A Spatial Discrete Choice Model of Crime

Álvaro J. Riascos Villegas Douglas Newball-Ramírez Mateo Dulce Rubio
Andrés Hoyos

May 23, 2022

Abstract

Overall, the causal relationship between proactive policing (in the sense of more time of police presence) and the incidence of crime is not yet well established. In this work, we use a unique experimental data set tailored to identify the causal impact of police patrolling on crime. Using this data, set we exploit an identification strategy based on a random utility model of crime location choice. The model allows us to identify agents' utilities from observable data. We were able to estimate own-and cross-elasticities of crime to patrolling time and we were able to evaluate alternative patrolling strategies (i.e, interventions or policy scenarios). To the extent of our knowledge, both the identification of elasticities and the counterfactual analysis are novel features in this literature. Our estimates show that 1% more time patrolling reduces crime an average of 0.19%. Cross-price elasticities show little support to negative spillover effects of police patrolling. We also evaluate four different patrolling time allocation strategies without increasing the total police time available. Our results show that allocating police time according to crime incidence and the elasticities of each quadrant, could potentially reduce violent crime by 4.13% and property crime by 6%.

Keywords— Crime prediction, Crime location, Discrete choice model, Logistic regression, Two-stage least squares, Counterfactual analysis, Elasticity estimation

1 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: Shoaib et al. (2018), Guangyin et al. (2020), Moreno, Quintero, Riascos, Nonato, and Sanchez (2021), Hu, F., C., and H. (2018), Jin et al. (2019), Kupilik and Witmer (2018), Moreno, Dulce, Riascos, Castaño, and Rodriguez (2020), G. O. Mohler, Short, Brantingham, Schoenberg, and Tita (2011), G. Mohler (2014), Reinhart and Greenhouse (2017), Stec and Klabjan (2018), 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 the 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 reduce crime, but at the same time crime incidents call for police presence at specific sectors of a city. Untangling this casual relationship is an instance of what Holland (1986) has called the fundamental problem of causal inference: What would have happened had the police not made presence at 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 and especially for US cities.

The evidence on the effects of proactive policing is mixed (see National Academies of Sciences and Medicine (2018) for a comprehensive study for the

US). Braga, Welsh, and Schnell (2015) 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, Weisburd, Gill, Vitter, and Teichman (2014) 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, Taniguchi, Groff, and Wood (2011) 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, Fox, Carr, and Spade (2016) 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. In a recent study, Fitzpatricka, Gorra, and Neill (2020) conducted a controlled field experiment of police placed-based interventions on violent crime. Their study spans a 12-month period of intervention in 0.5% of the city of Pittsburgh's area. They found statistically significant reductions in serious violent crime counts within treatment hot spots as compared to control hot spots, with an overall reduction of 25.3% in violent crimes such as homicides, rape, robbery, and aggravated assault. Only foot patrols, not car patrols, had statistically significant crime reductions in hot spots. They found no evidence of crime displacement, but a weakly statistically significant spillover of crime prevention benefits to adjacent areas.

Finally we highlight Blattman, Green, Ortega, and Tobón (2021), a placed-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 this study, we used the Blattman et al. (2021) data set and propose a different identification strategy.[1] We used a discrete choice model of spatial selection of crime to study the impact of police time exposure to different sectors of the city. Discrete choice modeling of crime phenomena has not been fully exploited in the literature. The pioneer applications of spatial discrete choice modeling of crime are Bernasco and Nieubeerta (2005), Xue and Brown (n.d.) and Bernasco (n.d.). In Bernasco and Nieubeerta (2005), 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 sociodemographic data of 290 burglars who committed 548 burglaries in the city during the period 1996-2001. They estimated a random utility model with burglars' and burglaries' characteristics by means of a conditional logit model. Their results support and quantify some of the working hypothesis but there is no real causal identification of the effect of police patrolling on crime. In fact, they provide very indirect evidence suggesting that "*the likelihood of a neighborhood being selected for burglary is positively influenced by its supposed lack of guardianship as indicated by ethnic heterogeneity, by its physical accessibility as measured by the percentage of single-family dwellings, and by the number of potential objects in the neighborhood*" Bernasco and Nieubeerta (2005). Given that the the i.i.d. hypothesis used in Bernasco and Nieubeerta (2005)

[1] We are deeply in debt and thankful to the authors for allowing us to use their data set for this study.

for maximum likelihood estimation is controversial due to spatial correlation, Bernasco (n.d.) studied the role of spatial resolution and alternative modeling strategies. The authors did not address the problem of identifying the causal effects of police patrolling on crime. Finally Xue and Brown (n.d.) 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.

To summarize, our contribution is fivefold: (1) We use a unique large data set of experimental data that allows for the identification of the causal effect of police patrolling on crime. (2) As opposed to the original experimental study of Blattman et al. (2021), we used a different identification strategy that is based on random utility selection of spatial locations for crime. (3) Given the high dimensional nature of our feature space, we used double selection techniques for a more agnostic data-driven model specification and robustness check of our results. (4) Provided, our estimates of the structural parameters that determine the crime location choice of the offenders, we computed the police own- and cross-elasticity of crime for each of the quadrants (i.e., we computed where police patrolling is more effective). (5) Based on the same structural parameters and the computed elasticities, we evaluated different police patrol counterfactual strategies without increasing the total police time available. In particular, we tested 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 either a 10, 20, 30, 40, 50, 60, 70, 80, 90, or 100% increase, and (d) recursively increasing 1% of patrol time for the most insecure and elastic quadrant. From these counterfactual exercises we further contribute in the understanding of which police patrol deployments are more effective. We found that increasing police time to the more elastic and insecure quadrants resulted in significant reductions of crime occurrences across Bogotá.

2 The model

Our model follows closely the standard logistic discrete choice model, except for our estimation strategy that recognizes the possibility of endogenous explanatory variables.[2,3]

Consider N potential criminal offenders with symmetric preferences, each of them deciding between $J+1$ locations in the city 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, is given by

$$u_{ij} = \alpha P_j + X_j\beta + \xi_j + \varepsilon_{ij} \tag{1}$$

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 characteristics of location j, α and β are coefficients that have to be estimated and ε_{ij} is the idiosyncratic error term.

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

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

where $\delta_j = \alpha P_j + X_j\beta + \xi_j$. We denote this probability as $s_{ij}(P_j, X_j, \xi_j; \alpha, \beta)$.

Assuming $\varepsilon_{ij}, \varepsilon_{ij'}$ are i.i.d. extreme value type I distributed, $\varepsilon_{ij} - \varepsilon_{ij'}$ follows a logistic distribution and, by equation 1, location choice probabilities

[2] Details on the standard discrete choice model can be found in Train (2009). There are many models for casual analysis, mainly a different approach based on casual graphs Pearl (2009). However, we have chosen our approach based on simplicity (discrete choice theory is well founded on random utility theory McFadden (2001)), interpretability (we recover agents utility for crime, patrolling time own- and cross-elasticities,), statistical inference (we know how to estimate unbiased casual effects) and finally, the ability to do interventions or policy scenarios (utilities are considered a structural characteristic of agents and therefore invariant under the different policy scenarios considered). We do not argue against the benefits of a different approach, we just believe it is valuable to learn how far this simple approach can take us in understanding the causal relationship between policy patrolling and crime.

[3] Regarding our empirical strategy, there are many options. First, identification is guaranteed because we use a randomized control trial experiment Blattman et al. (2021). For estimation, we impose additional structure motivated by utility theory, mainly the fact that criminals make rational choices based on the perceived benefits (utility) of committing a crime. In turn, this utility is partially determined by observable covariates that we discuss below.

have a closed-form expression given by

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

where option $j = 0$ is assumed to be the *outside option* and, thus, δ_0 is normalized to zero.[4] 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.[5]

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

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

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.

Now, from equation (4) we can 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

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

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

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

[4]The 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 Hausman and McFadden (1984)'s specification test; the results are reported in Appendix A.

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

3 Empirical strategy

3.1 Main specification

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

$$\delta_j \equiv \mathtt{log}(S_j) - \mathtt{log}(S_0) = \alpha P_j + X_j\beta + \xi_j, \tag{7}$$

and thus equation (7) represents our ideal specification. In this equation, the definitions of the variables are the same as those mentioned in section 2. In particularly, P_j is a measure of police presence at location j and X_j is the set of measured characteristics (including an all-ones vector) for the same location. ξ_j in this case represents the error term.

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

$$S_j = \frac{C_j}{N} \tag{8}$$

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*.[6] However, we exploit the fact that

$$\sum_{j=0}^{J} S_j = 1. \tag{9}$$

Therefore, S_0 is given by

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

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 (hopefully not far from real) assumption that the police force is assigned according to the amount of committed crimes in each of the locations, there exists simultaneity between δ_j and P_j.[7] Thus, OLS estimates from equation (7) are plausibly biased and inconsistent.

[6]We thank an anonymous referee for asking us to clarify the meaning of the outside option and what it means to choose. In particular, note that aggregate crimes can change under different exposures to police presence.

[7]δ_j is not directly a measure of crimes but it is an increasing function of them. So if there exists simultaneity between C_j and P_j, there exists simultaneity between δ_j and P_j.

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

In 2016, Blattman et al. (2021) 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 ligthing and cleanup) Blattman et al. (2021). In particular, 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) Blattman et al. (2021). Also, in March 2016, 201 of the 1,919 hot spots received more intensive street light repair and cleaning Blattman et al. (2021).[8] In this work, we exploited the assignment to the first type of treatment to instrument the police presence P_j and identify the structural parameters of interest. That is, we estimate equation (7) by TSLS.

Following Imbens and Angrist (1994) and Angrist and Imbens (1995), the necessary assumptions for the treatment assignment, Z, to be a valid instrument 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., $1\{j \text{ is a crime hot spot}\} \in X_j$) Blattman et al. (2021). 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 Blattman et al. (2021) the *"police complied with their new orders for the full 8 months."* Thus, 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.[9] Furthermore, the experimental designed in Blattman et al. (2021) intended that the patrol time of untreated streets remained the same, which was empirically verified by the authors.

Therefore, TSLS estimates of θ are consistent. In particular, α is identified by $\hat{\alpha}^{TSLS}$ and it 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 Angrist and Imbens (1995).[10]

3.3 Selecting location characteristics - Double Selection

In general, city locations might have several characteristics that can influence the presence of crime. Also, potential criminal offenders might base their crime-location choice not only on the individual location characteristics, but on combinations or interactions of them. That makes the number of possible candidates to be included in X_j vast and, thus, ad hoc or intuition-based variable selection unfeasible. Hence, to select the variables that should be included in X_j we implemented the double selection methodology of Belloni, Chernozhukov, and Hansen (2014).

Following Belloni et al. (2014), we first (separately) ran a regularized lasso over the following two equations

$$\delta_j = \tilde{X}_j\gamma + \mu_j, \tag{11}$$
$$P_j = \tilde{X}_j\vartheta + \lambda_j, \tag{12}$$

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 and all their second degree interaction terms (i.e., $x_{kj} \times x_{rj}\ \forall\, k, r$). γ and ϑ are vectors of coefficients associated to $\tilde{X}_j$ 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 predict the outcome variable δ_j and the police presence P_j, respectively. Given the selected variables from both estimations, we then estimated 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).

[8] For further information about the different treatment arms see Blattman et al. (2021).

[9] That is, only treated segments received increased patrol time. Control segments remained the same.

[10] For more technical details check Angrist and Imbens (1995).

4 Data

In this work, we used the data set used by Blattman et al. (2021) 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 2016. We also took information for 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 segments in use by services, 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 and the proportion of street segments assigned to increased patrol time within each quadrant.

Provided with this data set, we estimated three versions of equation (7): one for violent crimes, another for property crimes, and a last one for total crimes. To construct each crime-specific δ_j, we assumed 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).[11] On the other hand, P_j is given by average minutes of police patrol time (from Blattman et al. (2021)), Z_j is the proportion of treated street segments, and X_j is selected from the list of variables (and their quadratic interactions) 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, 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 11.98 violent crimes, 24.07 property crimes, and a total of 36.05 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 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. Lastly, panel C reports summary statistics of the variables to be added to the vector X_j.

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

	N	Mean	SD	Min	Max
A. Reported crimes					
Violent crimes	1,050	11.98	8.61	0	70
Property crimes	1,050	24.07	17.92	0	149
Total crimes	1,050	36.05	21.45	0	157
B. Police presence					
Avg. police patrol time (minutes)	1,050	46.46	51.89	4.98	711.23
C. Quadrant characteristics					
Avg. dist. to nearest shopping center	1,050	740.03	708.45	21.72	4,686.01
Avg. dist. to nearest education center	1,050	320.15	237.79	55.98	2,767.30
Avg. dist. to nearest park/recreational center	1,050	682.03	392.87	37.57	2,855.38
Avg. dist. to nearest religious center	1,050	502.66	367.24	28.90	3,829.53
Avg. dist. to nearest health center	1,050	952.91	716.71	55.56	7,499.75
Avg. dist. to nearest additional services center	1,050	700.01	577.79	39.38	5,193.83
Avg. length of street segments	1,050	64.33	23.35	21.52	263.38
Avg. built meters per meter of street segment	1,050	22,728	13,601	2.66	148,750
Prop. of paved street segments	1,050	0.98	0.04	0.51	1.00
Prop. of street segments zoned for industry/commerce	1,050	0.25	0.22	0.00	1.00
Prop. of street segments zoned for service sector	1,050	0.09	0.15	0.00	1.00
Prop. of high income street segments	1,050	0.08	0.26	0.00	1.00
Prop. of middle income street segments	1,050	0.45	0.48	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. Dist.: distance

Figure 4 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 as a first naive approximation to the estimation of the police elasticity of crime.[12] The figure shows that for violent crimes and total crimes, there is no relationship between the police presence and the reported crimes. OLS estimates for these two types of crimes are -0.02 and 0.02, respectively, and are not statistically significant. For property crimes, a

[11]Data available at `https://www.dane.gov.co/`.

[12]Since both dependent and independent variables are in logarithms, the slope of the relationship captures an elasticity

marginally significant positive relationship 0.06 is found. 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 discussed in section 3.

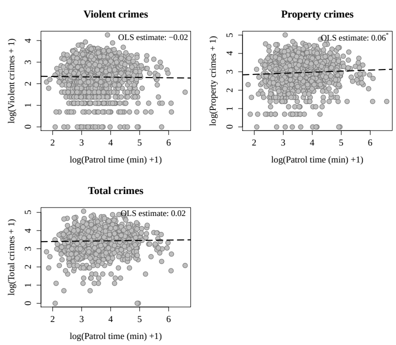


Figure 1: Scatter plots of the logarithm of each type of crime against the logarithm of the average patrol time within each quadrant. Linear fit displayed by the black dashed line. These graphs display a naive calculation of the police elasticity of crime. OLS coefficient estimates reported at the top-right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

5 Results

5.1 α estimates

Table 2 presents the estimated α parameter of equation (7) from different econometric methodologies for the three types of crimes mentioned in the previous section. Panel A and C display α estimates from OLS and TSLS that control for covariates, respectively. Panels B and D display estimates from OLS and TSLS that control for the selected covariates from the double selection methodology. As can be seen, the estimates are fairly consistent across methodologies. We find that, on average, police presence has no effect on violent crimes. Second, we found that police presence has a negative impact on the utility of committing property and total crimes, as was expected. In particular, following our TSLS + double selection preferred methodology, we found that an increase in 1 minute of police patrol time in location j reduces, on average, 0.004 units of the utility of committing a property crime in that location. This effect is statistically significant at the 1% significance level. For total crimes, the coefficient is -0.004 and is statistically significant at the 5% significance level. However, it has to be noted that, given the lack of effect in violent crimes, this result is fundamentally driven by the effect found for property crimes.

Table 2: α estimates from different econometric methodologies

	Violent Crimes (1)	Property Crimes (2)	Total Crimes (3)
A. OLS			
α estimate	-0.001	-0.003***	-0.002***
	(0.001)	(0.001)	(0.001)
N	1,026	1,040	1,047
Adj. R^2	0.282	0.348	0.194
B. Double Selection			
α estimate	-0.001	-0.002***	-0.002***
	(0.001)	(0.001)	(0.001)
N	1,018	1,018	1,018
Adj. R^2	0.316	0.404	0.225
C. TSLS			
α estimate	-0.002	-0.005***	-0.003**
	(0.001)	(0.002)	(0.001)
N	1,026	1,040	1,047
Adj. R^2	0.277	0.338	0.189
D. TSLS + Double Selection			
α estimate	-0.003	-0.004***	-0.004**
	(0.002)	(0.002)	(0.001)
N	1,018	1,018	1,018
Adj. R^2	0.307	0.396	0.216

Notes: *p<0.1; **p<0.05; ***p<0.01. Heteroskedasticity robust standard errors in parentheses (HC3 for panels A to C, HC1 for panel D (see Zeileis (2004) for technical details)). All four estimations included a vector of covariates. For panels A and C the following covariates were included: proportion of paved street segments, proportion of street segments in use by industry and commerce, proportion of street segments in use by services, 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. For panels B and D double-selection-selected covariates were included. These are discussed in section 5.2.

These results contrast with those found by Blattman et al. (2021). In particular, they found that doubling the police patrol time has no impact on property and overall crimes. Also, in their basic TSLS specification, using as dependent variable the level of crime, they found no effect of police patrol time on overall crimes unless they interact it with the baseline crime. These differences between their and our results might be driven either by: 1) differences in the definition of the dependent variable, given that they used the levels of crime while we used a log-ratios of crime shares that depend on the definition of N; 2) differences in the statistical power, given that they only estimated the impact for hot spot street segments, while we estimated it for all quadrants in the city, or 3) differences in the unit of analysis, given that they estimated the impact for street segments, where the crime reports might be low, and we estimated the impact for quadrants, where crime reports are greater by definition. It has also to be noted that they use inverse probability weighting in their estimations to correct for endogenous exposure to spillovers, as well as randomization inference to correct for fuzzy clustering Blattman et al. (2021). Our results, however, are fairly robust to specifications and strongly suggest a negative impact of police patrol time on crime.

It is to be noted that when TSLS estimates are invalid, OLS estimates provide suggestive information of 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.[13] Therefore, panels A and B of Table 2 report a lower bound of the real impact of police presence on the utility - and thus on the occurrence - of committing a crime. Also, note that in this work we assumed that the number of potential criminals is given by the 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 and general conclusions can be obtained from them.

5.2 Selected variables

Table 3 shows the estimated dependence of utility on observables before any variable selection. Some results are quite intuitive. For instance, the coefficients on the average distance to a shopping center, education center or nearest health center are negative. These might reflect the fact that close to these places there is usually more mobility and human interaction. 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 of certain norms. As expected, violent crimes depend negatively on middle- and high- income street segments, while property crimes depend positively on these same features.

Now, when double selection is used to select the most predictive variables of δ_j for property crimes (i.e., utility of crime) only two variables are chosen: average distance to the nearest shopping center and average distance to the nearest shopping center interacted with proportion of paved segments (see Table 4). However, neither is statistically significant. Moreover, in the case of violent crimes or total crimes, none are selected. These results highlight the importance of pursuing a data-driven specification appropriate for our problem.

In contrast, double selection of the variables that predict our measure of police patrolling selects 89 variables: one linear variable and 88 quadratic variables. These selected variables are reported in Table 5. The selection of that great amount of variables suggests that police patrolling is assigned strategically. It does not only depend on crime incidence, but also on many other complex features such as those quadratic terms reported in Table 5. This result further motivates future work to determine not only which quadratic terms, but which more complex features (third-, fourth-, etc. degree variables), might determine police patrol deployment and crime.

[13]Note 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 3: TSLS β estimates for covariates in their base form

	Violent Crimes (1)	Property Crimes (2)	Total Crimes (3)
Avg. dist. to nearest shopping center	-0.00004	-0.0005***	-0.0003***
	(0.00005)	(0.00005)	(0.00004)
Avg. dist. to nearest education center	-0.001***	-0.0001	-0.0002*
	(0.0001)	(0.0001)	(0.0001)
Avg. dist. to nearest park/recreational center	0.0001	0.0001	0.0001
	(0.0001)	(0.0001)	(0.0001)
Avg. dist. to nearest religious center	0.0002***	0.0002***	0.0003***
	(0.0001)	(0.0001)	(0.0001)
Avg. dist. to nearest health center	-0.00001	-0.0001*	-0.0001**
	(0.00004)	(0.00005)	(0.00004)
Avg. dist. to nearest additional services center center	-0.00001	-0.0001	-0.00004
	(0.00005)	(0.0001)	(0.00005)
Avg. length of street segments	-0.003**	-0.002	-0.005***
	(0.001)	(0.002)	(0.002)
Avg. built meters per meter of street segment	0.00001***	0.00000	0.00000*
	(0.00000)	(0.00000)	(0.00000)
Prop. of paved segments	0.609	0.801	1.319
	(0.716)	(0.790)	(0.841)
Prop. of street segments zoned for industry/commerce	0.149	0.416**	0.295*
	(0.191)	(0.178)	(0.165)
Prop. of street segments zoned for service sector	-1.140***	0.077	-0.130
	(0.211)	(0.215)	(0.192)
Prop. of high income street segments	-0.719***	0.403***	0.087
	(0.115)	(0.129)	(0.114)
Prop. of middle income street segments	-0.171**	0.423***	0.228***
	(0.076)	(0.080)	(0.071)
Prop. of hot spot street segments	-0.129	1.287	0.953
	(0.615)	(0.937)	(0.832)
Constant	-11.608***	-11.163***	-11.089***
	(0.759)	(0.828)	(0.873)
N	1,026	1,040	1,047
Adjusted R^2	0.277	0.338	0.189

Notes: *p<0.1; **p<0.05; ***p<0.01. Heteroskedasticity robust standard errors in parentheses (HC3). Parameter estimates reported for covariates in their base form regardless of their selection by the double selection methodology. Avg.: Average. Prop.: Proportion. dist.: distance

Table 4: TSLS β estimates for double-selection selected covariates that predict δ_j

	Violent Crimes (1)	Property Crimes (2)	Total Crimes (3)
Avg. dist. to nearest shopping center	-	-0.001	-
	-	(0.001)	-
Avg. dist. to nearest shopping center×	-	-0.0001	-
Prop. of paved segments	-	(0.001)	-
N	1,018	1,018	1,018
Adj. R^2	0.307	0.396	0.216

Notes: *p<0.1; **p<0.05; ***p<0.01. Heteroskedasticity robust standard errors in parentheses (HC1). Double-selection selected variables that explain P_j excluded since there were 89 selected variables and their interpretation would be cumbersome. Avg.: Average. Prop.: Proportion. Dist.: distance

5.3 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 TSLS + double selection methodology, we follow equation (6) to calculate own- and cross-elasticities.

The first panel of Figure 2 reports the own- elas-

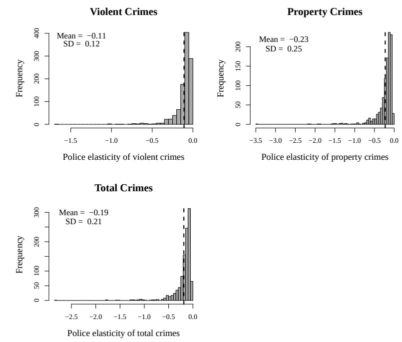
Table 5: Double-selection selected covariates that predict P_j

	Prop. of paved segments	Prop. of street segments zoned for industry/commerce	Prop. of street segments zoned for service sector	Prop. of high income street segments	Prop. of middle income street segments	Avg. dist. to nearest shopping center	Avg. dist. to nearest education center	Avg. dist. to nearest park/recreational center	Avg. dist. to nearest religious center	Avg. dist. to nearest health center	Avg. dist. to nearest additional services center center	Avg. length of street segments	Avg. built meters per meter of street segment
A. Linear selected variables			✓										
B. Quadratic selected variables													
Prop. of paved segments	✓	✓	✓			✓	✓	✓	✓	✓	✓		
Prop. of street segments zoned for industry/commerce	✓	✓	✓	✓		✓	✓	✓		✓	✓		✓
Prop. of street segments zoned for service sector	✓	✓	✓		✓			✓	✓	✓	✓		✓
Prop. of high income street segments	✓	✓			✓	✓	✓					✓	
Prop. of middle income street segments			✓		✓	✓		✓	✓	✓			✓
Avg. dist. to nearest shopping center	✓				✓	✓		✓					✓
Avg. dist. to nearest education center				✓						✓	✓	✓	✓
Avg. dist. to nearest park/recreational center	✓	✓	✓			✓							✓
Avg. dist. to nearest religious center	✓		✓		✓								✓
Avg. dist. to nearest health center		✓			✓		✓				✓		✓
Avg. dist. to nearest additional services center center	✓	✓	✓				✓			✓	✓		
Avg. length of street segments				✓			✓					✓	
Avg. built meters per meter of street segment		✓	✓		✓	✓	✓	✓	✓	✓			✓

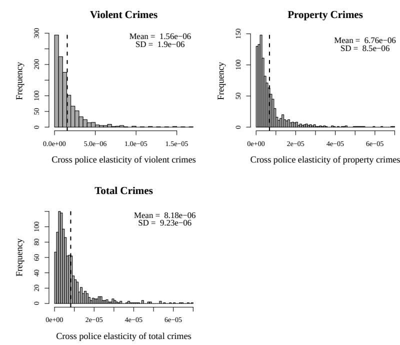
Notes: Panel A presents check marks (✓) for the linear double-selection selected variables. This is the variables in their base form. Panel B presents check marks (✓) for the quadratic double-selection selected variables. In this case, a check mark in row ℓ and column m indicates that the interaction between variables ℓ and m was selected. No coefficients are displayed to avoid cumbersome interpretations. Panel B is a symmetric matrix of check marks.

ticity distribution across all locations. For violent crimes the elasticity is, on average, -0.11. That is, a 10% increase in patrolling time in a quadrant reduces crime, on average across all quadrants, by 1.1% in that quadrant. The effect on property crime is twice as large, which seems consistent with intuition. In contrast to the results reported Blattman et al. (2021) our results do show a significant and public policy relevant causal effect impact of police patrolling intensity on crime reduction.

The second panel of Figure 2 is also interesting. By construction, it implies that crime location are substitutes in response to police patrolling (positive average police cross-elasticity). 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 respond to police presence. Our results are probably a biased estimate of this hypothesis and suggest that, if anything, crime displacement is negligible (not only on average).



(a) Police elasticity of crime



(b) Cross police elasticity of crime

Figure 2: Distribution of the own and police cross-elasticity of crime across quadrants for Bogotá. Elasticities separately displayed for violent crimes, property crimes, and total crimes. Vertical dashed lines represent mean elasticities. Mean and standard deviation of elasticities reported in each graph. Source: Own elaboration.

6 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 (*doing*, in the language of computer scientists, causal theory, see Pearl (2019)). The question we now want to answer is: What would have been the number of crimes had police patrols pursued four different strategies?[14]: (1) uniform time (each segment receives 33.82 minutes of patrol time independently of any characteristic)[15]; (2) time spent proportional to historic crime rates per segment, (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 $x \in \{0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100\}$[16]; and

[14]Note that we did not study optimal patrolling routes. We focused on the impact of policy presence measured as the time spent at quadrants which is the only effect we can study due to our experimental data and identification strategy. For a review study on the patrol routing problem see Dewinter, Vandeviver, Vander Beken, and Witlox (2020).

[15]33.82 is the total observed patrol time divided by the number of quadrants.

[16]This 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.

(4) police patrol time set to a minimum for all the quadrants and the residual time is reassigned using a recursive algorithm such that a 1% increase of additional patrol time is given to the quadrant, with the highest level of elasticities and crime. Then the new elasticities and crime are recalculated and the process repeated until no residual time is left over. For a detailed explanation of these last two reassignment methodologies, see Appendix B. An important characteristic of all policy scenarios is that, by construction, the economic costs of all of them is roughly the same as the baseline (see Table 6, Base 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.[17]

Panel A of Table 6 shows the average observed crimes and the in-sample fit 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.

Panel B of Table 6 shows 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 probably the strategy most frequently used by the police in the city of Bogotá to allocate patrolling time.

The third policy scenario has 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 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.

Finally, the fourth allocation strategy results in the lowest levels of crimes. This shows that there is room for much improvement in police patrolling allocation. For the sake of illustration, take the predicted base scenario as the base case. Switching from this scenario to the fourth policy scenario results in an average reduction of 0.43 in violent crimes per quadrant (out of 10.39) and of 1.29 (out of 21.33) of property crimes per quadrant. Since we have $1,050$ quadrants in the city, this implies a reduction of 451 violent crimes and $1,354$ property crimes for the whole city for the period of study (1 year). That is, informing the decision of police patrol deployment has important impacts on the security of the city without incurring in additional cos ts. Therefore, the results suggest that the gains from optimal strategies, such as those based in machine learning techniques, may be substantial and highly cost-effective.

7 Conclusions and Discussion

In this paper we used a unique experimental data set Blattman et al. (2021) 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 toward the identification of the casual 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, an underexploited model in the criminology literature. The model added sufficient structure to allow for the identification of agents' fundamental parameters, such as their perceived utility of committing a crime. Once preferences were identified from observable data, we computed the own- and cross-elasticities of crime to time patrolling and we were able to evaluate alternative patrolling strategies. To the extent of our knowledge, both exercises are novel features in this literature, and are very relevant from a public policy perspective.

In this regard, our estimations showed an av-

[17] We thank an anonymous referee for asking us to clarify our measure of costs and benefits of policy scenarios. This motivated some of the exercises we did.

Table 6: Counterfactual analysis of different types of police patrol assignments

	Violent Crimes			Property Crimes			Total Crimes		
	Benefited Q. N	Benefited Q. %	Predicted # Mean (SD)	Benefited Q. N	Benefited Q. %	Predicted # Mean (SD)	Benefited Q. N	Benefited Q. %	Predicted # Mean (SD)
A. Base scenario									
Observed	-	-	12.31	-	-	24.61	-	-	36.92
	-	-	(8.51)	-	-	(17.89)	-	-	(21.18)
Predicted	-	-	10.39	-	-	21.33	-	-	32.90
	-	-	(4.12)	-	-	(9.90)	-	-	(9.44)
B. Counterfactual scenarios									
Uniform	544	53.43	10.68	544	53.43	22.64	544	53.43	34.37
			(4.12)			(11.12)			(10.22)
Proportional time	514	50.49	10.40	511	50.19	21.73	516	50.68	33.39
			(4.15)			(11.03)			(11.03)
Reassignment 3									
0% increase (Base case)	0	0.00	10.39	0	0.00	21.33	0	0.00	32.90
			(4.12)			(9.90)			(9.44)
10% increase	809	79.47	10.36	776	76.23	21.23	742	72.89	32.89
			(4.06)			(9.70)			(9.37)
20 % increase	664	65.23	10.36	621	61.00	21.20	585	57.47	32.93
			(4.06)			(9.63)			(9.47)
30 % increase	565	55.50	10.36	514	50.49	21.20	475	46.66	33.01
			(4.05)			(9.64)			(9.64)
40 % increase	475	46.66	10.38	427	41.94	21.25	400	39.29	33.08
			(4.09)			(9.71)			(9.81)
50% increase	421	41.36	10.38	359	35.27	21.29	339	33.30	33.15
			(4.09)			(9.75)			(9.99)
60 % increase	371	36.44	10.39	313	30.75	21.35	293	28.78	33.22
			(4.12)			(9.85)			(10.12)
70% increase	329	32.32	10.41	277	27.21	21.40	259	25.44	33.29
			(4.16)			(9.94)			(10.20)
80 % increase	294	28.88	10.44	252	24.75	21.44	232	22.79	33.41
			(4.21)			(10.01)			(10.51)
90 % increase	272	26.72	10.44	221	21.71	21.53	214	21.02	33.46
			(4.23)			(10.13)			(10.64)
100% increase	245	24.07	10.45	200	19.65	21.60	192	18.86	33.56
			(4.26)			(10.24)			(10.83)
Reassignment 4	513	50.39	9.96	538	52.84	20.04	636	62.47	32.00
			(2.80)			(5.97)			(5.43)

Notes: For each type of crime in each counterfactual scenario, three columns are reported. The first one reports the number of quadrants that receive more police patrol time than in the base scenario (benefited quadrants from now on). The second one reports the percentage of benefited quadrants. These two columns are mainly useful to assess reassignment 3, for the rest are just informative of how many quadrants are better off in terms of time provided the new reassignment. The third column reports the mean and standard deviation in parentheses of the predicted number of crimes that result from each police patrol time allocation scenario (first row corresponds to the observed number of crimes). Panel A reports the observed and predicted crimes for the base scenario (i.e., no reallocation). Panel B reports the predicted crimes for each of the four reassignment scenarios: uniform police patrol time, proportional police patrol time, reassignment 3 (see Appendix B), and reassignment 4 (see Appendix B). Reassignment 3 corresponds to the predicted number of crimes when the police patrol time is redistributed such that it is increased x% for quadrants that report both the highest police elasticity of crime and the highest share of crimes, where x is set to be either 0, 10, 20, 30, 40, 50, 60, 70, 80, 90, or 100. Thus, for the third reassignment methodology, each row reports a different percentage increase. The first row corresponds to a 0% increase, which is equivalent to the base scenario prediction reported in Panel A. These rows of the third reassignment report the exact numerical values depicted in Figure 3. Reassignment 4, on the other hand, corresponds to the predicted number of crimes when the police patrol time is redistributed following a recursive algorithm of 1% increments in time for quadrants that report both the highest police elasticity of crime and the highest share of crimes. The step-by-step of the third and fourth methodologies of redistribution are explained in Appendix B.

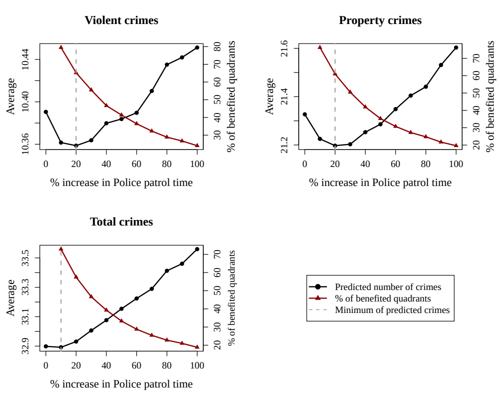


Figure 3: Relationship between the predicted average number of crimes and the percentage increase in police patrol time that receive the most elastic and insecure quadrants. Each of the graphs presents the results of implementing the first reassignment methodology explained in Appendix B. Black lines with dots report the predicted number of crimes. Red lines with triangles report the percentage of quadrants that received the x percentage increase in police patrol time. Vertical gray dotted lines indicate which percentage increase results in the minimum of predicted crimes. It seems that it is more effective to increase the police patrol time by a 10-20% for about 70% of the quadrants than increase it by 100% for < 20% of the quadrants. Source: Own elaboration.

erage location own-elasticity of total crime to police patrolling (measured in minutes of presence at a particular quadrant) of -0.19. That is, 1% more patrolling time reduces crime an average of 0.19%. This is in sharp contrast to the estimation of Blattman et al. (2021) that hardly reports any statistical significance of the effects of police patrolling on crime. Also, cross-elasticities show little support to spillover negative effects of police patrolling, another highly contested effect and discussed in the literature.

Taking advantage of our estimation of a structural model, we evaluated four different patrolling strategies: random, proportional (naive), a one-time time reassignment (reassignment 3), and a dynamic reassignment (reassignment 4) see 6 for details. Results show that allocating police time according to crime incidence and the elasticities of each quadrant (e.g., reassignment 4), could potentially reduce violent crime by 4.13% percent and property crime by 6%.

Our results show the value of using even simple models to allocate police in the city. These suggests that there is considerable space to improve the efficiency of police patrolling using state-of-the-art machine learning prediction models. We believe this is a promising extension of our paper and we leave it for future work.

References

Angrist, J. D., & Imbens, G. W. (1995). Two-stage least squares estimation of average causal effects in models with variable treatment intensity. *Journal of the American statistical Association*, *90*(430), 431–442.

Belloni, A., Chernozhukov, V., & Hansen, C. (2014). High-dimensional methods and inference on structural and treatment effects. *Journal of Economic Perspectives*, *28*(2), 29–50.

Bernasco, W. (n.d.). Modeling micro-level crime location choice: Application of the discrete choice framework to crime at places. *Quantitative Journal Criminology*, *26*.

Bernasco, W., & Nieubeerta, P. (2005). How do residential burglars select target areas. *British Journal Criminology*, *44*(4), 296–315.

Blattman, C., Green, D. P., Ortega, D., & Tobón, S. (2021). Place-based interventions at scale: The direct and spillover effects of policing and city services on crime. *Journal of the European Economic Association*, *19*(4), 2022–2051.

Braga, A. A., Welsh, B. C., & Schnell, C. (2015). Can policing disorder reduce crime? a systematic review and meta-analysis. *Journal of Research in Crime and Delinquency*, *52*(4), 567-588. Retrieved from `https://doi.org/10.1177/0022427815576576` doi: 10.1177/0022427815576576

Dewinter, M., Vandeviver, C., Vander Beken, T., & Witlox, F. (2020). Analysing the police patrol routing problem: A review. *ISPRS International Journal of Geo-Information*, *9*(3). Retrieved from `https://www.mdpi.com/2220-9964/9/3/157` doi: 10.3390/ijgi9030157

Fitzpatricka, D. J., Gorra, W., & Neill, D. B. (2020). Policing chronic and temporary hot spots of violent crime: A controlled field experiment. *British Journal Criminology*, *44*(4), 296–315.

Guangyin, J., Qi, W., Cunchao, Z., Yanghe, F., Jincai, H., & Jiangping, Z. (2020). Addressing crime situation forecasting task with temporal graph convolutional neural network approach. *12th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA)*, 474–478.

Hausman, J., & McFadden, D. (1984). Specification tests for the multinomial logit model. *Econometrica: Journal of the econometric society*, 1219–1240.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, *81*(396), 945–960. Retrieved from `http://www.jstor.org/stable/2289064`

Hu, Y., F., W., C., G., & H., Z. (2018). A spatio-temporal kernel density estimation framework for predictive crime hotspot mapping and evaluation. *Applied Geography 99*, 89-97.

Imbens, G. W., & Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, *62*(2), 467–475. Retrieved from `http://www.jstor.org/stable/2951620`

Jin, G., Wang, Q., Xia, Z., Feng, Y., Qing, C., & Jincai, H. (2019). Crime-gan: A context-based sequence generative network for crime forecasting with adversarial loss. *2019 IEEE International Conference on Big Data (Big Data)*, 1460-1469.

Kupilik, M., & Witmer, F. (2018). Spatio-temporal violent event prediction using gaussian process regression. *Crime and place: Crime prevention studies*(1(2)), 437-451.

McFadden, D. (2001, June). Economic choices. *American Economic Review*, *91*(3), 351-378. Retrieved from `https://www.aeaweb.org/articles?id=10.1257/aer.91.3.351` doi: 10.1257/aer.91.3.351

Mohler, G. (2014). Marked point process hotspot maps for homicide and gun crime prediction in chicago. *International Journal of Forecasting*, *30*(3), 491–497.

Mohler, G. O., Short, M. B., Brantingham, P. J., Schoenberg, F. P., & Tita, G. E. (2011). Self-exciting point process modeling of crime. *Journal of the American Statistical Association*, *106*(493), 100–108.

Moreno, J., Dulce, M., Riascos, A., Castaño, Y., & Rodriguez, P. (2020). A manifold learning data enrichment methodology for homicide prediction. *7th International Conference on Behavioral, Economic and Socio-Cultural Computing*.

Moreno, J., Quintero, S., Riascos, A., Nonato, L., & Sanchez, C. (2021). Homicide prediction using sequential features from graph signal processing. *21st Computing Conference*.

National Academies of Sciences, E., & Medicine. (2018). *Proactive policing: Effects on crime and communities* (D. Weisburd & M. K. Majmundar, Eds.). Washington, DC: The National Academies Press. Retrieved from `https://www.nap.edu/catalog/24928/proactive-policing-effects-on-crime-and-communities` doi: 10.17226/24928

Novak, K., Fox, A., Carr, C., & Spade, D. (2016, 09). The efficacy of foot patrol in violent places. *Journal of Experimental Criminology*, *12*. doi: 10.1007/s11292-016-9271-1

Pearl, J. (2009). Causal inference in statistics. an overview. *Statistics Surveys*, *3*, 96–146.

Pearl, J. (2019, feb). The seven tools of causal inference, with reflections on machine learning. *Commun. ACM*, *62*(3), 54–60. Retrieved from `https://doi.org/10.1145/3241036` doi: 10.1145/3241036

Ratcliffe, J., Taniguchi, T., Groff, E., & Wood, J. (2011, 08). The philadelphia foot patrol experiment: A randomized controlled trial of police patrol effectiveness in violent crime hotspots. *Criminology*, *49*, 795 - 831. doi: 10.1111/j.1745-9125.2011.00240.x

Reinhart, A., & Greenhouse, J. (2017). Self-exciting point processes with spatial covariates: modeling the dynamics of crime. *arXiv preprint arXiv:1708.03579*.

Shoaib, K., Fariha, S., Tianlu, Q., Yikang, R., Arezu Imran, B., Muhammad, S., ... Jiechen, W. (2018). Network constrained spatio-temporal hotspot mapping of crimes in faisalabad. *Applied Spatial Analysis and Policy 11*, 599–622.

Stec, A., & Klabjan, D. (2018). Forecasting crime with deep learning. *arXiv preprint arXiv:1806.01486*.

Telep, C., Weisburd, D., Gill, C., Vitter, Z., & Teichman, D. (2014). Displacement of crime and diffusion of crime control benefits in large-scale geographic areas: a systematic review. *Journal of Experimental Criminology*, *10*, 515–548.

Train, K. E. (2009). Logit. In *Discrete choice methods with simulation* (2nd ed., p. 34–75). Cambridge University Press. doi: 10.1017/CBO9780511805271.003

Xue, Y., & Brown, D. E. (n.d.). Spatial analysis with preference specification of latent decision makers for criminal event prediction. *Decision Support Systems*, *41*.

Zeileis, A. (2004). Econometric computing with hc and hac covariance matrix estimators.

A Haussman-McFadden (1984) IIA test

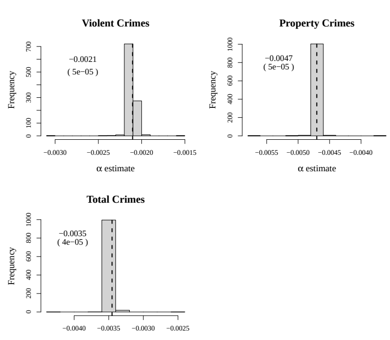


Figure 4: Haussman-McFadden (1984) (Hausman & McFadden, 1984) specification test of the independence of irrelevant alternatives (IIA) assumption. α estimated 1,018 times excluding in each iteration one different quadrant. Given that parameter estimates remain stable, IIA assumption seems to hold. Mean (SD) displayed for each graph. Source: Own elaboration.

B Intervention Scenarios

In this appendix we explain the reassignment strategies 3 and 4. The average predicted number of violent, property, and total crimes that result from these two reassignment strategies are displayed in Panel B of Table 6.

B.1 Reassignment 3

The third reassignment strategy considered tries to reallocate police patrol time from the quadrants that are both the less elastic and report less crimes toward the quadrants that are more elastic and report more crimes. In particularly, from the time that is taken away from the former, we increase by $x\%$ the time that was originally received by the latter. In this case, we try different values of $x \in \{10, 20, 30, 40, 50, 60, 70, 80, 90, 100\}$ for each type of crimes. This reassignment strategy is explained step by step below:

1. Obtain the minimum observed patrol time received by a quadrant:

$$P_{min} = \min\{P_1, P_2, \ldots, P_J\} \tag{13}$$

where P_j is the observed police time in quadrant j in our data set.

2. Set a *baseline* (BL) police patrol time for each of the J quadrants of the sample as follows:

$$P_j^{BL,c} = \min\{P_j, P_{min} \times \frac{w_j}{w_{min}} \times \frac{S_j^c}{S_{min}^c}\} \tag{14}$$

where w_j is the number of street segments covered by quadrant j, w_{min} is the number of street segments covered by the quadrant that reported the minimum patrol time, S_j^c is the share of crimes of type $c \in \{Violent, Property, Overall\}$ that occurred in quadrant j, and S_{min}^c is the share of crimes of type c that occurred in the quadrant that reported the minimum patrol time. That is, we set the *baseline* police patrol time as the minimum between the original police patrol time and the *minimum* police patrol time that is proportional to the number of street segments and the occurrence of crime at each quadrant. This guarantees that the total patrol time assigned is less than or equal to the total patrol time observed. Note that $P_j^{BL,c}$ is indexed by c, which means that we implemented this algorithm three different times for each of the three types of crime.

3. Obtain the residual (RES) time between the observed patrol time and the *baseline* time that will be reassigned, which is given by

$$P^{RES,c} = \sum_{j=1}^{J} \left(P_j - P_j^{BL,c} \right). \quad (15)$$

4. Distribute $P^{RES,c}$ across the first K_x quadrants that report both the highest police elasticity of crime (given the *baseline* time assigned) and S_j^c, such that each of the K_x quadrants receive an increase of $x\%$ in the police patrol time that they originally received, while the remaining $J-K_x$ quadrants receive $P_j^{BL,c}$. That is,

$$P_j^{RED,c} = \begin{cases} (1+x\%)P_j & \text{if} \quad j \in \{1,2,\ldots,K_x\} \\ P_j^{BL,c} & \text{if} \quad j \in \{K_x+1, K_x+2, \ldots, J\} \end{cases} \quad (16)$$

where $P_j^{RED,c}$ denotes the redistributed police patrol time. Note that K_x is indexed by x. It means that the number of quadrants that receive a $x\%$ increase, K_x, depends on that increase, since $P^{RES,c}$, the available time to be redistributed, is fixed. The greater the percentage increase x, the less quadrants are benefited by the redistribution. Therefore, a trade-off exists in this algorithm between the intensive and extensive margins of police presence across quadrants.

As was noted above, this algorithm was implemented 30 times, one for each $x \in \{10, 20, 30, 40, 50, 60, 70, 80, 90, 100\}$, for each of the three types of crimes that we consider in our work–violent, property, and total crimes. For each crime, we select the $x\%$ increase which results in the lowest level of crimes, and we compare it to the original scenario. Results of this algorithm are presented and discussed in Section 6.

B.2 Reassignment 4

The fourth reassignment strategy has the same spirit as the third one. However, in this case, we do not impose a uniform x percentage increase on the police patrol time across all the K_x quadrants that can receive it. Instead, we perform a recursive algorithm which iteratively assigns a 1% increase in police patrol time in the quadrant that reports both the highest elasticity and share of crime, and then recompute the elasticities and shares of crime for the whole sample. The algorithm iterates until $P^{RES,c}$ is fully redistributed. That is, steps 1 to 3 are the same as those from the previous methodology. The methodology step by step goes as follows:

1. Obtain the minimum observed patrol time received by a quadrant:

$$P_{min} = \min\{P_1, P_2, \ldots, P_J\} \quad (17)$$

where P_j is the observed police time in quadrant j in our data set.

2. Set a *baseline* (BL) police patrol time for each of the J quadrants of the sample as follows:

$$P_j^{BL,c} = \min\{P_j, P_{min} \times \frac{w_j}{w_{min}} \times \frac{S_j^c}{S_{min}^c}\} \quad (18)$$

3. Calculate the residual time to be redistributed as

$$P^{RES,c} = \sum_{j=1}^{J} \left(P_j - P_j^{BL,c} \right). \quad (19)$$

4. Rank the quadrants according to both their estimated police elasticity of crime, $E_{S_j^c, P_j^{BL,c}}$, and their predicted share of crime, S_j^c.

5. Increase by 1% the baseline police patrol time of the quadrant that is ranked first (i.e., the one with both the highest elasticity and share of crime) in the previous step. That is,

$$P_j^{RED,c} = \begin{cases} (1.01)P_j^{BL,c} & \text{if} \\ \quad j \in \text{argmax}_{j\in\{1,\ldots,J\}}\{E_{S_1^c, P_1^{BL,c}} \times S_1^c, \ldots, \\ \qquad E_{S_J^c, P_J^{BL,c}} \times S_J^c\} \\ P_j^{BL,c} & \text{otherwise.} \end{cases} \quad (20)$$

6. Update $P^{RES,c}$ such that

$$P^{RES,c} = P^{RES,c} - (1.01)P_j^{BL,c} \quad (21)$$

7. Reestimate $E_{S_j^c, P_j^{BL,c}}$ and S_j^c with the new distribution of police patrol times.

8. Repeat steps 5) through 8) until $P^{RES,c} = 0$.