Relationship between crime and characteristics of urban parks in Medellín, Colombia

December 1, 2022

Juliana Restrepo Tobar[1]

Advisors:
Sara Arango Franco[2]
Henry Laniado Rodas[3]

Research practice 2
Final Report
Mathematical Engineering
Department of Mathematical Sciences
School of Sciences
Universidad EAFIT

NOVEMBER 2022

[1]jrestrepot@eafit.edu.co - Mathematical Engineering student
[2]sarangofc@gmail.com - The Edgelands Institute
[3]hlaniado@eafit.edu.co - Department of Mathematical Sciences, Universidad EAFIT.

Abstract

Environmental criminology theories analyze the relationship between crime rates and features of bounded geographic areas, such as neighborhoods, census tracts, cities, counties, states, or nations. Crime Prevention Through Environmental Design (CPTED) theory represents a multifaceted approach to crime reduction that outlines six principles: Surveillance, defensible space, access control, territorial reinforcement, target hardening, space management, and activity support. Although there are studies on the relationship between crime and urban characteristics in areas of Medellín, Colombia, none use parks as their spatial unit of analyisis. We ran four models to analyze the relationship between crimes (thefts to people and homicides) in urban parks and their socioeconomic, surveillance and urban design characteristics. We found that covariates that relate to the floating population in the parks are significant and correlate positively to their crime rates. Density of trees does not appear to have correlation with crimes in parks, and unexpectedly, the density of cameras and light poles in the park correlates positively to the robberies within it. Finally, the number of sports facilities (i.e. courts, fields, stadiums, etc.) within and around a park seems to have a negative correlation to its robberies.
Keywords: Public Spaces, Security, Surveillance, Urban Design.

1 Introduction

As Pratt & Cullen (2005) describe, a “macro-level” or “ecological” analysis examines how features of delimited geographic areas —such as neighborhoods, census tracts, cities, counties, states, or nations— are related to crime rates. Macro-level theories seek to explain why certain characteristics of ecological areas, but not others, account for the distribution of crime. In particular, the concept Crime Prevention Through Environmental Design (CPTED), coined by Jeffery (1977), explores the assumption that the proper design and effective use of built environment can improve the quality of life while reducing the fear and incidence of crime.

The influence of the environment and physical space on crime has been widely addressed in the literature. Variables including weather (Butke & Sheridan, 2010; Cohn & Rotton, 2000), urban layout (Patino *et al.*, 2014), land use and built environment (Inlow, 2021), physical structures (Wilcox *et al.*, 2004), tree cover (Troy *et al.*, 2012; Donovan & Prestemon, 2012), littering (Weaver, 2015), temperature (Sorg & Taylor, 2011) and housing quality (Armitage *et al.*, 2013) have been proved to relate to crime. For instance, Cerdá *et al.* (2012) observed that improving public spaces and creating new institutions in Medellín provided more opportunities for neighbors to interact, develop trust, and become willing to intervene when the social order was threatened, thus reducing crime.

Medellín was once known as the most violent city in the world due to an urban war set off by the drug cartels at the end of the 1980s. In the last few years, though homicides have stayed relatively low, the number of reported robberies has been increasing. To illustrate, Medellín has a population of 2 569 million and according to SISC (2019), a total of 2 277 homicides and 117 113 thefts were recorded in Medellin from 2016 to 2019. 63.19% of them were robberies, 15.33% thefts of motorcycles, 3.14% thefts of cars, 12.71% thefts to commerces, 5.58% thefts to residencies, and 0.05% thefts to financial entities.

Theft is an endemic problem in the region- arguing its disproportionately common occurrence, the significant growth in theft rates, and the recurrent use of violence (Jaitman *et al.*, 2017). As Casa de las Estrategias (2021) stated, robbery modalities show that violence in Medellín continues to be firmly rooted: violent robbery or “atraco” quintuples snatching or pickpocketing. Moreover, they found that in 57% of all theft reports in 2015, a weapon was involved. Therefore, in the last few years, Medellín's City Hall has sought to deter criminals using surveillance strategies. The National Police has 2 886 working security cameras in Medellín and has recently implemented a camera system called Robocop, which allows monitoring of any point of the city in real-time without physical networks or electricity, as stated by Arango (2021).

Data from SISC reveals that most robberies to people occur in public places. However, few studies on crime in Medellín have considered parks as their spatial unit of analysis. Therefore, we used data gathered from Medellín's public areas to develop various regression models to assess how variables relating to illumination, vegetation, built environment, surveillance, and land use correlate to crime in parks.

The remainder of this paper is organized as follows. Section 2 presents the state of the art of environmental analysis of crime, as well as studies of crime in Medellín. Section 3 describes the methodology. Section 4 shows the results of the study, and conclusions and future work are presented in Section 5.

2 State of the art

At the beginning of the 20th century, a macro-level analysis of crime started to emerge, shifting the focus from individuals to geographical units. Shaw *et al.* (1929) began studying social disorganization and argued how macro-level processes explained why crime was endemic to inner-city neighborhoods. Pratt & Cullen (2005) affirm that in the 1970s, these macro-level theories took off with four primary school contributions: Cohen & Felson (1979) development of routine activity theory; the seminal work of Blau & Blau (1982) on inequality and violent crime; the rediscovery in the 1980s of Shaw and McKay's social disorganization theory by scholars; and, finally, Blumstein *et al.* (1978)'s interest in deterrence theory at the macro-level.

CPTED is a famous theory of macro-level analysis of crime. As described by Armitage & Monchuk (2019), it represents a multifaceted approach to crime reduction that draws upon theories from environmental criminology, architecture, and urban design. Its importance as a crime reduction approach has been formalized through strategy, policy, and regulation and its effectiveness has been confirmed in evaluations. Cozens & Love (2015) outline six principles of CPTED: **Surveillance** (e.g. street lighting, large windows that look to the street, police stations), **defensible space**, **access control** (deny access to potential targets by creating a heightened perception of risk in offenders), **territorial reinforcement** (signage, fences, clear delimitation of private, semi-private, and public spaces), **target hardening** (use of locks or burglar alarms to increase the effort and risk of offending), **image/space management** (routine maintenance of built environment, giving a sense of "ownership"), and **activity support** (encouraging the use of built environment for safe activities).

A famous scholar of the relationship between urban design and arising human behaviors (such as crime) is Jane Jacobs. Jacobs (1993) explained that public peace (the sidewalks and street peace) of cities is not kept primarily by the police, but rather is kept primarily by ' 'an intricate, almost unconscious network of voluntary controls and standards among the people themselves, and enforced by the people themselves." She proposed that there must be "eyes on the street", mixed land use, and small blocks to create and strengthen a sense of community, among others.

The relationship between crime and urban characteristics in Medellín has been studied. Patino *et al.* (2014) used remote sensing to assess the relationship between crime and the urban layout. They tested which land cover, structure, and texture descriptors were significantly related to intra-urban homicide rates in Medellín while controlling for socioeconomic confounders. They found that areas with higher homicide rates tended to be more crowded and cluttered, with small dwellings with different roofing materials located close to one another, and often lacked other homogeneous surfaces such as open green spaces, wide roads, or large facilities.

Studies show that interventions in neighborhood physical infrastructure can reduce violence. In 2004, municipal authorities in Medellín built a public transit system to connect isolated low-income neighborhoods to the city's urban center. Framing such interventions as a natural experiment, Cerdá *et al.* (2012) compared the intervention neighborhoods and control neighborhoods before 2003 and after 2008. Results show that the decline in homicide rate was 66% greater in intervention neighborhoods than in control neighborhoods. Furthermore, the researchers found that improving public spaces and creating new institutions promoted interaction and trust between neighbors to interact, which made them willing to intervene when the social order was threatened. This aligns with Jacobs (1993) view, in the sense that building a community is important in crime prevention.

As for the deterrent effect of surveillance cameras on crime in Medellín, Gómez *et al.* (2021) found that quasi-random allocation of cameras led to a decrease in crimes and arrests, with data provided by Medellín's Security Secretariat. The results suggest offenders were deterred rather than incapacitated, and there is no evidence of crime displacement or diffusion of benefits to surrounding locations.

3 Methodology

3.1 Data

Our spatial units of analysis are 1044 parks in Medellín. The original data included 3743 public places, but since the polygons' sizes were drastically different, we removed the 5% largest and the 60% smallest public places. We also discarded the parks that were within Medellín's *corregimientos*[4]. Figure 1 is a map of the parks we analyze.

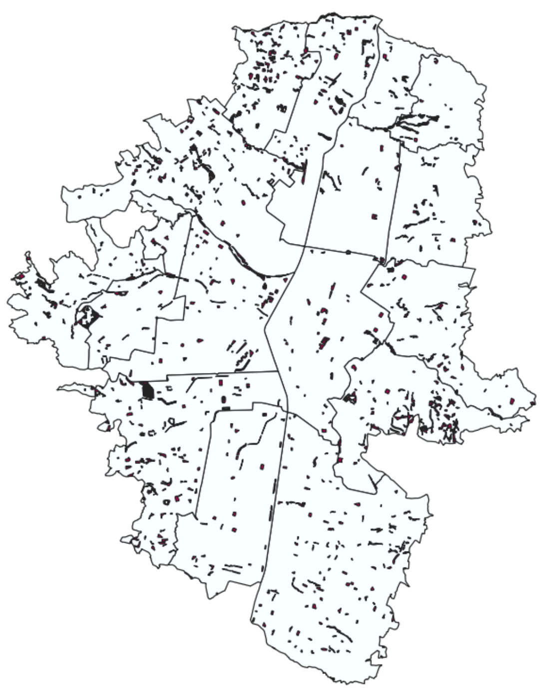

Figure 1: Spatial Units of Analysis (parks in Medellín)

The selected variables to quantify crime are robberies to people and homicides from 2018 to 2022. Therefore, robbery and homicides are the dependent variables in the models. We chose robbery because the urban design could easily affect it due to its nature (i.e., it depends directly on the offender's and target's place and circumstances). Homicide was selected because it is the best-documented crime, meaning that its reporting is homogenous across all geographies and sociodemographic characteristics, unlike robbery which is severely under-reported. To aggregate the crimes, we computed a 10-meter buffer for each park. We took two approaches: The first was counting all the crimes that occurred within each 10-meter buffer, as equation (1) shows.

$$\text{Count}_i = \sum_{j \in J} I(j \in B_i) \tag{1}$$

Where Count_i is the count of crimes in park i, J is the set of all crimes in Medellín, and B_i is the 10-meter buffer of park i. The second was counting the crimes within each park but assigning a

[4] Smaller jurisdictions/townships within the city government and territory but not confined by its urban footprint

weight $\alpha_j < 1$ to crimes within the 10-meter buffer but not within the park. Equation (2) explains the approach more clearly.

$$\text{Intensity Value}_i = \sum_{j \in J} I(j \in P_i) + \alpha_j I(j \in (B_i - P_i)) \tag{2}$$

Where Intensity Value$_i$ is the intensity value of crimes in park i, J is the set of all crimes in Medellín, B_i is the 10-meter buffer of park i, P_i is the park i, and $\alpha_j = \frac{1}{\text{distance}(j, \text{centroid}(P_i))}$.

The intensity value is a useful tool used to study many phenomena, among them crime (see Mccord & Ratcliffe (2009)). It assigns a weight to a data point based on how close it is to the centroid of the spatial unit of analysis. We use a slight variation since we compute the counts $j \in P_i$ and only use the intensity value for $j\ in B_i - P_i$. We chose this approach because the buffer choice was based only on intuition so it might cause biases. Therefore, we take away weight from crimes inside the buffer but not inside the park.

Figures 2, 3, 4, 5 show the heatmaps of the dependent variables.

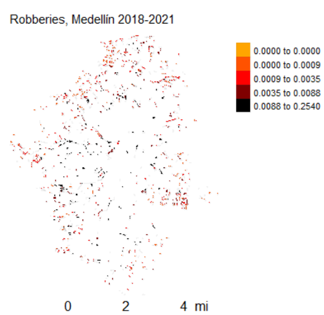

Figure 2: Robbery counts heatmap

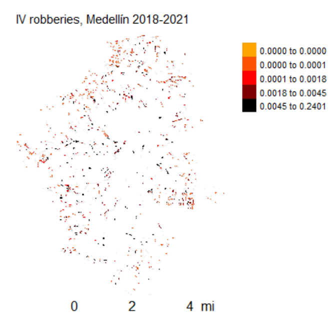

Figure 3: Robbery intensity values heatmap

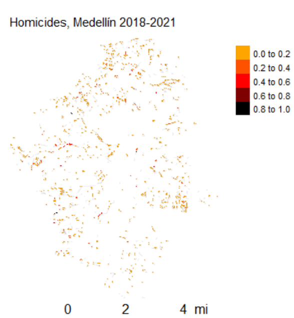

Figure 4: Homicide counts heatmap

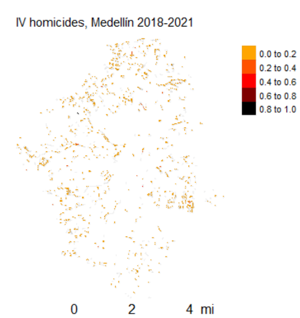

Figure 5: Homide intensity values heatmap

Table 1 shows the covariates we extracted for each public place using various georeferenced data sources.

Varible	**Name**	**Link to environmental criminology theory**
Density of light poles within 10 meters	DENS.LIG	Illumination
Density of public CCTV (closed-circuit television) cameras within 10 meters	DENS.CAM	Surveillance
Density of trees within 10 meters	DENS.TREE	Vegetation
Distance to nearest sports facility (e.g. badges, courts, and fields)	DIST.SPORT	
Distance to nearest educative institution	DIST.EDU	Sense of community and encouragement of the use of built environment for safe activities
Distance to nearest cultural facility (e.g. libraries, museums, etc)	DIST.CULT	
Sports facility count within 100 meters	NUM.SPORT	
Educative institution count within 100 meters	NUM.EDU	
Cultural facility count within 100 meters	NUM.CULT	
Distance to the nearest public force equipment (e.g. CAIs, police stations)	DIST.PUB.FORCE	Access control (CPTED) and surveillance
Distance to the nearest justice institution (e.g. police departments, houses of justice)	DIST.JUS	
Justice institution count	NUM.JUS	
Mode of land use within 200 meters (Dummy)	USE_<particular use>	
Distance to nearest bar	DIST.BAR	
Distance to nearest Metro station	DIST.STATION	Land use and floating population proxy
Bar count within 100 meters	NUM.BAR	
Store/business count within 100 meters	NUM.STORE	
Metro station count within 100 meters	NUM.STATION	
Buffer area	AREA.BUFF	
Mean population density within 200 meters	POB.DENS	
Downtown (Dummy)	DOWNTOWN	
Mode of stratus within 200 meters	MOD.STRATUS	
Mode of educative level within 200 meters	MOD.EDU	Socioeconomic factors
Mean % of young men within 200 meters	YOUNG.MEN	
Mean % of young women within 200 meters	YOUNG.WOMEN	

Table 1: Covariates

The covariates DENS.LIG, DENS.CAM, and DENS.TREE are crucial in the CPTED theory since they relate to illumination, surveillance, and vegetation. To determine their densities, we calculated their counts in the 10-meter buffers and divided them by the buffers' areas. We computed the variables DIST.SPORT, DIST.EDU, DIST.CULT, DIST.PUB.FORCE, DIST.JUS, DIST.BAR, and DIST.STATION by selecting the minimum distance from the centroid of the park to the centroid of these sites. NUM.SPORT, NUM.EDU, NUM.CULT, NUM.BAR, NUM.STORE, NUM.STATION, and NUM.JUS are covariates that quantify how many instances of these sites are in a 100-meter buffer of each park. The buffer selection was pseudo-random, which may bias the models. For the variable AREA.BUFF we calculated the areas of the 10-meter buffers we used to count the crimes, DENS.LIG, DENS.CAM, and DENS.TREE. DOWNTOWN is a dummy variable that takes the value 1 when the park is in Medellín's downtown and 0 otherwise. It is a relevant covariate to the models since the floating population and crimes in the city's downtown are very high compared to other parts. Medellín's Census of 2018 is the data source of the covariates POB.DENS, MOD.STRATUS, MOD.EDU, YOUNG.MEN, YOUNG.WOMEN. The Census contains socioeconomic details of every block in Medellín, but oftentimes some blocks have missing information. Thus, to prevent biases from missing information, we used the mean population densities instead of the sums of population densities in the blocks within a 200-meter buffer of the parks. We also used the mean to compute YOUNG.MEN and YOUNG.WOMEN for this reason, and also because they are proportions so summing them did not make sense. MOD.STRATUS and MOD.EDU are the modes of the stratus and the educative level of the blocks within the 200-meter buffer of each park. Finally, land use is a categorical covariate that can take the values USE_dotacional, USE_areas.de.baja.mixtura, USE_areas.de.alta.mixtura, and USE_areas.de.media.mixtura. USE_dotacional refers to areas dedicated to public facilities. USE_areas.de.baja.mixtura (i.e., low mixed areas) are predominantly residential areas where housing predominates, allowing the mixture with economic activities of daily use, where the citizen accesses based on the proximity of his residence. USE_areas.de.media.mixtura (i.e., medium mixed areas) correspond to urban spaces with agglomeration or medium concentration of economic activities in relation to the residence. USE_areas.de.baja.mixtura (i.e., areas and corridors of high mixed use) are dominated by land use based on economic activities and the provision of public services, with lower proportions of residential use.

3.2 Models

Since crime is a spatial phenomenon, we tested for spatial dependencies in the aggregated crimes using Moran's I test and confirmed that there were clusters of thefts and homicides in parks in Medellín (parks close to each other have similar crime rates). Therefore, spatial auto-correlation must be considered when building a model.

We use the classical approach to spatial econometric modeling. First, we estimate the following linear model with N observations and k variates:

$$y = X\beta + \epsilon$$

Where y is a $Nx1$ random variable, X is a Nxk matrix of variates, β is a $kx1$ vector of weights, and ϵ is a $Nx1$ error vector.

Then, we use two Lagrange Multiplier tests to determine which spatial model to use: LM_ρ and LM_λ, where ρ and λ are scalar autoregressive parameters, and:

$$LM_\rho = \frac{(\hat{\epsilon}'Wy/\hat{\sigma}^2)^2}{NJ}$$

with

$$J = \frac{1}{N\hat{\sigma}^2}[(WX\hat{\beta})'M(WX\hat{\beta}) + T\hat{\sigma}^2]$$

where W is a row-standardized (NxN) matrix of exogenously determined elements representing the spatial morphology, M is $I - X(X'X) - 1X'$, and T is the trace of the matrix, and:

$$LM_\lambda = \frac{(\hat{\epsilon}'W\hat{\epsilon}/\hat{\sigma^2})^2}{T}$$

If $LM_\rho > LM_\lambda$, then we estimate the spatial lag model

$$y = \rho W y + X\beta + \mu$$

by including a spatially lagged dependent variable (MLLAG). If $LM_\rho < LM_\lambda$, we estimate the spatial AR error model

$$y = X\beta + (I - \lambda W)^{-1}\epsilon$$

using the maximum likelihood estimators for the spatially autoregressive error model (MLERROR). If both tests are significant, we use the robust version of the tests and estimate the specification by the more significant one, and if none of them are, we use a Geographically Weighted Regression (GWR).

In our case, we create several spatial econometric models. For both robberies and homicides, we consider their counts and their intensity values. As such, there are 4 models in total.

4 Results

4.1 Robbery counts

We ran a baseline OLS model with all the covariates that were explained in Table 1, its results are displayed in Table 2. However, its VIF evidenced important multicollinearity between YOUNG.MEN and YOUNG.WOMEN, so the latter was removed from the model, fixing the problem. Table 3 shows the results of the model without YOUNG.WOMEN (Fixed Model).

Variable	**p-value**
(Intercept)	0.0000 ***
DENS.LIG	0.0000 ***
DENS.CAM	0.0002 ***
DENS.TREE	0.2854
DIST.SPORT	0.5107
DIST.EDU	0.0783 .
DIST.CULT	0.6837
NUM.SPORT	0.0520 .
NUM.EDU	0.2393
NUM.CULT	0.2398
DIST.PUB.FORCE	0.5886
DIST.JUS	0.5326
NUM.JUS	0.5046
DIST.BAR	0.1236
DIST.STATION	0.1236
NUM.BAR	0.0000 ***
NUM.STORE	0.0000 ***
NUM.STATION	0.0003 ***
AREA.BUFF	0.0000 ***
POB.DENS	0.6564
DOWNTOWN	0.0000 ***
MOD.STRATUS	0.4461
MOD.EDU	0.0174 *
YOUNG.MEN	0.3658
YOUNG.WOMEN	0.5030
USE_dotacional	0.9538
USE_areas.de.baja.mixtura	0.5326
USE_areas.de.alta.mixtura	0.6372
USE_areas.de.media.mixtura	0.5929

R^2	**p-value**
0.4546	< 2.2e-16 ***

Table 2: Baseline OLS model

Varible	**p-value**
(Intercept)	0.0000 ***
DENS.LIG	0.0000 ***
DENS.CAM	0.0003 ***
DENS.TREE	0.3001
DIST.SPORT	0.4958
DIST.EDU	0.0772 .
DIST.CULT	0.6454
NUM.SPORT	0.0541 .
NUM.EDU	0.2308
NUM.CULT	0.2398
DIST.PUB.FORCE	0.6114
DIST.JUS	0.5140
NUM.JUS	0.4805
DIST.BAR	0.1291
DIST.STATION	0.1381
NUM.BAR	0.0000 ***
NUM.STORE	0.0000 ***
NUM.STATION	0.0003 ***
AREA.BUFF	0.0000 ***
POB.DENS	0.6663
DOWNTOWN	0.0000 ***
MOD.STRATUS	0.4688
MOD.EDU	0.0143 *
YOUNG.MEN	0.3468
USE_dotacional	0.9389
USE_areas.de.baja.mixtura	0.4736
USE_areas.de.alta.mixtura	0.7059
USE_areas.de.media.mixtura	0.5844

R^2	**p-value**
0.4543	< 2.2e-16 ***

Table 3: Fixed Model

Though not interpretable, the Fixed Model has an excellent p-value and a decent R^2, indicating that the selected covariables do affect theft in public places. We ran the Lagrange Multiplier tests based on the Fixed Model's fit and a matrix W of distances to the nearest neighbors (using the parks' centroids). After a sensitivity analysis, we set the upper bound for W to be 1 km, which means two parks are considered to be neighbors if their distance is less than 1 km. Only the MLLAG test was significant (p-value = 0.0001), thus we implemented a spatial lag model. Table 4 shows the model's results.

Varible	**Estimate**	**p-value**
(Intercept)	-5.1681e-03	0.0035 **
DENS.LIG	1.1884e-02	0.0000 ***
DENS.CAM	1.0254e-02	0.0004 ***
DENS.TREE	-4.7988e-03	0.2619
DIST.SPORT	2.0520e-03	0.5919
DIST.EDU	-5.7704e-03	0.0752 .
DIST.CULT	1.3090e-03	0.6884
NUM.SPORT	-5.0619e-03	0.0402 *
NUM.EDU	3.2409e-03	0.1907
NUM.CULT	7.5979e-03	0.2080
DIST.PUB.FORCE	1.2460e-03	0.4592
DIST.JUS	1.8001e-03	0.4525
NUM.JUS	-1.6802e-03	0.5344
DIST.BAR	5.3396e-03	0.0816 .
DIST.STATION	-3.5214e-03	0.2240
NUM.BAR	1.2534e-01	0.0000 ***
NUM.STORE	4.4289e-02	0.0000 ***
NUM.STATION	1.0930e-02	0.0004 ***
AREA.BUFF	1.3653e-02	0.0000 ***
POB.DENS	-0.00637018	0.5955
DOWNTOWN	1.6535e-02	0.0000 ***
MOD.STRATUS	-2.4035e-03	0.3691
MOD.EDU	3.9179e-03	0.0386 *
YOUNG.MEN	-4.4650e-03	0.2247
USE_dotacional	-3.8728e-04	0.9210
USE_areas.de.baja.mixtura	-8.3159e-04	0.3603
USE_areas.de.alta.mixtura	-8.9302e-05	0.9566
USE_areas.de.media.mixtura	-1.0399e-03	0.5718

ρ	**LR test value**	**p-value**
0.18553	9.2165	0.0023984

Table 4: Spatial Lag Model (Model 1)

The covariates' significance barely changes between the Fixed Model and Model 1, but only the latter considers the spatial autocorrelation. Thus, only the spatial model's coefficient estimates are interpretable. Model 1's p-value is still significant, which indicates that the covariates explain theft to some extent even when considering spatial autocorrelation. DENS.LIG appears to be very significant in the model, but contrary to CPTED's theory, parks with a greater density of light poles have more thefts. DENS.CAM also has a positive correlation with thefts The mismatch of these results and the theory signals that the variables might not explain well the phenomenon, or that, rather than bringing information about the activities/surveillance in public places, they are a proxy of the floating population in said places. NUM.SPORT is also a significant covariate, and its coefficient suggests that parks with more badges, courts, fields, or sportive institutions in their surroundings present fewer thefts. This result agrees with the environmental criminology theory that states that encouraging the use of built space for safe activities reduces crime rates. This result agrees with the criminology theory which states that encouraging the use of built space for safe activities reduces crime rates. The number of stores, bars, and Metro stations surrounding a public place is positively correlated with its robbery count, which could occur because these variables explain the occupancy of that place. The dummy variable DOWNTOWN has a positive estimated coefficient, as we had previously predicted. Finally, the mode of the highest educational level achieved in the blocks surrounding a park has a positive correlation to its robbery count. We hypothesize that this coefficient is positive because people living in blocks with higher educational attainment might report thefts more than people living in other areas.

The coefficients estimated in the Model 1 are global estimates, but the behaviors of certain features are not necessarily homogeneous throughout Medellín. Therefore, we also fitted a Geographically Weighted Regression to estimate the robbery count, which lets us visualize local behaviors. The regression yielded a Quasi-Global R^2 of 0.7674, and Figures 6, 7, 8, and 9 show the results for DENS.LIG, DENS.CAM, NUM.SPORT, DENS.TREE respectively. We selected to visualize these variables since they are directly related to urban design and most of them showed to be significant in the spatial lag model.

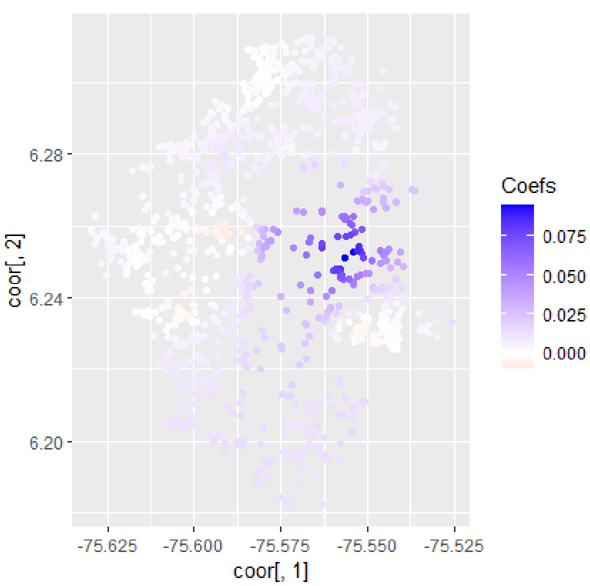

Figure 6: DENS.LIG local coefficients

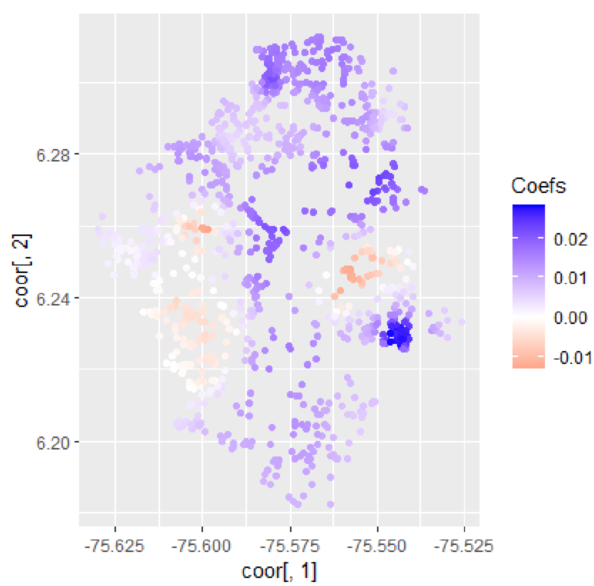

Figure 7: DENS.CAM local coefficients

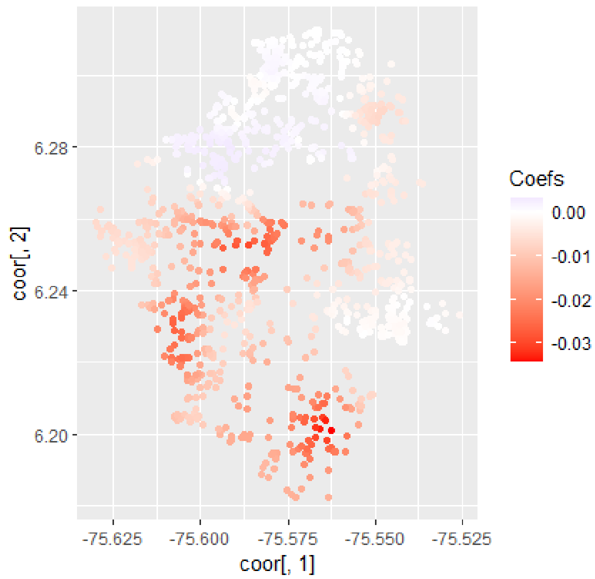

Figure 8: NUM.SPORT local coefficients

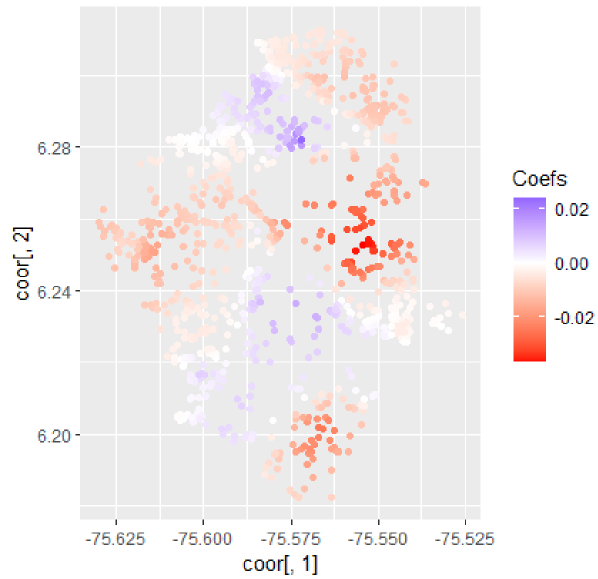

Figure 9: DENS.TREE local coefficients

Figure 6 shows that the density of light poles has a strong positive correlation with thefts to people, especially in Medellín's downtown, which agrees with the results of the Spatial Lag Model. Figure 7 reveals that for most parks in Medelín, a greater density of cameras translates into more robberies. However, parks in Medellín's downtown, Belén, and La América show fewer thefts when their camera density is high. The number of sports facilities surrounding the park has an interesting result: it has a negative correlation with robberies for most parks in the city, but the correlation does not hold for North Medellín. Finally, Figure 9 shows an ambiguous result for the tree density, since it can either have a positive or negative coefficient depending on the zone of the city.

4.2 Robbery intensity values

For the robbery intensity value models, we adapted the same workflow as before. First, we fitted an OLS baseline model to test which spatial model to use. Table 5 shows the results of the baseline model. Only the LMLAG test was significant with a p-value of 0.0102, so we implemented the spatial lag model shown in Table 6.

Estimate	p-value
(Intercept)	0.0091 **
DENS.LIG	0.0000 ***
DENS.CAM	0.0038 **
DENS.TREE	0.3824
DIST.SPORT	0.0231 *
DIST.EDU	0.1399
DIST.CULT	0.9053
NUM.SPORT	0.0461 *
NUM.EDU	0.0984
NUM.CULT	0.0552
DIST.PUB.FORCE	0.9688
DIST.JUS	0.7637
NUM.JUS	0.0203 *
DIST.BAR	0.0475 *
DIST.STATION	0.5481
NUM.BAR	0.0000 ***
NUM.STORE	0.0001 ***
NUM.STATION	0.0807 .
AREA.BUFF	0.0000 ***
POB.DENS	0.5623
DOWNTOWN	0.0000 ***
MOD.STRATUS	0.9227
MOD.EDU	0.5713
YOUNG.MEN	0.8307
USE_dotacional	0.9668
USE_areas.de.baja.mixtura	0.1746
USE_areas.de.alta.mixtura	0.0252
USE_areas.de.media.mixtura	0.2288

R^2	**p-value**
0.3788	< 2.2e-16 ***

Table 5: Baseline OLS Model

Variable	**Estimate**	**p-value**
(Intercept)	-4.4597e-03	0.0082 **
DENS.LIG	1.1778e-02	0.0000 ***
DENS.CAM	5.9306e-03	0.0313 *
DENS.TREE	-5.0869e-03	0.2115
DIST.SPORT	3.2406e-03	0.3733
DIST.EDU	-5.5673e-03	0.0710 .
DIST.CULT	1.9097e-03	0.5381
NUM.SPORT	-5.6475e-03	0.0160 *
NUM.EDU	3.6992e-03	0.1161
NUM.CULT	8.2453e-03	0.1506
DIST.PUB.FORCE	1.7114e-03	0.2850
DIST.JUS	9.3502e-04	0.6811
NUM.JUS	-4.1825e-03	0.1037
DIST.BAR	3.8723e-03	0.1849
DIST.STATION	-1.4390e-03	0.6001
NUM.BAR	1.2562e-01	0.0000 ***
NUM.STORE	1.7985e-02	0.0167 *
NUM.STATION	7.2012e-03	0.0148 *
AREA.BUFF	1.0292e-02	0.0000 ***
POB.DENS	-3.8684e-03	0.4905
DOWNTOWN	1.6922e-02	0.0000 ***
MOD.STRATUS	2.1040e-04	0.9344
MOD.EDU	6.9150e-04	0.6998
YOUNG.MEN	-5.9003e-04	0.8857
USE_dotacional	-6.4818e-05	0.9861
USE_areas.de.baja.mixtura	-5.5532e-04	0.5198
USE_areas.de.alta.mixtura	-2.1155e-05	0.9892
USE_areas.de.media.mixtura	-1.1842e-03	0.4981

ρ	**LR test value**	**p-value**
0.1339	4.0484	0.044214

Table 6: Spatial Lag Model (Model 2)

The baseline OLS model yielded an outstanding p-value, exactly the same as the OLS model of robbery counts, however, the R^2 decreased. The p-value of model 2 is higher than that of model 1, so the performance of this model could be worse. All covariates that were shown to be significant in Model 1 (see Table 4), are significant in Model 2, except MOD.EDU, whose p-value is considerably higher now. The signs of the covariates' coefficients are also the same in both models, so their interpretations are identical.

4.3 Homicide counts

We performed the same procedure for the homicide counts. We first computed a baseline OLS model, then we ran Lagrange Multiplier tests to choose a spatial model. None of the LM tests was significant, but the Morans' I test indicated spatial autocorrelation, so we ran a GWR.

Variable	p-value
(Intercept)	0.8500
DENS.LIG	0.0001 ***
DENS.CAM	0.1635
DENS.TREE	0.3113
DIST.SPORT	0.0440 *
DIST.EDU	0.6954
DIST.CULT	0.8743
NUM.SPORT	0.2541
NUM.EDU	0.9754
NUM.CULT	0.1498
DIST.PUB.FORCE	0.0012 **
DIST.JUS	0.4014
NUM.JUS	0.2920
DIST.BAR	0.2872
DIST.STATION	0.0429 *
NUM.BAR	0.2996
NUM.STORE	0.0768 .
NUM.STATION	0.2395
AREA.BUFF	0.0007 ***
POB.DENS	0.9560
DOWNTOWN	0.0085 **
MOD.STRATUS	0.0554 .
MOD.EDU	0.7726
YOUNG.MEN	0.9289
USE_dotacional	0.9714
USE_areas.de.baja.mixtura	0.2153
USE_areas.de.alta.mixtura	0.6065
USE_areas.de.media.mixtura	0.8421
R^2	**p-value**
0.08339	$< 1.253e-08$ ***

Table 7: Baseline OLS model

Variable	Global Estimate
X.Intercept.	-0.0058
DENS.LIG	0.1361
DENS.CAM	0.0444
DENS.TREE	-0.0324
DIST.SPORT	-0.1001
DIST.EDU	0.0161
DIST.CULT	-0.0064
NUM.SPORT	0.0382
NUM.EDU	0.0010
NUM.CULT	-0.0502
DIST.PUB.FORCE	0.1136
DIST.JUS	0.0307
NUM.JUS	-0.0314
DIST.BAR	-0.0472
DIST.STATION	0.0853
NUM.BAR	-0.0347
NUM.STORE	0.0906
NUM.STATION	0.0380
AREA.BUFF	0.1092
POB.DENS	0.0020
DOWNTOWN	0.1069
MOD.STRATUS	-0.1219
MOD.EDU	0.0167
YOUNG.MEN	0.0038
USE_dotacional	0.0011
USE_areas.de.baja.mixtura	0.0474
USE_areas.de.alta.mixtura	0.0202
USE_areas.de.media.mixtura	-0.0061
Quasi-Global R^2	
0.09343634	

Table 8: Geographically Weighted Regression Model (Model 3)

Even if the baseline OLS model's R^2 is substandard, its p-value is excellent, which indicates the covariates correlate with the homicide counts. Unfortunately, GWR does not yield p-values for each covariate, since they could be misleading, however, it does supply global estimated coefficients and a Quasi-Global R^2. The latter is slightly higher than the OLS model's R^2, nevertheless, it is still deficient, which indicates that Model 3's functional form or the choice of independent variables poorly represent relevant aspects of the actual causality of homicides. Hence, the global coefficients should not be trusted and thus we will not interpret them.

4.4 Homicide intensity values

The homicide intensity value was the only dependent variable that did not show evidence of spatial autocorrelation since Moran's I test did not reject the null hypothesis. However, we tried to run a GWR, but its bandwidth converged to the upper bound of the line search (i.e. the regression contained little pattern). Therefore, we only fitted an OLS model.

Variable	**Estimate**	**p-value**
(Intercept)	-0.0076	0.3283
DENS.LIG	0.0223	0.0430 *
DENS.CAM	0.0189	0.1373
DENS.TREE	-0.0007	0.9691
DIST.SPORT	-0.0103	0.5410
DIST.EDU	0.0094	0.5113
DIST.CULT	-0.0072	0.6135
NUM.SPORT	0.0174	0.1077
NUM.EDU	0.0103	0.3441
NUM.CULT	-0.0305	0.2502
DIST.PUB.FORCE	0.0083	0.2608
DIST.JUS	0.0105	0.9557
NUM.JUS	-0.0091	0.4430
DIST.BAR	0.0048	0.7227
DIST.STATION	0.0105	0.4070
NUM.BAR	-0.0005	0.9910
NUM.STORE	0.0099	0.7721
NUM.STATION	0.0139	0.3102
AREA.BUFF	0.0191	0.0351 *
POB.DENS	0.0084	0.7455
DOWNTOWN	0.0391	0.0003 ***
MOD.STRATUS	-0.0071	0.5486
MOD.EDU	-0.0045	0.5837
YOUNG.MEN	0.0009	0.9618
USE_dotacional	-0.0016	0.9251
USE_areas.de.baja.mixtura	-0.0016	0.6824
USE_areas.de.alta.mixtura	-0.0033	0.6423
USE_areas.de.media.mixtura	-0.0035	0.6675

R^2	**p-value**
0.04413	0.01148

Table 9: OLS Model (Model 4)

Though Model 4's R^2 is low, its p-value is significant, so the coefficients and the significances of the independent variables are interpretable. Only three variables are significant in this model:

DENS.LIG, AREA.BUFF, and DOWNTOWN. All of them correlate positively with homicides.

5 Conclusions and future research

Both robbery models yielded excellent metrics, and some of their results were intriguing. For example, only one socioeconomic variable was significant (and only in Model 1), whereas one would expect, based on the literature, greater significance of the socioeconomic variables than the others. Another interesting result is that of surveillance camera and light pole densities, as it appears that the more surveillance and lighting in a park, the more thefts. On the other hand, there are strong positive correlations between indicators of the floating populations and thefts in parks. For example, the number of bars, stores and Metro stations, which are good proxies for park occupancy, have a positive correlation with thefts. The only significant variable that correlates negatively with thefts is the number of fields, courts, etc. around the parks, which is very interesting as it is consistent with the activity support principle of CPTED. As per the homidide models, we hypothesize that the sparcity of the homicide data in parks caused their bad fits.

There were three main limitations to this study. First, the data is from 2018 to 2022, and the Covid-19 pandemic hit in 2019. Since the pandemic restrictions in Colombia were strict, people were in quarantine during a large part of those years, which favored the diminution of crime. Second, the selection of public places as analysis units for our model was a difficult endeavor. Even though the local government keeps a detailed hierarchy and structure of public spaces, these categories were not enough. We found size to be a better filtering category, but we are aware that more refining needs to be done, and we need to establish an urban-rural cutoff (that may not be the same as the city limits) point to distinguish the kinds of parks that we are after from more rural ones. Third, we selected the buffers based on intuiton and based on what fitted best the data. However, their size might cause biases in the model and needs to be studyed further.

Acknowledgments

We thank the Edgelands Institute and SISC for making this research possible.

According to the internal regulation on intellectual property within Universidad EAFIT, the results of this research practice are product of *Juliana Restrepo Tobar* and *Sara Arango Franco*

In case further products, beside academic articles, that could be generated from this work, the intellectual property distribution related to them will be directed under the current regulation of this matter determined by Universidad EAFIT (2017).

References

Arango, Jaime Horacio. 2021. Medellín tiene su propio "Robocop" para vigilar las calles. *El Colombiano*.

Armitage, Rachel, & Monchuk, Leanne. 2019. What is CPTED? Reconnecting theory with application in the words of users and abusers. *Policing: A Journal of Policy and Practice*, **13**(3), 312–330.

Armitage, Rachel, Rogerson, Michelle, & Pease, Ken. 2013. What is good about good design? Exploring the link between housing quality and crime. *Built Environment*, **39**(1), 140–161.

Blau, Judith R, & Blau, Peter M. 1982. The cost of inequality: Metropolitan structure and violent crime. *American sociological review*, 114–129.

Blumstein, Alfred, Cohen, Jacqueline, & Nagin, Daniel. 1978. *Deterrence and incapacitation: Estimating the effects of criminal sanctions on crime rates.* National Academy of Sciences Washington, DC.

Butke, Paul, & Sheridan, Scott C. 2010. An analysis of the relationship between weather and aggressive crime in Cleveland, Ohio. *Weather, Climate, and Society*, **2**(2), 127–139.

Casa de las Estrategias, SISC. 2021. *Un análisis institucional del hurto a personas: recuperar el derecho a la ciudad y vidas perdidas.* Tech. rept. Alcaldía de Medellín.

Cerdá, Magdalena, Morenoff, Jeffrey D., Hansen, Ben B., Tessari Hicks, Kimberly J., Duque, Luis F., Restrepo, Alexandra, & Diez-Roux, Ana V. 2012. Reducing Violence by Transforming Neighborhoods: A Natural Experiment in Medellín, Colombia. *American Journal of Epidemiology*, **175**(10), 1045–1053.

Cohen, Lawrence E, & Felson, Marcus. 1979. Social change and crime rate trends: A routine activity approach. *American sociological review*, 588–608.

Cohn, Ellen G, & Rotton, James. 2000. Weather, seasonal trends and property crimes in Minneapolis, 1987–1988. A moderator-variable time-series analysis of routine activities. *Journal of environmental psychology*, **20**(3), 257–272.

Cozens, Paul, & Love, Terence. 2015. A review and current status of crime prevention through environmental design (CPTED). *Journal of Planning Literature*, **30**(4), 393–412.

Donovan, Geoffrey H, & Prestemon, Jeffrey P. 2012. The effect of trees on crime in Portland, Oregon. *Environment and behavior*, **44**(1), 3–30.

Gómez, Santiago, Mejía, Daniel, & Tobón, Santiago. 2021. The deterrent effect of surveillance cameras on crime. *Journal of Policy Analysis and Management*, **40**(2), 553–571.

Inlow, Alana R. 2021. A comprehensive review of quantitative research on crime, the built environment, land use, and physical geography. *Sociology compass*, **15**(7), e12889.

Jacobs, Jane. 1993. *The death and life of great American cities.* Vintage Books.

Jaitman, Laura, Caprirolo, Dino, Granguillhome Ochoa, Rogelio, Keefer, Philip, Leggett, Ted, Lewis, James Andrew, Mejía-Guerra, José Antonio, Mello, Marcela, Sutton, Heather, & Torre, Iván. 2017. *The costs of crime and violence: New evidence and insights in Latin America and the Caribbean.* Vol. 87. Inter-American Development Bank Washington, DC.

Jeffery, Clarence Ray. 1977. *Crime prevention through environmental design.* Vol. 524. Sage Publications Beverly Hills, CA.

Mccord, Eric S., & Ratcliffe, Jerry H. 2009. INTENSITY VALUE ANALYSIS AND THE CRIMINOGENIC EFFECTS OF LAND USE FEATURES ON LOCAL CRIME PATTERNS.

Patino, Jorge E., Duque, Juan C., Pardo-Pascual, Josep E., & Ruiz, Luis A. 2014. Using remote sensing to assess the relationship between crime and the urban layout. *Applied Geography*, **55**, 48–60.

Pratt, Travis C, & Cullen, Francis T. 2005. Assessing macro-level predictors and theories of crime: A meta-analysis. *Crime and justice*, **32**, 373–450.

Shaw, Clifford, & McKay, Henry. 1929. Delinquency Areas. *Chicago: University of Chicago Press.*

SISC. 2019. *Indicadores de Seguridad y Convivencia Enero del 2016 a Octubre del 2019.*

Sorg, Evan T, & Taylor, Ralph B. 2011. Community-level impacts of temperature on urban street robbery. *Journal of Criminal Justice*, **39**(6), 463–470.

Troy, Austin, Grove, J Morgan, & O'Neil-Dunne, Jarlath. 2012. The relationship between tree canopy and crime rates across an urban–rural gradient in the greater Baltimore region. *Landscape and urban planning*, **106**(3), 262–270.

Universidad EAFIT. 2017. *Reglamento de propiedad intelectual.*

Weaver, Russell. 2015. Littering in context (s): using a quasi-natural experiment to explore geographic influences on antisocial behavior. *Applied Geography*, **57**, 142–153.

Wilcox, Pamela, Quisenberry, Neil, Cabrera, Debra T, & Jones, Shayne. 2004. Busy places and broken windows? Toward defining the role of physical structure and process in community crime models. *Sociological Quarterly*, **45**(2), 185–207.