention (SCPT) theory both focuses on the specific situations and individual behaviors that lead to criminal activities. Both theories incorporate elements of the opportunity theory, emphasizing the influence on criminal behavior by either increasing or decreasing crime opportunities. While SCPT emphasizes managing crime opportunities in specific environments, the RAT highlights the creation or hindrance of crime opportunities in individuals' daily activities. In this study, crime opportunities' dynamic changes are primarily measured through the ambient population and the activities of different income groups within it.

In the following sections, more discussion of this conceptual framework will be provided in the data and methods, variable analysis, regression model results, and conclusion and discussion parts.

Dependent variable. The dependent variable in this study is the number of thefts in each community, which was obtained from the ZG City Police Bureau in 2019. Theft is defined as the unlawful act of taking possession of another person's property without their permission, including items such as cash, mobile phones, and other belongings, while evading the control of others. The study focuses primarily on street theft excluding cases of burglary. As the data are official and some of the victims may not have reported the thefts to the police, the problem of dark figures in crime records could not be avoided in this study. However, as this is a relatively small proportion of the situation, the crime dataset for this study is still reliable.

Our dataset for thefts contains detailed records on the timing and geographic coordinates of each crime event. We use ArcGIS to link the point data of thefts with the corresponding communities. Subsequently, we aggregate the count of theft incidents at the community level, serving as the dependent variable.

Ambient population data. The estimation of the ambient population was conducted using mobile user data provided by a prominent telecommunications company in China. This telecom operator ranks among the top three in the country, serving a vast user base of over 300 million individuals and maintaining a market share upwards of 20%. In 2019, the data platform recorded a total of approximately 400 million mobile users.

The estimation of the ambient population adheres to certain rules: The company provides a 500 × 500 m grid layer, and whenever a user makes a phone call, sends a text message, or requests data services, the location of the user at the time of the service request is recorded. It is essential to clarify that an individual is considered part of the ambient population only if they stay in the same location for at least half an hour. This means that if a person stays within a single grid for an hour, they will be counted only once for that grid. Similarly, if they move across multiple grids, they will be included in the calculations for each grid only if their stay exceeds half an hour in each. Each person can be counted a maximum of two times per hour. If a user's phone is turned off midway or the signal is lost, their location will be assumed to be the last recorded location, and subsequent periods will not be recorded again.

In addition to user location information, the company utilizes multiple data sources, encompassing attributes tied to residential and behavioral characteristics of the users, such as housing prices in the residential district, mobile terminal prices, the number of

entertainment venues visited, travel volume to other cities (domestic and international), mode of transportation (airplane, high-speed trains, or self-driving), and phone expenses. These datasets are used to construct a more reliable and comprehensive indicator that measures individuals' socioeconomic backgrounds and daily activities. Machine learning methods are employed to model income profiles, which categorize users into eight income and economic status levels. In this study, these eight income levels are classified into five categories: low-income (annual income <50,000 yuan), lower-middle-income (annual income 50,000–100,000 yuan), middle-income (annual income 100,000–180,000 yuan), upper-middle-income (annual income 180,000–300,000 yuan), and high-income (annual income >300,000 yuan). Specific criteria for the different income levels and the description of the scenarios can be found in the supplementary material. All these datasets are processed by the telecommunications company.

Given that the company supplies grid-based ambient population data with a resolution of 500 m × 500 m, this study overlays the grid data with a community layer. It assumes a uniform distribution of the ambient population within each grid and calculates the proportion of the grid area to the community area. This allows for the aggregation of grid-level population data into community-level data, yielding the total number of ambient populations and different income groups within each community.

Covariates

Population Heterogeneity Index: In the early stages of research, the focus on racial heterogeneity primarily revolved around quantifying the proportions of White and Black populations, reflecting the prevailing social context of that time. Subsequently, Gibbs and Martin (1962) introduced the concept of social heterogeneity, and Blau (1977) incorporated this index into the realm of crime research. The formula for this index is as follows:

$$H = 1 - \sum_{i=1}^n P_i^2 \quad (1)$$

In this equation, H represents the social heterogeneity index, n denotes the number of distinct racial groups in the community, and P_i indicates the proportion of each racial group within the community.

As societal conditions evolved and grew increasingly complex, the inclusion of other ethnic minorities, particularly the Hispanic population, became pertinent in the calculation of the Herfindahl Index (Bernasco and Block, 2011). Subsequently, this index gained widespread acceptance and found application in the realm of criminological research.

$$H_r = 1 - \frac{B^2 + H^2 + W^2}{(B + H + W + O)^2} \quad (2)$$

The variables B , H , W , and O represent the numbers of the Black, Hispanic, White, and Other ethnic groups residing in each neighborhood, respectively. These values range from 0 to 1, with lower values indicating an uneven distribution of races and higher values indicating a more balanced racial distribution within the area.

However, conventional measures of racial heterogeneity encounter subsequent issues: primarily, they only consider the impact arising from intracommunity interactions. The ambient populations within regions are in constant flux due to various activities such as work, shopping, entertainment, education, and other purposes, which ultimately influence the dynamics of the local community's environment. Besides, in regions where racial/ethnic heterogeneity distinctions are less apparent, alternative indices such as socioeconomic attributes can be employed more effectively for substitution in computations. Lastly, comparisons

between numerical values may potentially engender misunderstanding due to differences in magnitude, yet proportional calculations accentuate the relative significance of elements.

Therefore, this study introduces a novel heterogeneity index building upon the formula (2). Firstly, this index derives its data from mobile signaling data and the socioeconomic profiles of users. Secondly, rather than relying on absolute numbers, this index employs the proportions of distinct income groups as the basis for its calculation. The new formula is as follows:

$$H_{APr} = 1 - \frac{A_{APj}^2 + B_{APj}^2 + C_{APj}^2 + D_{APj}^2 + E_{APj}^2}{(A_{APj} + B_{APj} + C_{APj} + D_{APj} + E_{APj})^2} \quad (3)$$

H_{APr} represents the racial heterogeneity, A_{APj} , B_{APj} , C_{APj} , D_{APj} and E_{APj} respectively represent the proportions of ambient populations of low-income, lower-middle-income, middle-income, upper-middle-income, and high-income groups in region j . The resulting score reflects the level of population heterogeneity, with higher scores indicating greater population variety and lower scores suggesting that the population is predominantly composed of a specific income group.

Social disorder: As mentioned earlier, crime analysis studies in China frequently use the proportion of migrant populations as an indicator of social disorganization (Song et al., 2018b). These individuals have complex social backgrounds and a lower sense of community identification. Moreover, their unfamiliarity with the social environment increases social anonymity, thereby limiting the effectiveness of informal social control mechanisms and ultimately exacerbating the level of social disorder. This study employs the proportion of nonlocal populations in each community, acquired from the seventh national population census, as an indicator of social disorganization.

Transportation accessibility: Transportation plays a significant role in influencing criminal decision-making. On one hand, it can reduce the costs associated with criminal activities by providing easy access to various locations. On the other hand, during peak travel periods, public transportation stations become crowded, offering more potential opportunities for offenders. Moreover, these stations also offer convenient escape routes for criminals after committing a crime. Empirical studies have indicated that crime tends to increase in areas with higher concentrations of transportation facilities (Xiao et al., 2021). In this study, the accessibility of transportation is measured by considering the presence of subway and bus stations in each community.

Crime attractors and crime generators: According to crime pattern theory (Brantingham and Brantingham, 1995), the concept of crime attractors is used to identify locations that offer favorable conditions for certain types of crimes, thus attracting potential offenders. For instance, establishments like bars (Groff, 2011) and internet cafes (Song et al., 2018b) can be attractive to motivated offenders since they provide opportunities where individuals' attention is often distracted, making them susceptible to theft. On the other side, crime generators are typically identified by the presence of places that individuals frequently visit as part of their daily activities. Such locations include shopping malls (Steenbeek et al., 2011), hospitals, banks (Haberma and Ratcliffe, 2015), and other such frequented sites. Interesting to note is that in certain studies, police stations are also identified as crime generators (Helbich and Jokar Arsanjani, 2015). This is because police stations are often located in densely populated areas, which inherently generate numerous opportunities for crime, increasing the likelihood of potential interactions between victims and perpetrators.

Method. The negative binomial regression model is a common choice for analyzing count data that exhibit overdispersion, where the variance is greater than the mean. This is often the case with criminal incident counts, which are non-negative integer variables with high variability. The negative binomial model accommodates this overdispersion and provides more accurate estimates compared to the standard Poisson regression model.

The negative binomial distribution is a continuous mixture of the Poisson distribution, which is suitable for modeling crime counts. It can be represented by the following equation:

$$\ln(\lambda_i) = \sum_{k=0}^k \beta_{ki} X_{ki} + \varepsilon$$

In the equation, λ_i represents the count of thefts in community unit i , β_k represents the regression coefficient, and X_k represents the value of the k th independent variable. When $k=0$, β_0 represents the intercept of the model, and ε represents the model residuals.

This study has established both a baseline model and an interaction model. The conceptual formula for the baseline model is as follows:

$$\ln(\lambda_i) = \beta_i P_i + \beta_i L_i + \sum_{a=1}^a \beta_{ai} T_{ai} + \sum_{b=1}^b \beta_{bi} A_{bi} + \sum_{c=1}^c \beta_{ci} G_{ci} + \varepsilon$$

In the equation, λ_i represents the count of thefts in community unit i , β represents the regression coefficient, P_i represents the ambient population in the i -th community, L_i represents the proportion of population heterogeneity in the i -th community, T_a represents the a -th transportation accessibility variable, A_b represents the b -th crime attractor variable, G_c represents the c -th crime generator variable, and ε represents the model residuals.

An interaction model is a statistical model for exploring the interaction between two or more independent variables. It allows us to test whether there is an interaction between the independent variables, i.e., how they together affect the dependent variable. In the interaction model, we multiply the ambient population and heterogeneity index, creating interaction variables. The interaction variables used in the study undergo a process of centering to reduce multi-collinearity between the interaction variables and the original variables.

In this study, the standardized regression coefficients (St.Coef.) are used to measure the strength of the impact of independent variables on thefts. These coefficients are calculated by standardizing the variables before running the negative binomial regression. Standardization eliminates the influence of variable scale, allowing for a more meaningful comparison of the magnitude of effects between different variables.

The incidence rate ratio (IRR) coefficients are also employed in the analysis. These coefficients indicate the multiplicative change in the probability of the dependent variable occurring for each unit increase in the independent variable. For example, an IRR of 1.5 means that a one-unit increase in the independent variable is associated with a 50% increase in the probability of the dependent variable occurring.

To compare the goodness of fit among different models, this study utilizes the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). These criteria provide a measure of the relative quality of different models. Lower AIC and BIC values indicate better model fit.

Result

Descriptive statistics. Table 1 presents the mean, standard deviation, and range of the dependent and independent variables. From the table, it can be observed that theft cases, ambient population, and nonlocal population rate exhibit the highest

Table 1 Descriptive statistics of variables.

Reference theory	Variables	Mean	Std. dev.	Min	Max	
Situational crime prevention theory	Dependent variables					
	Theft crime count (case)	32.040	48.496	0	674	
	Independent variables:					
	Population heterogeneity	0.704	0.016	0.550	0.740	
	Low income (%)	17.109	3.179	6.855	44.954	
	Lower-middle income (%)	29.631	2.960	0.338	47.914	
	Middle income (%)	28.245	2.160	8.114	45.322	
Routine activity theory	Upper-middle income (%)	16.798	2.153	5.407	26.688	
	High income (%)	8.217	1.853	0.404	16.117	
	Covariates					
	Target					
	Ambient population(k)	307.021	387.077	0.386	5937.670	
	Social disorganization theory	Social Disorder				
	Nonlocal rate (%)	31.981	23.056	0	97.442	
Crime pattern theory	Accessibility					
	Subway station	0.061	0.261	0	3	
	Bus station	2.929	4.258	0	102	
	Crime attractor					
	Internet cafés	0.423	0.924	0	11	
	Crime generator					
	Police station	1.131	1.550	0	11	
	Hospital	0.340	1.077	0	17	
	Shopping mansion	1.669	3.278	0	49	
	Convenient store	2.218	3.103	0	31	
	Wholesale market	0.227	1.171	0	19	
	Banks	3.841	5.488	0	71	

standard deviation related to high levels of dispersion in the spatial distribution. In contrast, facilities such as subway stations and other POIs show a relatively balanced distribution, Although the number and distribution of these POIs are fixed, the different types of POIs can have different effects on crime as the crowds they attract change over time. For example, communities with a higher distribution of bus and metro stops will attract more people at peak times, providing more opportunities for crime and leading to a higher risk of crime.

Comparing the proportions of different income groups, it can be noted that the proportion of middle-income individuals is the largest. Lower-middle and upper-middle income groups bear roughly equivalent proportions, while high-income individuals constitute the smallest segment. The ambient population is reported in thousands for statistical convenience.

Figure 2 provides the distribution of theft cases across ZG City in 2019. The graph highlights that theft incidents are somewhat scattered throughout the city, but certain areas within the city exhibit a more significant concentration of theft cases. In Fig. 3, which showcases the distribution of ambient population heterogeneity across different communities in ZG City, a noticeable pattern emerges. It becomes evident that communities characterized by higher population heterogeneity tend to display relatively higher numbers of theft cases. This observation seems to imply a potential association between population heterogeneity and thefts.

Regression results. In Table 2, Model 1 illustrates that all independent variables, with the exception of the high-income population, significantly and positively influence theft. Among the variables, the ambient population has the strongest positive impact on crime, as indicated by the standardized regression coefficient of 0.309. The incidence rate ratio (IRR) coefficient for the ambient population stands at 1.362, meaning that with each percent increase in the ambient population within a community unit, the probability of theft occurrence increases by 36.2%. When considering the impact of distinct income groups on crime, it appears that the upper-middle

population exerts the most significant upward pressure on thefts, as indicated by a coefficient of 0.068. The influence of the low-income population on crime is relatively smaller, with a coefficient of 0.023. These findings suggest that higher-income groups have a greater likelihood of attracting potential offenders to some extent.

The coefficient for the proportion of the nonlocal population is 0.137, illuminating its role in promoting crime. In terms of accessibility, the coefficient for bus stops is 0.186, ranking second to the impact of the ambient population. On the other hand, the impact of subway stations on crime is comparatively modest. Upon comparison of crime attractors and crime generators, it becomes apparent that places characterized by a higher frequency of monetary transactions, such as banks, convenience stores, and shopping malls, demonstrate a notably stronger allure for crime. Furthermore, it is worth observing a positive correlation between the police station and thefts. In Chinese cities, police stations are typically established in densely populated areas, which might lead to an elevated rate of reported theft incidents in its surrounding regions.

Comparing Model 2 with Model 1, the effects of the control variables on thefts have remained largely unchanged, except for the different income groups. In Model 2, neither the proportion of the low-income group nor the upper-middle-income group exhibits a significant correlation with thefts. However, the percentage of the high-income group becomes significant and negatively correlated with theft, with a coefficient of -0.072 . Of particular concern is the promotional effect of population heterogeneity. The IRR coefficient indicates that for every 1% increase in ambient population heterogeneity, the probability of theft occurrence increases by 33.4%.

Notably, when conducting interaction analysis, it is common to encounter issues of high collinearity among the variables. To address this, Model 3 excludes the ambient population variable due to its high collinearity with other variables. In Model 3, population heterogeneity continues to significantly influence thefts. The interaction term between ambient population and heterogeneity has the strongest impact on crime, marked by a coefficient of 0.301.

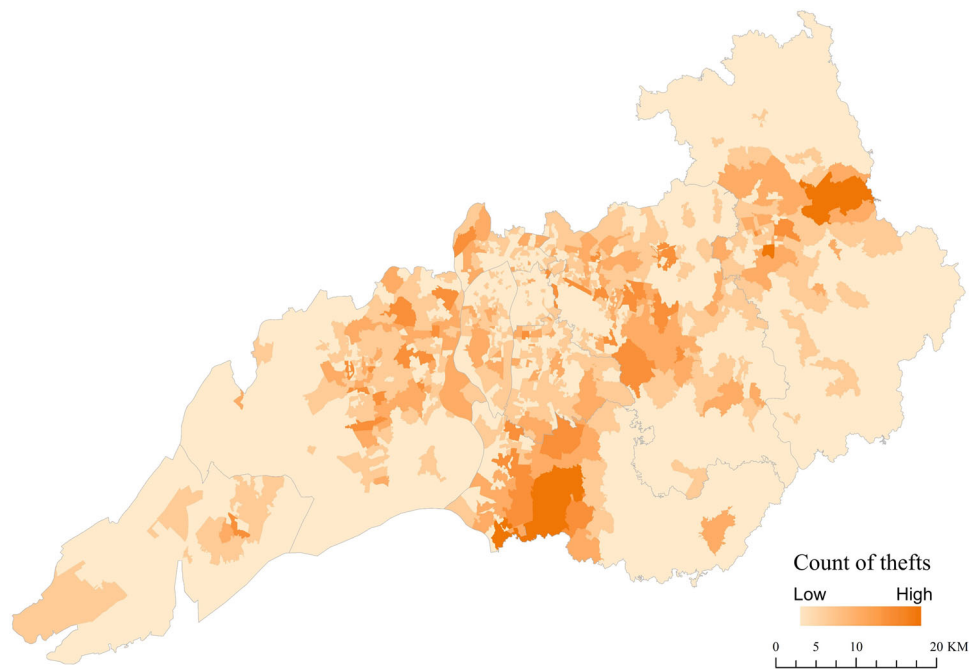


Fig. 2 Distribution of theft cases in ZG City in 2019. There exist theft hotspots in terms of space.

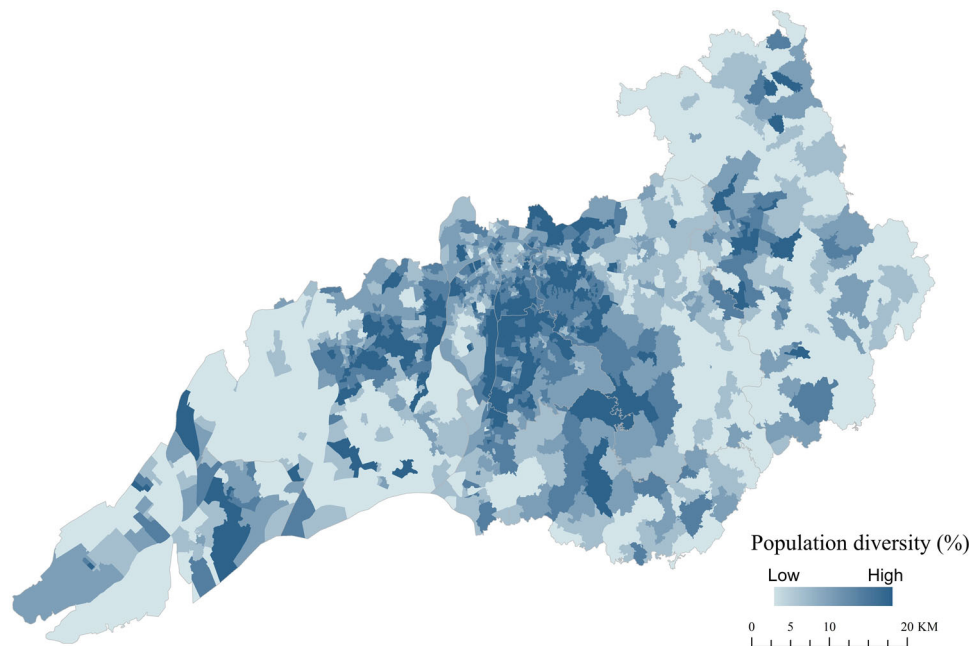


Fig. 3 Distribution of population heterogeneity in ZG City. In the central city, there is a higher level of population heterogeneity.

In the comparison of different income level groups, only the middle-income and high-income categories show significance. The middle-income group demonstrates a positive correlation with thefts, while the high-income group exhibits a negative correlation. Among all three models, Model 3 presents the lowest AIC (16543.569) and BIC (16644.873) values, indicating that it is the best-fitting model.

Conclusion and discussion

This study relies on mobile big data to measure the ambient population with different income levels and constructs a heterogeneity index to investigate the correlation between ambient

population heterogeneity and thefts, aiming to gain deeper insights into the relationship.

The primary finding of this study reveals that an increased incidence of thefts corresponds with a higher level of ambient population heterogeneity. According to the social disorganization theory, this association can be ascribed to the amplified anonymity inherent in population-diverse activities, making it easier for them to carry out thefts and escape undetected. Conventional static racial heterogeneity solely considers the impact of internal community composition on theft, disregarding the inevitable environmental change when residents leave for work or school during the day. In areas like CBD characterized by high commercialization, residents are fewer, and the social environment is

Table 2 Negative binomial regression results.

Thefts in 2019	Model 1			Model 2			Model 3		
	St. coef.	IRR	P	St. Coef.	IRR	P	St. Coef.	IRR	P
Ambient population (k)	0.309***	1.362	0.000	0.301***	1.351	0.000	-	-	-
Population Heterogeneity	-	-	-	0.288***	1.334	0.000	0.288***	1.333	0.000
Ambient population* Population Heterogeneity	-	-	-	-	-	-	0.301***	1.351	0.000
Low income (%)	0.023*	1.023	0.047	-0.014	0.986	0.216	-0.015	0.985	0.203
Middle income (%)	0.034*	1.035	0.024	0.044**	1.045	0.002	0.044**	1.045	0.002
Upper-middle income (%)	0.068***	1.07	0.000	-0.003	0.997	0.875	-0.003	0.997	0.861
High income (%)	0.009	1.009	0.595	-0.071***	0.932	0.000	-0.072***	0.931	0.000
Nonlocal rate (%)	0.137***	1.146	0.000	0.133***	1.142	0.000	0.133***	1.143	0.000
Subway station	0.042*	1.043	0.019	0.049**	1.050	0.006	0.049**	1.050	0.006
Bus stop	0.186***	1.204	0.000	0.197***	1.218	0.000	0.197***	1.218	0.000
Internet cafés	0.079***	1.083	0.000	0.076***	1.079	0.000	0.077***	1.080	0.000
Police station	0.068***	1.070	0.000	0.060***	1.062	0.001	0.060***	1.062	0.001
Hospital	0.068***	1.070	0.000	0.062***	1.064	0.001	0.062***	1.064	0.001
Shopping mansion	0.104***	1.109	0.000	0.109***	1.115	0.000	0.110***	1.116	0.000
Convenient store	0.134***	1.143	0.000	0.118***	1.126	0.000	0.118***	1.126	0.000
Wholesale market	0.053**	1.054	0.002	0.049**	1.050	0.004	0.049**	1.050	0.004
Banks	0.159***	1.173	0.000	0.159***	1.173	0.000	0.159***	1.173	0.000
AIC	16587.584			16543.788			16543.569		
BIC	16683.261			16645.092			16644.873		

***p < 0.001, **p < 0.01, *p < 0.05.

characterized by diverse ambient populations with varying economic backgrounds, rendering the context more intricate. Hence, employing dynamic ambient population heterogeneity is crucial for a more comprehensive explanation and prediction of crime (Gu et al., 2023). In addition to considering the dynamic perspective of population structure, this study suggests that, compared to the racial/ethnic diversity index, indices based on different socioeconomic backgrounds seem to provide a more direct explanation for property crimes such as thefts.

Regarding the moderating effect of ambient population heterogeneity, it amplifies the attraction of the ambient population to criminal activities, implying that thefts are more likely to occur in more complicated environments. As mentioned in situational crime prevention theory, the occurrence of crime is influenced by the characteristics of the crime situation and environment. In the research conducted by Boessen and Hipp (2015), it was found that when highly segregated blocks are situated within high racial/ethnic heterogeneity block groups, crimes will increase. However, Wenger (2019) discovered that as urban diversity increases, the positive association between neighborhood-level diversity and crime becomes less positive. This could potentially indicate that when the diversity within a city reaches a certain level, ethnic attributes might transform into racial harmony/tolerance.

Unlike the preceding research, the population heterogeneity constructed based on income in this study implies that a higher heterogeneity in the ambient population offers potential offenders a wider array of targets to choose from. Simultaneously, the more complex environment also poses challenges to informal surveillance. According to “The Eyes on the Street”, in communities with higher residential stability, residents exhibit a stronger informal supervisory effect, acting as a potent deterrent against criminal activities (Jacobs, 2016). However, in settings with more complex population compositions and higher mobility, the effectiveness of the “eyes on the street” diminishes. Meanwhile, the difficulty in establishing shared values and informal social control mechanisms further aggravates this issue, leading to a higher incidence of thefts within such neighborhoods.

The study also reveals variations in the impact of different income groups on thefts. The middle-income group and the middle-upper-income group demonstrate a positive correlation

with thefts. These groups offer relatively higher potential gains for criminals with lower risks and costs, making them more attractive targets. In contrast, the high-income group shows a negative correlation with thefts. This is likely due to the implementation of advanced security measures and higher risks associated with targeting affluent individuals. Moreover, if the neighborhood in which a high-income group lives has a high crime rate, property values in the neighborhood will depreciate accordingly. Therefore, the willingness of high-income groups to provide informal regulation may be higher in order to safeguard personal or public property interests. Regarding the low-income group, the study shows a positive correlation with crime in Model 1, although the impact is relatively small. Consistent with the findings of Song et al. (2023), this group is more exposed to the risk of crime during weekday mornings while commuting for work.

Consistent with prior research, the magnitude of the ambient population facilitates thefts. This is because thefts, unlike violent offenses, particularly benefit from a heightened ambient population, as it provides more opportunities (Song et al., 2018a). In addition, the proportion of nonlocal inhabitants also exhibits a statistically significant positive impact on thefts, which is perceived as a societal disorder (Feng et al., 2019). These individuals bring diverse social norms, considerable disparities in income and educational attainment, and a notable degree of population heterogeneity to these communities. As a result, inter-resident communication within these communities becomes limited, hindering the formation of community cohesion.

Regarding transportation accessibility, bus stations, and subway stations provide convenient and cost-effective modes of transportation for thieves. Additionally, areas with high traffic volume offer a larger pool of potential targets for criminals, thus an increased presence of public transportation stations within a community promotes theft occurrences (Liu et al., 2020). Internet cafes, as crime attractors, tend to attract the youth influenced by factors such as online gaming and exposure to negative information, who are more susceptible to engaging in criminal activities. Research has also demonstrated a higher likelihood of theft incidents occurring in the vicinity of Internet cafes and similar entertainment establishments. Crime generators, such as shopping malls, convenience stores, wholesale markets, and banks,

pose an increased risk of crime due to the concentration of crowds and monetary transactions (Haberman and Ratcliffe, 2015).

Like all research endeavors, this research has certain limitations that should be acknowledged. To begin with, the issue of data coverage arises. The data used in this study comes from big data on mobile phones, for older people or teenagers who may not have a mobile phone and cannot access data for this group. In addition, there is some bias in the data processing process when aggregating gridded data to communities. Therefore, while the data employed in this study partially reflects the ambient population more accurately than traditional methods, potential biases are still present. Furthermore, although this study analyzed the relationship between different income groups and crime using more detailed data, it did not further distinguish between weekdays and weekends, nor did it delve into the impact of different time periods within a day on theft behavior. Activities of the population exhibit regular patterns throughout the day, and these patterns may influence thefts differently. Lastly, the findings of this study only reveal correlations and do not apply to proving causal mechanisms; causality needs to be proved by randomized controlled trials or causal inference methods. Therefore, future research should use more refined data as well as more integrated approaches, (e.g., combining machine learning with causal inference), to deepen the comprehensive understanding of this research questions.

Notwithstanding the limitations, this study represents a noteworthy attempt to explore the impact of different income groups and ambient population heterogeneity on thefts. Leveraging more refined and dynamic big data, the research emphasizes the significance of the ever-changing social environment in crime analysis and further enriches traditional crime geography theories. Furthermore, the findings derived from this study hold substantial implications for the formulation of local police patrol strategies. It is recommended that the police prioritize areas exhibiting pronounced ambient population heterogeneity and promptly respond to the activities within these regions.

Data availability

Due to police department and telecommunications company requirements, the data that have been used are confidential.

Received: 15 August 2023; Accepted: 4 January 2024;

Published online: 15 January 2024

References

- Andresen MA (2007) Location quotients, ambient populations, and the spatial analysis of crime in Vancouver, Canada. *Environ Plan A Econ Space* 39:2423–2444
- Andresen MA (2011) The ambient population and crime analysis. *Prof Geogr* 63:193–212
- Avison WR, Loring PL (1986) Population diversity and cross-national homicide: The effects of inequality and heterogeneity. *Criminology* 24:733–749
- Bernasco W, Block R (2011) Robberies in Chicago: a block-level analysis of the influence of crime generators, crime attractors, and offender anchor points. *J Res Crime Delinq* 48:33–57
- Bernasco W, Luykx F (2003) Effects of attractiveness, opportunity and accessibility to burglars on residential burglary rates of urban neighborhoods. *Criminology* 41:981–1002
- Bernasco W, Nieuwbeerta P (2005) How do residential burglars select target areas? A New approach to the analysis of criminal location choice. *Br J Criminol* 5:296–315
- Blau PM (1977) A macrosociological theory of social structure. *Am J Sociol* 83:26–54
- Boessen A, Hipp JR (2015) Close-ups and the scale of ecology: Land uses and the geography of social context and crime. *Criminology* 53:399–426
- Boivin R, Felson M (2018) Crimes by visitors versus crimes by residents: the influence of visitor inflows. *J Quant Criminol* 34:465–480
- Brantingham P, Brantingham P (1995) Criminology of place: crime generators and crime attractors. *Eur J Crim Policy Res* 13:5–26
- Bursik Jr RJ (1988) Social disorganization and theories of crime and delinquency: problems and prospects. *Criminology* 26:519–552
- Cloward RA, Ohlin LE (2013) *Delinquency and opportunity: a study of delinquent gangs*. Routledge
- Cohen LE, Felson M (1979) Social change and crime rate trends: a routine activity approach. *Am Sociol Rev* 44:588–608
- Eck JE, Clarke RV (2019) Situational crime prevention: theory, practice and evidence. In: *Handbook on crime and deviance*. Krohn MD, Hendrix N, Penly Hall G, Lizotte AJ (eds.). Springer International Publishing, Cham. 355–376
- Felson M (1995) Those who discourage crime. *Crime Place* 4:53–66
- Felson M (2006) *Crime and nature*. Sage publications
- Feng J, Liu L, Dongping L, Liao W (2019) An examination of spatial differences between migrant and native offenders in committing violent crimes in a large Chinese city. *Int J Geo-Inf* 8:119
- Fleisher BM (1963) The effect of unemployment on juvenile delinquency. *J Polit Econ* 71:543–555
- Freilich JD, Newman GR (2017) *Situational crime prevention*. Oxford University Press
- Gibbs JP, Martin WT (1962) Urbanization, technology, and the division of labor: International patterns. *Am Sociol Rev* 27:667
- Groff E (2011) Exploring 'near': characterizing the spatial extent of drinking place influence on crime. *Aust NZ J Criminol* 44:156–179
- Gu X, Liu L, Lan M, Zhou H (2023) Measuring perceived racial heterogeneity and its impact on crime: an ambient population-based approach. *Cities* 134:104188
- Haberman CP, Ratcliffe JH (2015) Testing for temporally differentiated relationships among potentially criminogenic places and census block street robbery counts. *Criminology* 53:457–483
- Haleem MS, Do L, Ellison M, Bannister J (2021) The 'Exposed' population, violent crime in public space and the night-time economy in Manchester, UK. *Eur J Crim Policy Res* 27:335–352
- Hanaoka K (2018) New insights on relationships between street crimes and ambient population: Use of hourly population data estimated from mobile phone users' locations. *Environ Plan B Urban Anal City Sci* 45:295–311
- He L, Antonio P, Jiao J, An P, Lu C, Mao W, Long D (2020) Ambient population and larceny-theft: a spatial analysis using mobile phone data. *ISPRS Int J Geo-Inf* 9:342
- Helbich M, Jokar Arsanjani J (2015) Spatial eigenvector filtering for spatio-temporal crime mapping and spatial crime analysis. *Cartogr Geogr Inf Sci* 42:134–148
- Hipp JR (2007) Income inequality, race, and place: does the distribution of race and class within neighborhoods affect crime rates? *Criminology* 45:665–697
- Hipp JR (2011) Spreading the wealth: the effect of the distribution of income and race/ethnicity across households and neighborhoods on city crime trajectories*. *Criminology* 49:631–665
- Hipp JR (2016) Collective efficacy: how is it conceptualized, how is it measured, and does it really matter for understanding perceived neighborhood crime and disorder? *J Crim Justice* 46:32–44
- Jacobs J (2016) *The death and life of great American cities*. Vintage
- Jones RW, Pridemore WA (2019) Toward an integrated multilevel theory of crime at place: routine activities, social disorganization, and the law of crime concentration. *J Quant Criminol* 35:543–572
- Kontokosta CE, Johnson N (2017) Urban phenology: toward a real-time census of the city using Wi-Fi data. *Comput Environ Urban Syst* 64:144–153
- Kwan MP (2012) The uncertain geographic context problem. *Ann Assoc Am Geogr* 102:958–968
- Lan M, Liu L, Hernandez A, Liu W, Zhou H, Wang Z (2019) The spillover effect of geotagged tweets as a measure of ambient population for theft crime. *Sustainability* 11:6748
- Liu L, Lan M, Eck JE, Kang EL (2020) Assessing the effects of bus stop relocation on street robbery. *Comput Environ Urban Syst* 80:101455
- Malleson N, Andresen MA (2016) Exploring the impact of ambient population measures on London crime hotspots. *J Crim Justice* 46:52–63
- Metz N, Burdina M (2018) Neighbourhood income inequality and property crime. *Urban Stud* 55:133–150
- Pare PP, Felson R (2014) Income inequality, poverty and crime across nations. *Br J Sociol* 65:434–458
- Patterson EB (1991) Poverty, income inequality, and community crime rates. *Criminology* 29:755–776
- Pratt TC, Cullen FT (2005) Assessing macro-level predictors and theories of crime: a meta-analysis. *Crime Justice* 32:373–450
- Reynald DM (2010) Guardians on guardianship: factors affecting the willingness to supervise, the ability to detect potential offenders, and the willingness to intervene. *J Res Crime Delinq* 47:358–390

- Rice KJ, Csmith WR (2002) Socioecological models of automotive theft: Integrating routine activity and social disorganization approaches. *J Res Crime Delinq* 39:304–336
- Roncek DW, Maier PA (1991) Bars, blocks, and crimes revisited: Linking the theory of routine activities to the empiricism of “hot spots”. *Criminology* 29:725–753
- Sampson RJ, Groves WB (2017) 8 Community structure and crime: testing social-disorganization theory 1. In: *Social, ecological and environmental theories of crime*. Routledge. pp. 93–122
- Shaw CR, McKay HD (1942) *Juvenile delinquency and urban areas*. University of Chicago Press
- Song G, Cai L, Liu L, Xiao L, Wu Y, Yue H (2023) Effects of ambient population with different income levels on the spatio-temporal pattern of theft: a study based on mobile phone big data. *Cities* 137:104331
- Song G, Liu L, Bernasco W, Xiao L, Zhou S, Liao W (2018a) Testing indicators of risk populations for theft from the person across space and time: the significance of mobility and outdoor activity. *Ann Am Assoc Geogr* 108:1370–1388
- Song G, Liu L, Bernasco W, Zhou S, Xiao L, Long D (2018b) Theft from the person in urban China: Assessing the diurnal effects of opportunity and social ecology. *Habitat Int* 78:13–20
- Song G, Zhang Y, Bernasco W, Cai L, Liu L, Qin B, Chen P (2021) Residents, employees and visitors: effects of three types of ambient population on theft on weekdays and weekends in Beijing, China. *J Quant Criminol* 39:385–423
- Steenbeek W, Völker B, Flap H, Oort FV (2011) Local businesses as attractors or preventers of neighborhood disorder. *J Res Crime Delinq* 49:213–248
- Sun IY, Triplett R, Gainey RR (2004) Neighborhood characteristics and crime: a test of Sampson and Groves’ model of social disorganization. *W Criminol Rev* 5:1
- Tao S, He SY, Kwan M-P, Luo S (2020) Does low income translate into lower mobility? An investigation of activity space in Hong Kong between 2002 and 2011. *J Transp Geogr* 82:102583
- Taylor I, Walton P, Young J (2013) *The new criminology: for a social theory of deviance*. Routledge
- Taylor RB, Ratcliffe JH, Perenzin A (2015) Can we predict long-term community crime problems? The estimation of ecological continuity to model risk heterogeneity. *J Res Crime Delinq* 52:635–657
- Tsushima M (1996) Economic structure and crime: the case of Japan. *J Socio-Econ* 25:497–515
- Vold GB (1958) *Theoretical Criminology*; Oxford University Press
- Wang D, Li F (2016) Daily activity space and exposure: a comparative study of Hong Kong’s public and private housing residents’ segregation in daily life. *Cities* 59:148–155
- Wenger MR (2019) Clarifying the relationship between racial diversity and crime: neighborhoods versus cities. *Crime Delinq* 65:1513–1536
- Wikström P-OH (2006) Individuals, settings, and acts of crime: situational mechanisms and the explanation of crime. *The explanation of crime: context, mechanisms and development*. Cambridge University Press. pp. 61–107
- Wilson WJ (2011) *When work disappears: the world of the new urban poor*. Vintage
- Wo J, Rogers E, Berg M, Koyle, C (2022) Recreating human mobility patterns through the lens of social media: using twitter to model the social ecology of crime. *Crime Delinquency* 1–28
- Xiao L, Liu L, Song G, Zhou S, Long D, Feng J (2017) Impacts of community environment on residential burglary based on rational choice theory. *Geogr Res* 36:2479–2491
- Xiao L, Ruiters S, Liu L, Song G, Zhou S (2021) Burglars blocked by barriers? The impact of physical and social barriers on residential burglars’ target location choices in China. *Comput Environ Urban Syst* 86:101582
- Xu C, Yang Y, Song G, Liu L, Lan M, Chen X (2022) The impact of civil registration-based demographic heterogeneity on community thefts. *Habitat Int* 129:102673
- Zhang Y, Chai Y (2011) The spatio-temporal activity pattern of the middle and the low-income residents in Beijing, China. *Sci Geogr Sin*, 31:1056–1064
- Zhou S, Deng L (2010) Spatio-temporal pattern of residents’ daily activities based on T-GIS: a case study in Guangzhou, China. *Acta Geogr Sin* 65:1454–1463

Acknowledgements

This research was funded by the National Natural Science Foundation of China (Grant Nos. 42171218, 42071184, 41771096).

Author contributions

All authors contributed significantly to an article and approved the submitted version. CX: conceptualization, formal analysis, and writing. ZH: methodology, software, and writing. GS: supervision, writing, reviewing, and editing. DC: software, model calculation.

Competing interests

The authors declare no competing interests.

Ethical approval

This article does not contain any studies with human participants performed by any of the authors.

Informed consent

This study contained data from an official statistics agency and involved no personal or identifiable information.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1057/s41599-024-02610-8>.

Correspondence and requests for materials should be addressed to Guangwen Song.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024