University of Sydney
Masters in Biostatistics

Workplace Project Portfolio (WPP)

Ecological study of factors associated with homicide rate variability in El Salvador, 2016.

Victor Rodolfo Patino

December 2020

Supervisor: Professor Armando Teixeira-Pinto

Contents

Glossary

Standardised mortality rate - (Direct standardisation) rates are adjusted to reflect the overall mortality rate that would have been expected in a 'standard' population if it had the same age-specific rates as the study population; El Salvador's national population was used as "standard"

Palma inequality index[1] - the ratio of the top 10% of population's share of gross national income (GNI), divided by the poorest 40% of the population's share of GNI

Part A.

Preface

Introduction

The project presented in this portfolio was undertaken at the convenient intersection of my long-held interest in the dynamics of homicidal violence in Central America and the requirements of the Masters in Biostatistics program.

Latin America has consistently been reported as "the world's most dangerous region"[2], a claim highlighted by the region's epidemic-level rates of homicide. Particularly striking is the situation in Central America's so-called Northern triangle (Guatemala, Honduras, El Salvador), with deaths by homicide on par with situations in war-afflicted areas .

The World Health Organisation[3] classifies a homicide rate above 10 per 100,000 as "epidemic". In El Salvador, homicide rate peaked in 2015 at just over 100. While a general downward trend has been observed in the region, the situation is by no means flattering, with yearly rates above 50 still commonplace. The contrast is sharper when one considers that Nicaragua, which neighbours Central America's Northern triangle and has a considerably lower GDP per capita than the region, has a homicide rate orders of magnitude lower than its Northern counterparts.

This biostatistics project examines local-area (municipality-level) homicide data from El Salvador for 2016 to assess whether differences in homicide rates show any association with a host of socio-demographic factors including income, inequality, and household family structure, among others.

My Role

I completed this project under the supervision of Professor Armando Teixeira-Pinto as an individual, unaffiliated effort to contribute towards pinpointing factors involved in the dynamics of an 'endemic' epidemic such as homicidal violence in Central America.

As such, my role in the project involved data preparation and management, statistical analysis, and presentation and interpretation of results. Professor Teixeira-Pinto provided guidance and advice throughout the course of the project.

My hope is that results from this project will form part of a broader, longitudinal analysis and report capable of informing academics and policy-makers about the nature of homicidal violence in El Salvador as well as possible areas of intervention.

Teamwork

Most of my work was conducted independently. I liaised with Professor Armando Teixeira-Pinto for specific advice on statistical issues encountered.

Working within timelines

Initially the project was due for completion by end of November 2020; however, flexibility was needed due to (I suspect) higher-than-usual workloads for statisticians during the pandemic resulting in lack of availability of statistical supervisors. Timelines had to be adjusted accordingly as delays became inevitable.

Delays in supervisory arrangements resulted in reduced time-frames for report analysis and writing.

Reflections on learning

Communication skills

With little exposure to statistical discussions/analysis outside of the course, sharpening my communication skills, both verbal and written, became crucial to the progress of the project. In particular, it became essential to recognise and effectively summarise issues and advances in statistical analysis to Professor Teixeira-Pinto. The early recognition of concerns with data management and modelling had to be communicated promptly and efficiently.

Being part of one-man team did not exclude me from other communication-related issues. For instance, the importance of adequate and standardised naming conventions for data files was discovered shortly after starting the project. For instance, given that the project involved dealing with multiple datasets from various institutions, the process of selecting variables and merging them into a single, workable file, was gradual and painstaking. An initial working file named *unified.csv*, would quickly spiral into *unified_7_maps.csv*. It became apparent eventually that naming should be both informative (numbers seldom suffice) and succinct, eventually settling into a convention of using a combination of the names of the datasets which had been merged, as well as any important data cleaning which they had undergone. Of equal importance was adhering to sufficiently consistent and descriptive variable naming and labelling.

Work patterns/planning

The challenging nature of the (mostly) private project was met with the combined difficulty of the lengthy Melbourne lockdown of 2020, which resulted in strains in the availability of adequate space and time for working in a crowded young family home. In this scenario, constant work patterns became both essential as well as hard to come by.

Given my nature as a statistical novice, data management demanded a greater proportion of time/work than anticipated and required setting aside time to gain much-needed dexterity in the statistical package (R) than was provided by the introductory course. This came as both a challenge and a proverbial blessing in disguise because it became a gateway to an abundance of open resources readily available online (*stackexchange*, *stackoverflow, statmethods.net*, as well as numerous worldwide university pages).

This effectively put into perspective the introductory nature of much of what was learnt as well as underscoring the importance of self-directed learning in further developing a working statistical skillset.

Statistical principles, methods, computing

The project involved the analysis of a diverse set of publicly available/routinely collected datasets, including homicide data for El Salvador for 2016, a yearly socio-economic survey for most of the country's 262 municipalities (county-level equivalent), as well as demographic projections obtain from the country's ministries.

Given the scope of the available data, an ecological study was considered the best approach. A brief review of quantitative studies into crime and violence suggested the inclusion of potential factors that influence the outcomes. Summary measures of these were obtained for our unit of analysis.
R statistical software package was used for data management and analysis, as well as for the production of graphs.

Ethical consideration

All the datasets used are publicly available and did not contain any personal identifiable information. NHRMC guidelines suggest that population-wide datasets are particularly inclusive and intrinsically uphold core principles of justice, as well as providing a more even spread of benefits/burdens than when studies are based on selected participants.

Professional responsibility

Throughout the project I was keenly aware of my responsibility to maintain a diligent and careful approach towards the study, as well as the need to enhance my knowledge and skills to be able to fulfil the project to the highest standard possible.

Workplace Project Portfolio.

Ecological study of factors associated with homicide rate variability in El Salvador, 2016.

Location and dates:

Homicide victims in El Salvador throughout 2016.

Context:

The preface provides a brief summary of the general context for this project; it is included again here to provide an adequate overview of the study. Latin America in general, and Central America in particular, are renowned for their epidemic levels of homicide rates. Inter-national as well as intra-national variability in violent crimes may provide questions and answers aimed at pinpointing key elements involved in these dynamics. It was hoped that an ecological study of municipality-level homicide data could provide insights to contributing factors, as well as guidance and evidence to inform possible interventions. A group of often-reported socio-economic factors were included as potential factors associated with the outcome.

Contribution of student.

- Data management (coding and re-coding of variables, documentation, reproducibility, statistical quality assurance, data cleaning)
- Gaining proficiency in the use of statistical package R
- proposal of a parsimonious model for homicide rate variability using appropriate modelling
- summarising and presentation of results
- report writing

Statistical issues involved:

- standardisation of homicide rates for diverse population clusters
- regression models for count and/or continuous data
- model diagnostics - overdispersion, heteroscedasticity, AIC, Pearson's Chi-squared
- probability theory – expectation, distribution and transformation of count/rate data, unequal cluster sizes,

Declaration:

I declare this project is evidence of my own work, with direction and assistance provided by my project supervisor. This work has not been previously submitted for academic credit.

Victor Patino

Declaration

I acted as the statistical supervisor for Victor Patino, on the project that he has completed to satisfy the requirement of the Workplace Project Portfolio for his Master of Biostatistics.

Victor approached me with a research project that aimed to investigate the differences in criminality rates among counties of El Salvador.

Victor worked mostly independently and we had a few meetings to discuss the progress and direction of the work. The student revealed maturity in his knowledge of statistical methods and a clear understanding of their application. I am pleased with the work submitted and, with some editing, it has a clear potential to be published in a specialty journal.

Armando Teixeira-Pinto

armando.teixeira-pinto@sydney.edu.au

Mar 22th, 2021

School of Public Health, Faculty of Medicine

The University of Sydney

Introduction

Latin America makes up just 8% of the world's population but accounted for more than 30% of the global homicides in 2016; in other words, the homicide rate in the region quadruples the global average.[4] Levels of homicide are, of course, non-uniform, not only between countries, but also *within* countries, and even *within* regions and cities.

Once so-called violence hot-spots are identified, they tend to become the target of government interventions, with varying degrees of success. The relative success (or lack of) hinges, to a greater or lesser degree, on an accurate understanding of the causes and factors associated with crime. However, a great deal of the insights into the dynamics of violence in the region are obtained from descriptive or anecdotal accounts [4].

Violence from Central America has, in particular, garnered widespread media attention, especially with the secondary consequence of mass immigration towards the relative safety of northern neighbours such as the USA. This attention, however, has not translated into widespread quantitative research into the phenomena.

Political scientists such as Trujillo Alvarez (2017) have turned their attention towards political corruption and weak states (insufficient and/or inefficient institutions) to explain the variability in rates within the Central American region, while (qualitatively) being unable to observe any trend in relation to inequality and other socio-economic factors[5].

The aim of this project was to investigate whether socio-economic factors are able to explain any of the variability in homicide rates for 2016 in one of the world's most violent countries, El Salvador. At a glance, macro/ecological level public health/criminology literature offers conflicting reports on salient factors, from those highlighting the primary importance of poverty/income [6][7] to those suggesting the preeminence of relative deprivation (inequality) [8].

Methods

El Salvador is the smallest republic in the Central American region, with a territory of just over 20,000 km^2, and a total population of roughly 6 million. It is subdivided into 14 administrative regions called *departamentos,* which are in turn subdivided into 262 municipalities; these municipalities are the unit of analysis of our project.

Data Management

Country-wide daily homicide count data for 2016 was used for this study. The data was publicly available through the country's Ministry of Justice and Security website [9]. It contains information on homicide victims, including age, sex, municipality, weapon used, and date [of the homicide]. No personal identifiable information is included. The database was obtained in Excel spreadsheet format.

Socio-economic information was obtained through the country's Ministry of Economy micro-data hub [10]; it contains annual de-identified surveys comprising a representative sample of ~70,000 people from most (n = 230) municipalities. Of note, there are 32 municipalities for which no data is available.

A *csv* file was obtained and loaded directly into R. Mean values of monthly income and proportion of people living in households with parent-less children were calculated from these. Inequality (Palma index) was calculated by recoding each municipality's income into top decile and bottom 4 deciles; the ratio of the sum of all incomes within the two groups constitutes the index.

Regional population age-structure data was sourced from the Ministry of Health's webiste [11]. It was only available in portable document format (PDF), which had to be extracted with the use of software package Tabula[12], and manually edited to correct issues of format.

A single outlier was removed from the dataset, corresponding to the municipality of San Francisco Javier, with a standardised homicide rate of 417 (per year per 100,000), from a population of municipalities (n=262) with median 64 (mean 77), and standard deviation = 64 (greater than 5 standard deviations).

Particularly cumbersome was the need to verify municipality names in all three datasets due to inconsistencies in the naming conventions used by the different ministries, as well as a number of municipalities (correctly) sharing the same name, although in different regions.

R Studio (version 4.0.3, 2020-10-10) was used for regression analysis, as well as the production of graphs and maps.

Statistical Methods

Data on El Salvador's country-wide homicides for 2016 (n=5276) were used to investigate factors associated with variability in homicide rates.

Aggregated local-level (municipality) data was used to carry out an ecological study using generalised linear model with negative binomial errors and log link function for homicide counts. Municipality homicide counts were age-adjusted, and total population was used as exposure (offset).

Geographical distributions of raw, age-standardised homicide counts and model residuals are depicted.

Variables are described using median, mean, and inter-quartile range. The outcome variable of interest is homicides (as count in the final model, easily transformed and also initially examined as a continuous rate). Non-uniform age and sex distribution of homicide deaths, as well as the disparities in age-structure across the country, warranted standardisation of the mortality. Homicide counts/rates were age standardised to adjust for this variability.

For the purpose of association, average monthly income, Palma inequality index, and proportion of people living in households with parent-less children were investigated. Two-sided 5% significance level was used for main effects. In this particular case, the inequality index was estimated from the survey data on monthly income.

A normal model was initially considered by relying on homicide rate as a continuous variable (i.e. number of homicides per 100,000 population). Beyond its simplicity and ease of interpretation of results, the model had the appeal of being able to include the population of the municipality as an additional variable, with its own coefficient of effect. Normally distributed errors proved incompatible with the data, mainly due to the constraints of constant variance of the errors.

Hence, Poisson, quasi-Poisson, and negative binomial regression models were used to explore associations, relying on the (count) outcome variable of homicide numbers.

We initially model homicides, Y, as:

$$\mu = t \times \exp^{\beta_0+\beta_1 x_1+\beta_2 x_2+\beta_3 x_3 \text{¿}}$$

where E(Y) = μ, with Y ~ Poisson(μ), t = population(exposure), β_i= coefficient effect of variable *i*, x_i=model variables (inequality, monthly income, proportion of people living in households with parent-less children).

Equivalently for rate data, this can be expressed:

$$\frac{\mu}{t} = \exp^{\beta_0+\beta_1 x_1+\beta_2 x_2+\beta_3 x_3 \text{¿}}$$

In particular, with the use the log-linear link function,

$$\log(\mu) - \log(t) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$$

Of note, the use of the exposure term *log(t)* in regression models in R, effectively results in a fixed coefficient effect of 1 for the exposure variable (in this case, municipality population size). Thus, while we initially considered including population size as an explanatory variable, our final choice of model prevented us from doing so.

The negative binomial distribution is a generalisation of the Poisson distribution through the introduction of a gamma noise variable which has a mean of 1, and a scale parameter, v. Thus, a negative binomial variable can be defined by two parameters:

Y ~ NegBin(r,p),

with E[Y] = pr/(1-p),

V[Y] =pr/(1-p) 2

The multi-parameter distribution allows for greater dispersion of the variance, as opposed to Poisson, where Y ~ Poisson(ƛ), with E[Y] = ƛ, V[Y] = ƛ.

Goodness of fit tests were conducted to assess the adequacy of the models. Dispersion parameters, variance versus fitted mean were plotted and assessed for Poisson and quasi-Poisson models, as well as AIC values for nested models with maximum likelihood estimators (ie. excluding quasi-Poisson).

Salient issues in the analysis were heteroscedasticity as well as unequal cluster size. Simple linear regression using (continuous) standardised homicide rates proved inadequate due to the heterogeneous variance.

Results

A total of 5276 homicides were reported in 2016. Victims were overwhelmingly male (90%) and below 40 years of age. Standardised homicide rates throughout the 262 municipalities range from 0 to 417(outlier), with a median of approximately 8 and a mean of roughly 20 (x100,000). The population size of municipalities was highly skewed, with a median of about 11,000 (mean= ~25,000). A handful of municipalities dominate in terms of sheer demography, with municipalities with more than 100,000 (n=13) accounting for over 35% of the total population. The non-uniform distribution of homicide deaths, as well as the disparities in age-structure across the country, warranted standardisation of the mortality.

Table 1 presents a summary of variables. Socioeconomic data was available for n=230 municipalities, with missing information predominantly from smaller municipalities (median population of municipalities with missing data = 4262; interquartile range = 2,396 – 5,396). The median monthly income was $115.00, with interquartile range $97-$121. Median Palma index for each municipality was 1.7, with an interquartile range of 1.4 -2.0, with a minimum of 0.83, and a maximum of 5.7. Figure 1 provides a graphical summary of variables, as well as Pearson correlation values.

The adequacy of a standard Poisson model for the data was assessed graphically using Figure 2, as well as quantitatively by estimating the dispersion parameter. Both of these tools revealed deficiencies in the model, particularly with a dispersion parameter of ~8. The assumption of variance equal to the mean (denoted by the straight line) was not met, a requirement for Poisson errors. Table 2 summarises the information used to select a negative binomial model as the most adequate for the data at hand.

Figures 3A and 3B show the geographic distribution of raw and standardised homicide rates, respectively. After adjustment for age and gender, standardised homicide rates generally followed the same patterns seen in the crude homicide rates, with noteworthy changes only in a handful of municipalities. In terms of geographic variations in homicide rates, there would appear to be a North-South gradient, with lower rates tending to occur in the Northern regions of the country.

Table 3 contains coefficient estimates following negative binomial regression analysis. Inequality (as a continuous variable through the Palma index) was highly significant in the model. Neither income nor proportion of people living in a household with parent-less children were significant. Interactions between the variables were assessed and not found to be significant.

According to the multivariate model, each unit increase in the Palma inequality index is associated with a rate ratio of 0.73 (95% CI: 0.62 - 0.87); equivalently, a one unit increase of inequality index is associated with a homicide rate drop of $e^{e\wedge-0.307}$, or ~25%. This association is depicted in the univariate scatterplot in Figure 4.

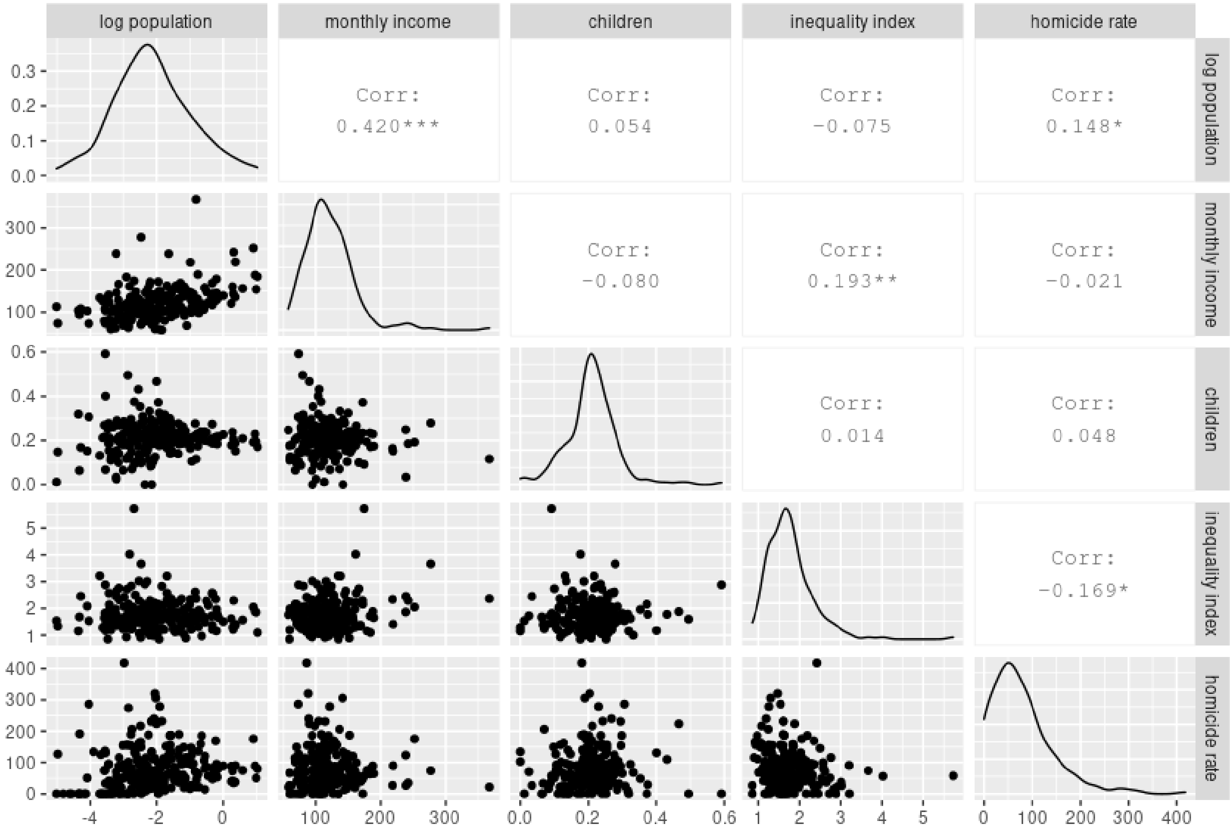

Figure 1. Scatterplot matrix, distribution curve, and Pearson correlation of model variables, including log-transformed population.

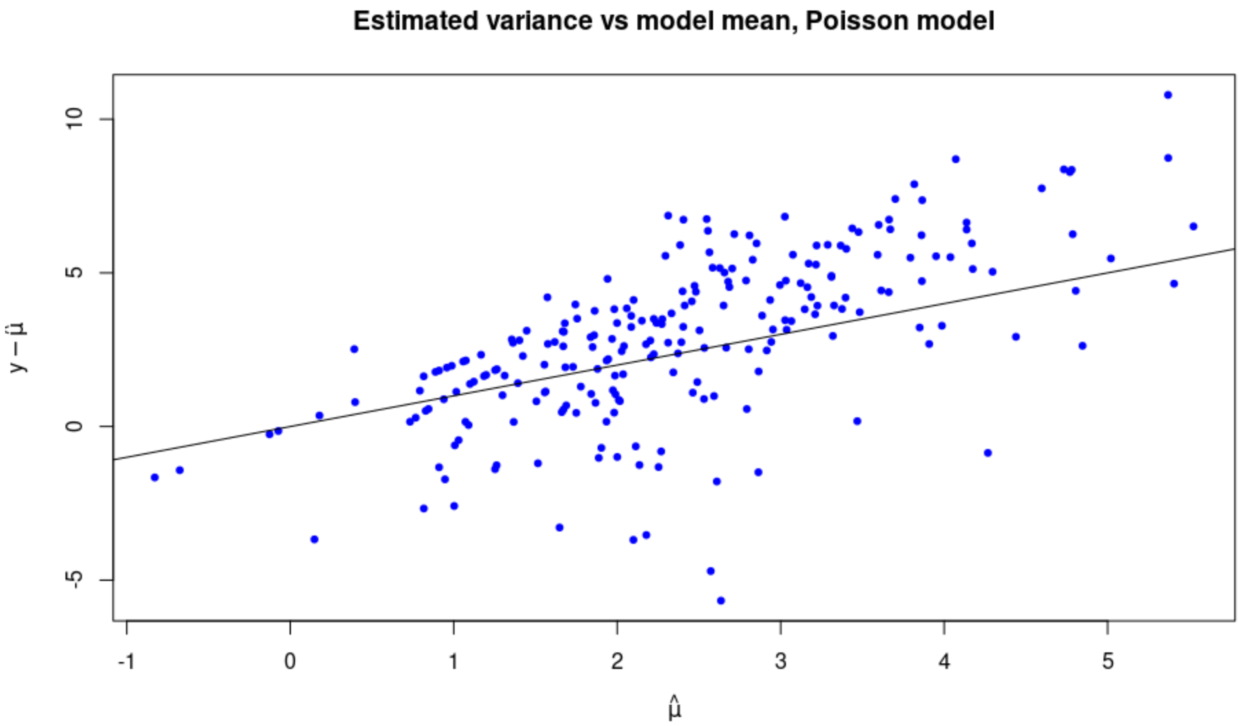

Figure 2. Estimated variance vs fitted mean for a multivariate Poisson model.

Table 1. Summary of variables

	N(%)	*Median(Mean)*	*Interquartile Range (min ; max)*
Standardised Rate	262	20(8)	2.74 ; 23.17 (0 ; 417)
Population	262	11,153 (24,888)	5,719 ; 23,952 (663 ; 281,996)
Income per capita (US$)	230	115.00 (121.08)	96.40 ; 139.07 (57.6 ; 367.6)
Children in parent-less households	230	0.2102 (0.2098)	0.1772 ; 0.2482 (0.0 ; 0.5918)
Palma index	230	1.6732 (1.7665)	1.39 ; 1.76 (0.85 ; 5.73)

Unexplained variation in SMR

Figure 4 presents the geographic distribution of unexplained variation in standardised homicide rates for a negative binomial model with no covariates (A) and our final model (B). Municipalities in blue/light blue colours indicate regions where the homicide rate is lower than predicted by the model; light blue/cyan/light yellow/ regions indicate areas where the model fits closest, whereas yellow/red/dark red colours represent regions where the homicide rates are considerably higher than the model would suggest. It appears the North-South pattern is also present in the variability unaccounted for by the model. Figure 5 displays the distribution (density) curve of the unexplained variation in standardised homicde rates (residuals) for both final model and one without covariates.

The figures show noticeable shifts in residual patterns, both in terms of geography as well as size and frequency. In particular, the absolute value of both the mean and median residuals are lower for full model than a model without covariates (mean: -2.11 vs -2.39; median: -12.96 vs -13.13)
Figure 7 plots the actual number of homicides for each municipality against the predicted homicides per municipality from the negative binomial model with all variables from Table 3. The model appears to provide a reasonable fit to the data with, as expected, notable increases in variability towards the higher ends of homicide rates.

Table 2. Selected assessment criteria for regression models of the form *homicide (count) ~ inequality + children + income*, with the log of population used as offset/exposure.

Model	**Dispersion parameter**	**AIC**	**Residual deviance**	**Degrees of freedom**	**p (χ^2 test)**
Poisson	1	2592.3	1686.1	226	>0.999
quasi-Poisson	7.961	-	1686.1	226	>0.999
Negative Binomial	2.854	1532.7	257.17	226	0.92

Table 3. Negative binomial model estimates

Variable	*β*	*p*	*95% CI*
Average monthly income per capita (US$)	<0.001	0.647	(-0.002, 0.003)
Proportion of people living in households with children with no parents	0.539	0.410	(-0.743, 1.830)
Palma index	-0.307	<0.001*	(-0.476, -0.138)

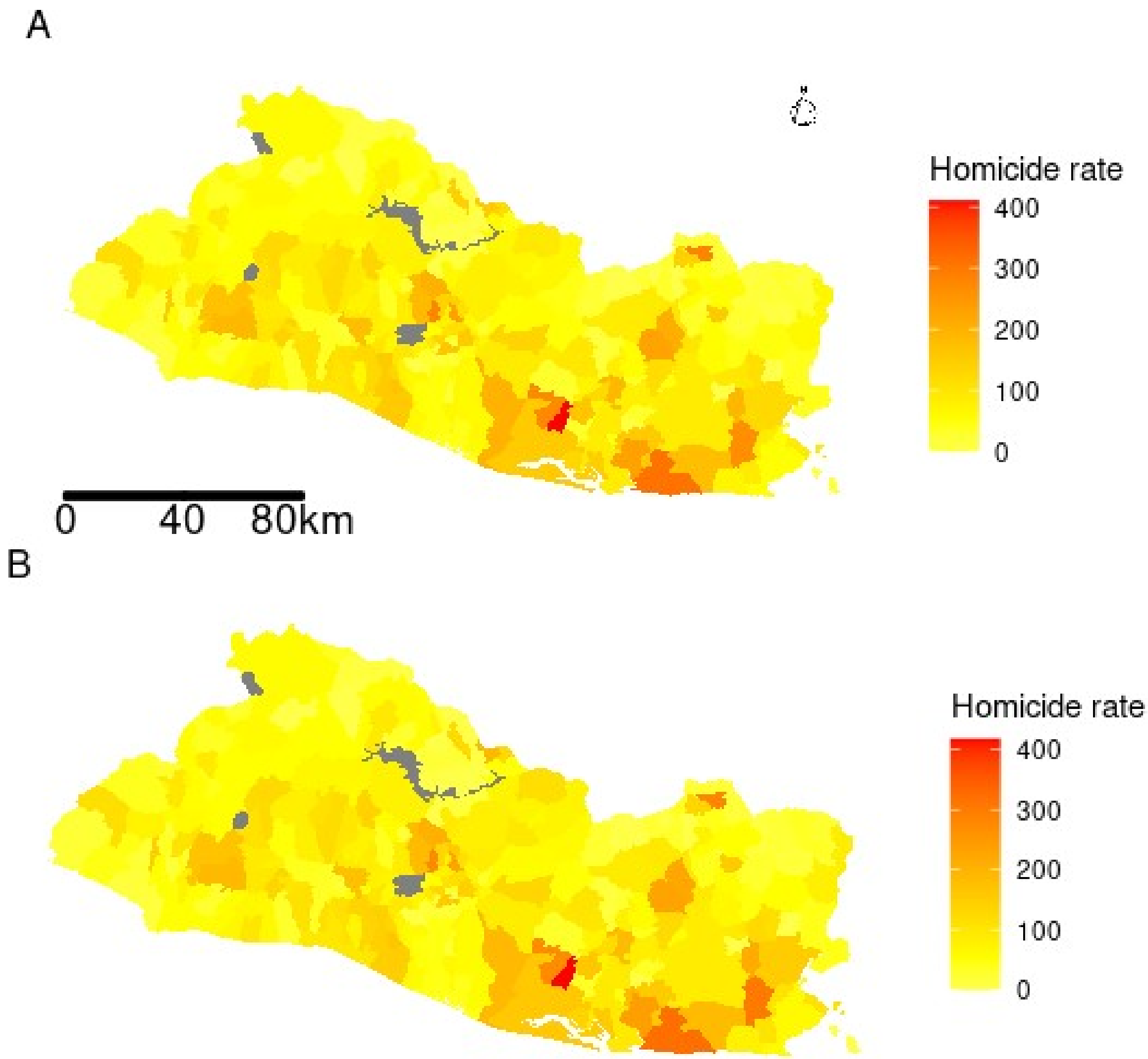

Figure 3. A. Raw homicide rates in El Salvador's municipalities, 2016. **B**. Standardised rates in El Salvador's municipalities, 2016. Grey areas represent inland waters/lakes.

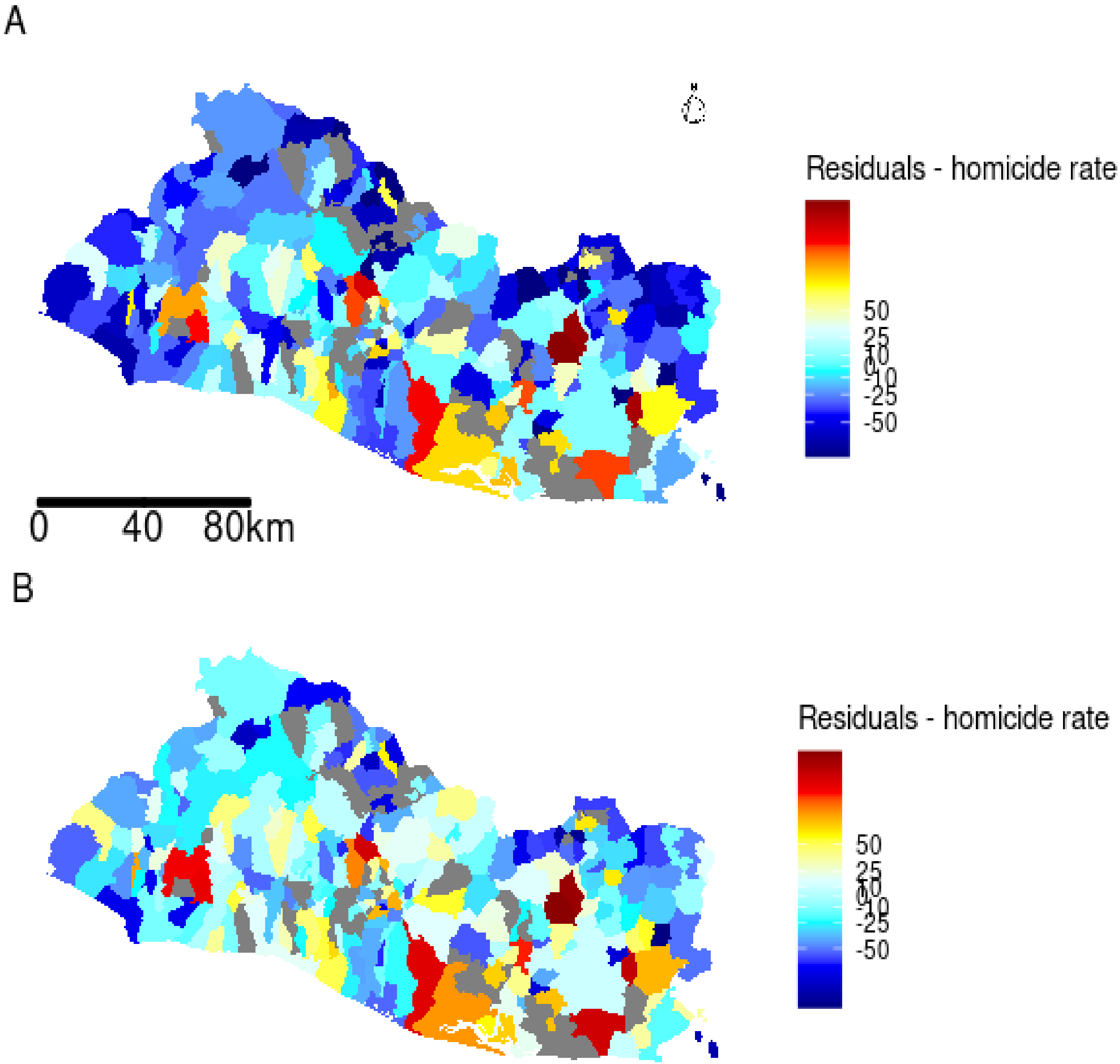

Figure 4. A. Geographical distribution of negative binomial model residuals with no covariates (**A**) and all three covariates (**B**). Grey areas represent inland waters and municipalities with missing socio-economic data.

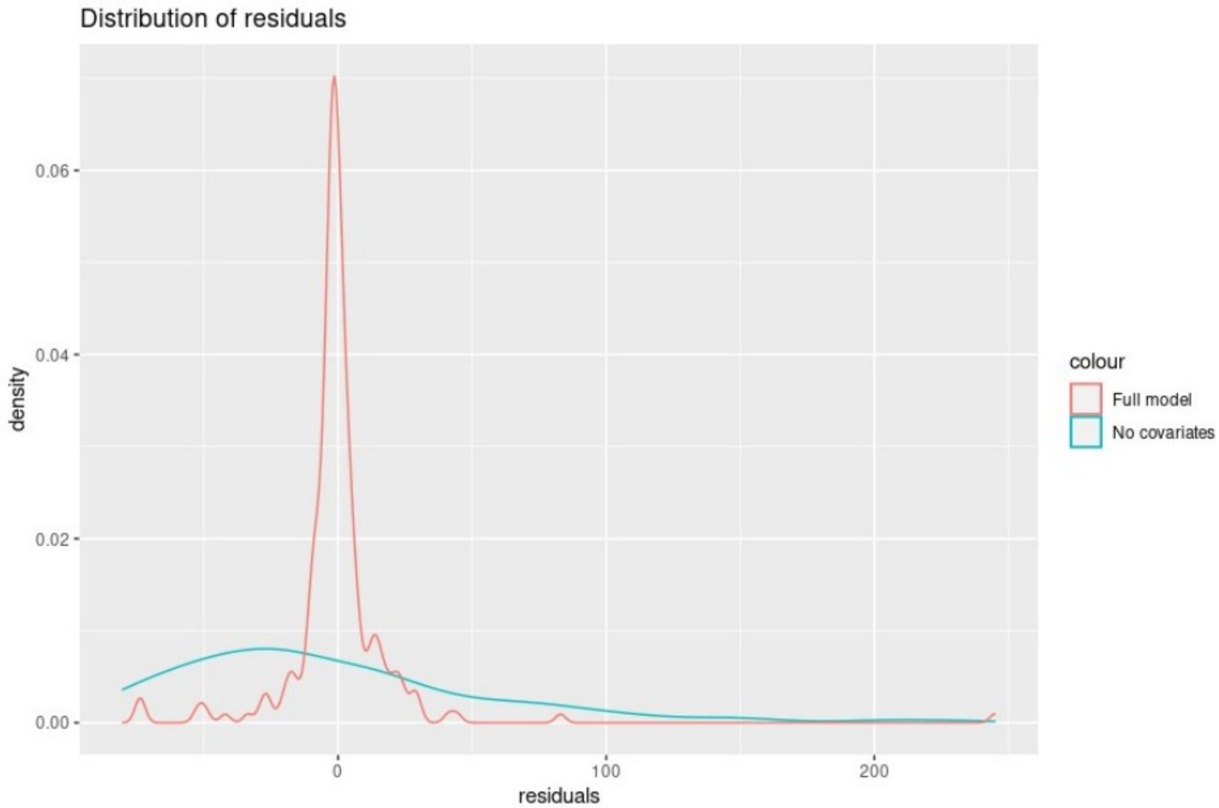

Figure 5. Distribution of residuals for negative binomial models with no covariates (blue) and full model (red).

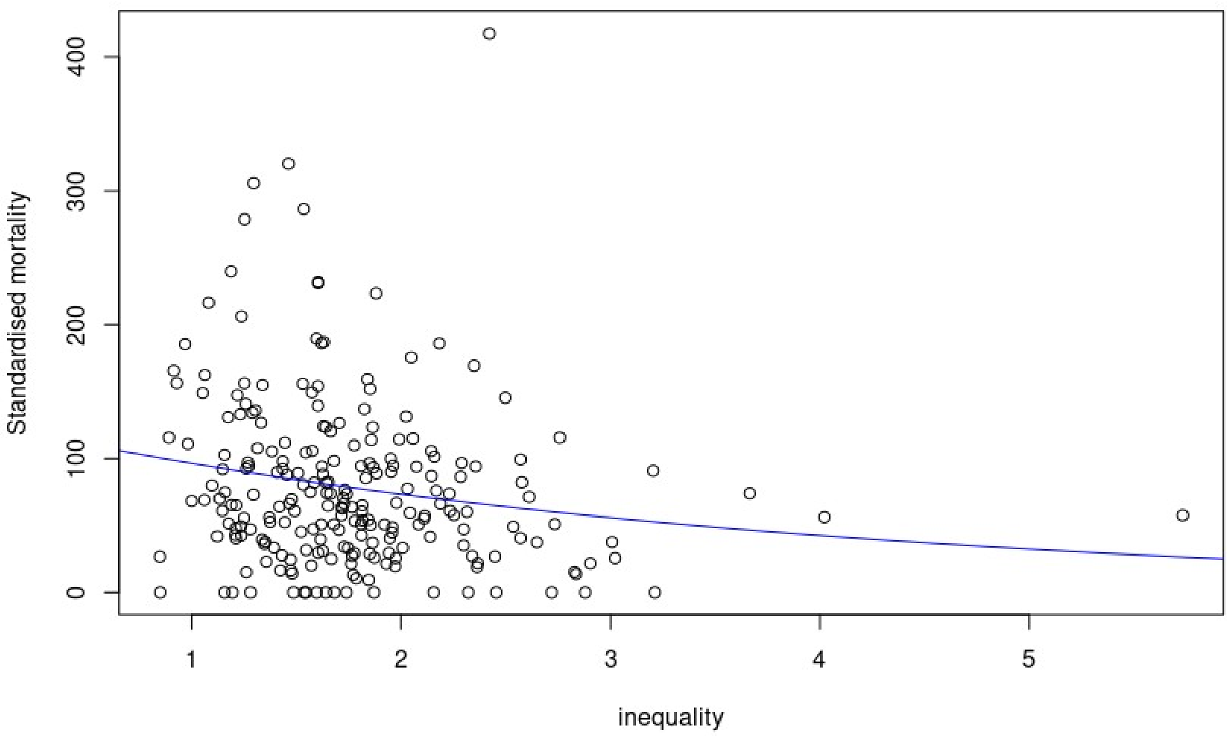

Figure 4. Plot of standardised homicide rate vs inequality index, including fitted univariate negative binomial regression (blue line).

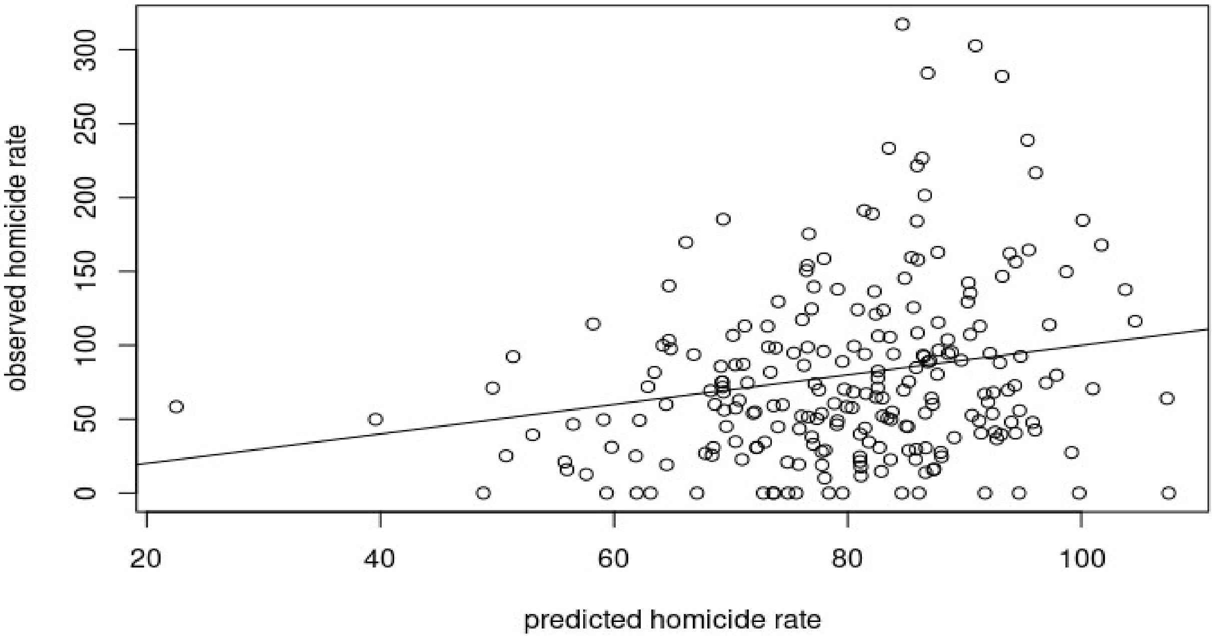

Figure 5. Observed vs predicted homicide rates for negative binomial regression model.

Discussion

High variability in homicide rates between municipalities was encountered, which was only slightly attenuated by adjusting the rates for sex and age. In light of some of the available evidence on factors associated with homicide rate variability [3][4][5], we included income-level indicators, as well as inequality, and the proportion of people living in households with parent-less children; of these variables, only a measure of inequality proved able to contribute to explain the differences in homicide rates. Nevertheless, inequality is not able to fully capture the regional variation in homicide rates.

What is most striking about the coefficients in Table 4, however, is the counter-intuitive effect of inequality. Whereas available evidence suggests that inequality is *positively* correlated with homicide rates [7][13] , we are faced with a strongly *negative* correlation. We posit that this is still able to be explained by existing theories of crime, in particular social disorganisation theory[14].

For most of the 20th century, El Salvador was basically owned and administered by an oligarchic elite in what is often described as semi-feudal nation-state with an agrarian mode of production, which resulted in, and was interrupted and transformed by, a 12-year bloody civil war [15] [16].

At the end of the conflict, the country had suffered a renewal of sorts, with the widespread adoption of neoliberal policies and the dismantling of the old export-oriented feudal-agricultural mode of production [15].

A possible explanation of the effects observed in this study, specifically concerning inequality, is that higher inequality indexes may be strongly associated with remaining enclaves of feudal-like agrarian zones, whereas lower Palma inequality indexes may be associated with regions where the effects of neoliberal reform (and rural-urban migration) have been more strongly felt. The relative anomy and disorganisation of the semi-industrial, semi-financial neoliberally-renewed areas contrast sharply with the highly unequal but highly structured and hierarchical agricultural society.

Also noteworthy is the significance of the Palma inequality index. The index is of relatively recent appearance, and other measures of inequality are purported to better describe the whole distribution of wealth (such as the well-known Gini); however the relatively constant proportion of income taken up by the deciles 2 to 6, has lead some to suggest that the the politics (and perhaps, the conflictivity?) of inequality are better described by the Palma [17].

Existing (qualitative) monographs on anthropological and economic history may also offer insights into the dynamics at play behind our data. Particularly, Polanyi[18] offers a glaring intuition when describing the period of chaotic transition in peri-Industrial Revolution England, ascribing much of the violence and "moral degradation" to the "lethal injury of [cultural] institutions" and the "disintegration of the cultural environment". That is, it would appear that in this case, as around Industrial Revolution England, cultural transformations spearheaded by economic reform has had a major impact on the milieu of violence.

There are, nonetheless, important limitations in the study. Of special concern is missing data on ~30 municipalities, which are overwhelmingly smaller and thus generally have lower levels of homicides; this could lead to a potential source of bias. Also of concern is the reliability of self-reported income data, used to estimate both municipality mean income, as well as the inequality index. The ecological nature of the study also carries inherent limitations; the adage "association does not imply causation" is specially relevant in this scenario, and we are limited to statistical interpretation of any associations. It is also important to consider the inherent ecological bias, even if all possible confounders were included in our model.[19]

Another important shortcoming is, of course, the limited amount of data used, specifically, a single year of homicide reporting (2016) from a single Central American country (El Salvador), as well as the (ingenuous) assumption of municipalities as independent observations. In regards to the former, the main driver of the shortcoming can be ascribed to the lack of readily available reporting from the region. In particular, access to raw data from socio-economic surveys has recently been limited in El Salvador, with 2016 being the most up-to-date open dataset; time-constraints as well as the exploratory nature of the project compounded these limitations. Ideally, a longitudinal, multi-country analysis would be used to further assess the findings in this ecological study.

In regards to the latter (assumption of independence of each municipality), geographic (spatial) autocorrelation could ideally be addressed through a number of mechanisms, such as the incorporation of a region covariate into the model, model stratification by region, or the use of mixed-effect models. The informally identified North/South gradient in homicide rates could also be formally assessed through measures such as Moran’s I or Mantel test. In general, autocorrelation tends to introduce biases and may lead to mis-estimations of model coefficients [20]

Despite these shortcomings, the results may still prove useful in as much as they suggest that levels of violence may be susceptible to cultural-behavioural interventions.

References

[1] Cobham, A., Schlögl, L. and Sumner, A. (2016), Inequality and the Tails: the Palma Proposition and Ratio. Glob Policy, 7: 25-36. https://doi.org/10.1111/1758-5899.12320

[2]Muggah, R., Carvalho, I. S., Aguirre, K. [2018] Latin America is the world's most dangerous region. But there are signs it is turning a corner. World Economic Forum, available from: https://www.weforum.org/agenda/2018/03/latin-america-is-the-worlds-most-dangerous-region-but-there-are-signs-its-turning-a-corner

[3]Lopez-Calva, L.F. [2019] Killing Development: The Devastating Epidemic of Crime and Insecurity in Latin America and the Caribbean. United Nations Development Program, available from: https://www.latinamerica.undp.org/content/rblac/en/home/presscenter/director-s-graph-for-thought/killing-development---the-devastating-epidemic-of-crime-and-inse.html

[4]Vilalta Perdomo, Carlos J.; Castillo, José G.; Torres, Juan A (2016), Violent Crime in Latin American Cities, Inter-American Development Bank. Available from: https://publications.iadb.org/publications/english/document/Violent-Crime-in-Latin-American-Cities.pdf

[5]Trujillo Álvarez, P 2017, 'Violence in Central America: on causes and consequences', *Anuario Latinoamericano – Ciencias Políticas y Relaciones Internacionales*, vol. 4, p. 21–, doi: 10.17951/al.2017.4.21.

[6]Dong B, Egger PH, Guo Y (2020) Is poverty the mother of crime? Evidence from homicide rates in China. PLoS ONE 15(5): e0233034. https://doi.org/10.1371/journal.pone.0233034

[7]Pare PP, Felson R. Income inequality, poverty and crime across nations. Br J Sociol. 2014 Sep;65(3):434-58. Doi: 10.1111/1468-4446.12083.

[8]Kawachi I, Kennedy BP, Wilkinson RG. Crime: social disorganization and relative deprivation. Soc Sci Med. 1999 Mar;48(6):719-31. doi: 10.1016/s0277-9536(98)00400-6

[9] Ministry of Security and Justice, El Salvador. [2020] *Repositorio de homicidios homologados*. Dirección de información y análisis. San Salvador, El Salvador. Available from: http://www.seguridad.gob.sv/dia/estadisticas-homologadas/repositorio-de-los-homicidios-desagregados-por-las-variables-homologadas/

[10]Ministry of Economy, El Salvador [2016] *Encuesta de Hogares de propósitos múltiples*. San Salvador, El Salvador. Available from: http://digestyc.microdatahub.com/index.php/catalog/16/data_dictionary

[11]Ministry of Health, El Salvador [2017],*Pirámides poblacionales año 2016*. San Salvador, El Salvador. Available from: https://www.salud.gob.sv/piramides-poblacionales-ano-2016-el-salvador/

[12] Tabula-java, extract tables from PDFs. Available from: https://github.com/tabulapdf/tabula-java/releases

[13]Kelly, Morgan. "Inequality and Crime." The review of economics and statistics 82.4 (2000): 530–539. Web.

[14] Shaw, C., McKay, H. ,Juvenile Delinquency and Urban Areas (Chicago: University of Chicago Press, 1942).

[15] Velasquez Carrilo, C. The Reconsolidation of Oligarchic Rule in El Salvador: The Contours of Neo-liberal Transformation — in *Dominant Elites in Latin America*. Cham: Springer International Publishing, 2017. 1–21. Web

[16] William I. Robinson (2002) Globalisation as a Macro-Structural-Historical Framework of Analysis: The Case of Central America, New Political Economy, 7:2, 221-250, DOI:10.1080/13563460220138853

[17] Cobham, A. (2013) Palma vs Gini: Measuring post-2015 inequality. Centre for global Development. Available from: https://www.cgdev.org/blog/palma-vs-gini-measuring-post-2015-inequality

[18] Polanyi, Karl [1944] *The Great Transformation: the political and economic origins of our time.* 2nd edition. Beacon Press, Boston

[19]Sander Greenland, Ecologic versus individual-level sources of bias in ecologic estimates of contextual health effects, *International Journal of Epidemiology*, Volume 30, Issue 6, December 2001, Pages 1343–1350, https://doi.org/10.1093/ije/30.6.1343
[20] Dormann, C.F. (2007), Effects of incorporating spatial autocorrelation into the analysis of species distribution data. *Global Ecology and Biogeography*, 16: 129-138. https://doi.org/10.1111/j.1466-8238.2006.00279.x