



Article

# Ambient Population and Larceny-Theft: A Spatial Analysis Using Mobile Phone Data

Li He <sup>1</sup>, Antonio Páez <sup>2,\*</sup>, Jianmin Jiao <sup>1,†</sup>, Ping An <sup>3</sup>, Chuntian Lu <sup>1</sup>, Wen Mao <sup>1</sup>  
and Dongping Long <sup>4</sup>

<sup>1</sup> School of Humanities and Social Science, Xi'an Jiaotong University, Xi'an 710049, China; lihegeo@xjtu.edu.cn (L.H.); jlb8582@xjtu.edu.cn (J.J.); luchuntian@xjtu.edu.cn (C.L.); maowen@stu.xjtu.edu.cn (W.M.)

<sup>2</sup> School of Geography and Earth Sciences, McMaster University, Hamilton L8S 4L8, Canada

<sup>3</sup> The School of Management, Xi'an Jiaotong University, Xi'an 710049, China; anping@stu.xjtu.edu.cn

<sup>4</sup> Center of GeoInformatics for Public Security, School of Geographic Sciences, Guangzhou University, Guangzhou 510006, China; longdp@gzhu.edu.cn

\* Correspondence: paezha@mcmaster.ca

† Xi'an Public Security Bureau, Xi'an 710002, China.

Received: 16 April 2020; Accepted: 20 May 2020; Published: 26 May 2020



**Abstract:** In the spatial analysis of crime, the residential population has been a conventional measure of the population at risk. Recent studies suggest that the ambient population is a useful alternative measure of the population at risk that can better capture the activity patterns of a population. However, current studies are limited by the availability of high precision demographic characteristics, such as social activities and the origins of residents. In this research, we use spatially referenced mobile phone data to measure the size and activity patterns of various types of ambient population, and further investigate the link between urban larceny-theft and population with multiple demographic and activity characteristics. A series of crime attractors, generators, and detractors are also considered in the analysis to account for the spatial variation of crime opportunities. The major findings based on a negative binomial model are three-fold. (1) The size of the non-local population and people's social regularity calculated from mobile phone big data significantly correlate with the spatial variation of larceny-theft. (2) Crime attractors, generators, and detractors, measured by five types of Points of Interest (POIs), significantly depict the criminality of places and impact opportunities for crime. (3) Higher levels of nighttime light are associated with increased levels of larceny-theft. The results have practical implications for linking the ambient population to crime, and the insights are informative for several theories of crime and crime prevention efforts.

**Keywords:** ambient population; larceny-theft; crime; mobile phone data; spatial analysis

## 1. Introduction

The residential population from census data is commonly used to measure the population at risk in the spatial analysis of crime, mostly due to the availability of data. Increasingly, however, it has become clear that the residential population fails to accurately capture the non-residential population and population patterns during the day [1]. It is well known that different types of crime have diverse types of population correlates [2], and many property crimes are sensitive to population types such as non-residential population. Of interest in this respect is larceny-theft, since such property crime is inherently a type of non-residential crime [3]. To better understand the spatial variation of such crime, it is important to accurately quantify the population at risk based on their demographic and routine activity characteristics. Risk, furthermore, is twofold: not only the population at risk of being victims,

but also the population who bring risk to others (the population at risk of offending), which is usually more difficult to quantify. Failing to identify such populations in the analysis of crime can result in misleading and biased results.

As an alternative to the census population, recent studies suggest the usefulness of the ambient population to quantify the population at risk [4,5]. A number of datasets have been used to measure the ambient population, including the Oak Ridge National Laboratory's LandScan Global Population Database, social media data, and spatially referenced mobile phone data [1,6,7]. These datasets have proved useful in quantifying the size, density, distribution, and mobility of the population at risk. The distribution, trips, and social activities of a population are embedded in the fabric of urban structure. Therefore, the interactions between urban layout and embedded human activity provide both static and dynamic risks and opportunities.

From a static perspective, the distribution of built environment and land use features reflects the way people use urban spaces [8]. As a consequence, city mosaics with various urban structures can shape people's daily activity patterns and further produce uneven crime opportunities and victimization risks through crime attractors, crime generators, and crime detractors [9]. Dynamically, the daily routines and social activities of the ambient population create an ebb and flow of intersections for motivated offenders, suitable targets, and a lack of capable guardianships [10], and routine activity theory indicates that crime usually occurs at those intersections. It is known that people with different demographic attributes exhibit divergent daily routines and social activities [11], and thus form different spatial and temporal intersections for crime opportunities and victimization risks in urban space. However, as Clarke [12] put it, the complexities of crime opportunity challenge the simple counting of available targets. Therefore, detailed demographic and social data of anonymous mobile phone users can more accurately measure the population at risk.

Considering this, in this paper, we intend to examine the link between the ambient population and larceny-theft. The research aims to identify the ambient population's demographic and social activity correlates impacting larceny-theft at the areal level. We also account for crime opportunities measured by crime attractors, generators, and detractors. The case study is larceny-theft in the city of Xi'an, Shaanxi province, China. The approach adopted is as follows. First, we test for global spatial autocorrelation in crime. Second, we collect spatially referenced mobile phone data with anonymous phone users' information, which allows us to extract more accurate demographic and social activity indices to measure the risk of offense at the areal level. Third, we collect Point of Interest (POI) data and categorize them into several types to quantify crime attractors, crime generators, and crime detractors. Fourth, we collect LuoJia 1-01 nighttime light data and develop an indicator to quantify regional socio-economic status. Fifth, to study the link between the ambient population and larceny-theft, a negative binomial regression model is estimated to explore the relationship between the count of larceny-theft and a series of socio-demographic variables of anonymous mobile phone users and crime opportunity variables. The results of this research are useful to (1) identify the static and dynamic larceny-theft correlates from the ambient population and urban structure, and (2) provide insights to inform crime reduction initiatives and the allocation of crime prevention resources.

## 2. Background

### 2.1. Ambient Population in Crime Analysis

The ambient population represents the diurnal movement of urban populations and is considered a better measure to understand urban crime [13]. The ambient population needs to include both the daytime and nighttime population, as people seldom limit their diurnal activities to the neighborhood of their place of residence. A challenge with the ambient population is that census data are not appropriate instruments for capturing activity patterns at higher temporal resolutions. Travel and activity surveys are more apt instruments, but they have more limited coverage and do not exist in many jurisdictions. To date, several kinds of datasets have been tested to generate ambient populations,

including mobile phone data [4,6], bus and metro smart card data [14], the LandScan Global Population Database (LSGPD) [1], social media location data [7,15], and footfall data [16]. Although they provide rich possibilities to generate an ambient population, they are not without limitations. For example, Malleson and Andresen [2] indicate that larger omissions and even errors are likely to be included in social media data. Li et al. [17] suggest that most smart card data only record a user's boarding location and time and lack a user's alighting information, which could lead to biased estimates of population.

Of the different sources of data, LSGPD data are widely used. The dataset provides an estimate of the 24 h average ambient population, with a fine spatial resolution of approximately 1km×1km [18]. Although the data are open and offer global coverage, a recent study by Deville et al. [6] suggests that the dataset remains largely constrained by population count data from censuses, although various sophisticated weighting or downscaling methods have been used to improve precision. In this case, the estimation of the global population may be biased due to the limitations embedded in census data, such as unreliability of estimates in less developed regions, the vague spatial resolution, and even the lack of contemporary data in resource-poor regions [6]. Moreover, research on crime hotspots has consistently suggested that crime tends to concentrate at micro places [19,20]. Thus, the spatial resolution of 1km×1km would tend to smooth the spatial heterogeneity at fine scales. More importantly, in the case of crime analysis at urban scale, the spatial resolution is not as precise as mobile phone data.

## 2.2. Mobile Phone Data for Crime Analysis

Understanding the dynamic distribution and movement of urban populations can help to improve our understanding of urban crime structure and to develop crime prevention initiatives. Due to the lack of robust systematic data collection to measure the rhythms of the population at appropriate temporal and spatial scales, the relationship between larceny-theft and the ambient population remains vague.

Data generated from communication tools such as mobile phones now provide valuable opportunities to explore the spatial patterns of human mobility and social behavior [21–23]. Mobile phone data have a high penetration rate, wide spatial coverage, and temporal continuity [24]. Studies have consistently demonstrated their usefulness in quantifying the ambient population with both high spatial and temporal resolutions [6]. For example, empirical studies using spatially referenced mobile phone data found that human mobility in urban space shows spatial and temporal regularity [25,26]. Such regularity also seems to apply to offenders, as shown in the research of Griffiths et al. [27]. Mobile phone data were recently used by Song et al. [26] to measure the impact of the ambient population's daily mobility flow on thieves' target location selection.

In light of this, the literature suggests that mobile phone data can more accurately portray the spatial and temporal dimensions of the city, and are more appropriate for crime analysis [4,28]. Ambient populations generated from mobile phone data have been characterized as “cellular census” data, and, when available, can replace conventional residential census data in the analysis of crime [29]. A limitation of most of the mobile phone data is that rich demographic and social attributes are removed to ensure privacy. Due to the general unavailability of a mobile phone user's attribute data, current studies mainly focus on quantifying the size, density, and spatial movement of the ambient population [30]. Rarely do studies examine the impact of the multidimensional demographic and social characteristics of the ambient population on crime. This is an understandable gap since, as noted by Malleson and Andresen [3], even simple measures of the ambient population such as population counts are difficult to obtain, as is other rich information like demographic and social activity attributes.

## 2.3. Routine Activity Theory

Our research builds on two pillar theories of environmental criminology: routine activity theory [10] and crime pattern theory [9]. These two theories are mutually supportive in various circumstances, and a combination of the two is needed to understand criminal events in many contexts [31].

Routine activity theory argues that crimes occur at the intersection of motivated offenders, suitable targets, and the absence of capable guardians [10]. It focuses on the role of “places” in inhibiting or encouraging crimes [32]. Sherman et al. [33] demonstrated the most important contribution of routine activity theory, that is, criminal activities are influenced not only by the numbers of offenders and targets, and the absence of guardians, but most importantly, by the “factors” influencing their convergence over space and time. Brantingham and Brantingham [9] suggested that these “factors” are places where people travel to and from routinely, such as work, school, entertainment districts, or shopping areas; therefore, they have the potential for becoming high crime places. Based on this argument, a large body of research has studied the relationship between criminal opportunity and specific features of urban design and urban architecture, and micro-scale features of the physical environment [31,34,35].

Recent studies suggest that elements of routine activity theory can also be measured by areal level variables, such as neighborhood and block group variables [36–38]. For example, Andresen [36] used census tract level variables to measure the average levels of motivated offenders, suitable targets, and absence of guardians in Vancouver. He et al. [31] used block group level variables to quantify the three elements of routine activity theory and found that they can significantly capture the spatial variation of crime in Columbus, Ohio. Hanaoka [4] also pointed out that the spatial and temporal characteristics of a population’s routine activities at the neighborhood level can impact an offender’s spatial decision making.

#### 2.4. Crime Pattern Theory

Crime pattern theory explains the spatial pattern of criminal events by combining routine activity theory and rational choice theory [32]. More specifically, it posits that opportunities for crime can be found at particular places with specific functions [39]. To measure opportunities at such places, Brantingham and Brantingham [9] suggested the use of crime attractors, crime generators, and crime detractors. Criminal activities are encouraged by crime attractors and generators but inhibited by crime detractors. The generators, attractors, and detractors are usually different types of land use or facility, which can significantly affect people’s routine activities and social interactions.

To be specific, crime generators are places that are highly accessible to the public and thus can concentrate specific types of population for non-criminal activities, and the concentration has the potential to transform people into victims of crime. Therefore, a crime generator has an indirect impact on criminal activities, and potential offenders can be both local insiders and outsiders. These places include commercial buildings, vacant buildings, sports facilities, public transport nodes, and cycle parking and renting facilities [9,40]. On the other hand, crime attractors are places that directly provide certain opportunities for deviant behaviors and crime. They may not attract a large number of people, but they can function well as suitable places where motivated offenders can easily hunt targets without capable guardianship [39]. These places include bars, entertainment areas, prostitution areas, drug markets, large shopping malls with poor security arrangements, and insecure parking lots [9,41]. Crime detractors are places that can deter potential offenders and thus inhibit criminal behaviors. These places are usually monitored by security guards or CCTV and less accessible to the public, such as places of worship, cemeteries, police stations, industrial plants, green areas, and universities [8,41].

A large body of research has emerged to empirically test crime pattern theory. For example, Roncek and Maier [42] found a positive relationship between bars (alcohol consumption) and levels of violent crime. Stucky and Ottensmann [43] found that commercial activity and high-density residential land uses were associated with higher levels of crime, whereas cemeteries and industry were associated with lower counts for some crimes. Browning et al. [44] found a non-linear association between commercial and residential density and crime in Columbus, Ohio. Song et al. [45] found a significant link between theft from the person in a large Chinese city and crime opportunity indices such as transit stations, catering services, net bars, and convenience stores. Feng et al. [46] also showed that size of bars, hotels, catering services, and parks is positively related to the level of crime. Importantly, the effects of certain facilities and land uses on crime appear to be conditioned by the socio-economic

characteristics of local neighborhoods [43]. The effect of certain land use types on crime remain the subject of debate. For example, although schools are identified as crime generators in studies by Bernasco and Block [39] and Kinney et al. [8], others such as Sypion-Dutkowska and Leitner [41] found that school buildings can deter crime due to that fact that they are under surveillance by security guards or receptionists. Such issues need further investigation in alternative urban contexts.

As stated by Brantingham and Brantingham [9], it is important to note that crime generators, attractors, and detractors can be measured at various scales, such as place, area, neighborhood, and even district. Similar to routine activity elements using areal level variables, recent empirical studies suggest the use of areal level indices to quantify neighborhood attractors, generators, and detractors [39,46]. A number of datasets have been used to capture these factors, and the majority are land use data [8,41]. More recently, Points of Interest (POI) data from open street map and other data providers have been widely used as a proxy of land-use data [3,15,26]. The rapid update rate and detailed spatial and attribute information has made POI data a popular alternative in the environmental criminology literature [26,46].

### 3. Study Area, Data, and Methods

#### 3.1. Study Area

The study area is the city of Xi'an, Shaanxi province, China. With a total population of 12,005,600 in 2018, Xi'an is the largest city and capital of Shaanxi Province, and one of the 13 emerging megacities in China. The prefecture spans 107°39' to 109°49'E longitude and 33°39' to 34°45' N latitude, with a total area of 10,752 km<sup>2</sup> and a built-up urban area of 700 km<sup>2</sup>. Xi'an has direct jurisdiction over 11 districts and two counties. The GDP of Xi'an was 932.12 billion Yuan (approximately 134.47 billion USD) in 2019, ranking among the top 20 cities in China, with a recent annual growth rate of 10 percent. The location and a digital elevation model of Xi'an are shown in Figure 1.

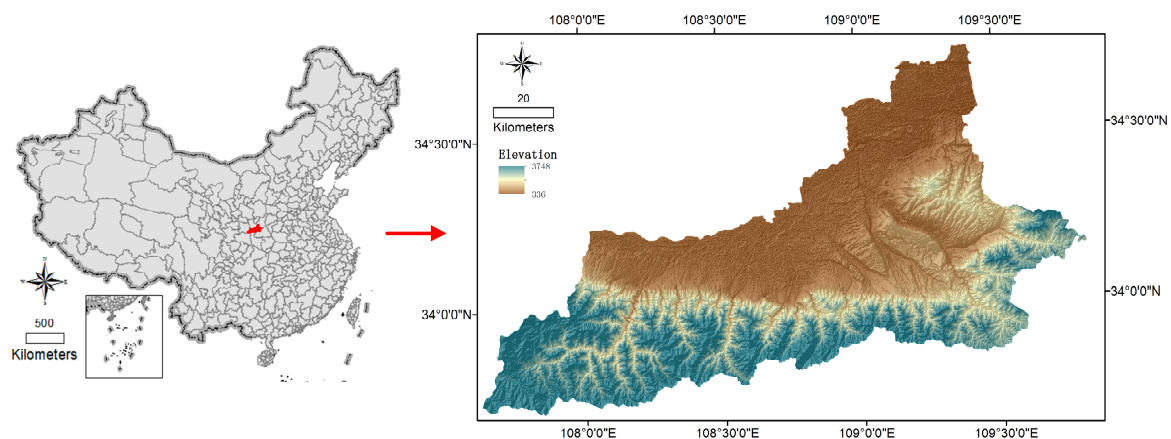


Figure 1. Location of Xi'an in China.

#### 3.2. Data

##### 3.2.1. Crime Data

The crime data used in this analysis are larceny-theft data for Xi'an over a one-year period, from 1 November 2018 to 1 November 2019. The larceny-theft data come in a pre-packaged size, and due to the unavailability of point level crime data, we collect areal level crime statistics from the Xian Public Security Bureau. This dataset includes 52,874 observations in total. The data fields include the crime type, date the crime occurred, and law enforcement agencies, among others.

The areal unit in this analysis is the Paichusuo territory (PCS in short). A PCS is a police sub-station, similar to a police precinct. There are 187 PCSs in Xi'an, with an average area of 54.34 km<sup>2</sup>. Larceny-theft



was defined by the FBI as the unlawful taking, carrying, leading, or riding away of property from the possession of another in which no use of force or fraud occurs. In our data, it mainly includes crimes such as pocket-picking and purse-snatching (78.23%), the theft of electric bicycles (13.88%), shoplifting (4.82%), bicycle thefts (1.12%), thefts of motorcycles (0.59%), thefts of goods (0.20%), and thefts of livestock (0.09%). Of larceny-theft arrests, 47.29% are for low value crimes (1000–3000 RMB); 25.22% of arrests are for amounts between 5000 and 10000 RMB.

Steffensmeier et al. [47] found comparable gender gaps (approximately 35%) for larceny arrestees in the U.S. Unlike in the U.S., in our study area, 94% of persons arrested for larceny-theft are men. Some female arrestees were pregnant and had multiple convictions for larceny-theft. They were making full use of the criminal procedure law, that is, pregnant female offenders can apply for bail pending trial and can avoid jail sentences. The city of Xi'an is the domicile place for 36.7% of the arrestees, 25.1% have their domicile in other cities in Shaanxi province, 21.1% are from Xinjiang province, and the remaining 18.1% are from other provinces of China. The arrestees have distinct age characteristics; 60% are between the ages of 36 and 69, 25.2% are between the ages 26 and 35, and 16.7% are between the ages of 18 and 25. It is worth noting that 90.4% of criminals have previous convictions, 63.4% of criminals are from the floating population, 89.4% are unemployed, and 77.4% have only junior high school and primary school education. In terms of victims, 61.2% are women, 32.7% are the elderly, and 6.1% are young and middle-aged men.

Although crime data availability determines the areal unit of our analysis, it is still important to consider the Zone MAUP issue, as different spatial arrangements may produce different results. An alternative spatial unit is the sub-district, which is administered by the sub-district office. A sub-district is one of the smaller political divisions in China, and there are 117 sub-districts in the city of Xi'an. However, sub-district offices are mainly responsible for community services and economic development within the jurisdiction. PCSs, in contrast, are the basic organizational units in Chinese police organization and fully responsible for local crime control and public security defense. Therefore, crime prevention strategies and police resources differ from PCS to PCS; as a result, crime levels and characteristics show more differences between PCSs, rather than between sub-districts. As argued by Cabrera-Barona et al. [48], to address MAUP, spatial units should be able to maximize homogeneity within each unit and heterogeneity between them. For this reason, the PCS is one of the most frequently used geographic units in spatial crime analysis in the background of China's cities, e.g., Chen et al. [49], Chen et al. [50], and Song et al. [45].

### 3.2.2. Spatially Referenced Mobile Phone Big Data

To measure the ambient population, we use spatially referenced mobile phone data, together with anonymous users' demographic and social activity data. These data are collected from a research institute under China's Ministry of Posts and Telecommunications, and cover all three state-run telecommunication operators in China (i.e., China Telecom, China Unicom, and China Mobile). Therefore, the dataset has full coverage of mobile phone users in the study area. The dataset shows that the average number of daily records of mobile phone data in Xi'an is at least 6.9 billion. On average, each mobile phone user has 350 records of location every day.

The anonymized and aggregated mobile phone user's information and activity include user's historical locations, time stamps of locations, native place, gender, date of birth, and call logs. Therefore, the dataset can be used to generate ambient population variables such as the size of the population in a specific zone, the size of the local and non-local (i.e., non-Xi'an) populations, people's social regularity (SR) in a specific PCS zone, and the diversity index of people's native place (DINP). These measures are defined next.

Social regularity (SR) is the number of phone users who exhibit regular social activities in terms of phone calls and SMS. It is calculated using the number of calls (CT) and users (CU) of phone calls "made" and "received" by each mobile phone in each day, and the number of times (ST) and users (SU) of SMS "received" and "delivered" by each mobile phone in each day. We identify outliers

using the “1.5 IQR rule” (IQR refers to the inter-quartile range,  $IQR = Q3$  (the third quartile) –  $Q1$  (the first quartile); outliers are defined as observations that fall below  $Q1 - 1.5 IQR$  or above  $Q3 + 1.5 IQR$ ). First, we identify the presence of “outliers” from CT and ST, which we label *OutlierTime*. We then find “outliers” from CU and SU, which we label *OutlierUser*. To identify irregular users, we calculate the intersection of the two sets of outliers, i.e.,  $OutlierTime \cap OutlierUser$ . Subtracting from the total number of cell phone users in a region, we finally have SR. A lower value of SR indicates a larger population that displays irregular social behavior in an area, which we hypothesize can lead to increased deviant behavior.

The DINP is defined as a Modified Simpson’s Diversity Index [51]:

$$DINP = -\ln\left(\sum_{i=1}^s \left(\frac{n_i}{N}\right)^2\right) \quad (1)$$

where  $n_i$  denotes the population from the  $i$ -th city in China,  $N$  is the sum of  $n_i$ , and  $S$  denotes the total number of origin cities in a PCS zone. The minimum value of the index is zero, which indicates maximum homogeneity of the population’s origin city. As the DINP increases, the heterogeneity of the people’s origin city increases.

Studies have proved the usefulness of using either short-term mobile phone data such as a week’s or long-term data such as a year’s in quantifying the ambient population [39,46]. Given the high computational cost of processing vast amounts of data, we employ four months’ data in total, namely that of January, April, July, and October in 2019. The dataset contains hourly counts of mobile phone activity for each day of the four months per 306m×306m grid cell. We calculate the four-month average to represent the average ambient population over a period of one year. The grid cells are then aggregated at the PCS level. This process helps with the temporal consistency between crime data and mobile phone data, and further considers the potential seasonal variation of the urban population.

### 3.2.3. Point of Interest (POI) Data

Point of interest (POI) data give additional information about urban space. These data contain information about local amenities, such as the name or description for the POI, longitude, latitude, address, and city. Importantly, POI data describe uneven opportunities for crime and the situational nature of crime from the perspective of the built environment.

Informed by the theoretical perspectives discussed previously, we collect various types of POI in 2019 from AutoNavi Holdings Ltd. (Alibaba’s mapping unit), also known as Gaode Map (amap.com). Following previous studies [41,45], we also calculate crime opportunity variables using internet bars, billiard rooms, bus stops, metro stations, bars, card rooms, bath centers, karaoke bars, convenience stores, supermarkets, shopping malls, restaurants, industrial plants, and public security organs, including police stations, procuratorate, and courts. All variables are aggregated to the level of the PCS for spatial analysis. We assume that the urban fabric depicted by the selected POIs indicates the spatio-temporal variation of crime opportunities, which further impacts potential offenders’ decision-making on crime location selection and potential targets’ victimization risks.

### 3.2.4. LuoJia 1-01 Nighttime Light Imaging

Nighttime lights (NTL in short), calculated from weather satellite recordings, are increasingly used as a proxy for social and economic activities, such as GDP, population size, and urban land use and expansion [52–54]. A recent study by Zhou et al. [55], for example, illustrates the usefulness of NTL in quantifying “urban edge areas”, regions that significantly affect the spatial patterns of crime. In our study, NTL is used as a proxy for regional socio-economic status. A higher mean value of NTL in a PCS zone indicates a larger population size and more economic activity.

The LuoJia 1-01 satellite, launched on 2 June 2018, represents a new generation of nighttime light (NTL) data source. It can cover global images within 15 days and has been widely used to evaluate

socio-economic indices, urban extent, human activity, and geographical characteristics at national and regional scales. Compared with conventional NTL data collectors such as DMSP-OLS and NPP-VIIRS, LuoJia 1-01 has on-board radiance calibration to avoid the problem of saturation, which provides NTL images at a finer spatial resolution of 130 m with 14-bit high quantization. We collected the NTL image of our study area from the Hubei Data and Application Network for High Resolution Earth Observation System ([http://59.175.109.173:8888/app/login\\_en.html](http://59.175.109.173:8888/app/login_en.html), accessed on 10 February 2020). The date of our image is 3 March 2019.

The original absolute radiance of LuoJia 1-01 is floating-point data. To store the data conveniently, they were amplified, stretched, and stored in INT32 format. The standard image needs a radiometric calibration, so that the digital numbers (DN) value of the image can be converted to a radiance value, and the lighting brightness and discrepancy can be analyzed accordingly. The exponential equation of the radiance conversion formula is as follows:

$$L = DN^{\frac{3}{2}} \cdot 10^{-10} \quad (2)$$

where  $L$  denotes the radiance value after absolute radiation conversion, the unit of  $L$  is  $W/(m^2 \cdot sr \cdot \mu m)$ , DN denotes the digital number gray value of the LuoJia image's pixel, and  $W$  is the bandwidth. The LuoJia image's radiometric range is between the value of 460 and 980 nm; therefore,  $W$  equals  $5.2 \times 10^{-7}$  m.

### 3.3. Methods

#### 3.3.1. Exploratory Spatial Data Analysis

The first method is Exploratory Spatial Data Analysis (ESDA). More specifically, Global Moran's  $I$  is used to test for the spatial autocorrelation of crime at the PCS level. Social science variables are usually positively spatially autocorrelated due to the way phenomena are geographically organized [56]. Therefore, it is necessary to explore the presence of spatial autocorrelation among our variables first. The global Moran's  $I$  statistic introduced by Moran [57] is the most popular measure of spatial autocorrelation, and it takes the form as follows [58]:

$$I = \frac{n}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij} (X_i - \bar{X})(X_j - \bar{X})}{\sum_i (X_i - \bar{X})^2} \quad (3)$$

where  $n$  is the total number of spatial features (PCS zones in this study),  $X_i$  and  $X_j$  are values for feature  $i$  and  $j$ ,  $\bar{X}$  is the mean of  $X$ ,  $(X_i - \bar{X})$  indicates the deviation of  $i$ 's feature value  $X_i$  from its mean  $\bar{X}$ , and  $w_{ij}$  are the elements of the spatial weight matrix, indicating the spatial weight between feature  $i$  and  $j$ .

Moran's  $I$  represents the linear association between a PCS's crime count and the spatially weighted average of the neighboring PCS's crime count. The expected value of Moran's  $I$  statistic is  $-1/(n-1)$ ; it tends to zero as the sample size increases. A value larger than  $-1/(n-1)$  indicates positive spatial autocorrelation, that less than  $-1/(n-1)$  indicates negative, and that approaching  $-1/(n-1)$  suggests no spatial autocorrelation [58]. Hypothesis testing against the null hypothesis of spatial independence can be evaluated by the  $z$ -score and  $p$ -value. The  $z$ -score values associated with a 95 percent confidence level are  $-1.96$  and  $+1.96$ , indicating that if the computed  $z$ -score falls between  $-1.96$  and  $+1.96$ , the  $p$ -value is larger than  $0.05$ , and the spatial autocorrelation is not statistically significant.

#### 3.3.2. Negative Binomial Regression

The dependent variable, the count of larceny-theft, is non-negative and right-skewed, and the variance is greater than the mean. Therefore, to accommodate the potential over-dispersion, we employ



a negative binomial regression model [59]. The correlates used for regression analysis are all selected based on routine activity theory and crime pattern theory. The regression model is defined as:

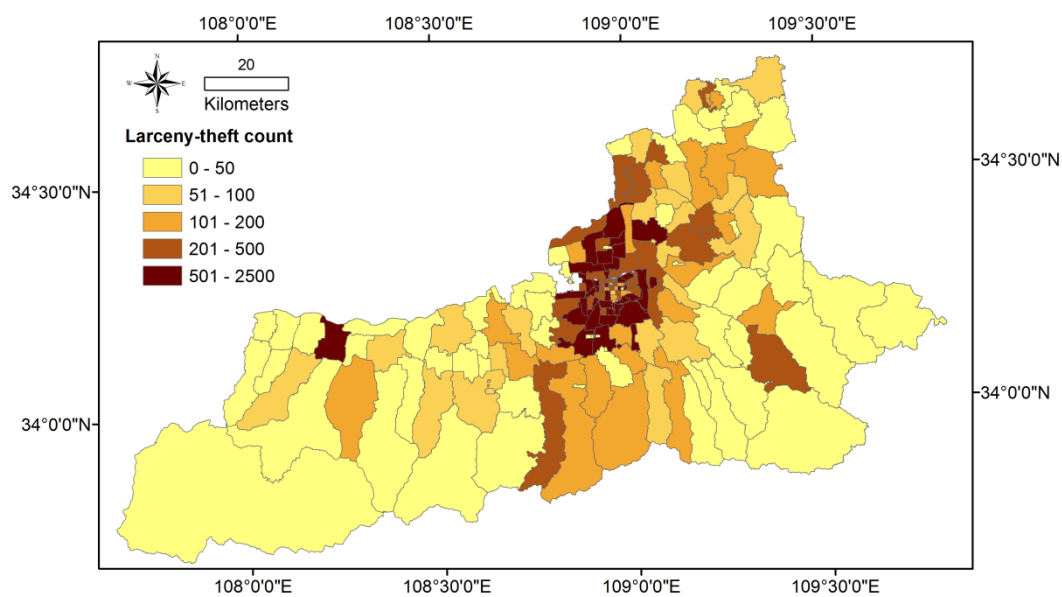
$$\ln(\mu_i) = \sum_{j=0}^j \beta_j X_{ij} + \varepsilon_i \quad (4)$$

where  $X$  denotes a matrix of ambient population variables and crime opportunity variables, a vector of coefficients  $\beta$  is to be estimated,  $\varepsilon$  denotes the random error term, and  $\ln(\mu_i)$  is the natural logarithm of the expected count of larceny-theft  $\mu_i$  in the  $i$ -th PCS.

## 4. Analysis and Results

### 4.1. Exploratory Spatial Data Analysis

To assess the spatial pattern of larceny-theft, we first map the count of crime as shown in Figure 2. PCSs in the downtown area (i.e., the central area of the map) generally experience very high volume of larceny-theft, particularly the northern and southern PCSs in town area, while crime occurs less frequently in the suburban areas. It should be note that in addition to the high volume of crime downtown, two PCSs in the town areas of two counties (one in the west and the other in the southeast of the city) experience more crime, while surrounding PCSs have considerably lower counts of crime. To further explore the spatial dependency in crime data, we calculate Moran's I Index and z-scores using GeoDa 1.12.1.161 [60]. This is also done by using Queen's 1st contiguity weight, which is also created using GeoDa. The result indicates a significant positive autocorrelation ( $I = 0.384$ , pseudo  $p$ -value  $\leq 0.001$ ) and the concentration of larceny-theft in Xi'an in a non-random pattern.



**Figure 2.** Spatial distribution of larceny-theft in Xi'an, China.

### 4.2. Regression Analysis

After an extensive review of the relevant literature [3,8,15,26,39,41], we use the larceny-theft count as dependent variable, and the spatial unit is the PCS. Socio-economic status is controlled using the radiance mean value of NTL. The ambient population variables at the PCS level include the percentage of the non-local population, percentage of the population with regular social activity (SR), and diversity index of the population's native place (DINP).

We use a series of PCS-level crime opportunity variables to capture crime attractors, crime generators, and crime detractors. More specifically, crime attractors are measured as the

number of internet bars and billiards rooms; and the number of bars, card rooms, bath centers, and KTV rooms. Crime generators are measured through the number of bus stops and metro stations; number of convenience stores, supermarkets, and shopping malls; and number of restaurants. We use the number of industrial plant and public security organs to capture crime detractors. The distribution of these variables is mostly long tailed, and as a matter of practice, some variables were log-transformed before correlation analysis and parameter estimation. The model is estimated using Stata 14 (StataCorp, College Station, TX). Table 1 provides descriptive statistics of dependent and all explanatory variables. Table 2 provides correlations of the dependent variables and independent variables. The collinearity test for all variables does not identify pair-wise correlation in excess of 0.8 or  $-0.8$ , a common threshold in general practice [31].

**Table 1.** Descriptive statistics for variables used in the analysis (N = 187 observations). NTL, nighttime lights.

Code	Variables	Mean	Std.	Min	Max
<i>Dependent variable</i>					
Y	# of larceny-theft	279.71	353.53	0.00	2491.00
<i>Ambient population variables</i>					
X <sub>1</sub>	% of non-local population (ln)	−0.61	0.19	−1.06	−0.27
X <sub>2</sub>	% of population with regular social activity (SR) (ln)	−0.02	0.01	−0.03	0.00
X <sub>3</sub>	Diversity index of population's native place (DINP)	2.79	0.47	0.79	3.67
<i>Crime attractors</i>					
X <sub>4</sub>	# of internet bars, billiards rooms (ln)	−0.68	4.92	−9.21	4.57
X <sub>5</sub>	# of bars, card rooms, bath centers, KTV rooms (ln)	−0.71	5.07	−9.21	4.49
<i>Crime generators</i>					
X <sub>6</sub>	# of bus stops, metro stations	40.71	38.74	0.00	239.00
X <sub>7</sub>	# of convenience stores, supermarkets, shopping malls (ln)	1.61	2.76	−9.21	4.33
X <sub>8</sub>	# of restaurants (ln)	4.65	1.91	−9.21	7.56
<i>Crime detractors</i>					
X <sub>9</sub>	# of industrial plants, public security organs (ln)	3.49	1.54	−9.21	6.00
<i>Socio-economic status</i>					
X <sub>10</sub>	Radiance mean value of NTL (ln)	−24.24	3.68	−39.14	−18.33

**Table 2.** Linear correlations of dependent variables and independent variables.

Y	1										
X <sub>1</sub>	0.427	1									
X <sub>2</sub>	−0.406	−0.393	1								
X <sub>3</sub>	0.181	−0.281	−0.325	1							
X <sub>4</sub>	0.488	0.398	−0.397	0.172	1						
X <sub>5</sub>	0.513	0.374	−0.511	0.292	0.615	1					
X <sub>6</sub>	0.487	0.162	−0.323	0.158	0.411	0.405	1				
X <sub>7</sub>	0.377	0.162	−0.290	0.019	0.495	0.471	0.398	1			
X <sub>8</sub>	0.600	0.411	−0.509	0.197	0.734	0.684	0.493	0.596	1		
X <sub>9</sub>	0.493	0.385	−0.386	0.137	0.592	0.646	0.510	0.551	0.682	1	
X <sub>10</sub>	0.493	0.608	−0.579	0.214	0.581	0.599	0.225	0.263	0.632	0.443	1
Y	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>	X <sub>8</sub>	X <sub>9</sub>	X <sub>10</sub>	

We use the “nlcor” package in R [61] to explore the potential non-linear correlation between variables. The results of the correlation test are shown in the Supplementary Information. The “nlcor” package works by adaptively identifying multiple local regions of linear correlations to estimate the overall nonlinear correlation, and returns a non-linear correlation estimate, adjusted *p*-value, and a plot visualizing the non-linear relationships. The correlation estimate is between 0 and 1. The higher the value, the more the nonlinear correlation is. All correlation coefficients are significant at least at the  $p < 0.05$  level. The results indicate that four out of the 10 independent variables (the non-local

population, SR, internet bars and billiard rooms, convenience stores/supermarkets/shopping malls) have piecewise linear correlations with larceny-theft. Particularly, the correlation between crime and  $X_1$  changes its direction after a point; the same applies to the correlation between crime and  $X_2$ . Non-uniform piecewise linear correlations also exist between  $X_1$  and  $X_2$ , and between  $X_3$  and  $X_4$ – $X_{10}$ .

A positive relationship with larceny-theft is expected for the non-local population, diversity index of the population's native places, crime attractor variables, crime generator variables, and mean of NTL. A negative relationship, on the other hand, is expected for the regularity of social activity and the crime detractor variable.

The results of the model estimation are shown in Table 3. A backward stepwise approach was used to specify the model, retaining all variables that were statistically significant at least at the conventional 0.05 level of significance. The goodness of fit of the negative binomial model is summarized by a Pseudo  $R^2$  of 0.071. The estimate of the dispersion parameter is given as  $\alpha$ . The likelihood-ratio test for  $\alpha$  is significant (the chi-square value is 19,000 with one degree of freedom), and we can reject the null hypothesis of no overdispersion. A non-significant value of Moran's  $I$  in the model residuals (0.037) indicates that there is no significant spatial autocorrelation in the model residuals, which allays concerns about misspecification.

**Table 3.** Parameter estimates from the negative binomial model (N = 187 observations).

Variables	Coefficient	p-Value
Constant	6.390	0.000
<i>Ambient population variables</i>		
% of non-local population (ln <sup>1</sup> )	0.929	0.047
% of population with regular social activity (SR) (ln)	−30.238	0.027
Diversity index of population's native place (DINP)	-	-
<i>Crime attractors</i>		
# of internet bars, billiards rooms (ln)	0.075	0.001
# of bars, card rooms, bath centers, KTV rooms (ln)	0.086	0.000
<i>Crime generators</i>		
# of bus stops, metro stations	0.005	0.016
# of convenience stores, supermarkets, shopping malls (ln)	0.132	0.000
# of restaurants (ln)	-	-
<i>Crime detractors</i>		
# of industrial plants, public security organs (ln)	−0.126	0.035
<i>Socio-economic status</i>		
Radiance mean value of NTL (ln)	0.046	0.047
<i>ln</i> $\alpha$	−0.224	
$\alpha$	0.799	
Pseudo $R^2$	0.071	0.000
AIC	2278.504	
BIC	2310.815	
Moran's $I$ in residuals <sup>2</sup>	0.037	0.190

<sup>1</sup> The ln in parentheses denotes natural log transformation. <sup>2</sup> Moran's  $I$  coefficients and z-scores are calculated using GeoDa 1.12.1.161; the calculation is performed by using Queen's 1st contiguity weight, which is also created using GeoDa 1.12.1.161.

The model consists of eight significant variables, all with signs as expected. The results generally indicate that the ambient population variables and crime opportunity variables are significant correlates of crime concentration at the PCS level. Concretely, two out of the three ambient population variables are significant; the percentage of the non-local population is significant and positive, while the percentage of population with regular social activity (SR) is significantly negative. This indicates that PCSs with higher non-local population rates tend to experience more larceny-theft, while PCSs with larger populations displaying regular social activities tend to have lower larceny-theft. The number of internet bars and billiards rooms has a significantly positive relationship with larceny-theft, and the

number of bars, card rooms, bath centers, and KTV rooms is also significantly positive. This suggests that PCSs with higher numbers of entertainment venues tend to see higher levels of larceny-theft. The number of bus stops, metro stations, convenience stores, supermarkets, and shopping malls is found to be significantly positive, which implies that PCSs with more said crime generators tend to have increased occurrences of criminal activities. By contrast, the number of industrial plants and public security organs has a significantly negative relationship with crime, which indicates their role as larceny-theft detractors.

## 5. Discussion

A synthesis of the results of the spatial patterns of larceny-theft and negative binomial regression leads to several relevant conclusions.

The higher levels of larceny-theft in the downtown areas of both the city (i.e., the central area of the map) and counties but lower levels in the suburban areas is consistent with previous studies [62,63]. Glaeser and Sacerdote [62] argued that town areas experience higher levels of crime due to higher pecuniary benefits and lower arrest probabilities. Ladbroke [64] stated that population density, a higher level of migration and population growth, and a larger young population are dominating factors influencing the high volume of crime in town areas.

Of importance is that two ambient population variables (the non-local population and social regularity) are significant. These variables indicate that it is not just population density, but the non-local population too, that influences larceny-theft. This result can be understood from the perspective of social disorganization [65]. A large number of empirical studies suggested that increased level of the non-local population in a neighborhood can lead to loss of informal social control [37,38,66], challenges to building collective efficacy [67], and an inability of local communities to attain common values [68].

The percentage of the population with regular social activity (SR) is a novel way to quantify the ambient population, thanks to new sources of data. We introduce this into our analysis since people's mobile social activity (calls and SMSs) is closely related to their routine activity in physical space. Furthermore, recent research has confirmed the validity of online routine activity in affecting people's potential of becoming a victim or offender in physical society [69–71]. This research indicates that a lower level of SR in a PCS zone indicates a higher population displaying irregular social activities, which could potentially lead to an increased potential of being victimized or deviant behaviors. It is worth noting that people's social regularity is negatively related to the percentage of the non-local population with a correlation of  $-0.406$  (Table 2), indicating that neighborhoods with more non-local population tend to have more mobile, socially irregular individuals.

The other key findings of this analysis are as follows. In agreement with previous studies, all significant crime attractor, generator, and detractor variables have the expected signs. To be specific, first, we find that entertainment venues can raise regional crime opportunities, which is in line with previous studies [8,45]. Internet bars and billiards rooms can be categorized as crime attractors, because of their potential in attracting idle personnel. In China, these establishments often serve as hotbeds for several types of deviant behavior such as pocket-picking, gang crime, assault, and drug-related crime. This is because, on the one hand, as Bax [72] argued, people attracted by these places may suffer from internet addiction, internet gaming disorders, and several other disorders. On the other hand, the significant financial costs in such places tends to push people to pursue quick monetary gains from illegal means. In areas with rich crime attractors and generators, it would be important to optimize land uses and modify the built environment to attract more eyes on the street and reduce opportunities to commit crimes. More mixed uses can attract more foot traffic to neighborhoods, and the increased street activity could further bring more "eyes on the street" and promote informal social control. Extensive literature has found that mixed commercial and residential land use areas are associated with lower levels of crime [73]. In this process, the characteristics of the ambient population need to be taken into account, as some found that the crime reducing effect of mixed land use usually depends on socio-demographic characteristics [74].

Second, the positive impact of bus stops and metro stations on larceny-theft confirms the crime-generating effect of public transport nodes, as supported by research on crime pattern theory. For one thing, according to routine activity theory, a large number of thefts of mobile phones, bags, and other personal belongings take place in transport stations because they are able to attract a high volume population lacking awareness of crime prevention [32]. On the other hand, criminals prefer to perpetrate theft in neighborhoods close to transit nodes, obviously because such neighborhoods are easy to enter and leave through public transport. The attracting effect of transport nodes on theft can reach up to 200 m [41].

Third, the negative sign of the crime detractor validates the crime-inhibiting effect of industrial plants and public security organs. These places are monitored by security guards or CCTV and are not generally used by the public. The guarded building's effect on crime has a spillover effect of around 50 m [41]. Empirical studies have shown that increased guarded buildings can help to reduce crime by up to seven percent in developing neighborhoods [14].

The positive effect of mean NTL on crime can be understood from the social disorganization and environmental criminology perspectives. NTL has consistently been proved to be an effective proxy of regional socio-economic status. A higher mean value of NTL in an area indicates a higher volume of population and level of economic activity. The literature consistently shows a relationship between the volume of property crime and population size [62,63]. Another reasonable guess regarding NTL is that NTL reflects street lighting conditions. On the one hand, Clark [75] argue that there is no reliable evidence to support the crime-detering effect of outdoor lighting. Night lights can be thought of not as inducers of crime, but places where people conduct their non-diurnal activities and whose presence makes them targets of potential victims of crime [76].

## 6. Conclusions

This study investigated the effect of ambient population measures and crime opportunity indices on the spatial pattern of larceny-theft in Xi'an. Our findings provide insights regarding quantifying the ambient population's social activities from mobile phone data. The results contribute to the literature in several ways.

First, our findings confirm the usefulness of mobile phone big data in providing reliable estimates of several important ambient population measures, such as the non-local population and the population's social regularity. Due to the general unavailability of high-precision attribute information and mobile social activity data of phone users, previous studies mainly focus on measuring the size, density, and spatial movement of the ambient population. Our study reveals a dimension of the social characteristics of the ambient population and enriches our understanding of the link between aggregated social behavior and the level of larceny-theft. Second, our findings shed light on the link between crime opportunity and larceny-theft from the perspective of crime pattern theory. Local geographies captured by crime attractors, generators, and detractors are critical in depicting the spatial variation of crime opportunities. In other words, we conclude that anonymous mobile phone user's aggregated demographic information and social activities, together with urban fabric proxies, can profoundly reveal extensive insights into the spatial heterogeneity of crime opportunities and victimization risks at the areal level.

It bears noting that the findings in this study are limited to larceny-theft at the PCS level. Hence, several avenues for future research can be discerned. First, the unavailability of point-level crime data limited our understanding of micro-level (i.e., street, community) built environmental correlates of larceny-theft in Xi'an, also impeding the mining of high precision ambient population characteristics as well as their link to larceny-theft. It is worth trying to resort to crowd sourcing to collect point level crime data in the future, which can also benefit police departments by uncovering potential unreported cases. In addition, the modifiable areal unit problem (MAUP) is an inevitable issue, which requires further investigation in future research. Given that point level crime data will be accessible in the foreseeable future, future research could replicate the present analyses using finer-scale spatial units, for



example, the 306m×306m grid cell, as fine as the grid size of aggregated mobile phone data. It could also further explore the existence of stability and radical change in the relationships between larceny-theft aggregated in various spatial units and a number of crime opportunity and socio-economic correlates. Second, further quantitative analysis is needed to explore the causal mechanisms between people's rich demographic attributes/mobile social activity (calls and SMSs) and larceny-theft. Rather than the PCS areal level used in this study, the causal analysis is expected to be performed at a micro or individual level.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2220-9964/9/6/342/s1>, Figure S1. The results of non-linear correlation test using “nlcor” package.

**Author Contributions:** Conceptualization, Li He and Antonio Páez; Data curation, Li He, Antonio Páez, Jianmin Jiao, Ping An, Chuntian Lu, Wen Mao and Dongping Long; Formal analysis, Li He; Funding acquisition, Li He; Investigation, Li He and Chuntian Lu; Methodology, Li He, Jianmin Jiao, Ping An and Chuntian Lu; Project administration, Li He; Resources, Jianmin Jiao and Ping An; Software, Wen Mao; Supervision, Li He; Validation, Antonio Páez, Jianmin Jiao and Dongping Long; Writing—original draft, Li He; Writing—review & editing, Antonio Páez. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Ministry of Education, Humanities and Social Sciences Research Youth Foundation, China (No. 20YJC840014), the China Postdoctoral Science Foundation (No. 2017M623151), the Major Project of The National Social Science Fund of China (No. 19ZDA149), the National Natural Science Foundation of China (No. 41901172), and the Fundamental Research Funds for the Central Universities (No. SK2020011).

**Acknowledgments:** The authors would like to acknowledge the editor and three anonymous reviewers for their insightful suggestions and constructive comments. The authors would also like to acknowledge Xi'an Public Security Bureau for making available the crime data, and the 10<sup>th</sup> Research Institute of the Ministry of Posts and Telecommunications of China for making available the mobile phone big data. The authors thank Pengfei Cheng, Min Liu, and Fang He for the assistance in processing mobile phone data. The first author also thanks Chen for her unwavering support.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Andresen, M.A. The Ambient Population and Crime Analysis. *Prof. Geogr.* **2011**, *63*, 193–212. [\[CrossRef\]](#)
- Malleon, N.; Andresen, M.A. The impact of using social media data in crime rate calculations: Shifting hot spots and changing spatial patterns. *Cartogr. Geogr. Inf. Sc.* **2015**, *42*, 112–121. [\[CrossRef\]](#)
- Malleon, N.; Andresen, M.A. Exploring the impact of ambient population measures on London crime hotspots. *J. Crim. Justice* **2016**, *46*, 52–63. [\[CrossRef\]](#)
- Hanaoka, K. 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.* **2016**, *45*, 295–311. [\[CrossRef\]](#)
- Song, G.; Liu, L.; Bernasco, W.; Xiao, L.; Zhou, S.; Liao, W. 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.* **2018**, *108*, 1370–1388. [\[CrossRef\]](#)
- Déville, P.; Linard, C.; Martin, S.; Gilbert, M.; Stevens, F.R.; Gaughan, A.E.; Blondel, V.D.; Tatem, A.J. Dynamic population mapping using mobile phone data. *Proc. Natl. Acad. Sci.* **2014**, *111*, 15888–15893. [\[CrossRef\]](#)
- Hipp, J.; Bates, C.; Lichman, M.; Smyth, P. Using Social Media to Measure Temporal Ambient Population: Does it Help Explain Local Crime Rates? *Justice Q.* **2018**, *36*, 1–31. [\[CrossRef\]](#)
- Kinney, J.B.; Brantingham, P.L.; Wuschke, K.; Kirk, M.G.; Brantingham, P.J. Crime Attractors, Generators and Detractors: Land Use and Urban Crime Opportunities. *Built Environ.* **2008**, *34*, 62–74. [\[CrossRef\]](#)
- Brantingham, P.L.; Brantingham, P.J. Criminality of place: Crime generators and crime attractors. *Eur. J. Crim. Pol. Res.* **1995**, *3*, 1–26. [\[CrossRef\]](#)
- Felson, L.E.C. Social Change and Crime Rate Trends: A Routine Activity Approach. *Am. Sociol. Rev.* **1979**, *44*, 588. [\[CrossRef\]](#)
- Morency, C.; Páez, A.; Roorda, M.; Mercado, R.; Farber, S. Distance traveled in three Canadian cities: Spatial analysis from the perspective of vulnerable population segments. *J. Transp. Geogr.* **2011**, *19*, 39–50. [\[CrossRef\]](#)
- Clarke, R.V. OPPORTUNITY-BASED CRIME RATESThe Difficulties of Further Refinement. *Br. J. Criminol.* **1984**, *24*, 74–83. [\[CrossRef\]](#)

13. Andresen, M.A. Location Quotients, Ambient Populations, and the Spatial Analysis of Crime in Vancouver, Canada. *Environ. Plan. A: Econ. Space* **2007**, *39*, 2423–2444. [[CrossRef](#)]
14. Liu, L.; Jiang, C.; Zhou, S.; Liu, K.; Du, F. Impact of public bus system on spatial burglary patterns in a Chinese urban context. *Appl. Geogr.* **2017**, *89*, 142–149. [[CrossRef](#)]
15. Lan, M.; Liu, L.; Hernandez, A.; Liu, W.; Zhou, H.; Wang, Z. The Spillover Effect of Geotagged Tweets as a Measure of Ambient Population for Theft Crime. *Sustainability* **2019**, *11*, 6748. [[CrossRef](#)]
16. Crols, T.; Malleson, N. Quantifying the ambient population using hourly population footfall data and an agent-based model of daily mobility. *GeoInformatica* **2019**, *23*, 201–220. [[CrossRef](#)]
17. Li, T.; Sun, D.; Jing, P.; Yang, K. Smart Card Data Mining of Public Transport Destination: A Literature Review. *Information* **2018**, *9*, 18. [[CrossRef](#)]
18. Dobson, J.E.; Bright, E.A.; Coleman, P.R.; Durfee, R.C.; Worley, B.A. LandScan: A global population database for estimating populations at risk. *Photogramm. Eng. Rem. S.* **2000**, *66*, 849–857.
19. Bowers, K.J.; Johnson, S.D.; Pease, K. Prospective Hot-Spotting: The Future of Crime Mapping? *Br. J. Criminol.* **2004**, *44*, 641–658. [[CrossRef](#)]
20. Shiode, S. Street-level Spatial Scan Statistic and STAC for Analysing Street Crime Concentrations. *Trans. GIS* **2011**, *15*, 365–383. [[CrossRef](#)]
21. Vespignani, A. Predicting the Behavior of Techno-Social Systems. *Science* **2009**, *325*, 425–428. [[CrossRef](#)]
22. Lazer, D.; Pentland, A. (Sandy); Adamic, L.; Aral, S.; Barabasi, A.-L.; Brewer, D.; Christakis, N.; Contractor, N.; Fowler, J.H.; Gutmann, M.; et al. SOCIAL SCIENCE: Computational Social Science. *Science* **2009**, *323*, 721–723. [[CrossRef](#)] [[PubMed](#)]
23. Liu, Y.; Liu, X.; Gao, S.; Gong, L.; Kang, C.; Zhi, Y.; Chi, G.; Shi, L. Social Sensing: A New Approach to Understanding Our Socioeconomic Environments. *Ann. Assoc. Am. Geogr.* **2015**, *105*, 512–530. [[CrossRef](#)]
24. Panigutti, C.; Tizzoni, M.; Bajardi, P.; Smoreda, Z.; Colizza, V. Assessing the use of mobile phone data to describe recurrent mobility patterns in spatial epidemic models. *R. Soc. Open Sci.* **2017**, *4*, 160950. [[CrossRef](#)] [[PubMed](#)]
25. González, M.C.; Pinheiro, F.L.; Barabasi, A.-L. Understanding individual human mobility patterns. *Nature* **2008**, *453*, 779–782. [[CrossRef](#)]
26. Song, G.; Bernasco, W.; Liu, L.; Xiao, L.; Zhou, S.; Liao, W. Crime Feeds on Legal Activities: Daily Mobility Flows Help to Explain Thieves' Target Location Choices. *J. Quant. Criminol.* **2019**, *35*, 831–854. [[CrossRef](#)]
27. Griffiths, G.; Johnson, S.D.; Chetty, K. UK-based terrorists' antecedent behavior: A spatial and temporal analysis. *Appl. Geogr.* **2017**, *86*, 274–282. [[CrossRef](#)]
28. Bogomolov, A.; Lepri, B.; Staiano, J.; Oliver, N.; Pianesi, F.; Pentland, A. Once Upon a Crime: Towards Crime Prediction from Demographics and Mobile Data. In *Proceedings of the 16th International Conference on Multimodal Interaction*; Istanbul, Turkey, 12–16 November 2014; Association for Computing Machinery: Istanbul, Turkey; pp. 427–434.
29. Reades, J.; Calabrese, F.; Ratti, C. Eigenplaces: Analysing cities using the space – time structure of the mobile phone network. *Environ. Plan. B Plan. Des.* **2009**, *36*, 824–836. [[CrossRef](#)]
30. Dhondt, S.; Beckx, C.; Degraeuwe, B.; Lefebvre, W.; Kochan, B.; Bellemans, T.; Panis, L.I.; Macharis, C.; Putman, K. Integration of population mobility in the evaluation of air quality measures on local and regional scales. *Atmospheric Environ.* **2012**, *59*, 67–74. [[CrossRef](#)]
31. He, L.; Páez, A.; Liu, D. Built environment and violent crime: An environmental audit approach using Google Street View. *Comput. Environ. Urban Syst.* **2017**, *66*, 83–95. [[CrossRef](#)]
32. Eck, J.E.; Weisburd, D. Crime Places in Crime Theory. In *Crime and Place, crime Prevention Studies*; Eck, J.E., Weisburd, D., Eds.; Criminal Justice Press: Monsey, NY, USA, 1995.
33. Sherman, L.W.; Gartin, P.R.; Buerger, M.E. HOT SPOTS OF PREDATORY CRIME: ROUTINE ACTIVITIES AND THE CRIMINOLOGY OF PLACE\*. *Criminology* **1989**, *27*, 27–56. [[CrossRef](#)]
34. Cozens, P.; Love, T.; Nasar, J.L. A Review and Current Status of Crime Prevention through Environmental Design (CPTED). *J. Plan. Lit.* **2015**, *30*, 393–412. [[CrossRef](#)]
35. Wilcox, P.; Quisenberry, N.; Cabrera, D.T.; Jones, A. Busy Places and Broken Windows? Toward Defining the Role of Physical Structure and Process in Community Crime Models. *Sociol. Q.* **2004**, *45*, 185–207. [[CrossRef](#)]
36. Andresen, M.A. Crime measures and the spatial analysis of criminal activity. *Brit. J. Criminol.* **2006**, *46*, 258–285. [[CrossRef](#)]

37. He, L.; Páez, A.; Liu, D.; Jiang, S. Temporal stability of model parameters in crime rate analysis: An empirical examination. *Appl. Geogr.* **2015**, *58*, 141–152. [\[CrossRef\]](#)
38. He, L.; Páez, A.; Liu, D. Persistence of Crime Hot Spots: An Ordered Probit Analysis. *Geogr. Anal.* **2016**, *49*, 3–22. [\[CrossRef\]](#)
39. Bernasco, W.; Block, R.; Rengert, G.; Groff, E.; Eck, J. Robberies in Chicago: A Block-Level Analysis of the Influence of Crime Generators, Crime Attractors, and Offender Anchor Points. *J. Res. Crime Delinquency* **2010**, *48*, 33–57. [\[CrossRef\]](#)
40. Mburu, L.W.; Helbich, M. Environmental Risk Factors influencing Bicycle Theft: A Spatial Analysis in London, UK. *PLoS ONE* **2016**, *11*, e0163354. [\[CrossRef\]](#)
41. Syption-Dutkowska, N.; Leitner, M. Land Use Influencing the Spatial Distribution of Urban Crime: A Case Study of Szczecin, Poland. *ISPRS Int. J. Geo-Information* **2017**, *6*, 74. [\[CrossRef\]](#)
42. Roncek, D.W.; Maier, P.A. BARS, BLOCKS, AND CRIMES REVISITED: LINKING THE THEORY OF ROUTINE ACTIVITIES TO THE EMPIRICISM OF "HOT SPOTS". *Criminology* **1991**, *29*, 725–753. [\[CrossRef\]](#)
43. Stucky, T.D.; Ottensmann, J.R. LAND USE AND VIOLENT CRIME\*. *Criminology* **2009**, *47*, 1223–1264. [\[CrossRef\]](#)
44. Browning, C.R.; Byron, R.A.; Calder, C.; Krivo, L.J.; Kwan, M.-P.; Lee, J.-Y.; Peterson, R.D. Commercial Density, Residential Concentration, and Crime: Land Use Patterns and Violence in Neighborhood Context. *J. Res. Crime Delinquency* **2010**, *47*, 329–357. [\[CrossRef\]](#)
45. Song, G.; Liu, L.; Bernasco, W.; Zhou, S.; Xiao, L.; Long, D. Theft from the person in urban China: Assessing the diurnal effects of opportunity and social ecology. *Habitat Int.* **2018**, *78*, 13–20. [\[CrossRef\]](#)
46. Feng, J.; Liu, L.; Long, D.; Liao, W. An Examination of Spatial Differences between Migrant and Native Offenders in Committing Violent Crimes in a Large Chinese City. *ISPRS Int. J. Geo-Information* **2019**, *8*, 119. [\[CrossRef\]](#)
47. Steffensmeier, D.; Harris, C.T.; Painter-Davis, N. Gender and Arrests for Larceny, Fraud, Forgery, and Embezzlement: Conventional or Occupational Property Crime Offenders? *J. Crim. Justice* **2015**, *43*, 205–217. [\[CrossRef\]](#)
48. Cabrera-Barona, P.; Pontón, G.J.; Melo, P. Types of Crime, Poverty, Population Density and Presence of Police in the Metropolitan District of Quito. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 558. [\[CrossRef\]](#)
49. Chen, J.; Liu, L.; Zhou, S.; Xiao, L.; Jiang, C. Spatial Variation Relationship between Floating Population and Residential Burglary: A Case Study from ZG, China. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 246. [\[CrossRef\]](#)
50. Chen, J.; Liu, L.; Xiao, L.; Xu, C.; Long, D. Integrative Analysis of Spatial Heterogeneity and Overdispersion of Crime with a Geographically Weighted Negative Binomial Model. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 60. [\[CrossRef\]](#)
51. Papa, G.L.; Palermo, V.; Dazzi, C. Is land-use change a cause of loss of pedodiversity? The case of the Mazzarrone study area, Sicily. *Geomorphology* **2011**, *135*, 332–342. [\[CrossRef\]](#)
52. Shi, K.; Yu, B.; Huang, Y.; Hu, Y.; Yin, B.; Chen, Z.; Chen, L.; Wu, J. Evaluating the Ability of NPP-VIIRS Nighttime Light Data to Estimate the Gross Domestic Product and the Electric Power Consumption of China at Multiple Scales: A Comparison with DMSP-OLS Data. *Remote. Sens.* **2014**, *6*, 1705–1724. [\[CrossRef\]](#)
53. Bruederle, A.; Hodler, R. Nighttime lights as a proxy for human development at the local level. *PLoS ONE* **2018**, *13*, e0202231. [\[CrossRef\]](#) [\[PubMed\]](#)
54. Yin, Z.; Li, X.; Tong, F.; Li, Z.; Jendryke, M. Mapping urban expansion using night-time light images from Luojia1-01 and International Space Station. *Int. J. Remote. Sens.* **2019**, *41*, 2603–2623. [\[CrossRef\]](#)
55. Zhou, H.; Liu, L.; Lan, M.; Yang, B.; Wang, Z. Assessing the Impact of Nightlight Gradients on Street Robbery and Burglary in Cincinnati of Ohio State, USA. *Remote. Sens.* **2019**, *11*, 1958. [\[CrossRef\]](#)
56. Griffith, D.A. *Spatial Autocorrelation and Spatial Filtering: Gaining Understanding through Theory and Scientific Visualization*; Springer: Berlin, Germany, 2003.
57. Moran, P.A.P. Notes on Continuous Stochastic Phenomena. *Biometrika* **1950**, *37*, 17. [\[CrossRef\]](#) [\[PubMed\]](#)
58. Cliff, A.D.; Griffith, D.A.; Amrhein, C.G. Multivariate Statistical Analysis for Geographers. *J. Am. Stat. Assoc.* **1999**, *94*, 654. [\[CrossRef\]](#)
59. Osgood, D.W. Poisson-Based Regression Analysis of Aggregate Crime Rates. *J. Quant. Criminol.* **2000**, *16*, 21–43. [\[CrossRef\]](#)
60. Anselin, L.; Syabri, I.; Kho, Y. GeoDa: An Introduction to Spatial Data Analysis. *Geogr. Anal.* **2006**, *38*, 5–22. [\[CrossRef\]](#)

61. Ranjan, C.; Najari, V. ProcessMiner/nlcor: Compute Nonlinear Correlations. 2019. Available online: <https://github.com/ProcessMiner/nlcor> (accessed on 16 April 2020).
62. Glaeser, E.L.; Sacerdote, B. Why is There More Crime in Cities? *J. Politi- Econ.* **1999**, *107*, S225–S258. [CrossRef]
63. Chang, Y.S.; Kim, H.E.; Jeon, S. Do Larger Cities Experience Lower Crime Rates? A Scaling Analysis of 758 Cities in the US. *Sustainability* **2019**, *11*, 3111. [CrossRef]
64. Ladbroke, D.A. Why are crime rates higher in urban than in rural areas? — Evidence from Japan. *Aust. N. Zealand J. Criminol.* **1988**, *21*, 81–103. [CrossRef]
65. Sampson, R.J.; Groves, W.B. Community Structure and Crime: Testing Social-Disorganization Theory. *Am. J. Sociol.* **1989**, *94*, 774–802. [CrossRef]
66. Chun, Y. Analyzing Space-Time Crime Incidents Using Eigenvector Spatial Filtering: An Application to Vehicle Burglary. *Geogr. Anal.* **2014**, *46*, 165–184. [CrossRef]
67. Sampson, R.J.; Raudenbush, S.W. Systematic Social Observation of Public Spaces: A New Look at Disorder in Urban Neighborhoods. *Am. J. Sociol.* **1999**, *105*, 603–651. [CrossRef]
68. Bursik, R.J. SOCIAL DISORGANIZATION AND THEORIES OF CRIME AND DELINQUENCY: PROBLEMS AND PROSPECTS\*. *Criminology* **1988**, *26*, 519–552. [CrossRef]
69. Marcum, C.D.; Higgins, G.E.; Ricketts, M.L. Potential Factors of Online Victimization of Youth: An Examination of Adolescent Online Behaviors Utilizing Routine Activity Theory. *Deviant Behav.* **2010**, *31*, 381–410. [CrossRef]
70. Pratt, T.C.; Holtfreter, K.; Reisig, M.D. Routine Online Activity and Internet Fraud Targeting: Extending the Generality of Routine Activity Theory. *J. Res. Crime Delinquency* **2010**, *47*, 267–296. [CrossRef]
71. Maimon, D.; Kamerdze, A.; Cukier, M.; Sobesto, B. Daily Trends and Origin of Computer-Focused Crimes Against a Large University Computer Network: An Application of the Routine-Activities and Lifestyle Perspective. *Br. J. Criminol.* **2013**, *53*, 319–343. [CrossRef]
72. Bax, T. Internet Addiction in China: The Battle for the Hearts and Minds of Youth. *Deviant Behav.* **2014**, *35*, 687–702. [CrossRef]
73. Anderson, J.M.; Macdonald, J.; Bluthenthal, R.; Ashwood, J.S. Reducing Crime by Shaping the Built Environment with Zoning: An Empirical Study of Los Angeles. *SSRN Electron. J.* **2012**, *161*, 699–756. [CrossRef]
74. Zahnw, R. Mixed Land Use: Implications for Violence and Property Crime. *City Community* **2018**, *17*, 1119–1142. [CrossRef]
75. Clark, B.A.J. Outdoor lighting and crime. Part 1: Little or no benefit. In Astronomical Society of Victoria, Inc. 2002. Available online: [https://www.researchgate.net/publication/239923212\\_Outdoor\\_Lighting\\_and\\_Crime\\_Part\\_1\\_Little\\_or\\_No\\_Benefit](https://www.researchgate.net/publication/239923212_Outdoor_Lighting_and_Crime_Part_1_Little_or_No_Benefit) (accessed on 16 April 2020).
76. Clark, B.A.J. Outdoor lighting and crime. Part 2: Coupled growth. In Astronomical Society of Victoria, Inc. 2003. Available online: [https://www.researchgate.net/publication/265248247\\_Outdoor\\_Lighting\\_and\\_Crime\\_Part\\_2\\_Coupled\\_Growth](https://www.researchgate.net/publication/265248247_Outdoor_Lighting_and_Crime_Part_2_Coupled_Growth) (accessed on 16 April 2020).



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).