Understanding the effect of compliance to quarantines and lockdowns on domestic violence occurrence in Bogotá

Exploring impact and response to the COVID-19 pandemic in Latin America and the Caribbean using mobility data

Presented to:

UNDP in Latin America and the Caribbean

Authors[1]:

Paula Rodríguez Díaz, Associate Subdirector at Quantil

Oscar Gómez, Researcher at Quantil

Douglas Newball, Researcher at Quantil

Johan García, Researcher at Quantil

Juan Sebastián Moreno, Associate Director at Quantil

Álvaro Riascos, Director at Quantil & Assistant Professor at Universidad de los Andes

John David Castro, Data Analyst of OAIEE at District Secretary of Security of Bogotá

Jose David Panqueva, Data Analyst of OAIEE at District Secretary for Security of Bogotá

[1] The opinions expressed in this publication are solely those of the authors. They do not purport to reflect the opinions or views of the District Secretary of Security of Bogotá, Quantil or its members.

Abstract

Quarantines and lockdowns due to COVID-19 have had multiple social effects in terms of family isolation, economic uncertainty and general stress that have a potential increase in domestic violence. In Bogotá, Colombia, official reports show an increase in the number of domestic violence reports in the period January-May 2020 of 20.3% (2.702) relative to the same period in 2019. We use information on imposed mobility restrictions and quantified mobility changes to develop a spatio-temporal domestic violence occurrence prediction model accompanied by machine learning interpretation techniques to understand the association of mobility patterns with the spatio-temporal behavior of domestic violence incidents in Bogotá. Additionally, we estimate different OLS-based-regression models to better understand the causal relationship between the quarantine lockdown duration and domestic violence, which we initially found to be significant from the predictive spatio-temporal approximation.

This research is presented as a joint work by researchers from Quantil | Matemáticas Aplicadas and the Office of Information Analysis and Strategic Studies of the District Secretary for Security, Coexistence and Justice of Bogotá. Both institutions have previously worked together as part of a research project funded by Colciencias on crime prediction models for robberies and homicides in Bogotá.

JEL Codes: C21, C23, J12, J18

Keywords: Domestic violence, lockdown, mobility data, COVID-19

Contents

1. Introduction 1

2. Related Research 1

3. Data 3

3.1. Data on domestic violence 3

3.2. Data on mobility 5

4. Methodology 6

4.1. Crime prediction models – Spatio-Temporal Generalized Additive Model 6

4.2. OLS-based Regression Models 8

OLS 8

Mediation analysis 8

Heterogeneous effects 9

5. Implementation 9

6. Results 12

6.1. Crime prediction models – Spatio-Temporal Generalized Additive Model 12

Out-of-sample performance evaluation 12

Model interpretability - partial dependence plots 13

6.2. OLS-based Regression Models 18

Regression analysis 18

Mediation analysis 19

Heterogeneous effects 20

7. Discussion and conclusions 22

References 24

1. Introduction

Domestic violence, particularly violence against women, has increased during COVID-19 related to quarantines and lockdowns around the world. In Bogotá, official reports show an increase of 20.3% (2,702 cases) in the number of domestic violence reports from January to May 2020 relative to 2019. In the subsequent months, reported incidents decreased and the cumulative variation in October relative to the number of reports in 2019 decreased by 1.3% (407 cases). During this period, mobility restrictions in the city affected the way citizens approached denunciation and reporting channels, as this meant barriers were imposed to in-person procedures. Additionally, mechanisms to protect domestic violence victims from physical or psychological abuse and provide personal support also faced challenges because of these barriers.

In addition to in-person report mechanisms, Bogotá has two domestic violence support hotlines (Helping 155 and the District's Purple Helpline Women Listening to Women) which reinforce, and in some cases substitute, the in-person assistance and respective official complaint. Understanding how lockdowns affect domestic violence vulnerability, how the reporting channels fluctuate in time and space, and if there is a way to prioritize assistance to the victims, will provide support for further socio-economic interventions by authorities that assist and prevent this type of crime.

Previous work on mobility changes in Bogotá during the COVID-19 outbreak such as (Dueñas, et. al.) have shown that higher levels of poverty are associated with higher mobility flows. Even more, this correlation becomes significant after the implementation of lockdown policies. Given this situation, our research incorporates geo-referenced socioeconomic and demographic data in order to account for variations in mobility data derived from socioeconomic and demographic protected features.

2. Related Research

Traditional criminological theories study the influence of biological factors, individual development, and social forces in the creation of criminals. In contrast, environment-centered theories study the immediate circumstances in which crime events occur. According to these theories, the immediate environment affects criminal behavior. In this sense, due to the mandatory confinement measures to combat the spread of COVID-19, the convergence of potential victims and potential perpetrators in the same place generates risk conditions, that is, the immediate environments for the occurrence of acts of domestic violence are materialized.

Regarding the measures taken by governments around the world to combat the spread of COVID-19, Brooks et al. (2020) perform a literature review that gathers the evidence related to the psychological effects that these measures produce on people. They analyzed 24 studies that were selected under the criteria of the rigorousness of the documents. In the 24 studies, the authors identified a series of stressors that can produce negative psychological effects such as post-traumatic stress, confusion, and anger. Stressors include prolonged quarantine, fear of contagion, frustration, boredom, lack of resources, lack of information, and uncertainty about quarantine.

In line with Brooks et al. (2020), different studies have identified that perpetrators of domestic violence experience anger or respond with anger to stressful or distressing situations. Norlander et al. (2005) conducted a meta-analysis that included 33 studies on the relationship between hostility, anger, and domestic violence committed by men against women. These authors found slightly higher levels of hostility and anger in men who committed domestic violence against women. Likewise, Stuart et al. (2006, 2008) and Shorey et al. (2011) found a significant association between anger and women arrested for domestic violence. In turn, these authors found that anger predicts domestic violence, even after controlling for alcohol abuse, relationship problems, and having been a victim of psychological violence.

Some recent studies are pointing towards an increase in domestic violence under the context of mobility restrictions related to the COVID-19 pandemic: First, Boserup et al. (2020) conducted a review of official figures on domestic violence in the United States after the entry into force of mandatory confinement measures. In San Antonio, Texas, there was an 18% increase in domestic violence calls during the month of March 2020, compared to the same period last year (Department of Government and Public Affairs, 2020). In Jefferson County, Alabama, a 27% increase in calls for domestic violence was evidenced when studying the same periods (Jefferson County Sheriff's Office, 2020). In Portland, Oregon, there was a 23% increase in arrests for domestic violence during the weeks following the declaration of the measure, compared to previous weeks (Portland Police Bureau, 2020). Finally, New York City registered a 10% increase in complaints of domestic violence during the month of March 2020, compared to the same period of the previous year (New York City Police Department, 2020).

Also, for the United States, Leslie and Wilson (2020) found that a 7.5% increase in domestic violence-related calls to police, in 14 large cities after March 9, was above the seasonal trends and could be attributed to the pandemic. The date was determined by observing a drastic reduction in mobility data and is before official stay-at-home orders. In a similar study, Sanga and McCrary (2020) determined that the pandemic generated a greater increase in domestic violence in neighborhoods with no recent history of violence.

In Latin America, Perez-Vicent and Carreras (2020) identified a 32% increase in calls to the domestic violence hotline in Buenos Aires, Argentina as a result of the mandatory lockdown. They also found evidence of substitution on the reporting channel as calls made by the police fell while direct calls from victims increased.

The COVID-19 pandemic is not the only context in which restricted mobility has been linked to a higher incidence of domestic violence. There are other experiences such as the Ebola outbreak in West Africa, which produced numerous studies regarding domestic violence. Although the literature has not been rigorous in the use of data, it has found a certain relationship of the stressful conditions resulting from some measure of confinement. Documents by non-governmental organizations (UNDP, 2015) show an increase in domestic violence during the Ebola outbreak in West Africa. Other documents also show a similar trend in sexual crimes (Onyango et al., 2019). These studies identified that the measures taken to deal with the outbreak increased the contact time between the possible victim and the possible aggressor (Onyango et al., 2019; IRC, 2015; UNDP, 2015). In turn, they identified that the fight against gender-based violence lost priority compared to the fight against the Ebola outbreak. In particular, they established that the police and the judicial system focused on the fight against the Ebola outbreak, which could contribute to creating an "environment of impunity". In contrast, some case studies found that given the conditions there was an increase in adolescent pregnancy as a consequence of the Ebola outbreak (Korkoyah et al., 2015), another little-explored form of violence.

A risk case that was explored by Zahran et al. (2009) where they studied criminal behavior in Florida after the occurrence of different natural disasters registered between 1991 and 2005. These authors used both geographic analyzes and binomial regressions of conditional effects to evaluate the impact of natural disasters on crime at a county level. It was found that while property crimes and violent crimes registered a significant decrease, gender-based violence registered a significant increase. Similar increases in gender-based violence were reported after the Mount St. Helens eruption in Othello, Washington, in 1980 (Adams et al., 1984); after Hurricane Katrina that affected southern Mississippi counties in 2005 (Schumacher et al., 2010); after the wildfires that hit Australia in 2019 (Parkinson, 2019); and after the earthquake that struck Haiti in 2010 (Weitzman et al., 2016).

Finally, Stark et al. (2011) carried out a meta-analysis of violence against women during complex emergencies. In this study, “complex emergencies” were defined as “humanitarian crises in a country, region or society where there is a considerable or total decomposition of authority due to internal or external conflicts” (Stark et al., 2011). These authors found that the reported intrafamily violence throughout the 10 studies included in the meta-analysis was high. Among the causal mechanisms hypothesized by these studies are an increase in stress, an increase in controlling behaviors on the part of men, an increase in unemployment, the closure of care centers for women, and an increase in alcohol consumption. These causal mechanisms are in line with the causal mechanisms highlighted in the gender-based violence literature during the Ebola outbreak in West Africa.

3. Data

3.1. Data on domestic violence

From our collaborative work with the Information Analysis Office of the District Secretary for Security of Bogotá we were able to access two main georeferenced data sources that record all the crimes occurring in Bogotá and all main personal assistance given to women in the city. In first place, we have the georeferenced information from the **Delinquential, Contraventional and Operative Statistical Information System of the National Police** (Sistema de Información Delincuencial, Operacional y Operativo de la Policía Nacional, SIEDCO). This information system reports all crime incidents in the city of Bogotá. The data available in SIEDCO describes each victimizing event of different criminal phenomena, including intra-family violence or in this context, domestic violence[2]. Particularly, we can identify the incident's date, hour, place of occurrence, as well as the gender and age of the victim.

[2] According to Colombia’s legislation, the following addresses intra-family violence: Articles 11-14, 42 and 43 of The Constitution, Law 82 of 1993 (which made public norms to support women heads of households), Law 294 of 1996 modified by Law 1257 of 2018 (which created public measures to prevent, rectify and punish intra-family violence), Law 599 of 2000 (which made public the crimes against what constitutes a family -articles 299 to 238 of Penal Code), Law 823 of 2003 (which made public norms for women's equal opportunities), Law 1098

Reporting domestic violence requires the government -in this case the town hall- to provide victims of domestic violence immediate protection from further physical or psychological abuse. The following steps are necessary to ensure that the victims are assisted and protected by authorities: (i) immediate protection for themselves and their children at Family Commissaries (Comisarías de Familia) if the aggressor is a part of the family, or at the Attorney General's Office (Fiscalía General de la Nación, FGN) if the aggressor is from outside the family unit. The victim can also call the National Police to provide protection. These authorities will determine whether the victim requires a shelter; (ii) a file police report must be presented to the authorities at the Comprehensive Assistance Centre for Intrafamily Violence (Centro de Atención Integral contra la Violencia Intrafamiliar, CAVIF), the Comprehensive Assistance Centre for Victims of Sexual Crimes (Centro de Atención Integral de Víctimas de Delitos Sexuales, CAIVAS), the Immediate Response Units (Unidades de Reacción Inmediata, URI) of the FGN, Family Commissaries or local police stations of the National Police (The Immigration and Refugee Board, 2017).

In the cases where violence is reported, it is possible that the victim reports the incidents with other authorities different from the FGN or the National Police and in this case the event will not be reflected in SIEDCO. Additionally, due to the quarantines and lockdowns, police reports were expected to decrease in most of the cases due to the lack of face-to-face assistance from the competent authorities and the mobility limitations of the population. Previous to the pandemic, the presence of the victims in the authorities' installations was required in most cases and no virtual assistance was immediately implemented or frequently used to report domestic violence.

One of the main domestic violence attention lines and incident report mechanisms used during the lockdowns, given its non-face-to-face dynamic, was the **District Purple Line** (Línea Púrpura Distrital) from the District Secretary of Women in Bogotá. To overcome the lack of completeness in the SIEDCO data due to mobility restrictions, we also work with aggregate information from the district attention line for women, the District Purple Line. This line offers all day attention since 2015 to women older than 18 years in any kind of situation where their rights are affected. This line is different to other hot lines such as the Emergency Line 123 (NUSE123), in the sense that it is a mechanism created to give support to women who have any kind of issues with their rights, making it the most frequent hot line for reporting domestic violence incidents. The diversity of problems received by the District Purple Line is managed by a team of psychologists, nurses and social workers that provide the necessary attention and response for the victims. The District Purple Line not only is a telephone line but can also be reached by email and a WhatsApp account to answer different requests that citizens make related to women vulnerability.

The District Purple Line is not a mechanism to officially report crimes with the law enforcement as in SIEDCO or emergencies as in NUSE123. The Purple Line’s goal is to provide attention and activate the institutional measures to prevent crimes or right affectations. However, it serves as a trustworthy source for domestic violence records. Additionally, the District Purple Line aims to reduce the impact of different crimes that have been perpetrated. According to the above mentioned, the District Purple Line attends situations or problems like:

---

of 2006 (which made public the laws on children and adolescents), Law 1639 (which defines the victim's protection measures when there is an acidic substance attack) and the Law 1719 of 2014 (which seeks to guarantee access to justice for victims of sexual abuse).

- Violence against women on public or private places.
- When women identify situations where they feel threatened.
- It provides different paths for attention according to law 1257 of 2008.
- It is the main channel to attend survivors of femicides and their families.
- The line provides information about public supply about women rights.
- It provides information about health, especially sexual and reproductive rights.

For the implemented model, we work with daily and spatially aggregated observations of domestic violence from SIEDCO and the District Purple Line. These observations are aggregated in regions (urban sectors) of approximately 0.5 km^2, with 623 spatial divisions over the urban region of Bogotá. We calculate, for each sector, if there was at least one reported incident from SIEDCO or the District Purple Line. Figure 1 shows the number of urban sectors with at least one occurrence of domestic violence from January 1st to October 17th of 2020. As Figure 1 shows, Bogotá had four general changes in terms of lockdown restrictions, starting with a drill and general lockdown in the third week of March. As seen in this figure, occurrences of domestic violence seem to decrease when the drill and general lockdown are implemented. However, this specific decrease and the apparent decrease in the following months correspond primarily to a possible shift in the reporting mechanisms as stated previously. This means that, as overall the number of reports in SIEDCO is much greater than the reports in the District Purple Line, the overall trend is that reports decreased heading into lockdown and increasing afterwards. This occurred even though the number of reports through the District Purple Line grew steadily since the beginning of March, however, never enough to change the overall trend dominated by SIEDCO reports.

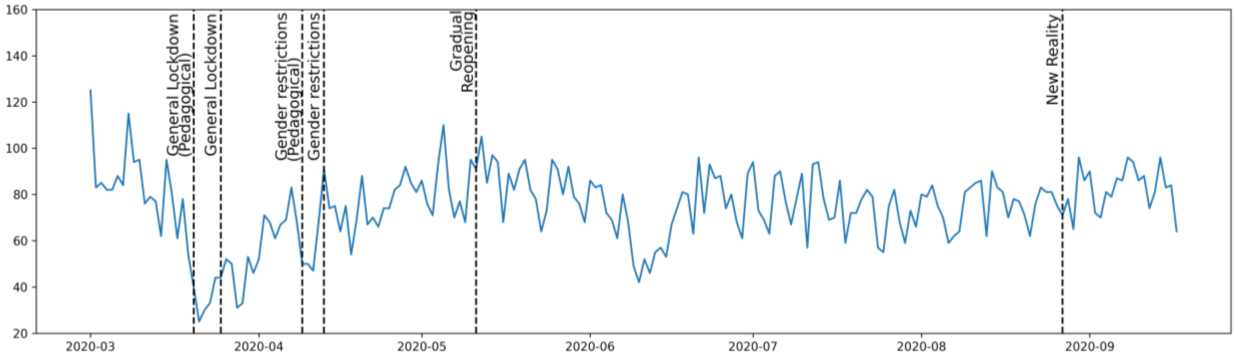

Figure 1. Daily number of sectors with domestic violence occurrence registered in SIEDCO or District Purple Line

3.2. Data on mobility

We use mobility data provided by GRANDATA to quantify the degree of compliance with quarantine and lockdown restrictions in Bogotá. This data corresponds to the daily percentage variation in mobility change with respect to the mobility registered on March 2nd for urban sectors in Bogotá. Urban sectors are spatial divisions with a mean area of 0.5 km^2, with 623 spatial divisions over the urban region of Bogotá. Figure 2 shows the average mobility change with respect to the corresponding day of the week from the first week of March and Figure 3 shows these same values averaged by week. These graphs enable a mobility change comparison that accounts for changes given by day of the week and removes the effect of decrease in mobility from the weekends, making it possible to better understand change in mobility due to lockdown restrictions rather than to a single day (March 2nd).

We use the stratum assigned to each urban sector as a proxy for socioeconomic status. Stratum is a government-assigned level from 1 to 6 according to socioeconomic characteristics such as housing conditions. Stratum is assigned progressively, having 1 as the level with poorest quality and 6 corresponding to the richest neighborhoods. The differences in the average mobility changes between different stratum suggest that lockdown restrictions were complied with in different degrees given the corresponding socioeconomic capacities, having lower stratum few or almost none decrease in mobility patterns.

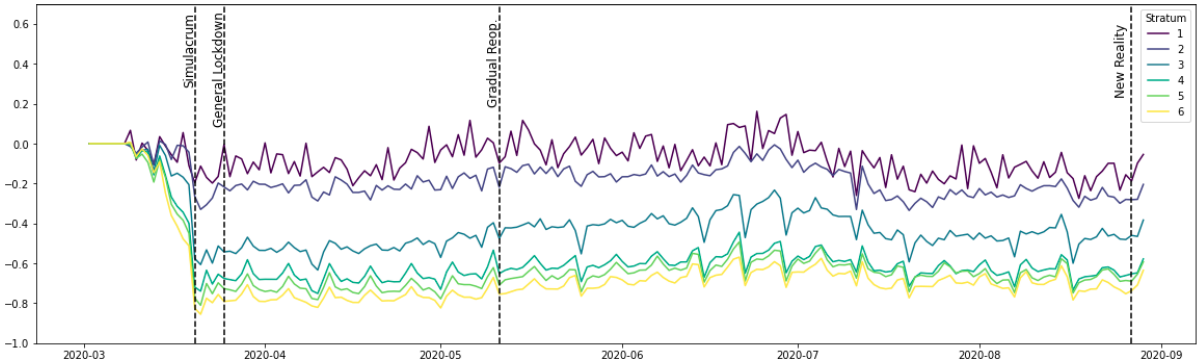

Figure 2. Average daily mobility change by stratum with respect to the corresponding day of the week for the first week of March

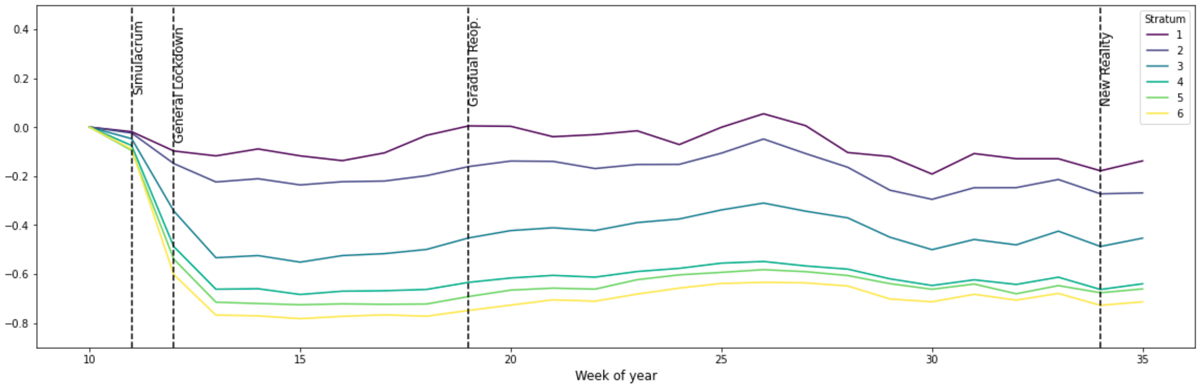

Figure 3. Average daily mobility change by stratum with respect to the corresponding day of the week for the first week of March – Weekly Average

4. Methodology

4.1. Crime prediction models – Spatio-Temporal Generalized Additive Model

State-of-the-art crime prediction models exploit the spatio-temporal clustering patterns and the self-exciting nature of criminality to predict vulnerable crime areas. However, specific types of crimes, such as domestic violence, do not necessarily present a self-exciting nature. Our methodological proposal corresponds to replacing the self-exciting component in well-known crime prediction models, such as the ones presented in (Wang, et. al.) for spatio-temporal mobility patterns that can better reflect the occurrence of domestic violence. This proposal is done under the assumption that mobility patterns have a higher correlation with domestic violence incidents than previous incidents occurred in a spatial neighborhood.

Models such as the ones presented in (Wang, et. al.) and (Mohler, et. al.) exploit the spatio-temporal clustering patterns of crimes in order to accurately predict future crime occurrences in a spatio-temporal manner. In particular, (Wang, et. al.) also incorporates socioeconomic and demographic features to discover underlying factors related to crimes and to predict future crime incidents. This model, which consists of a Spatio-Temporal Generalized Additive Model (S-T GAM) model, accounts for nonlinear relationships between the explanatory variables and the response variable and can incorporate different types of data, such as spatial, temporal, geographic, and demographic data, to make predictions. In this way, this type of crime prediction models can leverage mobility data to understand its effects on the occurrence of domestic violence.

The generalized additive models model non-linear relationships between the explanatory variables, in our case variables with spatiotemporal attributes, and the intensity of domestic violence incidence to be estimated. The foregoing allows potentially obtaining more precise predictions than those obtained with linear models. GAMs can be used to model continuous variables such as crime intensity or binary variables such as the occurrence or incidence of crime in defined spatio-temporal units. The generalized additive model is defined by a response variable that we want to predict - in our case the crime intensity λ - in a defined spatiotemporal location whose distribution belongs to the exponential family, and g as a link function that establishes the relationship between the mean and the linear predictor:

$$y \sim ExponentialFamily(X),$$
$$g(X) = \beta_0 + \sum_{j=1}^{p} f_j(x_j).$$

This model fits functions f_j to each variable $x_j$to capture the non-linear relationships between the predictors and the response variable. Adjusted functions f_j correspond, by default, to functions composed of splines, which allows to model non-linear relationships without having to manually test different transformations on each variable. For a classification task to estimate the incidence of domestic violence occurrences we consider a binomial distribution and a logistic link function.

We also propose to make use of existing model interpretability techniques in order to gain a better understanding of the models developed. The importance of model interpretability has gained the attention of researchers in recent years (Doshi-Velez et al., 2017), especially when it pertains to high-stakes decision making. Applying these techniques can help identify if a model is incorporating hidden biases, spurious correlations, or false generalizations (Hohman et al., 2019). Many of these methods treat the model as a black box and so rely on perturbations to the input data (Ribeiro et al., 2016). However, they are vulnerable and may provide unstable and misleading explanations (Slack et. al., 2020). In our analysis we choose to use the ST-GAM model because of its intrinsic interpretability properties, which allow us to make an interpretation of the model directly without recurring to ad-hoc techniques which might be misleading.

4.2. OLS-based Regression Models

To understand the causal relationship between the quarantine lockdown duration and domestic violence, different OLS-based-regression models are estimated. First, a multiple OLS model that relates the two variables and controls for possible confounding factors. Second, mediation analysis is run to evaluate the mobility change as a driver of the relationship. Third, an OLS model that includes interaction terms to explore some heterogeneous effects. These three models are explained below.

OLS

To estimate the impact of the quarantine lockdown duration and the occurrence of domestic violence the following OLS model is estimated

$$Total\ cases_s = \beta_0 + \beta_1 Lockdown\ days_s + X_s\Gamma + \varepsilon_s$$

where $Total\ cases_s$ is the count of reported domestic violence cases between March 1 and September 9 of 2020 in sector s of Bogotá D.C., $Lockdown\ days_s$ is the count of effective lockdown days faced by sector s in the same period, X_s is a vector of socioeconomic controls, and ε_s is the error term. β_0 is the constant term, β_1 is the coefficient that captures the relationship between lockdown duration and domestic violence, and Γ is the vector of coefficients specific to the variables included as controls.

Under the conditional independence assumption (i.e., exogeneity of independent variables), the OLS estimator of β_1 is interpreted as the causal impact of lockdown duration on domestic violence. Given that the government ordered lockdown in the different sectors of Bogotá D.C. based on the incidence of COVID-19 and not on the occurrence of domestic violence, the exogeneity assumption is plausible. However, it can be argued that: 1. the places where it is more probable to occur domestic violence are also the poorest, 2. that the poorest places are those where the lockdown enforcement might be less or the population density is higher so that the spread of the virus is greater in those places, and, therefore, 3. β_1 could be capturing more than the confinement effect on domestic violence. Thus, to control for confounding factors and avoid potential biases socioeconomic and poverty indicators such as household population density, social class, access to services, age group distribution, and gender distribution are included in the regression.

Mediation analysis

Given the estimation of a causal relationship, it is relevant to understand what are the channels through which the lockdown days impacts domestic violence. Due to the availability of mobility change data in the same period (March 1 to September 9 of 2020), mobility change is evaluated as a possible channel. To do that, we follow MacKinnon, Fairchild & Fritz's (2007) mediation analysis methodology. In particular, to evaluate a potential mediator, the following three equations are estimated:

1. $Total\ cases_s = \beta_0 + \beta_1 Lockdown\ days_s + X_s\Gamma + \varepsilon_s$
2. $Mobility\ change_s = \alpha_0 + \alpha_1 Lockdown\ days_s + u_s$

3. $Total\ cases_s = \gamma_0 + \gamma_1 Lockdown\ days_s + \gamma_2 Mobility\ change_s + X_s\theta + \varepsilon_s$

where $Mobility\ change_s$ is the average mobility change in sector s between March 1 to September 9 of 2020.[3]

The first equation is the same as the one proposed in the previous subsection to test whether the lockdown duration has an impact on domestic violence. The second equation tests whether lockdown days had an impact on mobility change too. And the third equation tests whether the impact of the lockdown days on domestic violence remains the same or it is less when one controls for the mediator (i.e. Mobility change). Therefore, mobility change can be interpreted as a channel if: 1. $\beta_1 \neq 0$ (we need an impact before evaluating mediators), 2. $\alpha_1 \neq 0$ (we need a relation between lockdown days and the mediator so that it can be considered a mediator), 3. $\gamma_2 \neq 0$ (the mediator should be active even after controlling for the treatment), and 4. $|\gamma_1| < |\beta_1|$ (the effect of lockdown days is reduced when one takes out the mediator through which the treatment affected domestic violence).

Heterogeneous effects

Lastly, to explore whether the effect of the lockdown on domestic violence varies with sector characteristics, the following OLS regression model that included an interaction term is estimated:

$$Total\ cases_s = \gamma_0 + \gamma_1 Lockdown\ days_s + \gamma_2 M_s + \gamma_3 Lockdown\ days_s + \gamma_4 Lockdown\ days_s \times M_s + X_s\Pi + \omega_s$$

where M_s is the moderator variable (either continuous or dichotomous) considered to study heterogeneous effects and $Lockdown\ days_s \times M_s$ is the interaction term which coefficient, γ_4, captures how the lockdown changes with M_s. The rest of the terms and variables are the same as those previously described.

Specifically, to understand how the lockdown effect varies with respect to 1. Gender composition of the sector, 2. number of persons by household, and 3. the dominant social class in sector, three different regressions are estimated (so that the reference category is easy to interpret). Each of the three regressions defines M_s as each of the three variables described before.

5. Implementation

In order to use the Spatio-Temporal GAM model, we merge data from GRANDATA (mobility), SIEDCO (domestic violence), and DANE (census). Mobility data is provided for each day between March 3, 2020 and August 29, 2020. It consists of percentage changes in mobility for each day with respect to a baseline set on March 2, 2020. The data is given for urban sectors within the city of Bogotá (623 in total). Census data is also available for each of these sectors. From this data we obtain the static, spatial characteristics used in the ST-GAM

[3] Mobility change of each weekday was calculated with respect to the corresponding weekday of the first week of March.

model, which include income, household, education, and other characteristics for the population in each urban sector (see Table 1 for the full list of variables). Finally, for each occurrence of domestic violence recorded in the SIEDCO and District Purple Line databases, we have the date and the urban sector in which it took place.

We also created a database containing the data for the dates and types of quarantine that were imposed in Bogotá over the same period. Bogotá underwent several types of quarantine over the months between February and August of 2020. From March 20 to March 24, the city went through a quarantine drill, which was immediately followed by a period of general quarantine that extended until May 10. From then on, a series of localized, strict quarantines were implemented in varying areas of the city. These strict quarantine decisions were made at the locality (20 in total) and Urban Planning Zone (UPZ) (125 in total) levels. We assigned a numerical value to each state of quarantine (see Table 2 for the list of states). However, UPZs are not exactly subdivided into Urban Sectors, as they are subdivisions with different purposes and defined by different entities. Therefore, some Urban Sectors have a significant portion of their area outside of one single UPZ. To be able to match the quarantine data to the rest of the datasets, we assigned the quarantine state corresponding to the UPZ with the largest share of area for a given Urban Sector, if such UPZ accounted for 90% or more of the area. Otherwise, we assigned the stricter state between the two UPZs with the highest share of the area according to the ordering in Table 2.

From the mobility and quarantine data we created additional temporal variables to use as input for the ST-GAM model. These variables provide better regressors for the model as they include historic information from the past for each observation. For mobility data, we also calculated minimum, maximum, and standard deviation over the past week for each observation, as well as percentage changes with respect to the past day, past week, and baseline for the day of the week (taken as the average mobility over the two weeks without quarantine). For quarantine data, we calculated consecutive days in quarantine, consecutive days without quarantine, and total time in quarantine since the beginning of March. Finally, we included a dummy variable encoding the number of days without a domestic violence incidence in the sector, following the original ST-GAM model (Wang, et. al.).

Variable Name	Description
movilidad	Mobility as a percentage of the mobility on March 2
min8d_mobility	Minimum mobility over the past week
max8d_mobility	Maximum mobility over the past week
std8d_mobility	Standard deviation of mobility over the past week
cambio_diario_mov	Percentage change in mobility over past day
cambio_semana_mov	Percentage change in mobility over past week
cambio_base_dia_mov	Percentage change in mobility over baseline days
seguido_sin_cuarentena	Consecutive days without quarantine
seguido_cuarentena	Consecutive days with quarantine
total_cuarentena	Total days in quarantine
indicador_cuarentena	Quarantine indicator
indicador_estricta	Strict quarantine indicator
seguido_sin_inci	Consecutive days without domestic violence incidents
AREA	Area of the urban sector in square meters
viviendas	Number of residences
personas	Number of people
energia_elec	Proportion of residences with electricity
parques_zverde	Number of parks / green zones
personas_por_hogar	People per household
personas_por_vivienda	People per residence
personas_por_area	People per area
estrato_n	Income level n
uso_vivienda	Residential units
uso_mixto	Mixed use units
uso_no_residencial	Non-residential units
tipo_vivienda_casa	Residence type – house
tipo_vivienda_apart	Residence type – apartment
tipo_vivienda_cuarto	Residence type – room
ocup_viv_viv_temporal	Residence type – temporary
ocup_viv_desocupada	Residence type – unoccupied
gas_1	Proportion of residences with natural gas
internet_1	Proportion of residences with internet
edad_n	Proportion of people of age $(n-1)*10$ to $n*10$
educacion_prim	Proportion of people with elementary education
educacion_secund	Proportion of people with secondary education
educacion_superior	Proportion of people with undergraduate education
educacion_post	Proportion of people with graduate education
educacion_ninguno	Proportion of people with no education
dia_5	Indicator for Saturdays
dia_6	Indicator for Sundays
fin_semana	Indicator for weekends

Table 1. Spatio-temporal variables extracted and used in the ST-GAM model

No Quarantine	0
New Reality	1
Gradual Reopening	2
Quarantine Drill	3
General Quarantine	4
Strict (Local) Quarantine	5

Table 2. Quarantine States in Bogotá between March 2, 2020 and August 29, 2020

6. Results

GAMs are powerful machine learning models not only because of their well performing predictive power, but also because they allow for human interpretability. By examining the individual contributions of different features to the model's predictions through Partial Dependence Plots (PDPs) we can gain an intuitive understanding of what the model is basing its predictions on. In this section we evaluate the model's performance on out-of-sample test data and analyze the PDPs that the model produces.

6.1. Crime prediction models – Spatio-Temporal Generalized Additive Model

Out-of-sample performance evaluation

Even though the goal of our methodology is to analyze the effect of higher compliance with quarantines with domestic violence, we check the robustness of our model on out-of-sample data to ensure the validity of our conclusions. We use two different validation methods, a temporal validation, where we train the model on data before August 1, 2020 and test it on the remaining data, and a spatial validation, where we randomly select 80% of the urban sectors for training and test on the remaining 20%. As an evaluation metric, we use a standard metric used to evaluate crime prediction models, which is a modification of the Hit Rate measure that controls by the percentage of area marked as a possible crime hotspot. The measure is calculated as follows: given a temporary unit in the test set, and a percentage of the total area to be monitored, we calculate the sectors with the highest predicted probability of having a crime according to the model. Then the Hit Rate is the number of crimes that occurred in these sectors, divided by the total number of crimes that occurred over the given time period. However, because the higher the percentage of the total area covered, increases the number of sectors to be monitored, and the Hit Rate as such, we must control for this percentage of area covered. The results can be averaged over the different possible percentages of area to be covered, resulting in a metric similar to the typical ROC-AUC scores. However, in crime detection it is usually desired to deploy available, limited units as efficiently as possible, so one usually focuses on Hit Rate measures below a given threshold of area covered. In the results, we give for reference what the metrics would be for the best possible predictors, in which all marked sectors would have crime occurrences. Figure 4 shows examples of these curves for different evaluation methods. Table 3 shows the full performance results, separated by current quarantine state and income level.

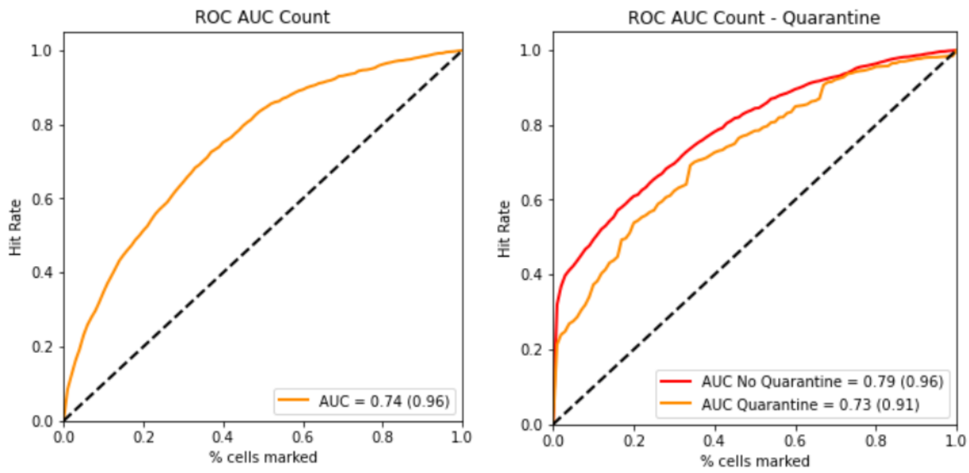

Figure 4. Hit Rate / Area Covered curves for different evaluation methods. Left: Overall performance on ST-GAM on the Temporal evaluation. The best possible total AUC is in parentheses. The Hit Rates at 5%, 10%, and 20% of area covered are 23%, 35%, and 51% respectively. Right: Performance of ST-GAM on the Spatial evaluation, separated by the current quarantine state. Best possible AUCs are in parentheses. Hit Rates at 5%, 10%, and 20% of area covered are 42%, 49%, and 61% for no quarantine and 28%, 37%, and 54% for quarantine, respectively.

	Count_HR_AUC (Best possible)	Count_HR_5	Count_HR_10	Count_HR_20
GAM - Temporal validation	0.742 (0.955)	0.231	0.347	0.514
GAM - T Low income	0.724 (0.948)	0.209	0.317	0.49
GAM - T High income	0.673 (0.978)	0.169	0.305	0.471
GAM - T No quarantine	0.726 (0.951)	0.21	0.328	0.491
GAM - T Quarantine	0.769 (0.965)	0.332	0.457	0.595
GAM - Spatial validation	0.716 (0.951)	0.231	0.326	0.485
GAM - S Low income	0.707 (0.946)	0.243	0.332	0.476
GAM - S High income	0.769 (0.967)	0.416	0.462	0.622
GAM - S No quarantine	0.788 (0.962)	0.423	0.493	0.61
GAM - S Quarantine	0.731 (0.909)	0.276	0.373	0.538

Table 3. Validation metrics on Temporal (T) and Spatial (S) test sets.

Model interpretability - partial dependence plots

Using mobility data, lockdown restrictions and other socioeconomic and demographic characteristics for each urban sector in the implemented prediction model, enables us to use interpretability techniques to further study the effects of the aforementioned characteristics in the occurrence of domestic violence. The interpretation is carried out firstly through the study of statistical significance which indicates, under a certain level of confidence, if the correlation between an explanatory variable and the occurrence of domestic violence is statistically different from zero.

Since the estimated relationship between the variables is non-linear in nature, the effect of each variable on the incidence of domestic violence depends on the value of the independent variable in which such effect is to be measured. That is, the effect of the socioeconomic stratum variable on domestic violence is different if it is calculated to go from stratum 2 to 3, than to increase from stratum 5 to 6. Therefore, we construct partial dependence plots that present the marginal effect value of each covariate on the response variable as a function of the different values that the explanatory variable to be analyzed takes, leaving the rest of the variables constant (*cetheris paribus*).

Formally, if X_S is an explanatory variable and X_C its complement (the other covariates), then the partial dependence of the function $f\ (X)\ =\ f\ (X_S, X_C)$ on X_S is given by:

$$pd_{X_S}(x_S) = E_{X_C}\left[f(x_S, X_C)\right] = \int f(x_S, x_C)p(x_C)dx_C,$$

with $f\ (x_S, x_C)$ the response function or decision function for a given sample whose values are defined by x_S for the variable X_S and x_c for the variables X_C . In the specific case, $f\ (x_S, x_C)$ is the expected domestic violence intensity function on the city and the set X are the different variables used to estimate this function. Computing this integral for various values of x_s produces the partial dependence plots.

The analysis of the partial dependence plots helps us understand the overall trends that the model is using to make its predictions, and they reflect the function parameters the models estimated in order to be as accurate as possible. However, they do not represent any causal relationships between the variables and the incidences of domestic violence, just the correlations and patterns the model exploited to make its predictions. As such, the interpretations made are with respect to the probability of the model predicting that an incident is going to occur, rather than the probability that an incident does actually occur.

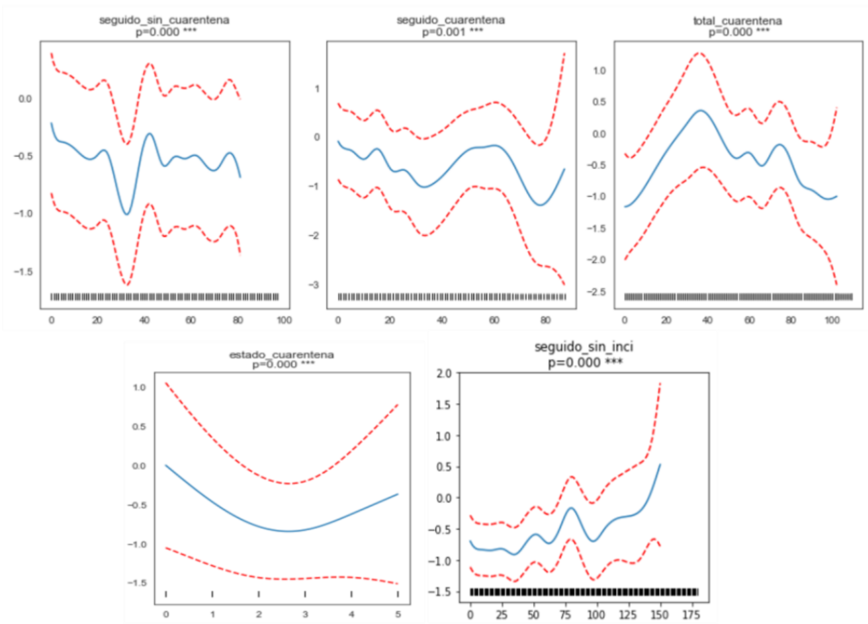

Figure 5. Partial Dependence Plots (PDPs) of a selection for ***lockdown-related variables and previous incidences***

Even though mobility change did not appear to be a significant variable when explaining the incidence of domestic violence, variables characterizing the number of consecutive days with and without lockdown restrictions (*seguido_sin_cuarentena*, *seguido_cuarentena and total_cuarentena*) are significant with 99% confidence (Figure 5). The partial dependence plot (PDP) for *seguido_sin_cuarentena* shows that the probability of domestic violence occurrence gets reduced when the consecutive days without quarantine increase from 0 to 30 days and stays almost constant for more than 30 consecutive days without quarantine. The partial dependence plots for consecutive days with restrictions and total days with restrictions (*seguido_cuarentena and total_cuarentena*) show that, in general, the probability of occurrence increases as the total days in restriction accumulate. In particular, we identify a significant increase in this effect when there are between 40 to 60 cumulative quarantine days.

The partial dependence plot for *seguido_sin_inci* (Figure 5, row 2) shows that the probability of domestic violence occurrence increases as more time since the last incident has occurred but does not increase when a short amount of time has passed since the last incident. In a sense, this result is in opposition to what is usually found in the literature on crime prediction models, where the probability of a crime is higher the shorter the time since the last crime occurred. This is usually attributed to the self-exciting nature of crime, however, in our case, we did not expect the same to hold given that we are focusing on domestic violence crimes which are expected to have a different spatio-temporal dynamic pattern than from other types of crimes such as robberies. Contrary to crimes such as theft, where criminals might feel confident in their mode of operation after being successful in a certain area and can be expected to return to it in the near future, we do not expect domestic violence to behave in the same way as an incident within one-household does not have the same influence on possible perpetrations in households nearby.

Other partial dependence plots for spatial characteristics (Figure 6) display expected behaviors, for instance the probability of a crime increases with the area of a sector, as the number of people in the sector increases with it. However, after a certain point the probability decreases, as most likely the largest areas correspond to rural sectors with a low population density. The probability also increases as the number of people increase, as well as the number of people per household. Similarly, as the percentage of properties used for housing increases, so does the probability of domestic violence, while the opposite happens as the percentage of properties that are unoccupied or have no residential use increases.

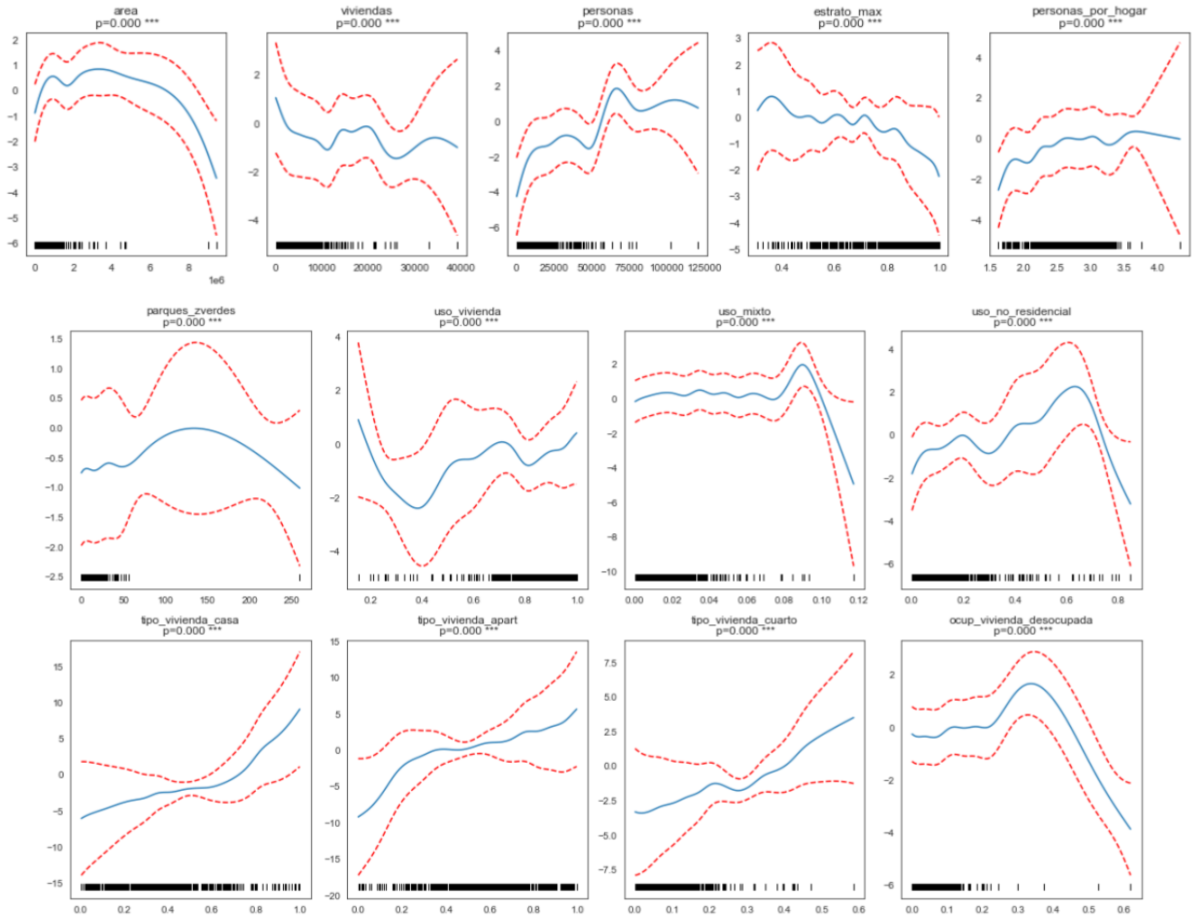

*Figure 6. Partial Dependence Plots (PDPs) for area, **number of households and number of people***

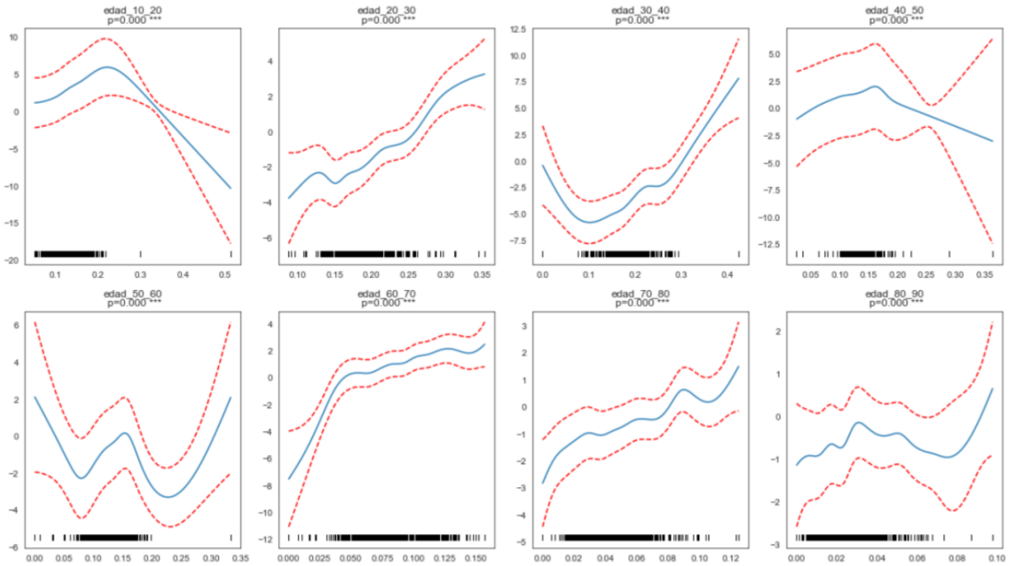

*Figure 7. Partial Dependence Plots (PDPs) for the proportion of different **age groups***

The PDPs in Figure 7. show that when the proportion of children of ages 0-10 increases, the probability of domestic violence decreases. For every other age group the opposite holds except for the age group between 40 and 50, however, it can be observed that this effect is not as clear as the confidence intervals widens at the end of the graph. Figure 8. shows that the PDPs for every single income group behaves in the same manner, as the proportion of households in a given income group increases, the probability of domestic violence decreases. We interpret this as meaning that areas where the income distribution is homogeneous have lower probability of domestic violence, while areas where households have heterogeneous income distributions have a higher probability.

The PDPs in Figure 9. show that as the proportion of people with elementary education, professional, and post-graduate education increase, so does the probability of domestic violence. The opposite only holds for the proportion of people with secondary education. This seems counterintuitive at first, however we notice that for instance, in Bogotá the number of people with post-grad education is highly concentrated on areas with very high income. In this case these results might be very similar to the results of the difference in income levels as they could indicate homogeneity or heterogeneity in income level.

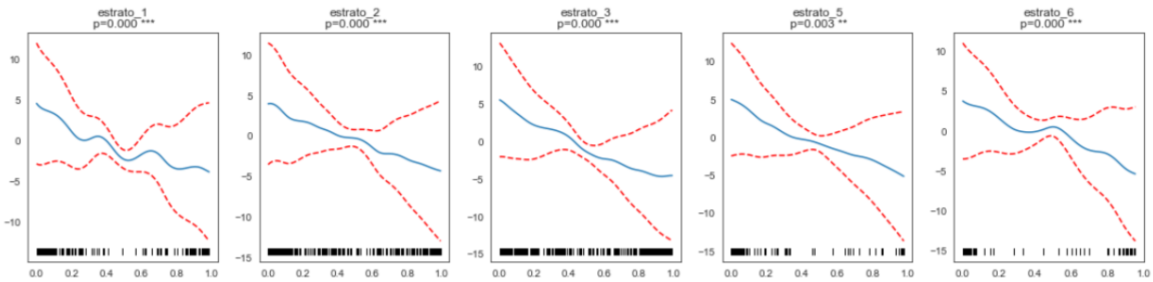

*Figure 8. Partial Dependence Plots (PDPs) for proportion of households at different **income levels***

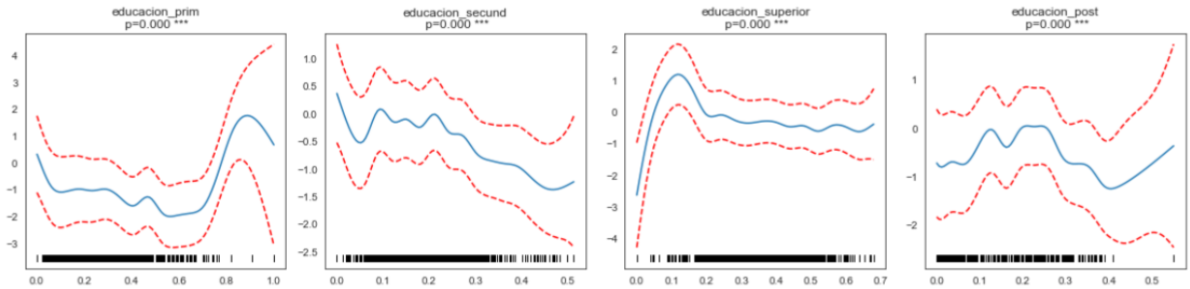

*Figure 9. Partial Dependence Plots (PDPs) for proportion of people with different **education levels***

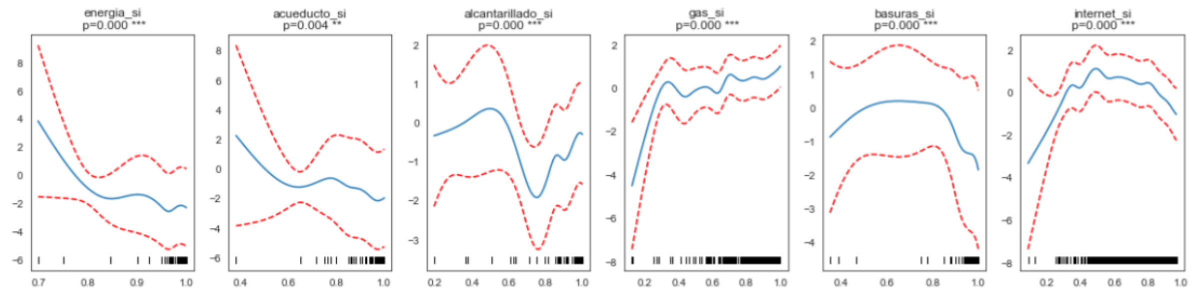

Figure 10. Partial Dependence Plots (PDPs) for proportion of households with ***access to different services***

In almost all cases, the proportion of households with access to public services such as energy, thrash collection and internet decreases the probability of domestic violence. The opposite only holds for access to natural gas, where there seems to be a strong effect of the proportion of households with access to it and an increase in the probability of domestic violence.

6.2. OLS-based Regression Models

Regression analysis

Table 4. OLS estimation of the impact of lockdown days on domestic violence cases

	Domestic violence cases	
	(1)	(2)
Lockdown days	1.014***	0.725***
	(0.202)	(0.195)
Constant	-39.859***	153.534
	(14.683)	(209.571)
Covariates	No	Yes
Observations	620	620
Adjusted R2	0.038	0.195
F Statistic	25.181*** (df = 1; 618)	8.147*** (df = 21; 598)

Notes: ***p<0.01, **p<0.05, *p<0.1. Domestic violence (dependent variable) is measured as the total number of reported cases of domestic violence in the period between March 1 to September 9 of 2020 in the sector. Lockdown days (treatment/independent variable) is measured as the total number of confinement days in the sector between March 1 to September 9 of 2020. Covariates included: First, second, fourth, fifth, and sixth class proportion of households in the sector; proportion of household access to energy, water, gas, sewerage, garbage collection, and internet; proportion of men in the sector; number of persons by household, proportion of people between the ages 0-10, 20-30, 30-40, 40-50, 50-60, 60-70, and 70+.

Table 4 shows the estimated conditional and unconditional impact of lockdown duration on the reports of domestic violence[4]. Both types of estimations produce statistically significant coefficients that suggest that increases in lockdown duration cause an increase in about 0.7-1 cases of domestic violence. Specifically, column (1) shows that an increase in one lockdown day causes, on average, an increase in the reported domestic violence cases by one. As was discussed before, however, this estimation might be overestimating the effect of lockdown duration. This since the variable can be related to other socioeconomic confounding factors that also affect domestic violence. Once it is controlled for socioeconomic characteristics the estimated impact is slightly lower. This indicates that the coefficient of column (1) is somewhat upward biased since it captures more than the impact of the lockdown duration alone. Still, the coefficient of column (2) remains strongly significant once the bias is plausibly corrected. It indicates that a *ceteris paribus*[5] increase in one lockdown day increases the reported number of domestic violence cases by 0.7 (equivalent to 0.01 SD of the dependent variable), on average.

Mediation analysis

Table 5. Mediation analysis of mobility change as a channel of the impact of lockdown days on domestic violence cases

	Domestic violence cases		Mobility change
	(1)	(2)	(3)
Lockdown days	0.725***	0.715***	0.002*
	(0.195)	(0.194)	(0.001)
Mobility change		37.611***	
		(13.749)	
Constant	153.534	87.957	-0.478***
	(209.571)	(209.818)	(0.073)
Covariates	Yes	Yes	No
Observations	620	620	620
Adjusted R2	0.195	0.204	0.003
F Statistic	8.147*** (df = 21; 598)	8.201*** (df = 22; 597)	3.096* (df = 1; 618)

Notes: ***p<0.01, **p<0.05, *p<0.1. Domestic violence (dependent variable) is measured as the total number of reported cases of domestic violence in the period between March 1 to September 9 of 2020 in the sector. Lockdown days (treatment/independent variable) is measured as the total number of confinement days in the sector between March 1 to September 9 of 2020. Mobility change (dependent/independent variable) is measured as the average of daily mobility change with respect to the corresponding weekday of the first week of March 2020. Covariates included: First, second, fourth, fifth, and sixth class proportion of households in the sector; proportion of household access to energy, water, gas, sewerage, garbage collection, and internet; proportion of men in the sector; number of persons by household, proportion of people between the ages 0-10, 20-30, 30-40, 40-50, 50-60, 60-70, and 70+.

[4] We emphasize our results on the *report* of domestic violence since we can't separate the impact on occurrence from the impact on the reports. The latter might be affected by other factors such as changes in daily behavior due to quarantine that increase the reports, but not necessarily the occurrence. We estimate the effect on reports that might be mixing those two effects.
[5] All else equal/constant.

Table 5 displays the results of the mediation analysis conducted to evaluate the mobility change as a mediator. Column (1) displays the results of the main regression without controlling for mobility, so the interpretation is the same as the one presented in Table 1. Column (2) shows the result of the same regression once it is controlled for mobility change. Column (3) presents the regression of mobility change against lockdown duration. To interpret mobility change as a mediator it must be that 1. the unconditional impact of lockdown duration on reported domestic violence is different from zero, 2. the relationship between mobility change and lockdown duration is different from zero, 3. the mediator has an impact on reported domestic violence once it is controlled for lockdown duration, and 4. the conditional impact of lockdown duration once it is controlled for the mediator is lower in absolute value to the unconditional one.

As can be seen in Table 5, the four conditions hold, which means that mobility change is effectively a channel or mediator through which lockdown duration impacts the reported domestic violence. However, the decrease in the estimated impact of lockdown once it is controlled for the mediator is just 0.01. This suggests that even if the mobility change is a possible mediator, it is not the main mediator and that the impact of lockdown duration might be driven by other factors such as economic and psychological stress. This result is consistent with García-Hernández, López, Cabra, Otálora & Arias (2021) findings.

Heterogeneous effects

Table 6. Heterogeneous effects analysis of the impact of lockdown days on domestic violence cases

	Domestic violence cases			
	(1)	(2)	(3)	(4)
Lockdown days	0.725***	-0.684	-2.894	-0.793
	(0.195)	(5.305)	(1.805)	(0.879)
Lockdown days*Proportion of men		2.934		
		(11.042)		
Lockdown days*Number of persons by household			1.283**	
			(0.636)	
Lockdown days*Predominant social class = 1				2.209**
				(0.932)
Lockdown days*Predominant social class = 2				1.371
				(0.928)
Lockdown days*Predominant social class = 3				0.751
				(1.124)
Lockdown days*Predominant social class = 4				1.014
				(1.531)
Lockdown days*Predominant social class = 5				0.425
				(1.548)
Constant	153.534	252.497	439.745*	275.692
	(209.571)	(427.449)	(252.660)	(218.307)
Covariates	Yes	Yes	Yes	Yes
Mean of moderator variable		0.479	2.757	

Observations	620	620	620	620
Adjusted R2	0.195	0.194	0.199	0.197
F Statistic	8.147*** (df = 21; 598)	7.768*** (df = 22; 597)	8.002*** (df = 22; 597)	6.840*** (df = 26; 593)

Notes: ***p<0.01, **p<0.05, *p<0.1. Domestic violence (dependent variable) is measured as the total number of reported cases of domestic violence in the period between March 1 to September 9 of 2020 in the sector. Lockdown days (treatment/independent variable) is measured as the total number of confinement days in the sector between March 1 to September 9 of 2020. Covariates included: First, second, fourth, fifth, and sixth social class proportion of households in the sector for columns (1)-(3) and predominant social class dummies in the sector for column (4); proportion of household access to energy, water, gas, sewerage, garbage collection, and internet; proportion of men in the sector; number of persons by household, proportion of people between the ages 0-10, 20-30, 30-40, 40-50, 50-60, 60-70, and 70+. All the variables used in the interaction terms were included individually. Reference categories: (2) sectors with zero proportion of men, (3) sectors with zero persons by household on average, (4) sectors which predominant social class is the highest (6th).

Table 6 displays the analysis of heterogeneous effects conducted for the main model. Column (1) displays the overall average effect of lockdown duration on reported domestic violence (i.e., the same displayed in table 1). The first row of column (2) presents the effect of lockdown on reported domestic violence for a sector that has no presence of men and the second row presents the gradient of the estimated effect as the men proportion increases in a sector. The first row suggests that, in a hypothetical sector where there is no presence of men, the effect of lockdown on domestic violence reports is null or even negative. However, as the proportion of men increases, the impact of lockdown increases. This might be related to the fact that the perpetrators of this kind of crime are generally men, so places with fewer men are less likely to suffer from domestic violence due to the confinement. Nonetheless, caution is needed to interpret this result since it was done only considering the coefficient and not the statistical significance. It must be noted that the interaction term (second row) is accompanied by a statistically non-significant coefficient.

On the other hand, the first row of column (3) shows the effect of lockdown on reported domestic violence for a hypothetical sector where the average number of persons by household is zero, which is statistically non-significant different from zero. The third row of column (3) presents the gradient of the effect as the proportion of people increases. As is expected, the more people are in one household, the greater the effect of lockdown duration on domestic violence reports. This result might be due to the greater probability of contact and conflict between two or more individuals that live in the same place when there is not much non-shared space.

Finally, the fourth column displays the differential effects of lockdown on reported domestic violence across sectors that have different predominant social classes. In particular, the first row displays the effect for the highest class, which is statistically non-significant. The fourth row (i.e., the second coefficient displayed in that column) shows how the coefficient changes for sectors that have a predominant lowest social class. Concretely, it indicates that while for the highest social class the impact of lockdown is null, for the lowest/poorest social class is positive and highly significant. The rest of the rows show how the coefficient changes as a sector moves from a predominant highest social class to a predominant second, third, fourth, and fifth social class, respectively. In general, it appears that the poorest the predominant the social class, the greater the effect of lockdown on reported domestic violence.

These three heterogeneous effects analyses shed light on which sectors of the city might be the more affected in terms of domestic violence by the lockdown. Particularly, the poorest and the more population-dense sectors are the ones that drive the overall positive relationship between the mentioned variables. This implies two things. The first is that the economic factors might be the main drivers of increases in domestic violence.[6] It has to be taken into account that the poorest sectors are the ones who face greater economic stress due to the pandemic and the lockdown. The second is that those places are the ones that must be targeted by policymakers, maybe by economic alleviation not only to reduce the overall effect of the pandemic, but also to reduce the domestic violence cases.

7. Discussion and conclusions

We use data science techniques along with interpretability techniques to understand the effect of compliance with quarantines and lockdowns on the spatio-temporal distribution of domestic violence occurrences. The predictive model, Spatio-Temporal GAM, achieves a good performance for predicting domestic violence crime events. We then study the behavior of the explanatory variables in the model through their partial dependence plots finding that the existence of lockdown restrictions and their cumulative effects are significant when explaining the incidence of domestic violence crimes in Bogotá. Furthermore, temporal variables on mobility changes end up not being significant as lockdown restrictions along with socioeconomic variables account for lockdown compliance patterns that, in terms of the predicted model implemented, are good predictors for the spatio-temporal occurrence of domestic violence. Additionally, we estimate different OLS-based-regression models to better understand the causal relationship between the quarantine lockdown duration and domestic violence and to further explain the non-significance obtained from the mobility changes in the Spatio-Temporal GAM approach.

Both the GAM and the OLS models found a positive relation between quarantine lockdown and the reported incidence of domestic violence. Specifically, it is found that longer confinements are causally related to greater reported incidence. However, both concluded that the relationship is not mainly driven by changes in mobility itself. This implies that reported domestic violence increases were due to other factors that were affected by the confinement such as the psychological and economic ones. Heterogeneous effects OLS model showed that the effects were driven by the lowest two social classes. This poses economic stress as one of the main drivers. On the other hand, given the lack of psychological factors information, no such channels could be evaluated. Recent work on domestic violence during COVID-19 done by García-Hernández et al., 2021 states, throughout the analysis of the longitudinal survey RECOVR, that households experiencing financial difficulties are more vulnerable to suffering violence during the COVID-19 pandemic. This aforementioned work was published after the implementation of our analysis and shows encouraging results given the common conclusion towards having economic stress as one of the main drivers for domestic violence while using different approaches from surveys and in our case, mobility and lockdown data.

Lockdowns are meant to reduce the number of infected patients, deaths, and relieve the health system. Through the findings of this paper, we recommend that, during the pandemic, the length of the lockdowns is reduced, or re-designed to be intermittent and rotative across regions, to decrease the exposure of households to psychological and economical stress. Local governments and police departments should increase the channels for reporting domestic violence, and eventually the design of these channels to ease the reporting while in a context

[6] See García-Hernández et al., 2021.

of stay-at-home orders. The threat itself of reporting may reduce the amount of violence a victim is exposed to, via either deterring the perpetrator or simply giving the victim the chance to seek aid earlier. As mentioned in Bullinger et al (2020), the rates of underreporting likely rose as a consequence of a lack of contact between victims and third parties, for example, children not being in school (rarely children are the reporter of domestic violence), or the closing of shelters not enabling face-to-face interactions between domestic violence advocates and responders with the victims. We also believe that a stay-at-home context makes the reporting of violence harder for the victims as they are enclosed all the time with their abusers, making it difficult to access the proper reporting channels. The challenge of designing these report channels becomes harder in Colombia as most vulnerable households, economically speaking, tend to have lower access to all sorts of services including mobile internet. Further research should focus on studying the underreport, its underlying causes and deterrents, and estimation to appropriately address and create policies towards reducing domestic violence.

References

Adams, P. R., & Adams, G. R. (1984). "Mount Saint Helens's ashfall: Evidence for a disaster stress reaction". *American Psychologist,* 39(3), 252.

Boserup, B., McKenney, M., & Elkbuli, A. (2020). "Alarming trends in US domestic violence during the COVID-19 pandemic". *The American Journal of Emergency Medicine*.

Brooks, S. K., Webster, R. K., Smith, L. E., Woodland, L., Wessely, S., Greenberg, N., & Rubin, G. J. (2020). "The psychological impact of quarantine and how to reduce it: rapid review of the evidence". *The Lancet.*

Bullinger, L. R., Carr, J. B., & Packham, A. (2020). COVID-19 and crime: Effects of stay-at-home orders on domestic violence (No. w27667). *National Bureau of Economic Research.*

Doshi-Velez, F., & Kim, B. (2017). "Towards a rigorous science of interpretable machine learning". *arXiv preprint arXiv:1702.08608*.

Dueñas, M., Campi, M., and Olmos, L. (2020). "Changes in mobility and socioeconomic conditions in Bogotá city during the COVID-19 outbreak", *arXiv preprint arXiv: 2008.1185*

Garcia-Hernandez, A., López Uribe, M., Cabra García, M., Otalora, M.J., and Arias, A.M.. (2021) "Violencia en el hogar durante COVID-19". *Nota de Política Pública, Innovations for Poverty Action (IPA).*

Hohman, F., Head, A., Caruana, R., Deline, R., & Drucker, S. M. (2019). "Gamut". *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19*. doi:10.1145/3290605.3300809

Immigration and Refugee Board of Canada (2017), "Colombia: Domestic violence, including legislation; state protection and support services available to victims in Bogotá, Santiago de Cali, and Medellín (2014-April 2017)", Canada. COL105771.E, available at: *https://www.refworld.org/docid/591615304.html* [accessed 13 November 2020]

Korkoyah Jr, D. T., & Wreh, F. F. (2015). "Ebola impact revealed: An assessment of the differing impact of the outbreak on the women and men in Liberia".

Leslie, E. & Wilson, R. (2020). "Sheltering in Place and Domestic Violence: Evidence from Calls for Service during COVID-19". Available at SSRN 3600646.

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.

Norlander, B., & Eckhardt, C. (2005). "Anger, hostility, and male perpetrators of intimate partner violence: A meta-analytic review". *Clinical psychology review*, 25(2), 11

Onyango, M. A., Resnick, K., Davis, A., & Shah, R. R. (2019). "Gender-Based Violence Among Adolescent Girls and Young Women: A Neglected Consequence of the West African Ebola Outbreak". *In Pregnant in the Time of Ebola* (pp. 121-132). Springer, Cham.

Parkinson, D. (2019). "Investigating the increase in domestic violence post disaster: an Australian case study". *Journal of interpersonal violence*, 34(11), 2333-2362.

Perez-Vincent, S. & Carreras, E. (2020). "Evidence from a Domestic Violence Hotline in Argentina". Washington, DC: Inter-American Development Bank.

Portland Police Bureau (2020). "Trends analysis: pre- and post-school closures".

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016, August). " Why should I trust you?" Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 1135-1144).

Sanga, S. & McCrary, J. (2020). "The Impact of the Coronavirus Lockdown on Domestic Violence". Available at SSRN 3612491.

Schumacher, J. A., Coffey, S. F., Norris, F. H., Tracy, M., Clements, K., & Galea, S. (2010). "Intimate partner violence and Hurricane Katrina: predictors and associated mental health outcomes". *Violence and victims*, 25(5), 588-603.

Shorey, R. C., Brasfield, H., Febres, J., & Stuart, G. L. (2011). "The association between impulsivity, trait anger, and the perpetration of intimate partner and general violence among women arrested for domestic violence". *Journal of Interpersonal Violence*, 26(13), 2681-2697.

Slack, D., Hilgard, S., Jia, E., Singh, S., & Lakkaraju, H. (2020, February). "Fooling lime and shap: Adversarial attacks on post hoc explanation methods". In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* (pp. 180-186).

Stark, L., & Ager, A. (2011). "A systematic review of prevalence studies of gender-based violence in complex emergencies". *Trauma, Violence, & Abuse*, 12(3), 127-134.

Stuart, G. L., Meehan, J. C., Moore, T. M., Morean, M., Hellmuth, J., & Follansbee, K. W. (2006). "Examining a conceptual framework of intimate partner violence in men and women arrested for domestic violence". *Journal of Studies on Alcohol*, 67, 102-112.

Stuart, G. L., Temple, J. R., Follansbee, K. W., Bucossi, M. M., Hellmuth, J. C., & Moore, T. M. (2008). "The role of drug use in a conceptual model of intimate partner violence in men and women arrested for domestic violence". *Psychology of Addictive Behaviors*, 22, 12-24.

Wang, X. and Brown, D. (2011). "The Spatio-Temporal generalized Additive Model for Criminal Incidents". *Proceedings of 2011 IEEE International Conference on Intelligence and Security Informatics, (pp. 42-47). IEEE.*

Weitzman, A., & Behrman, J. A. (2016). "Disaster, disruption to family life, and intimate partner violence: The case of the 2010 earthquake in Haiti". *Sociological science*, 3, 167-189.

Zahran, S., Shelley, T. O. C., Peek, L., & Brody, S. D. (2009). "Natural disasters and social order: Modeling crime outcomes in Florida". *International Journal of Mass Emergencies and Disasters*, 27(1), 26-52.