

A Location Discrete Choice Model of Crime: Police Elasticity and Optimal Deployment

Douglas Newball-Ramírez ¹, Álvaro J. Riascos Villegas ^{1,2,*}, Andrés Hoyos ², Mateo Dulce Rubio ³,

1 Quantil, Bogotá, Colombia

2 Facultad de Economía, Universidad de los Andes, Bogotá, Colombia

3 Department of Statistics & Data Science and Heinz College of Information Systems & Public Policy, Carnegie Mellon University, Pittsburgh, PA

4 Facultad de Ciencias, Departamento de Matemáticas, Universidad Nacional de Colombia, Bogotá, Colombia

* ariascos@uniandes.edu.co

Abstract

Despite the common belief that police presence reduces crime, there is mixed evidence of such causal effects in major Latin America cities. In this work we identify the casual relationship between police presence and criminal events by using a large dataset of a randomized controlled police intervention in Bogotá D.C., Colombia. We use an Instrumental Variables approach to identify the causal effect of interest. Then we consistently estimate a Conditional Logit discrete choice model with aggregate data that allow us to identify agents' utilities for crime location using Two Stage Least Squares. The estimated parameters allow us to compute the police own and cross-elasticities of crime for each of the spatial locations and to evaluate different police patrolling strategies. The elasticity of crime to police presence is, on average across spatial locations, -0.26 for violent crime, -0.38 for property crime and -0.38 for total crime, all statistically significant. Estimates of cross-elasticities are close to zero; however, spillover effects are non-negligible. Counterfactual analysis of different police deployment strategies show, for an optimal allocating algorithm, an average reduction in violent crime of 7.09% , a reduction in property crimes of 8.48% and a reduction in total crimes of 5.15% at no additional cost. These results show the potential efficiency gains of using the model to deploy police resources in the city without increasing the total police time required.

Introduction

Crime prediction is now ubiquitous in crime prevention and police resource planning. Studies on crime prediction and identification of risky spatial regions –the so-called “hot spots”– already exist in a vast quantity: CPFais, CPGCNN, CPglog, CPHu, CPJin2019, CPKupilik2018, CPkw, CPMohler2011, CPMohler2014, CPreinhart2017self, CPStec2018, to name a few. However, identifying and quantitatively characterizing the causal drivers of crime is a harsh scientific problem: the optimal allocation of police resources is guided by police deployment strategies (e.g., prediction models), which at the same time determine what crime incidents are reported or how crime is displaced from one sector to another in a city. In fact, one can hardly take a stance on the causal

relationship between police presence in a particular spot and crime incidents. On the one hand, the police presence at a particular spot tends to deter crime, but at the same time crime incidents call for police presence at specific sectors of a city. Thus, it is uncertain whether police presence is a cause or a consequence of the level of crime.

Untangling this causal relationship is an instance of what holland has called the fundamental problem of causal inference: *what would have happened had the police not been present in a particular sector of the city?* The gold standard to address this causal problem is the use of experimental data specially tailored toward this goal (i.e., a randomized controlled trial, or RCT). Running RCTs to evaluate the effect of patrols, time exposure or other intervention strategies in different regions of the city is difficult due to ethical concerns and high costs. Nevertheless, there are already several experimental or quasi-experimental studies addressing this issue, especially for US cities, but the evidence on the effect of increasing police presence in hotspots has been mixed, specially for major cities in Latin America (see NAP24928 for a comprehensive study of proactive policing in the US).

In this study, we use blattman2021place dataset which studies a placed-based police and city services intervention in Bogotá, Colombia.¹ blattman2021place found that increasing police presence has modest direct impacts, even when focusing on the highest-crime hot spots, and with crime displaced nearby ruling out total reductions in crime of more than 2%. We propose a different identification strategy using an Instrumental Variables approach and a discrete choice model of spatial selection of crime to study the impact of police time exposure on crime levels in different sectors of the city. In detail, we are interested in studying the average causal effect on crime reduction of increasing police patrol time. Moreover, we are interested in estimating the own- and cross-elasticities of crime to patrol time at different locations, which captures the percentage change in crime, if any, that results from a 1% increase in the police patrol time that such location receives. These can also be interpreted as heterogeneous treatment effects for each spatial location. Finally, we use the estimated elasticities to compare counterfactual policy scenarios in search of more efficient patrolling strategies.

To summarize, our contribution is: (1) We use a unique large dataset of experimental data that allows for the identification of the causal effect of police patrolling on crime, (2) We estimate a spatial discrete choice model with aggregate data that allow us to identify agents' utilities, (3) Provided our estimates of the structural parameters that determine the crime location choice of the offenders, we estimate the police own- and cross-elasticity of crime for each of the quadrants (i.e., we compute where police patrolling is more effective), and (4) Based on the same structural parameters and the estimated elasticities, we evaluate different police patrolling counterfactual strategies without increasing the total police time available. In particular, we test what would have happened had the police patrol time been allocated: (a) uniformly across quadrants, (b) proportional to the incidence of crimes, (c) such that the more insecure and elastic quadrants receive a 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100% increase, and (d) optimally: we solve an optimization problem for the best allocating strategy. From these counterfactual exercises we further contribute to the understanding of which police patrol deployment strategies are more efficient. Our results are in line with [?], but since we estimate a structural model we are able to evaluate counterfactual scenarios that significantly reduce crime.

The rest of the paper is organized as follows. In the next subsection we summarize the relevant Related Literature and frame our work within the literature investigating the causal effect of police presence on crime events. The Materials and Methods section details the spatial discrete-choice model we use to model criminals' choices of where to commit crimes. Thereafter, we present our Empirical Strategy to estimate the

¹We are thankful to the authors for allowing us to use their data set for this study.

parameters of interest along with our IV approach to identify the causal effect of police exposure on crime, while the Data section describes the experimental dataset used. In the Results section we analyze the estimations obtained regarding the causal effect of police presence on the level of crime, and the own- and cross- elasticities of such effect for each region of the city. Finally, in the Counterfactual Policy Scenarios section we exploit the structural model estimated to compare different counterfactual patrolling strategies. The Discussion section summarizes our findings and discusses limitations and next steps of our work.

Related Literature

Our work is mainly related to the literature on the causal effects of proactive police exposure on criminal activity. The evidence on such causal effect is mixed (see NAP24928 for a comprehensive study for the US). To provide positive evidence, braga used meta-analytic techniques to assess the impact of disorder policing on crime. They identified 30 randomized experimental and quasi-experimental tests of disorder policing suggesting that policing disorder strategies are associated with an overall statistically significant modest crime reduction. telep conducted a systematic review and a meta-analysis examining the extent to which there is crime displacement or a diffusion of crime in medium-sized or large geographic areas. They reviewed 19 publications covering 20 quasi-experimental studies. They found no significant overall evidence of displacement or a diffusion of benefits. Ratcliffe reported the results of a randomized controlled trial of police effectiveness across 60 violent crime hot spots in Philadelphia. Their results suggested a significant reduction in the level of violent crime for the treated area after 12 weeks. Moreover, they showed that targeted areas in the top 40% of pre-treatment violent crime counts (i.e., the most critical regions) had significantly less violent crime during the operational period. Targeted areas outperformed the control sites by 23%, resulting in a total net effect (once displacement was considered) of 53 violent crimes prevented. Novak examined the effectiveness of foot patrol in violent micro-places in Kansas City. They studied the effects of deployed foot patrol in hot spots for a period of 90 days. They employed a quasi-experimental design comparing four treatment to four control areas, estimating panel-specific autoregressive models for 30 weeks prior to and 40 weeks after the treatment. Their results reveal statistically significant short-run reductions in violent crime in the micro-places receiving foot patrol treatment, while no such reductions were observed in the control areas. At the same time they found no evidence of crime displacement to spatially contiguous areas. Similarly, andresen2019spatial study the spatial effect of police foot patrol in British Columbia, finding that the policy was effective in reducing crime when compared to the crime patterns in the same area before the program. This effect was concentrated in property crimes with no evidence of spatial crime displacement.

In sharp contrast, some studies in Latin America found no or even negative effects of increased policing on crime levels. Among these, we highlight blattman2021place, a place-based police and city services intervention at scale for Bogotá D.C., Colombia. The authors randomly assigned 1,919 streets to an 8-month treatment of doubled police patrols, greater municipal services, both, or neither. They studied the direct and spillover effects of such targeted state services. They found that increasing state presence has modest direct impacts, even when focusing on the highest-crime hot spots and with crime displaced nearby. Confidence intervals suggest they can rule out total reductions in crime of more than 2%. In the same line, collazos2021hot study the effect of an increase in police patrol by randomly assigning 384 out of 967 hotspots in Medellín, Colombia, to a six-months experiment finding no direct effects of increased police presence on crime reduction (besides car thefts). Similarly, galiani2022predictive study the effect on the level of crime of increased police presence in hotspots flagged by

a predictive policing software in Montevideo, Uruguay. Using data from an RCT, the authors found no statistically significant reduction in the overall crime or robberies. Despite the widespread belief that police presence significantly reduces crime, and the positive evidence in the literature for large cities, this does not appear to be the case for major cities in Latin America.

Finally, we contribute in a lesser extent to the literature on discrete choice modeling of crime phenomena. The pioneer applications of spatial discrete choice modeling of crime are bernasco2005, brown2006 and bernasco2010. In bernasco2005, the authors studied the selection of crime (burglary) locations in the city of The Hague, Netherlands. They evaluate several interesting hypotheses about crime: Are more valuable properties more attractive to burglars? Do higher mobility, neighborhood ethnicity, distance to burglar's home, distance to city center, etc., have a causal effect on crime? They used a conditional logit model with individual sociodemographic data of 290 burglars who committed 548 burglaries in the city during the period 1996-2001. Their results support and quantify some of the working hypothesis but there is no causal identification of the effect of police patrolling on crime. Given that the the i.i.d. hypothesis used in bernasco2005 for maximum likelihood estimation is controversial due to spatial correlation, bernasco2010 studied the role of spatial resolution and alternative modeling strategies. Finally, brown2006 used a spatial choice model coupled with clustering techniques to estimate a mixture model of crime. Their model focuses on crime prediction and does not address the causal problem that we address in this paper.

Materials and Methods

Proposed model

The model we estimate in this paper is an adaptation of the discrete-choice model of demand introduced by berry1994estimating. In his paper, berry1994estimating describes a –rather parsimonious– methodology that (1) allows one to estimate the structural parameters that drive the demand for differentiated products using aggregate data of product market shares, and (2) allows one to correct for the endogeneity of prices (and other product characteristics) by using Instrumental Variables (IV) and estimating by Two Stages Least Squares (TSLS). Thus, the discrete choice model we present in this paper is estimated using aggregate data of reported (and mainly unsolved) crimes in different locations of Bogotá D.C., police presence in each of the locations, and some other characteristics of the location that might potentially affect crime presence in each of the locations. Our model is described next.

Consider N potential criminal offenders with *symmetric preferences*,² each of them living in Bogotá D.C. (or in its surroundings) and deciding between $J + 1$ locations in the city where to commit a crime. Each potential offender bases her location choice on her perceived utility of committing a crime in each of the $J + 1$ locations. The associated utility u_{ij} of agent i of selecting location j to commit a crime is given by

$$u_{ij} = \alpha P_j + X_j \beta + \xi_j + \varepsilon_{ij} \quad (1)$$

where P_j is a measure of the police presence in location j , X_j is a vector of K observed characteristics of the location, ξ_j is the constant term and might be thought as the mean utility associated to unobserved (by the researcher) characteristics of location j , α and β are unknown parameters that determine the influence of P_j and X_j on u_{ij} , respectively, and ε_{ij} is the idiosyncratic error term.

²This assumption can be relaxed, but it is not under the scope of this paper.

It follows that the probability that potential offender i selects location j is given by

$$\begin{aligned}
P(i \text{ chooses } j) &= P(u_{ij} \geq u_{ik} \ \forall \ k \neq j) \\
&= P(u_{ij} \geq u_{i0}, \dots, u_{ij} \geq u_{iJ}) \\
&= P(\delta_j + \varepsilon_{ij} \geq \delta_0 + \varepsilon_{i0}, \dots, \\
&\quad \delta_j + \varepsilon_{ij} \geq \delta_J + \varepsilon_{iJ}) \\
&= P(\varepsilon_{ij} - \varepsilon_{i0} \geq -(\delta_j - \delta_0), \dots, \\
&\quad \varepsilon_{ij} - \varepsilon_{iJ} \geq -(\delta_j - \delta_J))
\end{aligned} \tag{2}$$

where $\delta_j = \alpha P_j + X_j \beta + \xi_j$ and $j = 0$ corresponds to the *outside option* of not committing a crime in any of the J locations of the city. We denote the probability described by equation (2) as $s_{ij}(P_j, X_j, \xi_j; \alpha, \beta)$.

Assuming that all the ε_{ij} terms are i.i.d and follow an extreme value type I distribution, we have, for any pair of different locations j and j' , that $\varepsilon_{ij} - \varepsilon_{ij'}$ follows a logistic distribution and, by equation (1), location choice probabilities have a closed-form expression given by:

$$s_{ij}(P_j, X_j, \xi_j; \alpha, \beta) = \frac{\exp(\delta_j)}{1 + \sum_{k=1}^J \exp(\delta_k)} \tag{3}$$

where, again, option $j = 0$ is assumed to be the *outside option* and, thus, δ_0 is normalized to zero.³ This *outside option* in this particular context can be interpreted as the choice of not committing a crime at all, or committing a crime but not at any of the street segments of our data set.⁴

From equation (3) it follows that the share of committed crimes at location j is:

$$\begin{aligned}
S_j(P_j, X_j, \xi_j; \alpha, \beta) &= \int_i s_{ij}(P_j, X_j, \xi_j; \alpha, \beta) \phi(i) di \\
&= s_{ij}(P_j, X_j, \xi_j; \alpha, \beta) \int_i \phi(i) di \\
&= s_{ij}(P_j, X_j, \xi_j; \alpha, \beta) \\
&= \frac{\exp(\delta_j)}{1 + \sum_{k=1}^J \exp(\delta_k)}.
\end{aligned} \tag{4}$$

This equivalence between $S_j(P_j, X_j, \xi_j; \alpha, \beta)$ and $s_{ij}(P_j, X_j, \xi_j; \alpha, \beta)$ is a result of the symmetric preferences assumption: $s_{ij}(P_j, X_j, \xi_j; \alpha, \beta)$ does not depend on individual characteristics.

An advantage of this model is that, contrary to the traditionally implemented Poisson Model in target-based studies of crime location, it allows us to estimate own- and cross-elasticities in a fairly direct way (see schmidheiny2011equivalence for a detailed explanation.) Particularly, from equation (4) we can easily derive the own- and cross-elasticities of crime with respect to police presence P_j (or any observed characteristic $x_{rj} \in X_j$), which captures the percentage change in crime when police presence is increased by 1%. In particular, the derivatives with respect to P_j are given by:

³The final logistic specification may be questionable since it implies the well-known assumption of independence of irrelevant alternatives (IIA): the relative odds of choosing location i over j is the same no matter what other alternatives are available. To test for the validity of this assumption, we follow hausman1984specification's specification test; the results are reported in Appendix ??.

⁴It can include some other choices such as committing a crime different from those we have in our data set such as a cybernetic crime or committing a crime at another (nearby) city or town.

$$\frac{\partial S_j}{\partial P_\ell} = \begin{cases} \alpha S_j(1 - S_j) & \text{if } j = \ell \\ -\alpha S_j S_\ell & \text{if } j \neq \ell \end{cases} \quad (5)$$

and thus, the police own- and cross-elasticities of crime are:

$$E_{S_j, P_\ell} \equiv \frac{\partial S_j}{\partial P_\ell} \frac{P_\ell}{S_j} = \begin{cases} \alpha(1 - S_j)P_j & \text{if } j = \ell \\ -\alpha S_\ell P_\ell & \text{if } j \neq \ell \end{cases} \quad (6)$$

The model described in this section contrasts both with the Conditional Logit (CL) model, which is commonly used in the crime location discrete choice framework, and the Poisson Model, mainly used in target-based crime location studies [?, 2]. Particularly, this model can be thought of as an aggregated version of the traditional Conditional Logit model that holds the advantages of the former and the latter. Regarding the advantages of the CL model, our model is also based on Random Utility Maximization (RUM) theory. This allows us to directly derive the statistical model to be estimated and, therefore, recover the structural parameters that drive the crime location choice [?, 1].

On the other hand, our model does not represent a “zero-sum” world where one more crime in one location represents exactly one less crime in another location, as is the case of the CL model. This means that our model allows for changes in attributes (such as police presence) to result in an overall change in the city’s level of crime, which is an advantage of the Poisson Model. Additionally, in line with the traditional CL model and in contrast with the Poisson model, our model allows for own-elasticities not to be constant and cross-elasticities to be different from zero (see schmidheiny2011equivalence for a detailed discussion). Lastly, our model can be easily estimated by Two Stages Least Squares (TSLS), which, in turn, allows us to deal with endogeneity with relative ease. That is not the case with Maximum Likelihood, which is more “parametrically” restrictive.

Empirical Strategy

Main specification

To estimate the structural parameters $\theta = (\alpha, \beta)$ from equation (1) we follow berry1994estimating and note that:

$$\delta_j \equiv \log(S_j) - \log(S_0) = \alpha P_j + X_j \beta + \xi_j, \quad (7)$$

and thus equation (7) represents our ideal specification. In this equation, the definitions of the variables are the same as those mentioned in section ???. Namely, P_j is a measure of police presence at location j and X_j is the set of measured characteristics for the same location (including an all-ones vector). ξ_j , in this case, represents the error term.

On the other hand, S_j for $j \in \{1, \dots, J\}$ is defined as the proportion of crimes committed at location j . That is,

$$S_j = \frac{C_j}{N} \quad (8)$$

where C_j is the number of crimes committed at location j and N is the number of potential offenders. S_0 , on the contrary, is not directly observed since it captures the

choice of the *outside option* – i.e., not committing a crime in any of the locations. However, we exploit the fact that

$$\sum_{j=0}^J S_j = 1. \quad (9)$$

And, therefore, we recover S_0 as

$$S_0 = 1 - \sum_{j=1}^J S_j \equiv 1 - \sum_{j=1}^J \frac{C_j}{N}. \quad (10)$$

Now that we know all the components from equation (7), we can easily estimate it by OLS. However, the OLS estimation faces one problem: P_j is endogenous. In particular, under the assumption that the police force is assigned according to the number of committed crimes in each of the locations, there exists simultaneity between δ_j and P_j .⁵

Thus, OLS estimates from equation (7) are plausibly biased and inconsistent.

Dealing with endogeneity: Two-Stage Least Squares (TSLS)

In 2016, blattman2021place, along with the National Police of Colombia and the Mayor's Office of Bogotá, designed and implemented a multi-arm security experiment at the level of street segments. Two different types of intervention were randomly delivered: 1) an increased police patrol time and 2) an improvement of the delivery of city services (street lighting and cleanup). Starting in January 2016 and during 8 months, 756 out of 1,919 street segments labeled as crime hot spots –out of the 136,984 street segments of the city– received a doubled patrolling time (92-167 minutes of police patrol per day) [?]. Also, in March 2016, 201 of the 1,919 hot spots received more intensive street light repair and cleaning [?].⁶ In this paper, we exploit the random assignment of the first type of treatment to instrument the police presence P_j , identify the structural parameters of interest, and consistently estimate them by TSLS. We take X_j as given and exogenously determined. Thus, we do not instrument them.

Following 10.2307/2951620 and angrist1995two, the needed assumptions for the treatment assignment, Z , to be a valid instrument for P_j are: 1) independence, 2) exclusion restriction, 3) rank condition (relevance), and 4) monotonicity (no defiers). First, given that the police patrol treatment was randomly assigned within the hot spot street segments, Z is plausibly independent of the potential outcomes (crime shares ratios) and the potential treatment status (police patrol time) once it is controlled for the hot spot label status (i.e., $1\{j \text{ is a crime hot spot}\} \in X_j$) [?]. Second, since the treatment arm that we consider in this paper only determines police patrol time, it is plausible that Z only affects δ_j through P_j . Third, according to the findings of blattman2021place the “police complied with their new orders for the full 8 months.” Thus, the rank condition is plausible and has been already verified. Fourth, given that the police officers were monitored via GPS every 30 seconds and police officers (and workers in general) plausibly double their efforts in a task only when they are ordered or incentivized to do so, it is reasonable that the monotonicity condition holds.⁷ Furthermore, the experimental design in blattman2021place intended that the patrol time of untreated streets remained the same, which was empirically verified by the authors. Thus, under the blattman2021place setting, Z is a valid instrument for P_j , when each j corresponds to a different street segment of the city.

⁵ δ_j is not directly a measure of crimes but it is a function of them. So, if there exists simultaneity between C_j and P_j , there exists simultaneity between δ_j and P_j .

⁶For further information about the different treatment arms see blattman2021place.

⁷That is, only treated segments received increased patrol time. Control segments remained the same.

As is explained in Section , however, we do not estimate our model with street-level data. Instead, we use quadrant-level data, which represents a higher level of aggregation unit. In this sense, the instrument we exploit for P_j is the proportion of street segments within quadrant j that were treated. Controlling now for the proportion of hot spots street segments within quadrant j , the assumptions that we discussed in the previous paragraphs naturally extend to this aggregated case and plausibly hold. Since the impact of the instrument on P_j depends on the initial level of police presence, we additionally control for the baseline (i.e., before the experiment) level of police presence to avoid possible endogenous relationships.

Under the aforementioned assumptions, which, as was discussed, seem to hold, the IV methodology allows us to correctly identify the parameter vector θ and to consistently estimate it by TSLS. In this case, the parameter α –which is our main focus in this paper– captures a weighted average of all the possible per-unit causal responses of δ_j to a marginal change in P_j caused by Z (angrist1995two).⁸

Selecting location characteristics - Double Selection

In general, city locations might have several characteristics that can influence the presence of crime. When the set of available variables that might potentially influence crime location is vast, including all of them in the model to be estimated might result in power loss issues and high multicollinearity. This, in turn, might reduce the traceability of what really matters for the crime location choice. A natural solution to this problem is to manually select the covariates to be included. However, this might be cumbersome when the list of available variables is large, and also might lack rigor since is a subjective decision. To avoid this issue, we follow belloni2014high and implement the double selection methodology to select the variables that should be included in X_j .

Following belloni2014high, we first (separately) run a regularized lasso over the following two equations

$$\delta_j = \tilde{X}_j \gamma + \mu_j, \quad (11)$$

$$P_j = \tilde{X}_j \vartheta + \lambda_j, \quad (12)$$

where δ_j and P_j are the dependent and independent variables from equation (7), respectively, and \tilde{X}_j is a vector of variables that includes all the available and exogenous location features that potentially affect crime location. γ and ϑ are vectors of coefficients associated with \tilde{X}_j in each equation, while μ_j and λ_j are the error terms.

Lasso algorithm applied to equations (11) and (12) selects the relevant location features that better predict the outcome variable δ_j and the police presence P_j , respectively. Given the selected variables from both estimations, we then estimate equation (7) by TSLS using $X_j = \tilde{X}_j^{DS,\delta} \cup \tilde{X}_j^{DS,P}$, where $\tilde{X}_j^{DS,\delta}$ and $\tilde{X}_j^{DS,P}$ are the set of Lasso-selected location characteristics from equations (11) and (12).

Data

In this work, we use the data set used by blattman2021place in their study. In particular, we use their data on violent, property, and total criminal official records at the quadrant level (i.e., an administrative and irregular subdivision of the city used for assigning police shifts; two police officers per quadrant per shift) that occurred in Bogotá D.C. during the experiment ran by them in 2016. We also use information of the following characteristics at the quadrant level: proportion of paved street segments, proportion of street segments in use by industry and commerce, proportion of street

⁸For more technical details check angrist1995two.

segments in use by housing, proportion of low income street segments, proportion of middle income street segments, proportion of high income street segments, average distance from each street segment to the nearest shopping center, average distance from each street segment to the nearest education center, average distance from each street segment to the nearest park or recreational center, average distance from each street segment to the nearest religious center, average distance from each street segment to the nearest health center, average distance from each street segment to the nearest services center (e.g., justice), average length of street segments, the average built meters per meter of street segment of length 100 meters around each street segment, and the proportion of street segments labeled as crime hot spots. Lastly, we take information on the average minutes of police patrol time that each segment received during the intervention and the proportion of street segments assigned to increased patrol time within each quadrant.

Provided this data set, we estimate three versions of equation (7): one for violent crimes, another for property crimes, and the last one for total crimes. To construct each crime-specific δ_j , we assume that the number of potential criminal offenders is given by the estimated number of unemployed (actively job-seeking or inactive) people aged 12 to 60 in the period of interest. N is calculated using the public data of unemployment and population projections of the (Colombian) National Administrative Statistics Department (DANE from its Spanish initials).⁹ On the other hand, P_j is given by average minutes of police patrol time (from blattman2021place), Z_j is the proportion of treated street segments, and X_j is selected from the list of variables mentioned above.

Descriptive statistics of these variables at the quadrant level (1,050 spatial units) are presented in Table 1. Panel A reports descriptive statistics of the reported crime within each quadrant during 2016, panel B reports descriptive statistics of the average police patrol time in minutes within each quadrant during and before the intervention, and panel C reports descriptive statistics of the average characteristics of the street segments within each quadrant. As can be seen, property crimes seem to be more frequent than violent crimes. On average there were reported 11.84 violent crimes, 23.86 property crimes, and a total of 35.07 crimes. These types of crimes ranged between 0 to 70, 0 to 149, and 0 to 157, respectively. This depicts the high heterogeneity that exists in terms of security in Bogotá D.C. Panel B, on the other hand, shows that on average the patrol time received by the street segments within each quadrant during the intervention is about 46 minutes. However, this time varies from 4.98 to 711.23. This again reflects the high heterogeneity of the security in the city. Also, this contrasts with the average patrol time at baseline, which is about 19.68 minutes (42% of the time during intervention). It seems that the intervention did effectively increase police presence. Lastly, panel C reports summary statistics of the variables to be added to the vector X_j .

Figure 1 reports scatter plots –and their respective linear regression fit line– of the logarithm of the reported crimes against the logarithm of the police patrol time. This can be thought of as a first *naive* approximation to the estimation of the police elasticity of crime.¹⁰ The figure shows no relationship between police presence and any of the reported crimes. OLS estimates are -0.02 for violent crimes, 0.05 for property crimes, and 0.01 for total crimes. None of them is statistically significant. This counterintuitive result suggests that, if any, the police elasticity of crime is positive. These results, however, might be driven by the simultaneity that exists between crimes and police presence, and therefore, justify the need to account for endogeneity as was

⁹ $N \approx 791,223$. Data available at [https://dane.gov.co/index.php/estadisticas-por-tema-mercado-laboral-historicos#marco-2005](https://dane.gov.co/index.php/estadisticas-por-tema-mercado-laboral/empleo-y-desempleo/mercado-laboral-historicos#marco-2005) and <https://www.dane.gov.co/index.php/estadisticas-por-tema/demografia-y-poblacion/proyecciones-de-poblacion>.

¹⁰Since both dependent and independent variables are in logarithms, the slope of the relationship captures an elasticity.

Table 1. Descriptive statistics of Bogotá D.C.'s quadrants (2016)

	N	Mean	SD	Min	Max
<i>A. Reported Crimes</i>					
Violent Crimes	1,050	11.84	8.56	0.00	70.00
Property Crimes	1,050	23.86	17.88	0.00	149.00
Total Crimes	1,050	35.70	21.42	0.00	157.00
<i>B. Police Presence</i>					
Avg. police patrol time (minutes)	1,050	46.58	52.09	4.98	711.23
Baseline avg. police patrol time (minutes)	1,050	19.68	18.47	3.79	248.03
<i>C. Quadrant characteristics</i>					
Avg. longitude of street segments (mt)	1,050	64.20	23.44	21.52	263.38
Avg. built squared meters per street segment meter	1,050	22,786.17	13,602.68	3.19	148,750.19
Avg. distance to nearest shopping center	1,050	739.00	706.64	21.72	4,686.01
Avg. distance to nearest educational center	1,050	319.18	234.42	55.98	2,767.30
Avg. distance to nearest park/recreational center	1,050	680.85	390.99	37.57	2,855.38
Avg. distance to nearest police station	1,050	603.29	362.46	60.47	2,943.86
Avg. distance to nearest religious/cultural center	1,050	501.72	365.22	28.90	3,829.53
Avg. distance to nearest health center	1,050	951.65	715.99	55.56	7,499.75
Avg. distance to nearest additional services office	1,050	698.72	575.83	39.38	5,193.83
Avg. distance to nearest transport infrastructure	1,050	102.69	65.70	17.13	1,091.14
Avg. distance to closest surveillance camera	1,050	592.81	522.62	17.72	6,051.21
Prop. of street segments zoned for industry/commerce	1,050	0.25	0.22	0.00	1.00
Prop. of street segments zoned for housing	1,050	0.66	0.27	0.00	1.00
Prop. of middle income street segments	1,050	0.45	0.48	0.00	1.00
Prop. of high income street segments	1,050	0.08	0.26	0.00	1.00
Prop. of hot spot street segments	1,050	0.03	0.06	0.00	1.00
Prop. of treated hot spot street segments	1,050	0.01	0.03	0.00	0.50

Notes: Variables at the quadrant level were constructed from the aggregation or averaging of the original data set variables at the level of street segment. Reported crimes were aggregated. The remaining variables were averaged. Thus, summary statistics for these latter types of variables can be thought as *the mean/SD/min/max of the average* of the variable. An average of a dichotomous variable is reported as a proportion. Distances measured in meters. Avg.: Average. Prop.: Proportion. SD: Standard Deviation. Min: Minimum. Max: Maximum.

discussed in the Empirical Strategy section.

332

Results

333

We divide our results in two main subsections. In the first subsection, we present the results of the estimation of our discrete choice model of crime (i.e., structural model). In the second subsection, we use our estimated structural model of crime to study the crime consequences of some counterfactual scenarios of police allocation.

334

335

336

337

338

Structural model

339

α estimates

340

Table 2 presents, for each of the three types of crimes, the TSLS estimates of α from four different specifications of equation (7). The first specification corresponds to one where no covariates were included in X_j apart from the proportion of hot spots in the quadrant, which is needed for the instrument to be valid. In the second specification, we add the quadrant characteristics that were selected by the Double Selection algorithm [?]. In the third one, we additionally control for the baseline police patrol time, and in the fourth one, we add locality fixed effect. A locality is the greater administrative division of Bogotá D.C., and each of them contains several quadrants within. It is known that the patterns of crime in each locality are different, and thus, by including the locality fixed effects, we control for those differential patterns.

341

342

343

344

345

346

347

348

349

350

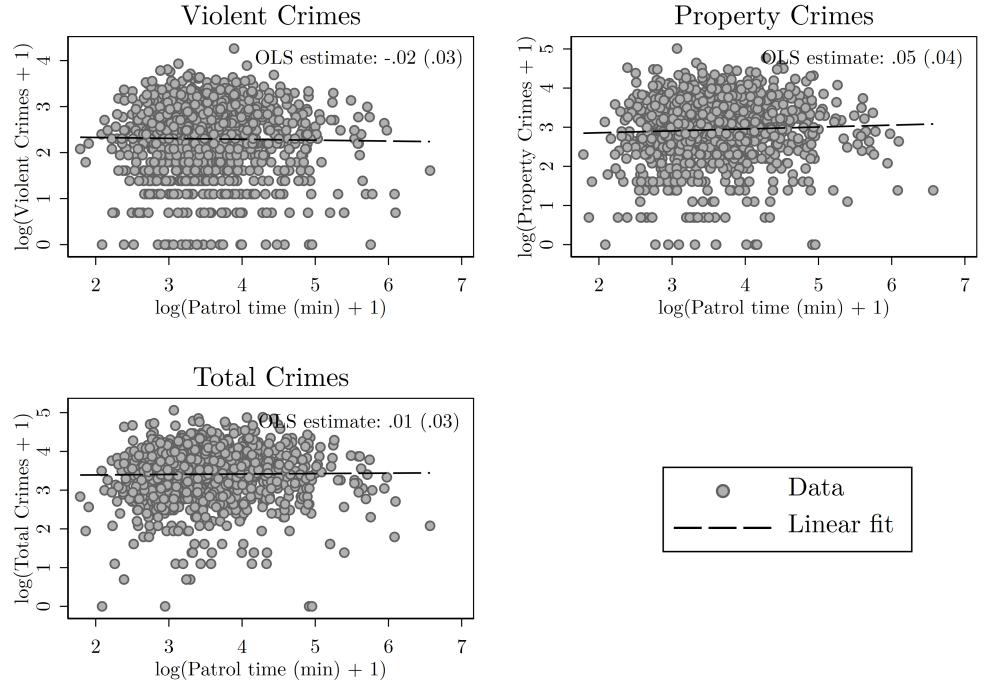


Fig 1. Scatter plots of the logarithm of each type of crime against the logarithm of the average patrol time within each quadrant. Linear fit displayed by the black dashed line. These graphs display a naive calculation of the police elasticity of crime. OLS coefficient estimates and their standard errors in parentheses reported at the top-right corner of each graph. * $p<0.1$; ** $p<0.05$; *** $p<0.01$. Source: Own elaboration.

As can be seen, the estimates are fairly consistent across specifications, and thus, we interpret the fourth one – our preferred specification. We find that police presence in a certain location has, on average, a negative impact on the utility of committing either a violent, a property, or any crime in it. An increase in one minute of police patrol time in location j reduces, on average, 0.006 units the utility of committing a violent crime and 0.008 units the utility of committing either property or any crime in that location. These effects are statistically significant at the 5% for violent crimes and 10% for property and total crimes. Thus, it follows from these estimates that police presence reduces the probability of a potential criminal committing a crime in a protected location.

Table 2. TSLS α estimates for the discrete spatial location choice model

α	Violent crimes				Property crimes				Total crimes			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
α	-0.004*	-0.003*	-0.005*	-0.006**	-0.004**	-0.005***	-0.009**	-0.008*	-0.005**	-0.005***	-0.008**	-0.008*
	(0.002)	(0.001)	(0.003)	(0.003)	(0.002)	(0.002)	(0.004)	(0.005)	(0.002)	(0.002)	(0.004)	(0.004)
Observations	1,050	1,050	1,050	1,050	1,050	1,050	1,050	1,050	1,050	1,050	1,050	1,050
Controls	No	Yes	Yes	Yes	No	Yes	Yes	No	Yes	Yes	Yes	Yes
Past police presence	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
Locality FE	No	No	No	Yes	No	No	Yes	No	No	No	No	Yes
R-squared	-0.023	0.309	0.310	0.380	0.000	0.339	0.340	0.414	-0.035	0.187	0.190	0.281

Notes: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Cluster robust standard errors at the level of locality in parentheses.

According to guimaraes2003tractable, the estimates of a traditional Conditional Logit model are equivalent to those of a Poisson Model with aggregated data. This is true since the first-order conditions of the likelihood maximization of both models are

equivalent [?]. Therefore, in order to compare our model to the Poisson and the traditional CL logit model, we additionally estimate a target-based Poisson Model using the number of reported crimes in each quadrant as the dependent variable. We, again, deal with the endogeneity by instrumenting P_j with the treatment assignment condition. However, in this case, estimates are recovered either by the Generalized Method of Moments (GMM) or the Control Functions (CF) Methodology.¹¹ Table 3 presents the α estimates of this Poisson Model. As can be seen, the estimates are very similar to those presented in Table 2. This similarity aligns with the fact that our model is an aggregated version of the CL model, and suggests that the methodology we implement in this paper is correct.

Table 3. α estimates from a target-based Poisson Regression Model

α	Violent crimes				Property crimes				Total crimes			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
α	-0.004 (0.003)	-0.003** (0.001)	-0.005* (0.002)	-0.006** (0.003)	-0.007 (0.005)	-0.006*** (0.002)	-0.010** (0.004)	-0.010* (0.005)	-0.006* (0.003)	-0.005*** (0.002)	-0.008*** (0.003)	-0.008** (0.004)
Observations	1,050	1,050	1,050	1,050	1,050	1,050	1,050	1,050	1,050	1,050	1,050	1,050
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Previous patrol time	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
Locality FE	No	No	No	Yes	No	No	No	Yes	No	No	No	Yes
Method	GMM	CF	CF	CF	GMM	CF	CF	CF	GMM	CF	CF	CF

Notes: *** p<0.01, ** p<0.05, * p<0.1. Cluster robust standard errors at the level of locality in parentheses.

Additionally, it is to be noted that when TSLS estimates are invalid, OLS estimates provide suggestive information about the true impact of police presence on crime. In particular, note that, given the simultaneity that exists between crime and police presence, OLS estimates are downward biased.¹² Therefore, if we estimate equation (7) by OLS we would obtain lower bounds (in absolute value) of the real average impact of police presence on the – and thus on the occurrence– of committing a crime in a certain location. As can be seen in Table ?? of Appendix ??, this holds and, therefore, provides confidence to our principal estimates.

Lastly, note that in this work we assume that the number of potential criminals is given by the total number of unemployed people aged 12 to 60, which results in a conservative share of crimes in each of the city locations. Thus, we are confident that the reported TSLS estimates, which are relatively close to the OLS estimates, are informative enough about the causal impact of police presence (and the rest of covariates) on crime location choice and, therefore, general conclusions can be obtained from them.

Selected covariates

Table 4 presents the coefficients associated with the covariates that were selected and ultimately included in our model estimation. Surprisingly, the Double Selection algorithm selected all the available control variables. Following our model, the coefficients depicted in the table reflect which attributes make a location more or less attractive to commit a crime. In general, the results are quite intuitive. Concretely, the coefficients on the average distance to a shopping center, education center, services offices, and nearest transportation infrastructure are negative, and the coefficients associated with building presence and housing presence are positive but small. These might reflect the fact that close to these places there is usually more mobility and human interaction, and thus, greater opportunity or ease for criminals to commit a crime.

¹¹See wooldridge2010econometric, chapters 8, 9, and 21 for technical details.

¹²Note that if the impact of police on crime is negative and the impact of crime on police is positive, OLS estimates that ignore the simultaneity just combine both effects into one. That is, a positive and a negative effect are averaged, yielding a less negative (or even a positive) result.

Table 4. TSLS β estimates of the selected variables

	Violent crimes (1)	Property crimes (2)	Total crimes (3)
Average built squared meters per street segment meter	0.0000*** (0.0000)	0.0000* (0.0000)	0.0000*** (0.0000)
Average distance to nearest shopping center	-0.0000 (0.0000)	-0.0003*** (0.0001)	-0.0001** (0.0001)
Average distance to nearest educational center	-0.0003*** (0.0001)	0.0001 (0.0001)	-0.0001 (0.0001)
Average distance to nearest park/recreational center	0.0001 (0.0001)	-0.0000 (0.0001)	-0.0000 (0.0001)
Average distance to nearest police station	-0.0003*** (0.0001)	-0.0002* (0.0001)	-0.0002*** (0.0001)
Average distance to nearest religious/cultural center	0.0001** (0.0001)	0.0001* (0.0001)	0.0002*** (0.0001)
Average distance to nearest health center	0.0000 (0.0001)	-0.0000 (0.0001)	-0.0000 (0.0001)
Average distance to nearest transport infrastructure	-0.0012*** (0.0004)	-0.0013** (0.0005)	-0.0014*** (0.0005)
Proportion of street segments zoned for industry/commerce	0.3433 (0.3190)	0.3073 (0.4207)	0.2764 (0.3905)
Proportion of street segments zoned for housing	0.4341* (0.2582)	-0.1778 (0.2954)	-0.0574 (0.2868)
Proportion of middle income street segments	0.0007 (0.0827)	0.5815*** (0.1103)	0.3855*** (0.1009)
Proportion of high income street segments	-0.6294*** (0.0832)	0.4551** (0.2190)	0.1491 (0.1757)
Average longitude of street segments (mt)	-0.0021 (0.0014)	-0.0012 (0.0036)	-0.0018 (0.0027)
Average distance to nearest additional services office	-0.0000 (0.0001)	-0.0001** (0.0001)	-0.0001 (0.0001)
Average distance to closest surveillance camera	0.0000 (0.0001)	-0.0001 (0.0001)	-0.0001 (0.0001)
Proportion of hotspot street segments	0.6522** (0.3028)	1.5188*** (0.5798)	1.3312*** (0.4215)
Observations	1,050	1,050	1,050
R-squared	0.3801	0.4138	0.2813
Controls	Yes	Yes	Yes
Previos patrol time	Yes	Yes	Yes
Locality FE	Yes	Yes	Yes

Notes: *** p<0.01, ** p<0.05, * p<0.1. Cluster robust standard errors at the level of locality in parentheses.

In contrast, the coefficient on the average distance to a religious center is positive suggesting the relevance of other determinants of crime associated with religious beliefs and respect for certain norms. As expected, criminals are less likely to select a high-income location to commit a violent crime, but are more likely to select a middle- or high-income to commit a property crime. We interpret these results as follows: the perceived benefit from a property crime is highest in richer places, while the perceived risk of committing a violent crime in such places is also higher. Therefore, violent crimes move towards poorer areas, while property crimes to richer ones.

Additionally, we find a negative coefficient for the average distance to the nearest police station, a positive coefficient for the proportion of hot spots within the quadrant, and a null coefficient for the average distance to the closest surveillance camera. The first result, although counterintuitive, might reflect the fact that the police stations are located in historically insecure places and that it is not the infrastructure itself, but the effective police presence that can affect crime levels. The second result reflects the fact that criminals tend to commit crimes in insecure places, and that there is constant feedback between past and future crime locations: crimes are more common where past crimes have been committed. The third result suggests that surveillance cameras have a low effect in deterring criminals to commit crimes in the watched places.

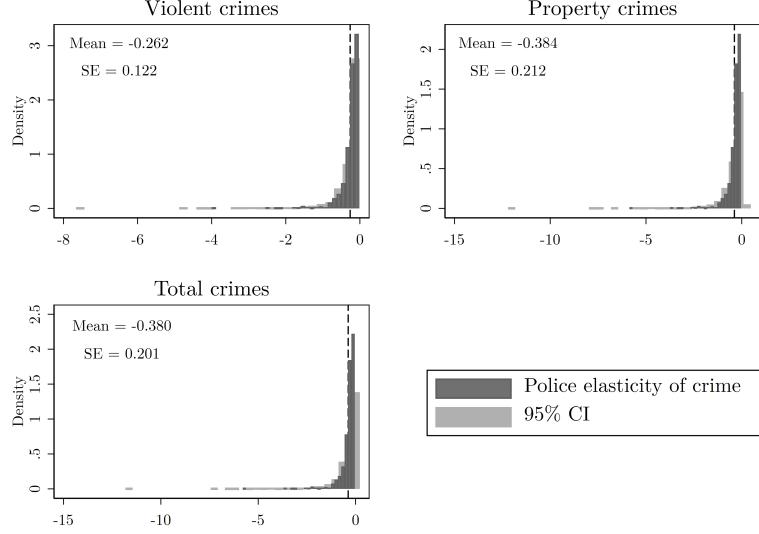
Lastly, no statistically significant effects are found for the average distance to parks or health centers, the average longitude of street segments, nor the proportion of street segments zoned for industry and commerce.

The police elasticity of crime

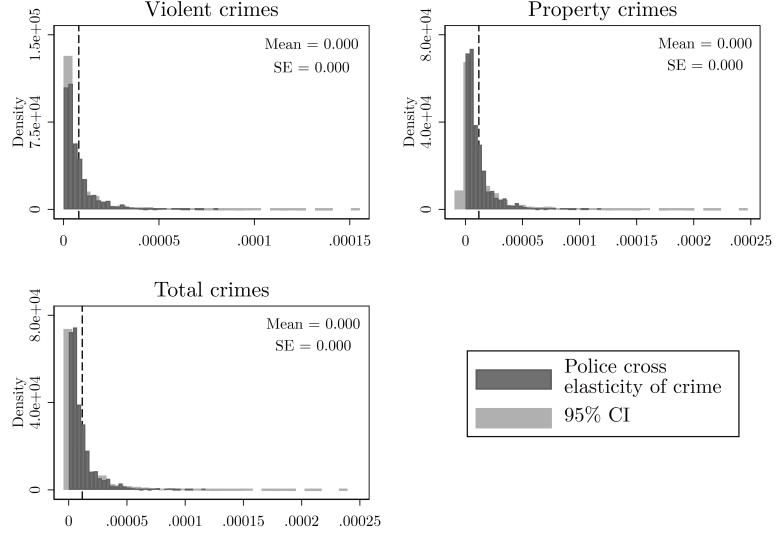
Our most important results from a public policy perspective are the estimation of the own- and cross-elasticities of crime to patrolling time at different locations, which captures the percentage change in crime that results from a 1% increase in the police patrol time that a location receives. To the extent of our knowledge, the estimation of these statistics in a properly identified structural model of crime location is new in the literature. Given the estimated parameters from our preferred specification, we follow equation (6) to calculate own- and cross-elasticities of crime with respect to the police presence (police elasticities of crime, from now on).

The first panel of Figure 2 reports the own- elasticity distribution across all locations. For violent crimes, the elasticity is, on average, -0.26 and is statistically significant with 95% confidence. That is, a 10% increase in patrolling time in a quadrant reduces crime, on average across all quadrants, by 2.6% in that quadrant. The effect on property crime is almost 1.5 times as large, which seems consistent with intuition. This result, therefore, holds for total crimes.

The second panel of Figure 2 shows the results in terms of crime displacement. By construction, our model takes every crime location as a substitute for the others. This is clear in equation (6), where the cross-elasticity has a positive sign and, by construction, a police presence increase in one location results in a crime displacement to other locations. Hence, our model is biased in favor of a hypothesis that has been at the center of the discussion in the criminology literature: crime is displaced rather than reduced in response to police presence. However, by construction, crime displacement in our model is, controlling for police presence, equally likely to all crime locations. This is an oversimplifying assumption that implies cross-elasticities may be very low although relevant. We discuss this further below when we compare our results with [?].



(a) Police elasticity of crime



(b) Cross police elasticity of crime

Fig 2

Our results are in line with those found by blattman2021place. In particular (see Table 3, page 2042 in [?]), they find that doubling the police patrol time causes an average reduction of crime of 13% on treated segments (direct effect). This is comparable to our police elasticity of crime (Panel (a), Figure 2). Now, they also report a small spillover effect (indirect effect) within street segments at less than 250 meters from treatment, suggesting that crime increases in the neighborhood of treated segments. This is comparable to our cross police elasticity (Panel (b), Figure 2). Although this indirect effect is small there are so many non-treated segments that are in the neighborhood of a treated segment, that when they aggregate and estimate the total effect on crime, the indirect effects override the direct effects and crime is estimated to increase (3%) as a consequence of police intervention. However, this aggregate result is

not statistically significant at a 10% level. In the next section we show that our results
are consistent with these results from [?]. However, since we estimate a structural
model, we are able to perform other types of policy scenarios that we show have an
important impact on crime reduction.

Counterfactual Policy Scenarios

A great advantage of using structural models of social phenomena is the ability,
conditional to the *prior* that the model is an accurate description of reality, to study
alternative policy scenarios. The question we now want to answer is: What would have
been the number of crimes had police patrols pursued different patrolling strategies? (1)
uniform time (each segment receives 33.82 minutes of patrol time independently of any
characteristic)¹³; (2) time spent proportional to historic crime rates per segment; and
(3) Police patrol time set to a minimum for all the quadrants and the residual time (i.e.,
the difference between the observed time and the minimum time) is reassigned such that
quadrants that report both the highest levels of crime and the highest police elasticity
of crime receive their initial police patrol time plus an $x\%$ increase, where
 $x \in \{0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100\}$.¹⁴ An important characteristic of all policy
scenarios is that, by construction, the economic costs of all of them is roughly the same
as the observed scenario (see Table 5: Observed scenario), the current police patrolling
strategy. Since the outcome is measured in terms of their average effect on crime, then
our policy scenarios measure efficiency gains, due to a better police-time reassignment,
on the city. We leave for a separate subsection the estimation of the optimal
reassignment policy and how it compares to blattman2021place study.

Basic policies

Table 5 shows the average observed crimes (i.e., Observed row) and the in-sample fit
(i.e., Predicted row) of the model in terms of the average predicted crimes, respectively.
As can be seen, our model underestimates by approximately three events the average
occurrence of crimes, and thus, can be regarded as a conservative and good enough
representation of reality.

The rest of the rows form Table 5 show what would have happened had the police
been deployed under the four different policy scenarios. In the first place, the
uniform-policing time distribution, which is the uninformed decision, results in more
crime occurrences (compared to the predicted number of actual crimes). In the second
place, proportional time results in basically the same levels of crime as those predicted
by the model in the base scenario. We argue, this is likely the strategy most frequently
used by the police in the city of Bogotá to allocate patrolling time.

The third policy scenario leaves a couple of lessons. First, provided that there are
limited time resources (i.e., the total available time used by all police officers patrolling
the city), there exists a trade-off between the number of quadrants that can benefit
from an increase in police patrolling time and the magnitude of the increase itself. That
is, the greater the percentage increase in time, the fewer quadrants receive such increase.
Second, given this trade-off, it is better to increase by small percentages the police
patrol time of most of the quadrants than increasing by a great percentage the police
patrol time of just a few quadrants. Specifically, it seems that an effective way to reduce
crime, compared to both the base scenario and the other percentage increases, is to
reassign patrolling time from 30% to 40% of the quadrants that are less insecure and

¹³33.82 is the total observed patrol time divided by the number of quadrants.

¹⁴This additional time is indirectly reduced in quadrants that are both less insecure and less elastic,
such that the total patrolling time across the city remains constant. For a detailed explanation of this
reassignment strategy, see Appendix ??.

less elastic (i.e., that are the least affected by police patrol time) and reassign it to the other 60% - 70% of the quadrants such that they receive 20% more police patrol time. These results are depicted in Figure 3, which better displays the existing trade-off and the optimal reallocation strategy to reduce each type of crime.

503
504
505
506

Table 5. Counterfactual policy scenarios

	Violent Crimes			Property Crimes			Total Crimes		
	Benefited N	Q. %	Predicted # Mean (SD) [SE]	Benefited N	Q. %	Predicted # Mean (SD) [SE]	Benefited N	Q. %	Predicted # Mean (SD) [SE]
Observed	-	-	11.84 (8.56)	-	-	23.86 (17.88)	-	-	35.70 (21.42)
Predicted	-	-	11.14 [0.91]	-	-	21.58 [1.75]	-	-	32.59 [1.89]
Uniform time	-	-	11.35 [1.61]	-	-	23.80 [2.54]	-	-	34.67 [2.54]
Proportional time	-	-	11.83 [1.84]	-	-	22.19 [2.38]	-	-	33.92 [2.67]
10% increase	1,000	95.2	11.10 [0.90]	1,000	95.2	21.41 [1.71]	999	95.1	32.36 [1.81]
20% increase	914	87.0	10.15 [0.97]	911	86.8	21.80 [1.66]	912	86.9	31.22 [1.73]
30% increase	855	81.4	10.30 [1.07]	854	81.3	22.42 [1.64]	854	81.3	31.53 [1.68]
40% increase	803	76.5	10.40 [1.13]	805	76.7	22.50 [1.66]	805	76.7	31.36 [1.67]
50% increase	756	72.0	10.46 [1.28]	756	72.0	22.77 [1.80]	756	72.0	31.57 [1.82]
60% increase	717	68.3	10.54 [1.40]	717	68.3	23.14 [1.85]	715	68.1	32.03 [1.85]
70% increase	682	65.0	10.59 [1.50]	682	65.0	23.33 [1.87]	680	64.8	32.26 [1.89]
80% increase	645	61.4	10.67 [1.61]	641	61.0	23.65 [1.94]	643	61.2	32.51 [1.89]
90% increase	600	57.1	10.75 [1.73]	601	57.2	24.16 [2.11]	601	57.2	33.34 [2.07]
100% increase	568	54.1	10.82 [1.85]	568	54.1	25.21 [2.40]	569	54.2	34.18 [2.44]

Notes: Mean observed and predicted crimes for each of the policy or time reassignment strategies of police patrol time. Standard errors at the level of locality in squared parentheses.

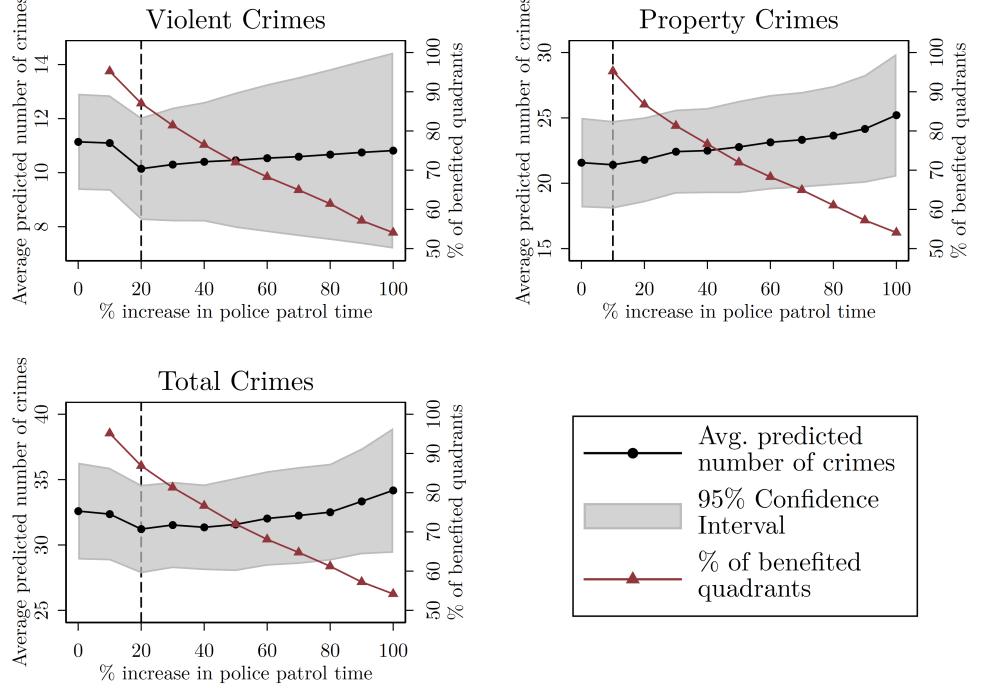


Fig 3. Third policy scenario: Increment in police patrol time vs proportion of quadrants benefited (red line) and predicted average number of crimes per quadrant with 95% confidence intervals (black line).

Optimal policy

We now estimate the optimal reassignment policy for this problem and we demonstrate its statistical significance. We also show that our model is consistent with blattman2021place by showing that, conditional on our model, the random reassignment of police patrolling conducted in their study implies a modest reduction of crime in line with their claim that they can rule out statistically significant aggregate effects of more than 2%.

To solve for the optimal policy we solve the following optimization problem:

$$\min_{P_1, \dots, P_J} \sum_{j=1}^J NS_j(P_j, X_j, \xi_j; \alpha, \beta) \quad (13)$$

$$\sum_{j=1}^J P_j = TP \quad (14)$$

where TP is the observed total amount of time of police patrolling. This problem is equivalent to:

$$\min_{P_1, \dots, P_J} \sum_{j=1}^J \exp(\alpha P_j + X_j \beta + \xi_j) \quad (15)$$

$$\sum_{j=1}^J P_j = TP \quad (16)$$

Intuitively, the optimization problem is equivalent to maximizing the outside option S_0 .

Using our model, we construct an additional counterfactual scenario in which we estimate the amount of crime in case there had not been the intervention reported in blattman2021place large scale experiment. For doing this, we identified the streets segments that where treated and set the police time to its original level (one half of the time observed for treated segments). Table 6 compares the results of three scenarios. The *Base scenario* corresponds to the prediction (fit) of the model to the actual observed data. The *Optimal scenario* reports the results of solving the optimization problem above, and the *No intervention scenario* reports the results of the scenario in which we estimate the amount of crime in case there had not been an intervention. The table also reports average crimes and total crimes in the city for violent, property and total crimes. For example for violent crimes, the table shows that the optimal scenario has on average per quadrant of 0.80 less crimes than the Base scenario and it is a statistically significant with a 99% confidence level (bootstrapped p -values reported in square brackets under the point estimates column). Aggregate violent crimes are 835.77 less than in the Base scenario, and statistically significant with a 99% confidence level. Note also that in the No intervention scenario there is a small increase in all crimes relative to the Base scenario, but it is not statistically significant. This last result approximately reproduces the main result in [?].

Table 6. Base Scenario, No Intervention and Optimal Policy

	2*Observed	2* Base scenario	2* Optimal scenario	Predicted		2* No intervention scenario	Difference		
				Difference					
				Point	Distribution				
<i>Panel A. Average crimes per quadrant</i>									
Violent crimes	11.84 (8.56)	11.14 (0.28)	10.35 (0.53)	-0.80*** [0.00]	-0.58*** [KS: 0.00; t: 0.00]	11.19 (0.30)	0.05 [0.65] [KS: 0.00; t: 0.00]		
Property crimes	23.86 (17.88)	21.58 (0.76)	19.75 (1.40)	-1.83*** [0.00]	-1.36*** [KS: 0.00; t: 0.00]	21.76 (0.86)	0.18 [0.65] [KS: 0.00; t: 0.00]		
Total crimes	35.70 (21.42)	32.59 (1.01)	30.91 (1.22)	-1.68*** [0.00]	-1.40*** [KS: 0.00; t: 0.00]	32.84 (1.19)	0.25 [0.68] [KS: 0.00; t: 0.00]		
<i>Panel B. Total crimes in the city</i>									
Violent crimes	12,435.00	11,698.57 (299.01)	10,862.81 (559.52)	-835.77*** [0.00]	-613.98*** [KS: 0.00; t: 0.00]	11,749.84 (316.67)	51.26 [0.65] [KS: 0.00; t: 0.00]		
Property crimes	25,053.00	22,657.88 (794.91)	20,739.53 (1,470.22)	-1,918.35*** [0.00]	-1,432.26*** [KS: 0.00; t: 0.00]	22,851.11 (906.93)	193.23 [0.65] [KS: 0.00; t: 0.00]		
Total crimes	37,488.00	34,222.57 (1,064.62)	32,459.56 (1,285.09)	-1,763.01*** [0.00]	-1,471.90*** [KS: 0.00; t: 0.00]	34,483.95 (1,244.58)	261.37 [0.68] [KS: 0.00; t: 0.00]		

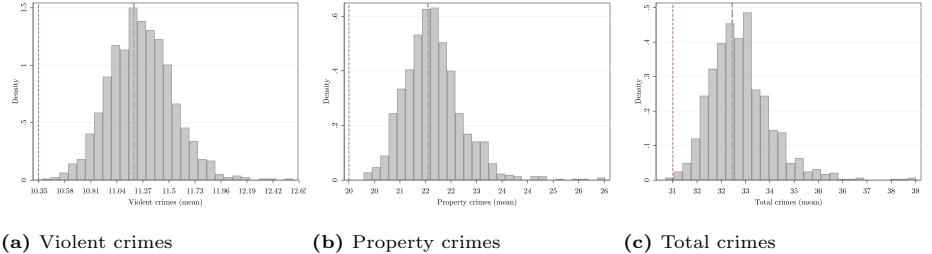
Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard deviation of the observed means and standard errors (SE) of the predictions in parentheses. Standard errors were obtained by a 1000-repetitions Bootstrap. p -values in squared brackets. KS: Kolmogorov-Smirnov test, t: t-test. Base scenario: crime predictions under the observed police patrol time. Optimal scenario: crime predictions under the optimal police patrol time distribution. No intervention scenario: crime predictions in the hypothetical case where no experiment was implemented. Difference corresponds to the rest between each of the last two scenarios and the base scenario. That is, Difference = Optimal - Base scenario, or Difference = No intervention - Base scenario. Two differences are presented, the point difference (Point), which corresponds to the difference between the predicted average (total) amount of crimes. The distribution difference (Distribution) correspond to the mean difference between the bootstrap distributions of predicted average (total) amount of crimes. For the Point difference, a one-sided p -value is reported. For the Distribution difference, p -values associated to a two-sample Kolmogorov-Smirnov test for equality of distributions and a one-sided mean difference t-test. For the Optimal scenario, the Point p -value is computed as: $p - value = \sum_{b=1}^{1000} 1\{\hat{C}^b \leq \hat{C}^*$ }/1000, where \hat{C}^b is the prediction under the base scenario for bootstrap iteration b and \hat{C}^* is the average (sum) prediction under the optimal assignment. For the No intervention scenario, the p -value is computed as $p - value = \sum_{b=1}^{1000} 1\{\hat{C}^b \geq \hat{C}^{NI}\}$ /1000, where \hat{C}^{NI} is the average (sum) prediction under the hypothetical case of no experiment.

Figure 4 shows the bootstrapped estimation of the uncertainty in the prediction of the average number of crimes per quadrant (Panel A) and total number of crimes (Panel B) in the base scenario. The vertical dash line shows the prediction of the model under blattman2021place intervention and the red dashed line the prediction under the optimal policy. One sided p -values are reported in table 6 in square brackets in the column of point estimates.

Figure 5 shows the bootstrapped estimation of the uncertainty for the three scenarios: No intervention policy with blue bars, base scenario (blattman2021place intervention) with gray bars and the optimal patrol scenario with red bars. Clearly, the optimal scenario implies a considerable shift of the distributions to the left of the average number of crimes per quadrant (Panel A) and total number of crimes (Panel B). Table 6 shows that all of these distributions are statistically significant to a 99% confidence level. Table 6 (see Distribution columns) shows p -values associated to a two-sample Kolmogorov-Smirnov test for equality of distributions and a one-sided mean

difference t-test. Although they are all significant differences the base scenario is marginally different to the no intervention scenario (still another confirmation of blattman2021place results). However the optimal scenario shows a significant reduction in crime relative to the no intervention and the base scenarios.

Panel A. Bootstrap distribution of the predicted average amount of crimes per quadrant



Panel B. Bootstrap distribution of the predicted sum of crimes

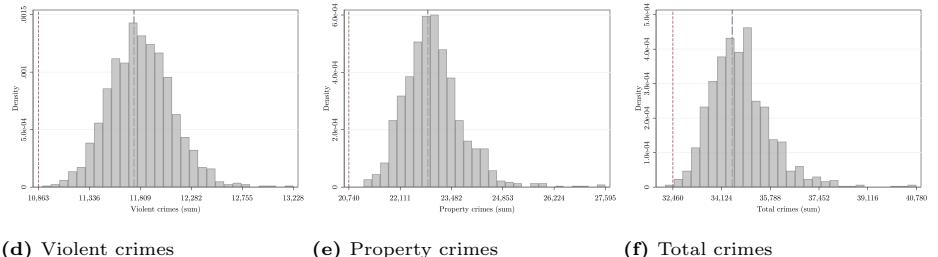
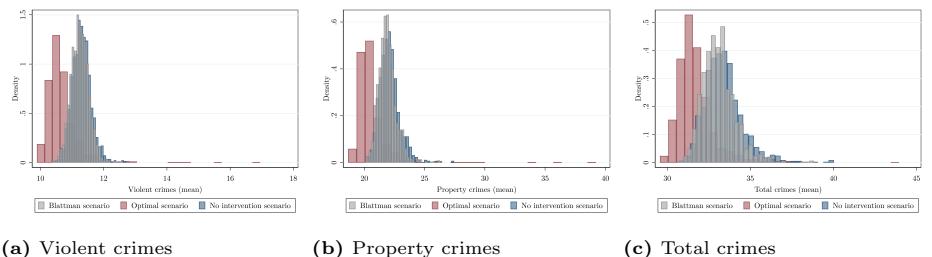


Fig 4. Each figure displays the bootstrap distribution of the predicted average amount of crimes per quadrant (Panel A) and predicted sum of crimes in the city (Panel B) under the base scenario. Vertical red (blue) dashed line displays the predicted mean and total number of crimes under the optimal (no intervention) police patrol scenario.

Panel A. Bootstrap distributions of the predicted average amount of crimes per quadrant



Panel B. Bootstrap distribution of the predicted sum of crimes

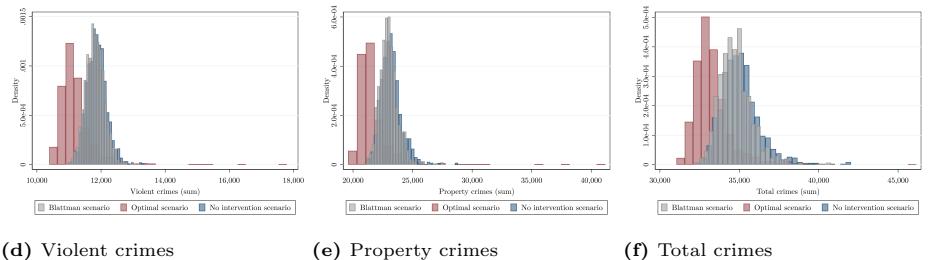


Fig 5. Each figure displays the bootstrap distributions of the predicted average amount of crimes per quadrant (Panel A) and predicted sum of crimes in the city (Panel B) under the base scenario (gray bars), under the optimal police patrol scenario (red bars), and under the no-intervention scenario (blue bars).

Discussion

In this paper we used a unique experimental data set blattman2021place at the scale of a big urban center, Bogotá, the capital city of Colombia with more than seven million people. This randomized controlled trial was specially tailored towards the identification of the causal effect of police patrolling on crime. Using this data set, we proposed an estimation strategy based on a random utility model of crime location choice. In particular, our model, a conditional logit model with aggregate data, allows us to identify agents' fundamental parameters, such as their perceived utility of committing a crime. Once preferences were identified from observable data, we estimated the own- and cross-elasticities of crime to patrolling time and we were able to evaluate alternative patrolling strategies with the same total time cost. To the extent of our knowledge, both exercises are novel features in the literature studying the causal effect of police presence on crime levels in major cities in Latin America, and are very relevant from a public policy perspective. Our model gives results consistent with those reported by blattman2021place.

We found that police presence in a certain location has, on average, a negative impact on the utility of committing either a violent, a property, or any crime in that location: a one minute of police patrol time in location j reduces, on average, 0.006 units the utility of committing a violent crime and 0.008 units the utility of committing either property or any crime in that location. These effects are statistically significant at the 5% for violent crimes and 10% for property and total crimes. Our most important results from a public policy perspective are the estimation of the own- and cross-elasticities of crime to patrolling time at different locations, which captures the percentage change in crime that results from a 1% increase in the police patrol time that a location receives. To the extent of our knowledge, the estimation of these statistics in a properly identified structural model of crime location is new in the literature. For violent crimes, the elasticity is, on average, -0.26 and is statistically significant at the 5% level. That is, a 10% increase in patrolling time in a quadrant reduces crime, on average across all quadrants, by 2.6% in that quadrant. The effect on property crime and all crimes is -0.38 , 1.5 times larger. Moreover, our cross-police elasticity suggest a negligible spillover effect.

Taking advantage of our estimation of a structural model, we solve for the optimal reassignment of police time in the city. We show that it is possible to obtain an average reduction in violent crime of 7.09%, a reduction in property crimes of 8.48% and a reduction in total crimes of 5.15%. If we aggregate over all quadrants these implies a reduction of 835 in violent crimes, 1,918 in property crimes and 1,471 in total crimes during the period of study. This optimal reassignment policy uses the same amount of aggregate time as the current allocation policy.

Our results are in line with those of found by blattman2021place. In particular, they find that doubling the police patrol time causes an average reduction of crime of 13% on treated segments (direct effect), comparable to our computation of police elasticity of crime. However, they also report a small spillover effect (indirect effect) within street segments at less than 250 meters from treatment, suggesting that crime increases in the neighborhood of treated segments. This is comparable to our cross police elasticity. Although this indirect effect is small there are so many non-treated segments that are in the neighborhood of a treated segment, that when they aggregate and estimate the total effect on crime, the indirect effects override the direct effects and crime is estimated to increase (3%) as a consequence of police intervention. However, this aggregate result is not statistically significant at a 10% level.

Our conclusions are limited by our model and identifying assumptions... We also leave as future research...

[3] [1]

References

1. Berry, Steven T Estimating discrete-choice models of product differentiation, *The RAND Journal of Economics*; 1994
2. Ruiter, Stijn Crime location choice *The Oxford handbook of offender decision making*; 2013
@articleschmidheiny2011equivalence, title=On the equivalence of location choice models: Conditional logit, nested logit and Poisson, author=Schmidheiny, Kurt and Brülhart, Marius, journal=Journal of Urban Economics, volume=69, number=2, pages=214–222, year=2011, publisher=Elsevier
@articleBernasco_{undated - uf}, title = "The Study of Crime and Place : A Methods Handbook", author = "Bernasco, W", editor = "Groff, E Rand Haberman, CP", year = 2023, publisher = Temple University Press
@articlehausman1984specification, title=Specification tests for the multinomial logit model, author=Hausman, Jerry and McFadden, Daniel, journal=Econometrica: Journal of the econometric society, pages=1219–1240, year=1984, publisher=JSTOR
@articlemcfadden1973conditional, title=Conditional logit analysis of qualitative choice behavior, author=McFadden, Daniel, year=1973, publisher=Institute of Urban and Regional Development, University of California ...
@articleblattman2021place, title=Place-based interventions at scale: The direct and spillover effects of policing and city services on crime, author=Blattman, Christopher and Green, Donald P and Ortega, Daniel and Tobón, Santiago, journal=Journal of the European Economic Association, volume=19, number=4, pages=2022–2051, year=2021, publisher=Oxford University Press
@articleangrist1995two, title=Two-stage least squares estimation of average causal effects in models with variable treatment intensity, author=Angrist, Joshua D and Imbens, Guido W, journal=Journal of the American statistical Association, volume=90, number=430, pages=431–442, year=1995, publisher=Taylor & Francis
@article10.2307/2951620, ISSN = 00129682, 14680262, URL = <http://www.jstor.org/stable/2951620>, author = Guido W. Imbens and Joshua D. Angrist, journal = Econometrica, number = 2, pages = 467–475, publisher = [Wiley, Econometric Society], title = Identification and Estimation of Local Average Treatment Effects, volume = 62, year = 1994
@articlebelloni2014high, title=High-dimensional methods and inference on structural and treatment effects, author=Belloni, Alexandre and Chernozhukov, Victor and Hansen, Christian, journal=Journal of Economic Perspectives, volume=28, number=2, pages=29–50, year=2014
@articleguimaraes2003tractable, title=A tractable approach to the firm location decision problem, author=Guimaraes, Paulo and Figueirido, Octávio and Woodward, Douglas, journal=Review of Economics and Statistics, volume=85, number=1, pages=201–204, year=2003, publisher=MIT Press 238 Main St., Suite 500, Cambridge, MA 02142-1046, USA journals ...
@bookwooldridge2010econometric, title=Econometric analysis of cross section and panel data, author=Wooldridge, Jeffrey M, year=2010, publisher=MIT press
@bookNAP24928, title=Proactive policing: Effects on crime and communities, author=National Academies of Sciences, Engineering, and Medicine and others,

year=2018, publisher=National Academies Press @articlebraga, author = Anthony A. Braga and Brandon C. Welsh and Cory Schnell, title =Can Policing Disorder Reduce Crime? A Systematic Review and Meta-analysis, journal = Journal of Research in Crime and Delinquency, volume = 52, number = 4, pages = 567-588, year = 2015, doi = 10.1177/0022427815576576,
URL = <https://doi.org/10.1177/0022427815576576>
, eprint = <https://doi.org/10.1177/0022427815576576>

@articletelep, author =Telep, C. and Weisburd, D. and Gill, Ch. and Vitter, Z. and D. Teichman, journal = Journal of Experimental Criminology, pages = 515–548, publisher = Springer, title = Displacement of crime and diffusion of crime control benefits in large-scale geographic areas: a systematic review, volume = 10, year = 2014

@articleRatcliffe, author = Ratcliffe, Jerry and Taniguchi, Travis and Groff, Elizabeth and Wood, Jennifer, year = 2011, month = 08, pages = 795 - 831, title = The Philadelphia Foot Patrol Experiment: A Randomized Controlled Trial of Police Patrol Effectiveness in Violent Crime Hotspots, volume = 49, journal = Criminology, doi = 10.1111/j.1745-9125.2011.00240.x

@articleNovak, author = Novak, Kenneth and Fox, Andrew and Carr, Christine and Spade, David, year = 2016, month = 09, pages = , title = The efficacy of foot patrol in violent places, volume = 12, journal = Journal of Experimental Criminology, doi = 10.1007/s11292-016-9271-1

@articlepearl, author = Pearl, Judea, title = The Seven Tools of Causal Inference, with Reflections on Machine Learning, year = 2019, issue_{date} = March2019, publisher = AssociationforComputingMachinery, address = NewYork, NY, USA, volume = 62, number = 3, issn = 0001 – 0782, url = <https://doi.org/10.1145/3241036>, doi = 10.1145/3241036, abstract = Thekindofcausalinferenceseeninnaturalhumanthoughtcanbe”algorithmitized”tohelpp Commun.ACM, month = feb, pages = 54–60, numpages = 7@articleholland, ISSN = 01621459, URL = <http://www.jstor.org/stable/2289064>, ab.

@articleneill2020, title=Policing Chronic and Temporary Hot Spots of Violent Crime: A Controlled Field Experiment, author=Dylan J. Fitzpatricka, and Wilpen Gorra, and Daniel B. Neill, journal=British Journal Criminology, volume=44, number=4, pages=296–315, year=2020, publisher=<https://arxiv.org/abs/2011.06019>

@articlebernasco2005, title=How do residential burglars select target areas, author=Bernasco, Win and Paul Nieuwbeerta, journal=British Journal Criminology, volume=44, number=4, pages=296–315, year=2005, publisher=Oxford University Press

@articlebernasco2010, title=Modeling Micro-Level Crime Location Choice: Application of the Discrete Choice Framework to Crime at Places, author=Bernasco, Win, journal=Journal of Quantitative Criminology, volume=26, pages=113–138, year=2010, publisher=Springer

@articlebrown2006, title=Spatial analysis with preference specification of latent decision makers for criminal event prediction, author=Xue, Yifei and Donald E. Brown, journal=Decision Support Systems, volume=41, pages=560–573, year=2006, publisher=Elsevier

@articleMohler2011, title=Self-exciting point process modeling of crime, author=Mohler, George O and Short, Martin B and Brantingham, P Jeffrey and Schoenberg, Frederic Paik and Tita, George E, journal=Journal of the American

- Statistical Association, volume=106, number=493, pages=100–108, year=2011, publisher=Taylor & Francis
- @articleMohler2014, title=Marked point process hotspot maps for homicide and gun crime prediction in Chicago, author=Mohler, George, journal=International Journal of Forecasting, volume=30, number=3, pages=491–497, year=2014, publisher=Elsevier
- @articleStec2018, title=Forecasting Crime with Deep Learning, author=Stec, Alexander and Klabjan, Diego, journal=arXiv preprint arXiv:1806.01486, year=2018
- @articleBuonanno2003, title=The socioeconomic determinants of crime. A review of the literature, author=Buonanno, P, journal=Working Paper Dipartimento di Economia Politica, Universita di Milano Bicocca, pages=63, year=2003
- @articleblattman2021place, title=Place-based interventions at scale: The direct and spillover effects of policing and city services on crime, author=Blattman, Christopher and Green, Donald P and Ortega, Daniel and Tobón, Santiago, journal=Journal of the European Economic Association, volume=19, number=4, pages=2022–2051, year=2021, publisher=Oxford University Press
- @articleangrist1995two, title=Two-stage least squares estimation of average causal effects in models with variable treatment intensity, author=Angrist, Joshua D and Imbens, Guido W, journal=Journal of the American statistical Association, volume=90, number=430, pages=431–442, year=1995, publisher=Taylor & Francis
- @article10.2307/2951620, ISSN = 00129682, 14680262, URL = <http://www.jstor.org/stable/2951620>, author = Guido W. Imbens and Joshua D. Angrist, journal = Econometrica, number = 2, pages = 467–475, publisher = [Wiley, Econometric Society], title = Identification and Estimation of Local Average Treatment Effects, volume = 62, year = 1994
- @micsala1997just, title=I just ran four million regressions, author=Sala-i-Martin, Xavier, year=1997, publisher=National Bureau of Economic Research Cambridge, Mass., USA
- @articlebelloni2014high, title=High-dimensional methods and inference on structural and treatment effects, author=Belloni, Alexandre and Chernozhukov, Victor and Hansen, Christian, journal=Journal of Economic Perspectives, volume=28, number=2, pages=29–50, year=2014
- @articlezeileis2004econometric, title=Econometric computing with HC and HAC covariance matrix estimators, author=Zeileis, Achim, year=2004, publisher=Institut für Statistik und Mathematik, WU Vienna University of Economics and ...
- @articlehausman1984specification, title=Specification tests for the multinomial logit model, author=Hausman, Jerry and McFadden, Daniel, journal=Econometrica: Journal of the econometric society, pages=1219–1240, year=1984, publisher=JSTOR
- @inbooktrain2009, place = Cambridge, edition = 2, title = Logit, DOI = 10.1017/CBO9780511805271.003, booktitle = DiscreteChoiceMethodswithSimulation, publisher = CambridgeUniversityPress, author = Train, KennethE., year = 2009, pages = 34–75
- @Articleijgi9030157, AUTHOR = Dewinter, Maite and Vandeviver, Christophe and Vander Beken, Tom and Witlox, Frank, TITLE = Analysing the Police

Patrol Routing Problem: A Review, JOURNAL = ISPRS International Journal of Geo-Information, VOLUME = 9, YEAR = 2020, NUMBER = 3, ARTICLE-NUMBER = 157, URL = <https://www.mdpi.com/2220-9964/9/3/157>, ISSN = 2220-9964

@articleMcFadden, Author = McFadden, Daniel, Title = Economic Choices, Journal = American Economic Review, Volume = 91, Number = 3, Year = 2001, Month = June, Pages = 351-378, DOI = 10.1257/aer.91.3.351, URL = <https://www.aeaweb.org/articles?id=10.1257/aer.91.3.351>

@articlePearl2009-PEACII, journal = Statistics Surveys, author = Judea Pearl, year = 2009, title = Causal Inference in Statistics. An Overview, volume = 3, pages = 96–146

@articlemohler2015randomized, title=Randomized controlled field trials of predictive policing, author=Mohler, George O and Short, Martin B and Malinowski, Sean and Johnson, Mark and Tita, George E and Bertozzi, Andrea L and Brantingham, P Jeffrey, journal=Journal of the American statistical association, volume=110, number=512, pages=1399–1411, year=2015, publisher=Taylor & Francis

@articlefitzpatrick2020policing, title=Policing Chronic and Temporary Hot Spots of Violent Crime: A Controlled Field Experiment, author=Fitzpatrick, Dylan J and Gorr, Wilpen L and Neill, Daniel B, journal=arXiv preprint arXiv:2011.06019, year=2020

@articlegaliani2022predictive, title=Predictive Policing in a Developing Country: Evidence from Two Randomized Controlled Trials, author=Galiani, Sebastian and Jaitman, Laura, journal=Available at SSRN 4080483, year=2022

@articleandresen2019spatial, title=The spatial effect of police foot patrol on crime patterns: A local analysis, author=Andresen, Martin A and Shen, Jen-Li, journal=International journal of offender therapy and comparative criminology, volume=63, number=8, pages=1446–1464, year=2019, publisher=SAGE Publications Sage CA: Los Angeles, CA

@articlecollazos2021hot, title=Hot spots policing in a high-crime environment: An experimental evaluation in Medellin, author=Collazos, Daniela and García, Eduardo and Mejía, Daniel and Ortega, Daniel and Tobón, Santiago, journal=Journal of Experimental Criminology, volume=17, pages=473–506, year=2021, publisher=Springer

3. Perry WL. Predictive policing: The role of crime forecasting in law enforcement operations. Rand Corporation; 2013.