Sharif University of Technology
School of Management and Economics

Exploring the Correlation between Income Inequality and Crime Rates in a Cross-National Context

Sadra Daneshvar 98104111[1]

Research Advisor: **Dr. Narges Tahmasebi**[2]

Course: **Econometrics**

November 2022

[1] Sharif University of Technology, Tehran, Iran: amirsadra.daneshvar@ie.sharif.edu
[2] Sharif University of Technology, Tehran, Iran: ntahmasebi22@gmail.com

Table of Contents

Table of Figures

Abstract

The intricate connection between income inequality and crime rates has long perplexed economists, politicians, and society at large. In our forthcoming research, we aim to elucidate, analyze, and scrutinize the variables at play in this complex issue. Subsequently, we will evaluate the hypothesis positing a correlation between them, utilizing data drawn from various countries.

In this pursuit, we have selected murder rates and theft rates as proxies for crime rates, and we have explored their relationship with the Gini coefficient—a measure of income inequality—across the years 2003 to 2020. To ensure robust findings, we have incorporated additional control variables, including education levels, urbanization rates, gross national income per capita, and economic growth. Finally, employing Stata software, we have constructed a multivariable regression model to address this multifaceted problem.

The results derived from our research indicate a general positive relationship between income inequality and crime when considering the data as a panel. However, it is worth noting that this relationship may exhibit variations in different countries.

1 Introduction

The relationship between income inequality and crime has long been a focal point in economic literature pertaining to criminal behavior [1]. Nobel laureate economist Gary Becker, half a century ago, posited a foundational theory in his influential article "Crime and Punishment: An Economic Approach," asserting that all criminal activity is fundamentally driven by economic considerations, with criminals engaging in rational decision-making. Becker contended that potential offenders weigh the expected benefits of unlawful actions against the probability of getting apprehended and facing penalties [6].

The crime rate, according to this perspective, hinges on factors such as the risks associated with arrest and the severity of penalties, as well as the disparity between the potential gains from criminal acts and the related opportunity costs [3]. In essence, this net benefit is theoretically represented by the wealth gap between affluent and disadvantaged populations [3].

Moreover, the nexus between inequality and crime has also been scrutinized through sociological lenses [2]. In broad terms, these sociological frameworks stem from the consistent observation that "lower-class individuals and residents of lower-class neighborhoods consistently exhibit higher official crime rates compared to other demographic groups" [2]. One prominent sociological paradigm, known as the "relative deprivation" theory, posits that inequality fosters social tensions [5]. The sense of deprivation and injustice compels economically disadvantaged individuals to seek redress and satisfaction through various means, including resorting to criminal activities against both their impoverished peers and the affluent [4].

Figure 1- Illustrating the effect of income inequality on the results of the survey

Recent Gallup polling data has offered partial validation of Becker's theory [6]. In a survey encompassing 148,000 respondents across 142 countries, individuals were queried about their perceptions of crime and their feelings of safety, based on four key criteria: trust in local law enforcement, personal safety when traveling alone, experiences of theft or property loss, and incidents of sexual assault within the past year. Correlating these responses with the level of income inequality, as gauged by the Gini coefficient, in each respective nation revealed a robust and positive correlation, as illustrated in Figure 1 [6].

These straightforward associations do not encompass the entirety of variations in individuals' perceptions of crime levels. Subsequently, scholars have expanded upon Becker's theory to explore the extent to which the affluent showcase their wealth [7]. In a study conducted in 2014, Daniel Hicks from the University of Oklahoma and Joan Hamory Hicks from the University of California at Berkeley revealed a compelling trend over a 20-year span: U.S. states characterized by the highest disparities in expenditure–pertaining to items like clothing, jewelry, automobiles, and dining out–also experienced the most pronounced instances of violent crime [8].

To delve more precisely into the mechanics of this phenomenon and elucidate the aforementioned correlation, we found it imperative to construct a regression model. For the examination of the crime rate, we opted for two variables, namely, the murder rate and robbery rate, as they serve as apt proxies. This selection was driven by their close association with overall crime levels, while sidestepping the endogeneity conundrum [1]. Moreover, we have employed the Gini index as a proxy to elucidate the degree of income inequality. Supplementary variables have been judiciously chosen to mitigate potential distortions and tailor the model, a comprehensive discussion of which will ensue in section 2.

* * *

2 Data

2.1 Data Sources, Structure, and Information

In our study, we focused on two dependent variables: HomicideRate and RobberyRate, both of which measure the occurrence of these crimes per 100 thousand people. The data source for HomicideRate was the World Bank, which provided data spanning from 1968 to 2020 for 140 countries. Additionally, we obtained RobberyRate data from United Nations sources, covering 88 countries in the period from 1990 to 2020.

For our independent variables, we initially conducted a regression analysis with only the Gini coefficient, which we also sourced from the World Bank. However, this preliminary analysis resulted in a weak model that failed to adequately explain variations in the dependent variables across all countries.

To enhance the model's predictive power and mitigate the impact of omitted variables, we incorporated two additional factors: economic growth (GDPGrowth) and gross national income per capita (GNIperCapita), both obtained from the World Bank. This multivariable regression yielded significantly improved results, although further refinements were still possible.

Finally, we introduced two more explanatory variables into our model: average years of education (AverageSchooling) and the rate of urbanization (Urbanization), covering the years 1980 to 2020, sourced from Our World in Data. These additions notably enhanced the model's accuracy.

Table 1 presents a summary of the data and their respective sources:

Table ١ - Data information

Variable Name	Available years	Countries available (count)	Source
Homicide Rate	1968-2020	140	World Bank
Rubbery Rate	1990-2020	88	UN
Gini Index	1968-2020	140	World Bank
GNI per Capita	1968-2020	140	World Bank
GDP Growth	1968-2020	140	World Bank
Average Schooling	1980-2020	122	OWID
Urbanization	1980-2020	122	OWID

2.2 Selecting and Cleaning Data

In handling the diverse data sources at our disposal, our primary objective was to combine them and rectify any discrepancies. Our initial course of action involved the exclusion of data preceding the year 2000, primarily due to the prevalence of numerous missing values. Additionally, we initially considered 32 countries from the global dataset, but ultimately narrowed it down to 11 countries whose data demonstrated greater consistency across various sources. These countries comprised Brazil, Colombia, France, Germany, Greece, Hungary, Italy, Mexico, Netherlands, Norway, and the United States.

However, even within this subset of 11 countries, there remained instances of missing data for murder and robbery rates until 2003. Consequently, we applied a filtering process to establish our final timeframe, encompassing the years 2003 to 2020. The conclusive dataset is provided as an appendix for further examination.

Nevertheless, some gaps persisted in the data, prompting us to employ interpolation techniques in Stata to supplement our dataset and ensure completeness.

Furthermore, an interesting aspect regarding our data was its suitability for panel analysis. Although this lay outside our formal training, driven by personal interest and curiosity, we ventured into panel analysis in addition to single-country analysis. Detailed insights into these aspects can be found in sections 3 and 4 of this study.

* * *

3 Single Country Analysis

3.1 United States

Prior to delving into the analysis of data from the United States, we conducted a preliminary examination to assess the extent of missing data. The findings of this assessment are presented in Table 2.

Table 2- United States data information

Variable	Missing	Total	Percent Missing
Country	0	18	0.000
Year	0	18	0.000
HomocideRate	3	18	16.670
RobberyRate	1	18	5.560
GiniIndex	1	18	5.560
GNIperCapita	0	18	0.000
GDPGrowth	0	18	0.000
AverageSch~g	0	18	0.000
Urbanization	0	18	0.000

You can observe that all variables, with the exception of HomicideRate, RobberyRate, and GiniIndex, are complete and devoid of any issues. These three variables have also been filled in and are prepared for further processing using interpolation. Before delving into a detailed analysis of this dataset, it would be beneficial to take a moment to review its summary information:

Table 3- United States data summary

Variable	Obs	Mean	Std. Dev.	Min	Max
GiniIndex	18	40.983	.43	40	41.6
RobberyRate	18	117.001	24.12	76.887	150.876
HomicideRate	18	5.306	.606	4.445	6.517
Urbanization	18	.811	.01	.796	.827
AverageSchooling	18	13.16	.264	12.825	13.683
GDPGrowth	18	.99	1.846	-3.698	2.896
GNIperCapita	18	55602.523	3562.465	50303.713	62173.485
Year	18	2011.5	5.339	2003	2020

Clearly, the data suggests that, on average, the American economy exhibits positive and substantial growth. Simultaneously, it is evident that the incidence of theft in the country is notably high.

Turning our attention to the relationship between the homicide rate and the Gini coefficient within this nation, our analysis indicates a weak positive correlation between these two variables, even without the control of other factors.

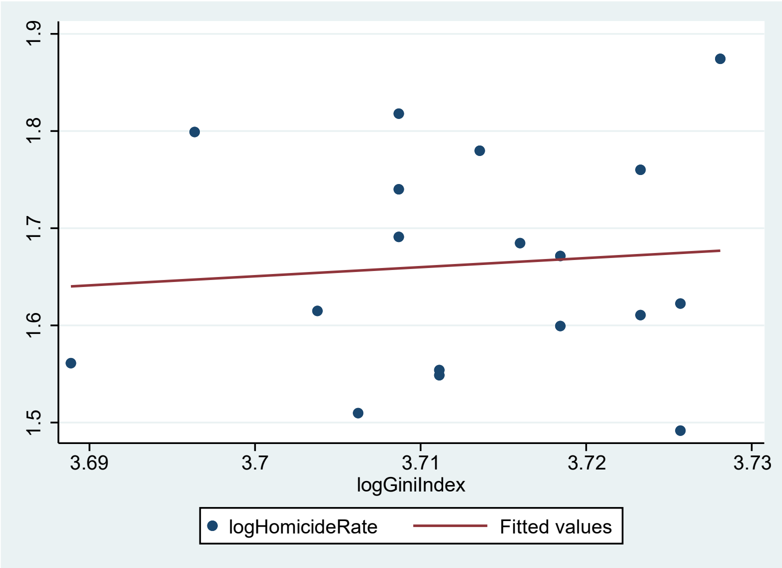

Figure 2- Correlation between the Gini coefficient and homicide rates in the United States

In the following analysis, we examined whether the data pertaining to this specific country adheres to a normal distribution. In cases where the data exhibited skewness, we endeavored to address this issue by applying the natural logarithm. This step was taken to ensure that the fundamental assumptions of Ordinary Least Squares (OLS) regression would not be compromised. In the context of the United States, our findings revealed the following:

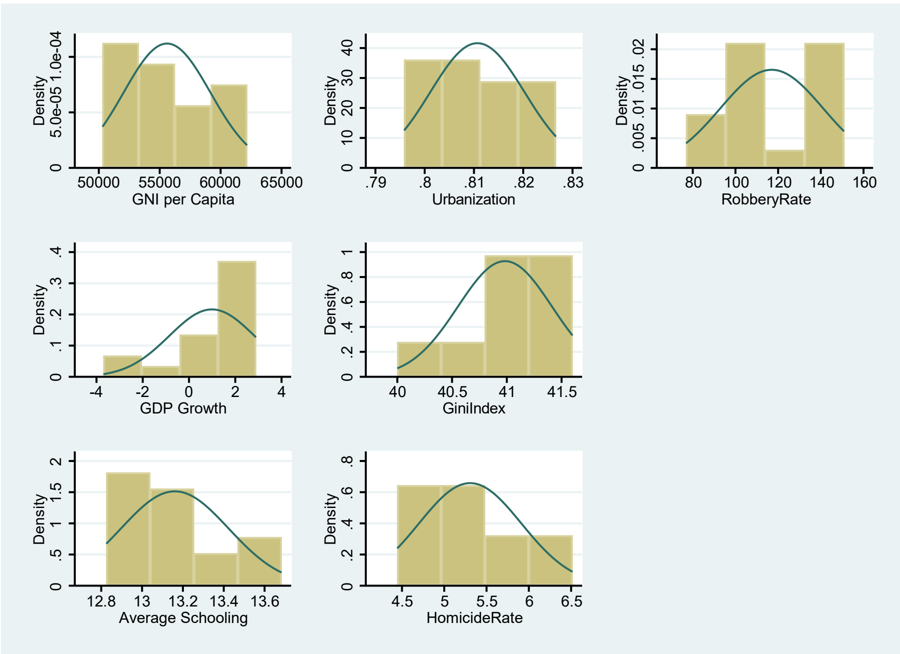

Figure 3- United States data distribution

We observe that the variables Homicide Rate, Average Schooling, Gini Index, and GDP Growth exhibit skewness. Consequently, we applied a logarithmic transformation to these variables and used their natural logarithms in our modeling.

Now, let's present the American regression models:

$$Model1: logHomicideRate = \beta_0 + \beta_1 logGiniIndex$$

$$\begin{aligned} Model2: logHomicideRate \\ = \beta_0 + \beta_1 logGiniIndex + \beta_2 GNIperCapita + \beta_4 GDPGrowth \end{aligned}$$

$$\begin{aligned} Model3: logHomicideRate \\ = \beta_0 + \beta_1 logGiniIndex + \beta_2 GNIperCapita + \beta_4 GDPGrowth \\ + \beta_5 logAverageSchooling + \beta_6 Urbanization \end{aligned}$$

You can find the regressions applied to this country, as detailed in section 3.1, in the table below:

Table 4- Regression of the homocide rate in America

VARIABLES	(1) Model One	(2) Model Two	(3) Model Three
logGiniIndex	0.937	4.408	-2.027
	(2.670)	(3.970)	(2.868)
GNIperCapita		-1.45e-05	6.14e-05**
		(1.17e-05)	(2.50e-05)
GDPGrowth		-0.00713	-0.0368**
		(0.0154)	(0.0150)
logAverageSchooling			15.77**
			(5.763)
Urbanization			-57.40***
			(11.85)
Constant	-1.815	-13.89	11.70
	(9.913)	(14.27)	(15.57)
Observations	18	18	18
R-squared	0.008	0.128	0.721

We can observe that in the third model, which incorporates a greater number of controlled variables, it explains approximately 72% of the variations in the dependent variable. This is indeed a substantial and desirable level of explanation. Furthermore, the positive coefficient associated with logGiniIndex suggests a direct relationship between this variable and the homocide rate. However, it's worth noting that this relationship is not statistically significant. The explanations and interpretations of the remaining variables are clear. Nonetheless, it's important to highlight that all of these variables are statistically significant in the third and last model.

These are also calculated for RobberyRate:

$$Model1: logRobberyRate = \beta_0 + \beta_1 logGiniIndex$$

$$Model2: logRobberyRate = \beta_0 + \beta_1 logGiniIndex + \beta_2 GNIperCapita + \beta_4 GDPGrowth$$

$$\begin{aligned} Model3: logRobberyRate \\ &= \beta_0 + \beta_1 logGiniIndex + \beta_2 GNIperCapita + \beta_4 GDPGrowth \\ &+ \beta_5 logAverageSchooling + \beta_6 Urbanization \end{aligned}$$

Table 5- Regression of the robbery rate in America

VARIABLES	(1) Model One	(2) Model Two	(3) Model Three
logGiniIndex	-1,217**	595.1	-220.0
	(485.5)	(403.5)	(221.4)
GNIperCapita		-0.00719***	0.00626***
		(0.00119)	(0.00193)
GDPGrowth		1.978	-4.533***
		(1.570)	(1.158)
logAverageSchooling			-1,098**
			(444.8)
Urbanization			-2,367**
			(914.9)
Constant	4,637**	-1,695	5,338***
	(1,803)	(1,451)	(1,202)
Observations	18	18	18
R-squared	0.282	0.803	0.964

In this regression analysis, it becomes evident that the third model offers a significantly superior explanation of the variations in the dependent variable when compared to the other two models. This enhanced performance can be attributed to the careful control and elimination of distortions caused by omitted variables. Notably, in this particular instance, the coefficient for logGiniIndex fails to reach statistical significance and, somewhat surprisingly, exhibits a negtive relationship, contradicting the expected evidence.

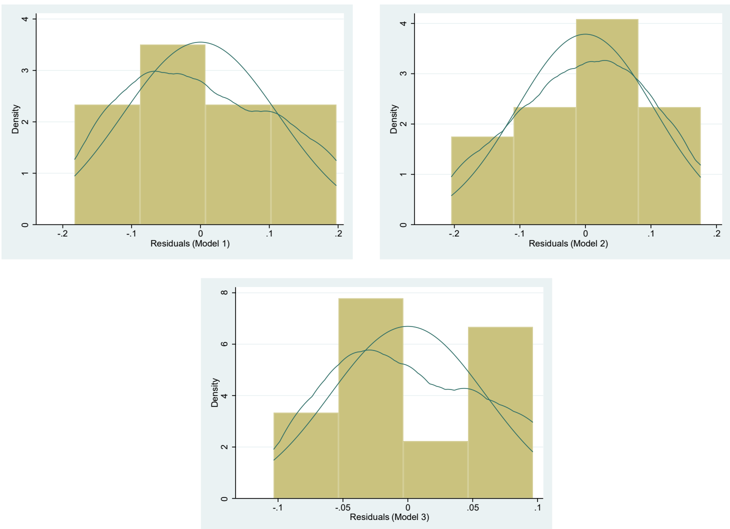

Figure 4- United States error distribution

It's evident that the error distribution in the regressions of all three models closely resembles a normal distribution, which is a favorable outcome. Generally, in the context of the United States, there doesn't appear to be a strong, precise correlation between the income inequality index and the prevalence of crime.

3.2 Mexico

Before delving into the Mexico data, we conducted an initial assessment to gauge the volume of available data. The findings from this evaluation are presented in Table 6.

Table 6- Mexico data information

Variable	Missing	Total	Percent Missing
Country	0	18	0.000
Year	0	18	0.000
HomicideRate	0	18	0.000
RobberyRate	3	18	16.670
GiniIndex	8	18	44.440
GNIperCapita	0	18	0.000
GDPGrowth	0	18	0.000
AverageSch~g	0	18	0.000
Urbanization	0	18	0.000

We can observe that all variables, except for Robberyrate and GiniIndex, are complete and free of any issues. These two variables have also been filled in and are now ready for processing using interpolation. Before delving into a detailed analysis of this dataset, it would be beneficial to take a moment to review its summary information:

Table 7- Mexico data summary

Variable	Obs	Mean	Std. Dev.	Min	Max
GiniIndex	18	48.292	1.379	45.4	50.1
RobberyRate	18	411.087	197.311	128.833	649.766
HomicideRate	18	18.293	7.276	8.122	29.071
Urbanization	18	.782	.016	.757	.807
AverageSchooling	18	8.246	.588	7.132	9.221
GDPGrowth	18	.362	3.138	-8.655	3.733
GNIperCapita	18	9146.963	412.863	8488.411	9841.469
Year	18	2011.5	5.339	2003	2020

We can observe that Mexico experiences consistently modest but positive economic growth. Conversely, the country faces a considerably high theft rate on average, nearly four times that of the United States.

Furthermore, an examination of the correlation between the homicide rate and the Gini coefficient in Mexico reveals an intriguing pattern. Notably, there appears to be a negative relationship between these two variables, even without controlling for other factors.

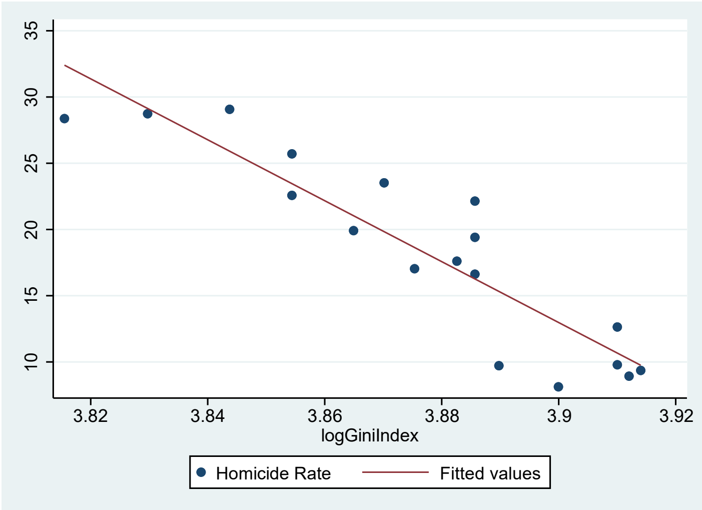

Figure ◦ - Correlation between the Gini coefficient and homicide rates in the Mexico

In the following analysis, we examined whether the data pertaining to this particular country conforms to a normal distribution. In cases where the data exhibited skewness, we endeavored to address this issue by applying the natural logarithm. This adjustment was made to ensure that the fundamental assumptions of Ordinary Least Squares (OLS) regression would not be compromised. In the context of Mexico, our findings are as follows:

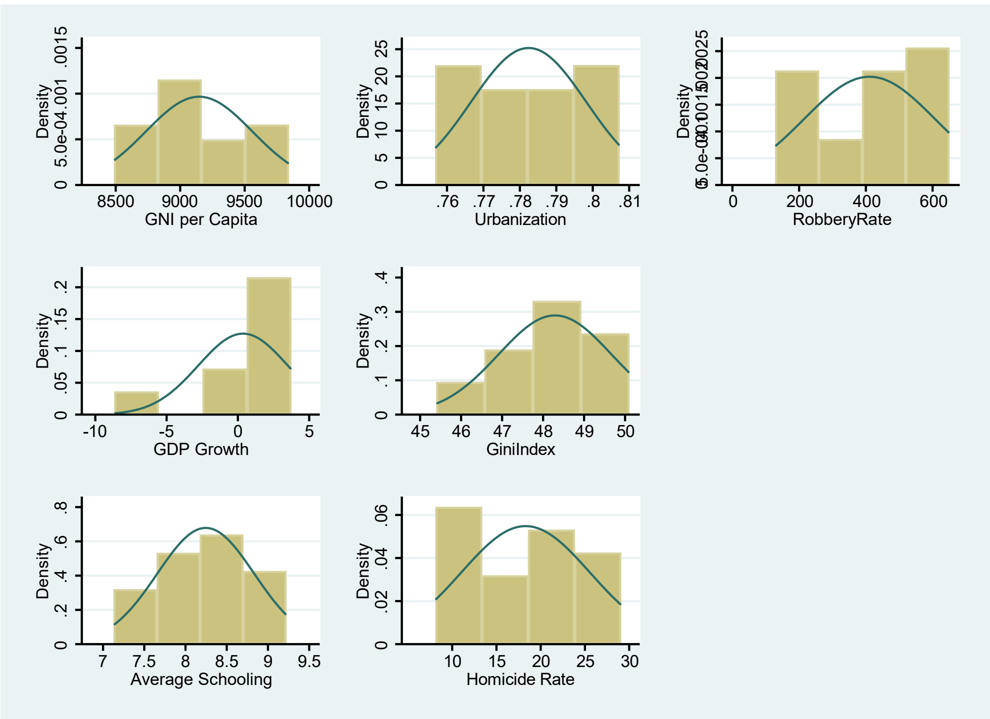

Figure 6- Mexico data distribution

We observe that the only variable exhibiting skewness is GiniIndex. Consequently, we applied a logarithmic transformation to this variable and used its logarithm in our modeling process.

The regression models for the Mexico dataset is outlined below:

$$Model1: HomicideRate = \beta_0 + \beta_1 logGiniIndex$$

$$Model2: HomicideRate = \beta_0 + \beta_1 logGiniIndex + \beta_2 GNIperCapita + \beta_4 GDPGrowth$$

$$\begin{aligned} Model3: HomicideRate \\ &= \beta_0 + \beta_1 logGiniIndex + \beta_2 GNIperCapita + \beta_4 GDPGrowth \\ &+ \beta_5 AverageSchooling + \beta_6 Urbanization \end{aligned}$$

You can find the regressions applied to this country in the table below:

Table 8- Regression of the homocide rate in Mexico

VARIABLES	(1) Model One	(2) Model Two	(3) Model Three
logGiniIndex	-230.0***	-213.2***	-128.1*
	(26.34)	(39.73)	(61.97)
GNIperCapita		0.00237	-0.00473
		(0.00269)	(0.00485)
GDPGrowth		0.00130	0.479
		(0.331)	(0.432)
AverageSchooling			2.231
			(6.790)
Urbanization			239.4
			(264.0)
Constant	909.8***	823.0***	352.2
	(102.1)	(171.7)	(337.0)
Observations	18	18	18
R-squared	0.827	0.840	0.873

All three models have effectively elucidated the variations in the dependent variable. However, it's worth noting that the third model outperforms the others in this regard. Additionally, it's noteworthy that the only variable exhibiting a significant coefficient is logGiniIndex, and interestingly, it carries a negative value.

Similar analyses were also conducted for RobberyRate:

$$Model1: RobberyRate = \beta_0 + \beta_1 logGiniIndex$$

$$Model2: RobberyRate = \beta_0 + \beta_1 logGiniIndex + \beta_2 GNIperCapita + \beta_4 GDPGrowth$$

$$\begin{aligned} Model3: RobberyRate \\ &= \beta_0 + \beta_1 logGiniIndex + \beta_2 GNIperCapita + \beta_4 GDPGrowth \\ &+ \beta_5 AverageSchooling + \beta_6 Urbanization \end{aligned}$$

Table 9- Regression of the robbery rate in Mexico

VARIABLES	(1) Model One	(2) Model Two	(3) Model Three
logGiniIndex	2,102	-1,471	-4,616*
	(1,632)	(1,999)	(2,290)
GNIperCapita		-0.396**	-0.174
		(0.135)	(0.179)
GDPGrowth		18.87	12.26
		(16.66)	(15.95)
AverageSchooling			770.3***
			(250.9)
Urbanization			-38,876***
			(9,755)
Constant	-7,740	9,723	43,956***
	(6,329)	(8,637)	(12,450)
Observations	18	18	18
R-squared	0.094	0.450	0.764

In this regression analysis, it becomes evident that the third model provides a notably superior explanation of changes in the dependent variable compared to the other two models. This improvement can be attributed to its effective control and elimination of distortions stemming from omitted variables. Notably, the coefficient associated with the logGiniIndex variable in the third model approaches significance and demonstrates an unexpected negative relationship, contradicting both our prior evidence and expectations.

Overall, the data reveals an intriguing and contentious pattern in Mexico, where there appears to be an almost inverse correlation between the income inequality index and the incidence of crime and criminal activities. This observation is both captivating and contentious, warranting further, more precise investigation in future studies.

3.3. Germany

Before starting to work with the German data, it was checked to determine how much empty data it has. The result of this study can be seen in Table 10.

Table 10- Germany data information

Variable	Missing	Total	Percent Missing
Country	0	18	0.000
Year	0	18	0.000
HomicideRate	0	18	0.000
RobberyRate	1	18	5.560
GiniIndex	2	18	11.110

Table 10- Germany data information (Cont.)

Variable	Missing	Total	Percent Missing
GNIperCapita	0	18	0.000
GDPGrowth	0	18	0.000
AverageSch~g	0	18	0.000
Urbanization	0	18	0.000

We can observe that all variables, except for RobberyRate and GiniIndex, are complete and free of any issues. These two variables have also been completed and are now prepared for processing through interpolation. Before delving into a detailed analysis of this dataset, it's a good idea to take a glance at its summary information:

Table 11- Germany data summary

Variable	Obs	Mean	Std. Dev.	Min	Max
GiniIndex	18	31.211	.685	30.1	32.7
RobberyRate	18	57.989	9.465	42.116	73.249
HomicideRate	18	1.02	.186	.746	1.362
Urbanization	18	.768	.006	.756	.775
AverageSchooling	18	13.901	.2	13.455	14.132
GDPGrowth	18	1.023	2.653	-5.455	5.87
GNIperCapita	18	39881.133	3234.015	34457.043	44758.513
Year	18	2011.5	5.339	2003	2020

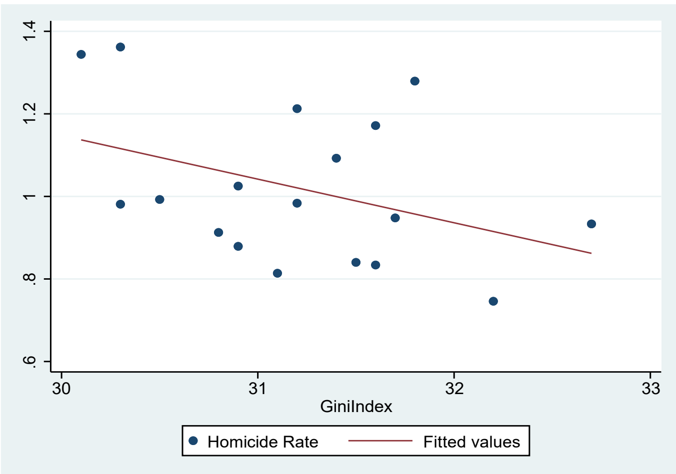

Figure 7- Correlation between the Gini coefficient and homicide rates in the Germany

We observe that Germany experiences consistently positive economic growth, which is noteworthy. Additionally, the incidence of theft in the country is relatively low on average.

Moving on, we delved into exploring the correlation between the homicide rate and the Gini coefficient in Germany. Interestingly, we found a negative relationship between these two variables when not accounting for other factors.

Subsequently, we conducted an assessment to determine if the data pertaining to this country adheres to a normal distribution. This step was taken to minimize any potential disruptions to the fundamental assumptions of Ordinary Least Squares (OLS) regression. In the context of our analysis concerning Germany, we have the following observations:

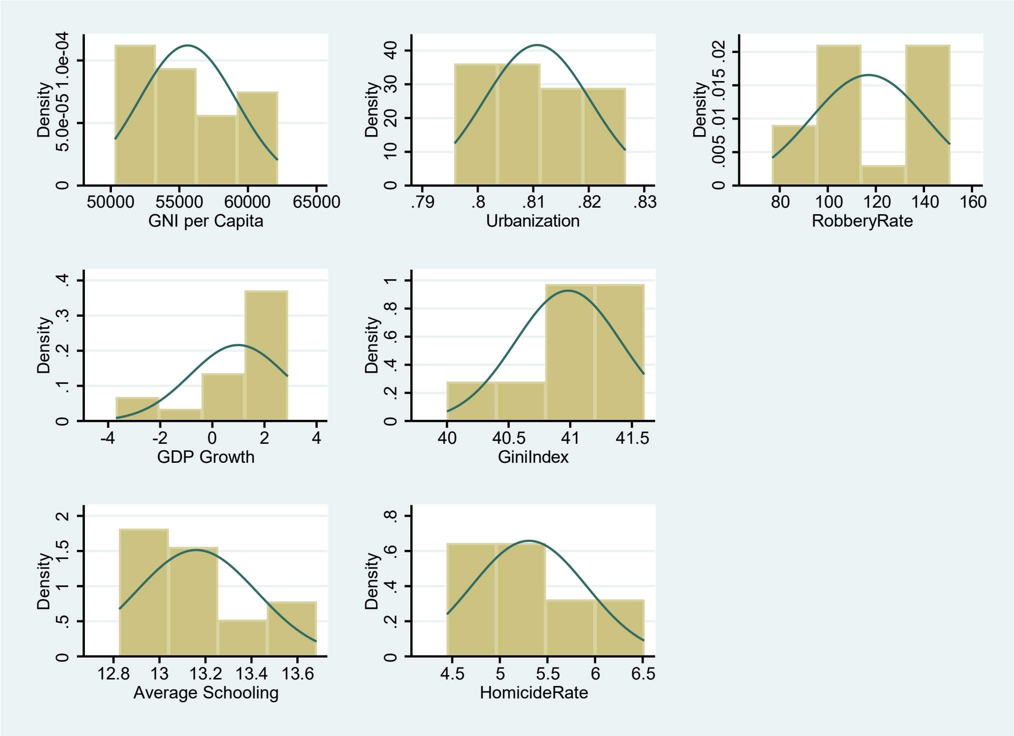

Figure ۸ - Germany data distribution

We observe that the Urbanization and Average Schooling variables exhibit skewness. Consequently, we applied a logarithmic transformation to these variables and used their logarithmic values in our modeling.

The regression model for the German dataset is as follows:

$$Model1: HomicideRate = \beta_0 + \beta_1 GiniIndex$$

$$Model2: HomicideRate = \beta_0 + \beta_1 GiniIndex + \beta_2 GNIperCapita + \beta_4 GDPGrowth$$

$$Model3: HomicideRate = \beta_0 + \beta_1 GiniIndex + \beta_2 GNIperCapita + \beta_4 GDPGrowth + \beta_5 logAverageSchooling + \beta_6 logUrbanization$$

You can find the regressions applied to this country in the table below:

Table 12- Regression of the homocide rate in Germany

VARIABLES	(1) Model One	(2) Model Two	(3) Model Three
GiniIndex	-0.106	0.0488	0.00220
	(0.0624)	(0.0671)	(0.0496)
GNIperCapita		-4.84e-05***	1.02e-05
		(1.39e-05)	(2.37e-05)

Table 12- Regression of the homocide rate in Germany (Cont.)

VARIABLES	(1) Model One	(2) Model Two	(3) Model Three
GDPGrowth		0.00651	0.00228
		(0.0131)	(0.00963)
logAverageSchooling			5.409
			(4.489)
logUrbanization			-34.02***
			(8.497)
Constant	4.319**	1.419	-22.67
	(1.949)	(1.776)	(12.83)
Observations	18	18	18
R-squared	0.152	0.546	0.808

We can observe that in the third model, which incorporates a greater number of controlled variables, it explains nearly 81% of the variations in the dependent variable. This percentage is quite substantial and considered favorable for our analysis. Additionally, the positive coefficient associated with logGiniIndex suggests a direct, albeit very slight, relationship between this variable and the murder rate. However, it's worth noting that this relationship is not statistically significant. It's important to emphasize that the explanation and interpretation of the remaining variables extend beyond the scope of our current discussion.

Similar calculations have been performed for RobberyRate:

$$Model1: RobberyRate = \beta_0 + \beta_1 GiniIndex$$

$$Model2: RobberyRate = \beta_0 + \beta_1 GiniIndex + \beta_2 GNIperCapita + \beta_4 GDPGrowth$$

$$\begin{aligned} Model3: RobberyRate \\ &= \beta_0 + \beta_1 GiniIndex + \beta_2 GNIperCapita + \beta_4 GDPGrowth \\ &+ \beta_5 logAverageSchooling + \beta_6 logUrbanization \end{aligned}$$

Table 13- Regression of the robbery rate in Germany

VARIABLES	(1) Model One	(2) Model Two	(3) Model Three
GiniIndex	-9.687***	-1.173	-1.231
	(2.463)	(1.278)	(1.336)
GNIperCapita		-0.00262***	-0.00264***
		(0.000265)	(0.000639)
GDPGrowth		0.530*	0.570**
		(0.250)	(0.259)
logAverageSchooling			155.5
			(121.0)
logUrbanization			-279.5
			(229.0)

Table 13- Regression of the robbery rate in Germany (Cont.)

VARIABLES	(1) Model One	(2) Model Two	(3) Model Three
Constant	360.3*** (76.89)	198.7*** (33.83)	-282.1 (345.7)
Observations	18	18	18
R-squared	0.492	0.936	0.946

In this regression analysis, we can clearly observe that the third model outperformed the other two models in explaining changes in the dependent variable. This superior performance can be attributed to its ability to control and eliminate distortions caused by omitted variables. Notably, in this case, the GiniIndex coefficient lacks statistical significance and even exhibits a negative relationship, which contradicts the expected evidence.

Furthermore, it's worth noting that the error distribution in the regressions of all three models closely approximates a normal distribution, which is a favorable characteristic for regression analysis.

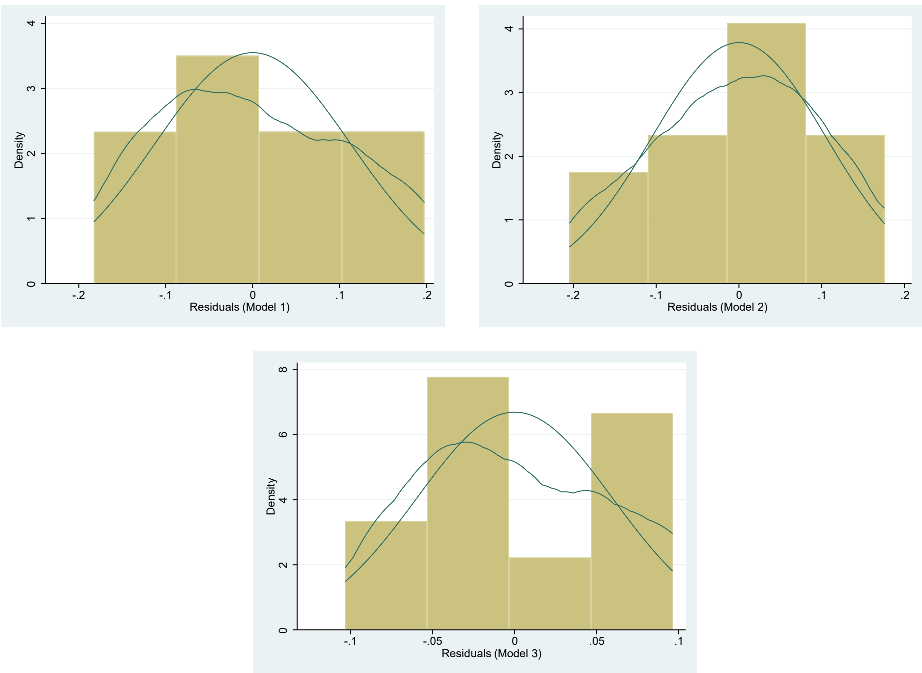

Figure 9- Germany error distribution

In summary, when examining the relationship between the index of income inequality and the incidence of crimes in Germany, it becomes evident that a precise and robust trend is not discernible.

* * *

4 Panel Data Analysis

In the upcoming section, we will introduce the regression model for panel data analysis. Before delving into the model itself, it's beneficial to first examine the summary of its information:

Table 14- Summary of panel data information

Variable		Mean	Std. dev.	Min	Max	Observations
Countr~e	overall	6	3.170294	1	11	N = 198
	between		3.316625	1	11	n = 11
	within		0	6	6	T = 18
Year	overall	2011.5	5.201279	2003	2020	N = 198
	between		0	2011.5	2011.5	n = 11
	within		5.201279	2003	2020	T = 18
GDPGro~h	overall	.720527	2.984468	-10.01628	6.522816	N = 198
	between		.9274805	-.6818621	2.209211	n = 11
	within		2.849747	-9.818751	6.94475	T = 18
Averag~g	overall	10.69311	2.219893	5.804	14.132	N = 198
	between		2.270063	7.071833	13.90067	n = 11
	within		.4684339	9.425273	11.83705	T = 18
Urbani~n	overall	.7819927	.0564561	.65489	.92236	N = 198
	between		.0559142	.6891517	.8746194	n = 11
	within		.0181848	.7107833	.8297333	T = 18
Homici~e	overall	8.365383	12.16919	.4615538	56.70396	N = 198
	between		12.19216	.7174441	34.17747	n = 11
	within		3.502582	-3.171863	30.89187	T = 18
Robber~e	overall	154.4513	180.7944	5.90248	806.491	N = 198
	between		172.3304	22.77389	556.7394	n = 11
	within		74.50891	-127.8025	404.2029	T = 18
GiniIn~x	overall	37.65985	9.457186	25.1	57.6	N = 198
	between		9.794773	27.36111	53.76111	n = 11
	within		1.334037	32.79874	42.72096	T = 18
GNIper~a	overall	30742.5	21639.6	4008.583	81611.71	N = 198
	between		22516.15	5319.622	75930.25	n = 11
	within		2245.945	18302.06	37313.46	T = 18

In this table, besides providing an overview, it also furnishes details about inter-country (between-country) and inter-time (within-time) factors, offering a more precise and comprehensive perspective.

It's worth highlighting that in section 3, we've analyzed a sample of three countries, but for the panel data analysis, all 11 countries were included in the calculations and regression. The outcome of the regression analysis for the murder rate and its explanatory variables is presented below:

Table 15- Panel regression on the homocide rate

HomicideRate	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
GiniIndex	.898	.072	12.43	0	.756	1.041	***
GNIperCapita	0	0	-1.14	.255	0	0	
GDPGrowth	.267	.129	2.08	.039	.014	.521	**
AverageSchooling	-.885	.333	-2.65	.009	-1.542	-.227	***
Urbanization	21.512	7.869	2.73	.007	5.991	37.033	***
Constant	-31.997	6.933	-4.61	0	-45.672	-18.322	***
Mean dependent var		8.365	SD dependent var			12.169	
R-squared		0.813	Number of obs			198	
F-test		166.403	Prob > F			0.000	
Akaike crit. (AIC)		1231.014	Bayesian crit. (BIC)			1250.744	

*** *p<.01,* ** *p<.05,* * *p<.1*

In panel mode, it's evident that a substantial portion, approximately 81%, of the variations in the dependent variable (murder rate) can be accounted for by the independent variables. Moreover, the Gini coefficient exhibits significance and a positive correlation. In practical terms, this means that for every one-unit rise in the Gini coefficient, assuming other factors remain constant, the number of murders per 100,000 people increases by approximately 0.9.

Table 16- Panel regression on the robbery rate

RobberyRate	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
GiniIndex	6.979	1.584	4.41	0	3.855	10.103	***
GNIperCapita	0	.001	-0.22	.826	-.001	.001	
GDPGrowth	-3.46	2.819	-1.23	.221	-9.021	2.1	
AverageSchooling	-30.176	7.303	-4.13	0	-44.58	-15.772	***
Urbanization	767.986	172.448	4.45	0	427.85	1108.121	***
Constant	-379.459	151.936	-2.50	.013	-679.138	-79.781	**
Mean dependent var		154.451	SD dependent var			180.794	
R-squared		0.592	Number of obs			198	
F-test		55.730	Prob > F			0.000	
Akaike crit. (AIC)		2453.521	Bayesian crit. (BIC)			2473.251	

*** *p<.01,* ** *p<.05,* * *p<.1*

The same holds true for theft rates! When examining the panel data, we observe that a substantial portion (60%) of the variations in the dependent variable (murder rate) can be accounted for by the independent variables. Additionally, the Gini coefficient shows statistical significance and a positive relationship. In practical terms, this means that for every one-unit increase in the Gini coefficient, holding all other factors constant, the number of thefts per 100,000 people increases by approximately 7 units.

* * *

5 Conclusion

In this research, our objective was to shed light on the intricate relationship between income inequality and crime levels, a long-debated issue among economists, politicians, and society at large. We aimed to explore the factors influencing this relationship.

By meticulously analyzing data gathered from various sources, we conducted individual country-level assessments. Specifically, we focused on 11 diverse countries spanning different regions of the world over the period from 2003 to 2020. Our goal was to elucidate how the Gini coefficient, a measure of income inequality, impacts crime rates.

The outcomes of our country-specific analyses occasionally diverged from our initial expectations and, in some cases, contradicted them. However, our comprehensive panel investigation revealed a robust and statistically significant correlation between the Gini coefficient (income inequality) and crime rates. Essentially, our findings provided substantial support for the ideas introduced in the research's introduction, particularly those related to Mr. Bakr's previous work on this topic.

* * *

6 Refrences

[1] P. Fajnzylber, D. Lederman, and N. Loayza, "Inequality and violent crime," *The journal of Law and Economics,* vol. 45, no. 1, pp. 1-39, 2002.

[2] M. D. Krohn, "Inequality, unemployment and crime: A cross-national analysis," *The Sociological Quarterly,* vol. 17, no. 3, pp. 303-313, 1976.

[3] L. G. Scorzafave and M. K. Soares, "Income inequality and pecuniary crimes," *Economics Letters,* vol. 104, no. 1, pp. 40-42, 2009.

[4] S. Harrendorf, M. Heiskanen, and S. Malby, *International statistics on crime and justice*. European Institute for Crime Prevention and Control, affiliated with the …, 2010.

[5] N. B. Adeleye, "The determinants of income inequality and the relationship to crime," *Yayınlanmış Yüksek Lisans Tezi,* 2014.

[6] T. Economist. "The stark relationship between income inequality and crime." The Economist Group Limited. https://www.economist.com/graphic-detail/2018/06/07/the-stark-relationship-between-income-inequality-and-crime

[7] M. Szalavitz, "Income inequality and homicide," *Scientific American,* vol. 319, no. 5, pp. 9-9, 2018.

[8] L. Zhang, "Violent Crime and Income Inequlaity," 2018.