The Effects of Inequality on Crime: A Cross-Sectional Analysis

Lucas P. G. de Barros
São Paulo School of Economics, FGV

Luiz Felipe de Vries
São Paulo School of Economics, FGV

Brief summary (in English)

The purpose of this paper is to discuss the impact of income inequality on crime at a municipal level in the state of São Paulo, Brazil. We use political alignment between state and city governments as an instrument for law enforcement, under the assumption that political interest could lead to the exogenous allocation of policing. This, combined with the use of a lagged dependent variable helps us soften the simultaneity bias associated with the joint determination of crime and law enforcement; which ultimately enables us to quantify the impact of inequality on crime.

Brief summary (in Portuguese)

O propósito deste artigo é discutir o impacto de desigualdade de renda sob crime em uma escala municipal no estado de São Paulo, Brasil. Usamos alinhamento político entre os governos municipais e estaduais como instrumento para policiamento, sob a hipótese de que interesse político poderia causar uma alocação exógena de policiamento. Isso combinado com o uso de um *lag* da variável dependente nos ajudou a amenizar o viés de simultaneidade associado à determinação conjunta de crime e policiamento; permitindo-nos quantificar o impacto de desigualdade sob crime.

***Keywords (in English)*): ** inequality, crime, instrument, PSDB, public policy

Keywords (in Portuguese): desigualdade, crime, instrumento, PSDB, política pública

ANPEC Area: *12) Economia Social e Demografia Econômica*

JEL CODE: D04, D63, H72

The Effects of Inequality on Crime: A Cross-Sectional Analysis

Lucas P. G. de Barros
São Paulo School of Economics, FGV

Luiz Felipe de Vries
São Paulo School of Economics, FGV

Abstract:

The purpose of this paper is to discuss the impact of income inequality on crime at a municipal level in the state of São Paulo, Brazil. Motivated by Becker's model as well as the abundant empirical work done on the determinants of crime, we propose a model which relates inequality and criminal behavior. Initially, we suggest that policing could impact inequality which should motivate its inclusion in our model explaining crime. This leads to a problem of jointly determined variables, the solution of which has been the main focus of criminal literature, and to which we propose the use of an instrumental variable, namely, the political alignment between local and state government (proxied by the city's mayor being from the same party as the governor's). Using a two stage least square approach and testing the robustness of our results for several subsamples in which one or more assumptions would most likely hold we found a consistently small, yet significant impact of inequality on crime.

1. Introduction

The theoretical background of most of the work done on crime, from an economic perspective, stems directly or indirectly from Becker's (1968) seminal paper. In it, the author postulates an economic model in which the supply of crime depends on penalties, potential gains and the opportunity cost associated with criminal behavior. In equilibrium, he suggests that an increase in law enforcement would lead to more probable penalties and, in terms, to a decrease in the supply of crime, in other words, one should expect a negative correlation between policing and crime. Later, Bourguignon (1999) further explores the model to account for the impact of income-inequality on crime (which will be later discussed) in an attempt to draw a connection between the unequal distribution of resources and economic development, using crime as the cause for social unrest and subsequent depressed economic growth. The author suggests that there should exist a positive correlation between income inequality and crime, particularly, crimes mainly motivated by financial gains such as robbery and theft.

From an econometric perspective, the joint determination of crime and law enforcement can create an endogeneity problem. In fact, most of the literature on the subject has attempted to solve this problem through various mechanisms. Corman and Mocan (2000) use monthly data under the assumption that hiring and training police officer takes time, which delays the government's response and thus crime in period $t+1$ is deterred by the police force decided based on the level of crimes in t, which softens simultaneity bias. Levitt (1997) uses an instrumental variable, arguing that in election periods policing increases arbitrarily, thus crime deterrence could be decomposed in an endogenous and an exogenous element, to which the author applied a two-stage least squared (2SLS) approach. McCrary (2002) later found a computational error in Levitt's work, to which the latter conceded in his reply in 2002 (see Levitt (2002)). In probably one of the most important papers to handle this endogeneity issue, Di Tella and Schargrodsky (2004) observe an increase in law enforcement in Buenos Aires after a terrorist attack, namely, a natural experiment of sorts, assumed to be exogenous from the level of other forms of crime, allowing the authors to successfully isolate the effect of policing on crime.

Lastly, we turn our attention to a paper more closely related to ours. Fajnzylber, Lederman and Loayza (2002) investigate the link between income inequality and crime across several countries. The authors resort to a theoretical model suggested in Glaeser, Sacerdote and Scheinkmann (1996), later discussed, which attributes inertial properties to crime and combine it with theoretical work of Arellano and Bond (1991) and Arellano and Bover (1995) on the use of appropriately lagged dependent variable as a mean to counter endogeneity caused by the joint determination of crime and policing.

In this paper, we will focus on a combination of techniques, particularly, we will use both a lagged dependent variable and an instrument. With data set from Brazil that has information on several city characteristic our two stage least square approach yielded a small, but significant, positive relation of inequality on crime, yet, the size of the impact might not be sufficient to justify public policy regarding inequality aiming to reduce crime on the short-run.

2. Theoretical Framework

Bourguignon (1999) explores an interesting and simple model, inspired by Becker (1968), to show that income inequality impacts crime rates positively. We shall briefly discuss a simple version thereof.

An agent chooses to commit a crime if the expected payoff of doing so is larger than the alternative. To analyze this choice, we consider a city of poor and rich agents. Rich agents initially have income w_r and poor agents, w_p. Consider also a probability p of getting caught if committing a crime, that crime pays a fraction of the victim's income xw_j if the offender is not caught, and a linear function of the agent's income, $-fw_i$ otherwise; and that the utility function of money is $u(.) = ln(.)$.

In that scenario, an agent k of type i commits crime against victims of type j if, and only if,

$$(1-p)\, ln\left(w_i + xw_j\right) + p(ln(w_i(1-f))) > ln(w_i) + h_k + \xi_k,$$

where h_k is a personal degree of honesty, and ξ_k accounts for any other unobservable conditions that affect the individual's choice to commit a crime.

$$\leftrightarrow (1-p)\, ln\left(w_i + xw_j\right) > (1-p)\, ln(w_i) + h_k + \xi_k\ - p\, ln(1-f)$$

$$\leftrightarrow \psi\left(p, w_j, w_i, x, f\right) := (1-p)\, ln\left(1 + \frac{xw_j}{w_i}\right) +\ p\ (ln(1-f)) > h_k + \xi_k,$$

This shows that for fixed levels of h_k and ξ_k, one chooses crime only when the ratio $\frac{w_j}{w_i}$ is large enough. In our model h_k is random variable with a distribution H that captures the different levels of honesty of individuals within a same city, whereas ξ_k , also assumed to be a random variable, captures other intrinsic city characteristics that could impact crime and are uniformly distributed among the citizens of a particular city and yet vary across municipalities. In other words, $h_k + \xi_k$ simultaneously captures unobservable city and individual characteristics, relevant to crime rates. Thus, a city with, ceteris paribus, larger inequality, and therefore, largerψ, should have more people k such that the unobservable characteristic from both city and individual is smaller than ψ, increasing crime rates. The purpose of this paper is to test whether such an effect exists.

However, other than different $\frac{w_j}{w_i}$ ratio, the model allows for variation of other variables, impacting crime rates directly. As p increases due to more or more efficient law enforcement, for example, we can see from the partial derivative:

$$\frac{\partial\psi}{\partial p} = -ln\left(1+\frac{xw_j}{w_i}\right) + ln(1-f) = ln\left(\frac{1-f}{1+\frac{xw_j}{w_i}}\right) < 0$$

that ψ is decreasing, which reduces the probability of one choosing crime, corroborating Becker (1968), Corman and Mocan (2000) and Di Tella and Schargrodsky (2004), all of which assumed that law enforcement curtails crime rates.

Note that f and x could also vary, but restricted to differences in penalties and types of crime. However, in Brazil, the federal government is the sole legislator over criminal laws, and our work concerns only burglary, not allowing for such kinds of variation.

Other than the parameters in ψ, we allow for variation of the distribution of $h_k + \xi_k$ among cities. As the inhabitants of a city are distributed differently in regards to their levels of honesty, ceteris paribus, the mass of probability of H corresponding to small values becomes different, making it either more or less probable for ψ to be larger than $h_k + \xi_k$, increasing or reducing crime rates. The idea to allow for varying degrees of honesty was introduced by Glaeser et al. (1996). In the paper, the authors see potential offenders in mainly three groups, independent from wealth; one group that rationally decides whether or not to commit a crime based on all the parameters; one consisting of extreme law-abiders (with acutely high levels of honesty) and, another, of extreme lawbreakers (with acutely low levels of honesty), both to which the choice of committing a crime is more biased by individual tendencies rather than by environmental conditions. Thus, capturing the effect of $h_k + \xi_k$ on crime rates are nearly impossible, since we cannot control for the many of the unobservable variables which shape its distribution. Further, we assume that most of these characteristics are relatively persistent over time, which causes the inertial behavior of crime, previously mentioned. Thus, this motivates us to include a lagged crime rate as an independent variable in our model that tries to capture the effect of $h_k + \xi_k$, since these unobservable variables impact crime rate in t as well as in $t-1$ rather homogenously.

Lastly, not considered in the model for sake of its simplicity, it is also important to account for factors such as the facility of getting jobs, reflected by both the urbanization and employment rate, as supported by Fajnzylber et al. (2002). Education and the overall wealth are also important factors, the first for its likely impact on h_k and the latter due to the concavity of $u(.) = ln(.)$ which implies that for larger levels of wealth, the marginal gain of criminal activity is smaller, hindering such behavior.

3. Data

3.1 Variables and Sources

Our model obviously is concerned with the interactions between more abstract variables such as wealth, inequality and crime. In this section we explore variables which we believe to be good enough proxies, more precisely, good measurements of these variables.

We begin with inequality, our main variable. Here, the GINI index is consistently used across literature as the best measurement of inequality, we obtained it a municipal level from IBGE ("*Instituto Brasileiro de Geografia e Estatística*"). It is an index ranging from 0 to 1 that measures how unequally distributed incomes are. Values closer to 0 indicate more equally distributed, and values closer to 1, more unequally distributed income among inhabitants. Our proxy for political alignment, later used as an instrumental variable, is a dummy variable that indicates whether the mayor of each city was of the political party PSDB in 2010, assumes value 1 if so, and 0 otherwise. This variable was constructed from the database from TSE ("*Tribunal Superior Eleitoral*") and its reasoning will be better explained in section 4.2.

For our measurement of crime, we chose the number of robberies per 100,000 inhabitants, since these types of offenses are the ones more likely to be motivated by income inequality as suggested by our theoretical framework. For both years (2009 and 2010) we use data from SEADE ("*Fundação Sistema Estadual de Análise de Dados*") through IMP ("*Informação dos Municípios Paulistas*"). Our measurement for policing was perhaps the most difficult one to define, since it ought to capture both the level of law enforcement as well as its quality. Had we chosen the number of police officers, for instance, we could miss out on quality, since we could have a corrupt or compliant body of law enforcement that did not actively combat crime. We decided to go for the number of police inquires which implies that an active investigation was opened and therefore captures the size of policing and its active deterrence of crime. This data was extracted from from SSP ("*Secretaria de Segurança Pública*").

For our control variables, we use the natural logarithm of the GDP per capita, the percentage of urbanization of the city, as a proxy for the levels of education we use high school dropout rate, namely, the percentage of high school dropouts in relation to students who had not dropped out by the end of the school year and, lastly, as a proxy for the ease of obtaining a job we used the formal employment rate measured by dividing the number of formal jobs by the population of the city. Note, however, that this ratio might be bigger than 1, as the number of jobs might be larger than the population. However, this does not damage the intuition of such measurement. All of these variables were taken from SEADE ("*Fundação Sistema Estadual de Análise de Dados*") through IMP ("*Informação dos Municípios Paulistas*").

Our measurement of wealth, employment and urbanization are pretty straightforward and used in most of the literature on the subject, the effect that wealth is logged surely interconnects with our assumption regarding our representative individual's utility function. With respect to education, we chose high school dropout rate since it is a major determinant of employment in posterior stages of life as well a deterrent of criminal behavior in adolescence, when most underprivileged children begin their ilicit activities.

3.2 Descriptive Analysis

Here, we summarize and present the correlation matrix and summary table of the variables included in regressions throughout the paper. Note, most importantly for the analysis later, that the covariance between the GINI index and police inquiries rates is different from zero and that police inquires and robbery rates are negatively correlated as well as GINI and crime are positively correlated. It is also relevant to turn our attention to the standard deviations of crime rates and GINI index later for our conclusion.

Variables	***Obs***	***Mean***	***Std. Dev.***	***Min***	***Max***
GINI Index	*645*	*0.458*	*0.057*	*0.334*	*0.686*
PSDB Mayor	*645*	*0.318*	*0.466*	*0*	*1*
Police Inquiry Rate 2010 (%)	*645*	*0.009*	*0.003*	*0*	*0.037*
Logged GDP per Capita	*645*	*9.699*	*0.546*	*8.488*	*12.060*
Urbanization Rate (%)	*645*	*84.317*	*14.282*	*24.920*	*100*
High School Dropout Rate (%)	*645*	*4.835*	*3.977*	*0*	*24.4*
Formal Employment Rate (%)	*645*	*0.219*	*0.147*	*0.054*	*1.938*
Robbery Rate 2009 (per 100 000 inh.)	*645*	*187.280*	*244.837*	*0*	*1710.303*
Robbery Rate 2010 (per 100 000 inh.)	*645*	*164.431*	*229.413*	*0*	*1590.643*

Variables	GINI Index	PSDB Mayor	Police Inquiry Rate 2010 (%)	Logged GDP per Capita	Urbanization Rate (%)	High School Dropout Rate (%)	Formal Employment Rate (%)	Robbery Rate 2009	Robbery Rate 2010
GINI Index	*1*								
PSDB Mayor	*-0.0024*	*1*							
Police Inquiry Rate 2010 (%)	*0.0830*	*0.1032*	*1*						
Logged GDP per Capita	*0.1519*	*-0.0746*	*-0.0128*	*1*					
Urbanization Rate (%)	*0.0933*	*-0.0817*	*0.0819*	*0.3400*	*1*				
High School Dropout Rate (%)	*0.0386*	*0.0395*	*0.1598*	*- 0.0462*	*-0.0718*	*1*			
Formal Employment Rate (%)	*0.0785*	*-0.0744*	*-0.0141*	*0.6546*	*0.1988*	*-0.0113*	*1*		
Robbery Rate 2009 (per 100 000 inh.)	*0.3851*	*-0.0505*	*-0.0423*	*0.3716*	*0.3518*	*-0.0240*	*0.1900*	*1*	
Robbery Rate 2010 (per 100 000 inh.)	*0.3891*	*-0.0656*	*-0.0379*	*0.3387*	*0.3511*	*-0.0171*	*0.1596*	*0.9639*	*1*

4. Econometric Model
4.1 Primitive Model and its Limitations

Having analyzed the microeconomic motivation behind the interaction between crime rates, inequality, law enforcement, education, the one-year lagged crime rates, urbanization and employment rates, we propose the following approximate linear model, proxied by the appropriate variables defined in section 3:

$$
\begin{aligned}
robbery\_rate_i^t &= \theta_0 + GINI_i\theta_1 + police\_inquiries\_rate_i^t\theta_2 + robbery\_rate_i^{t-1}\theta_3 \\
&+ GDP\_percapita_i\theta_4 + urbanization\_rate_i\theta_5 + dropout\_rate_i\theta_6 \\
&+ employment\_rate_i\theta_7 + \gamma_i
\end{aligned} \tag{1}
$$

OLS will not be consistent in estimating the above equation since including law enforcement could lead to a problem of simultaneous equations, in which $Cov(X_i, \gamma_i) \neq 0$, if and only if there was the joint determination of crime and policing in period t and in such case, in such case it would ill advised to simply use OLS.

However, one might argue that because security budgets are made on a yearly basis, how much governments both on state and municipal level spend with security is impacted by criminal rates from the past. In such case, the inclusion of a lagged dependent variable would more than suffice to counter any form of endogeneity. This relationship, however, is not entirely rigid on a municipal level, since most policing is done by the Military Police, which is controlled by the state government but allocated across cities with certain flexibility; thus, simply controlling by the one-year lag does not entirely solve our endogeneity problem, since cities that display higher crime rates in year t could see a larger allocation of resources in law enforcement throughout the year. And, equivalently, this could deter criminal behavior. Thus, our endogeneity problem persists if we include policing, even if controlled by the lagged crime rate.

It is important to mention that if we were to simply remove policing as an independent variable, we would most likely produce a consistent OLS estimator for our parameter of interest, namely θ_1, since there would be no joint determination of dependent and independent variables and thus $Cov(X_i, \gamma_i) = 0$ would hold. This would be true if and only if there was also no interaction between policing and inequality, otherwise we would have omitted variable bias and θ_1 would also be inconsistent, since policing would be captured by the error term and if policing was significant to inequality then again, we would have $Cov(X_i, \gamma_i) \neq 0$.

There are several paths through which inequality could interact with policing and although our microeconomic model is limited in that regard, we could argue, for instance, that higher rates of policing in an unequal city would be allocated to poorer boroughs, which are heavily dependent on the informal economy and which would see a drop of economic activity in the light of more ostensive law enforcement, deepening inequality. Equivalently, it is also reasonable to expect that more unequal cities would see political pressure from the richest citizens for more ostensive law enforcement, in terms of our model, we could expect that an individual w_r would be prone to financing a politician willing to increase law enforcement (and, thus, p) to discourage individual w_p from committing a crime against w_r. Although it is not the purpose of this paper to analyze such relationship (and there is bountiful literature on the subject of Conflict Theory regarding policing and inequality), we believe there to be one, which we will empirically test.

In such case, by including policing and acknowledging that it is somewhat codetermined with our dependent variable, we are forced to reject the hypothesis that $Cov(X_i, \gamma_i) = 0$ and we will proceed under the assumption that our primitive model is endogenous and OLS is inconsistent.

4.2 The use of an Instrumental Variable

To try to bypass this endogeneity problem, we will resort to an instrumental variable. The political party PSDB has successfully retained control of the state government of São Paulo for more than two decades; having elected the last 4 out of 5 governors since 1995 and being uninterruptedly in power for the past decade. Thus, the alignment between municipal power with the state's is somewhat important when it comes to the allocation of resources, such as law enforcement, especially, the Military Police. To overcome our endogeneity problem, we will use a dummy variable which assumes value 1 when the mayor of a city is also from PSDB and 0 otherwise.

Thus, we initially assume that a mayor being from PSDB is correlated with the level of policing in that city, measured by the number of police inquiries. However, it is also reasonable to believe that the state government allocates resources strategically, thus, a city politically fragmented or with a population insignificant for state elections could systematically receive less support even if the elected mayor is from the same party, which implies that the effect of PSDB on policing should be stronger in cities that do not present the previous conditions.

Moreover, our main identification assumption is that having a PSDB mayor is uncorrelated with any omitted variable that could impact crime rates. This is, of course, a bold and untestable assumption which could be easily undermined if, for instance, a city's ideology towards PSDB was accompanied by bigger investments in infrastructure and education by the state government that could raise incomes and mitigate crime through various channels. Although we cannot test such an assumption, we will test the robustness of our results by later restricting our sample to cities which we believe will simultaneously reduce the effect of any unobservable variable correlated with our instrument as well as where our instrument will be strongest (to compensate for the reduced numbers of observations), in an attempt to somehow curtail the effects of cities with extreme PSDB-leaning tendencies or extreme PSDB-aversion that could potentially violate our main identification assumption. However, it is noteworthy that the effect of any possibly unobservable variable is mitigated by the use of a lagged dependent variable, since the impact of being PSDB-leaning, for instance, is captured by crime rates in the previous year as well; thus, provided that there is no acute variation in PSDB support between 2009 and 2010 (a more reasonable assumption) the use of a lag should suffice to ensure our identification assumption.

5. Result Analysis

5.1 Main Findings

In this section, we will analyze our empirical results. We begin by trying to capture the relationship between policing and inequality mentioned in section 4.1. The four regressions below, columns 1 through 4, regress the GINI Index against policing and several subsets of our control variables. In all of them policing positively impacts inequality to a significance level of 5% and in three quarters of the regressions to a 1% level of significance. Even though we do not believe this model to be a good enough description of the causes of inequality (this was not, after all, the purpose of this paper), we believe that it provides sufficient evidence that there is a strong enough relationship between inequality and policing such that by omitting the latter from equation (1) would lead to omitted variable bias. Since law enforcement seems to be positively correlated with inequality and theory suggests that policing negatively impacts crime, by omitting policing we would most likely incur in a negative asymptotic bias, thus, underestimating the possibly positive impact of inequality on crime.

VARIABLES	(1) GINI Index	(2) GINI Index	(3) GINI Index	(4) GINI Index
Police Inquiry Rate 2010 (%)	***1.555****	***1.532****	***1.161***	***1.279***
	(0.565)	*(0.551)*	*(0.555)*	*(0.535)*
Robbery Rate 2009	***9.41e-05****	***9.11e-05****		
	(1.06e-05)	*(9.15e-06)*		
Logged GDP per capita	***0.00291***		***0.0172****	
	(0.00552)		*(0.00562)*	
Urbanization Rate (%)	***-0.000255***		***0.000165***	
	(0.000171)		*(0.000178)*	
High School Dropout Rate(%)	***0.000415***		***0.000528***	
	(0.000503)		*(0.000535)*	
Formal Employment Rate(%)	***-0.000665***		***-0.0137***	
	(0.0194)		*(0.0218)*	
Constant	***0.418****	***0.427****	***0.268****	***0.447****
	(0.0504)	*(0.00596)*	*(0.0508)*	*(0.00558)*
Observations	*645*	*645*	*645*	*645*
F-Statistic	*18.18****	*51.51****	*4.24****	*5.71****
R-squared	*0.163*	*0.158*	*0.034*	*0.007*

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Having included policing), we are now faced with the endogeneity problem previously mentioned. We proceed by estimating our two stage least square in the full sample and our main model given in equation (1) while omitting policing.

VARIABLES	(5) First Stage Police Inquiry Rate 2010 (%)	(6) Second Stage Robbery Rate 2010	(7) Main Regression (OLS) Robbery Rate 2010
GINI Index	***0.00732****	***172.0***	***90.37***
	(0.00261)	*(77.53)*	*(42.16)*
PSDB Mayor	***0.000805***		
	(0.000322)		
Robbery Rate 2009	***-1.82e-06****	***0.874****	***0.895****
	(6.19e-07)	*(0.0228)*	*(0.0164)*
Urbanization Rate (%)	***3.62e-05****	***0.866****	***-7.536***
	(1.08e-05)	*(0.306)*	*-6.549*
High School Dropout Rate (%)	***0.000148****	***2.008***	***0.485****
	(4.26e-05)	*(1.465)*	*(0.173)*
Formal Employment Rate (%)	***-0.000302***	***-32.29***	***0.350***
	(0.00136)	*(25.46)*	*(0.688)*
Predicted Police Inquiry Rate 2010 (%)		***-10996.8***	
		(7,916)	
Logged GDP per capita	***-6.83e-05***	***-8.406***	***-27.59***
	(0.000420)	*(7.451)*	*(25.44)*
Constant	***0.00292***	***29.17***	***-8.000***
	(0.00396)	*(73.16)*	*(59.14)*
Observations	*645*	*645*	*645*
F-Statistic	*5.41****		*710.85****
Wald chi-squared Statistic		*3622.48****	
R-squared	*0.062*	*0.902*	*0.931*

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

We immediately turn our attention to column 5, in which one can observe that our instrument, namely PSDB, is significant in explaining law enforcement to a 5% level of significance, as we had anticipated in section 4.2. Not only that, but our exogenous variables and our instrument have jointly explained policing so that the p-value of our F-test is kept well below the 1% level of significance. However, one must have in mind that our first stage regression has a low R^2 of merely 6%, which leads to a gain in variance in our second stage, hindering some of our inference work.

Turning our attention to column 6, one notices that the effect of law enforcement on crime is not entirely significant, but the estimator is distributed mostly over negative values, with a p-value of 16.5%, at least weakly corroborating our initial assumption. The purpose of using an instrumental variable was not to fully estimate the effect of policing on crime, but rather to break some of the endogeneity of jointly determined variables so that we could produce a consistent estimator for the effect of inequality on crime. For this effect, we believe that our instrument has served quite well, nearly doubling the effect of inequality on crime when compared to our model without law enforcement given in column 7, thus, corroborating our conjecture that omitting policing would lead to the underestimation of our coefficient of interest.

It is also important to mention that some of our control variables, such as the rates of formal employment and high school dropouts as well as the logged GDP per capita, are not much relevant statistically in columns 5 through 7, this is in part due to the use of our lagged dependent variable, which already captures some of those variables effects; but then again, the purpose of these controls was to help better isolate the magnitude of the effect of inequality on crime and this has been achieved.

Additionally, we are unable to reject the null hypothesis to a 5% level of confidence that our model is exogenous via a Hausman Test, however, since theory suggests that the model is simultaneously determined, we understand that by not rejecting this, we might be committing a type II error, which leads us to prefer the use of an instrumental variable anyhow.

5.2 Robustness Test

Lastly, the only assumption we are not yet satisfied with is the identification hypothesis regarding our instrument, namely, that the covariance between PSDB and our error term is equal to 0. Here, the best we can do is to attenuate the possible effect of a city being ideologically PSDB-leaning or extremely PSDB-averse by removing cities which were won on a land-slide. This obviously considerably reduces our sample, which would make our instrument particularly weak in the first stage, hindering any possible inference work in our second stage. To mitigate that, we also remove cities that have been won with very few votes, which implies that they are small cities without a second round (under 200.000 inhabitants) and politically fragmented in the sense that are multiple candidates; making these cities less likely to receive support from the state government as suggested in section 4.2. Thus, we turn our attention to two subsamples: cities where the winning candidate had between 40% and 60% of the valid votes (sum of every vote casted to every candidate) and between 38.16% and 72.56%, namely the average (55.36%) plus and minus a standard deviation (17.20%).

VARIABLES	(8) Between 40% - 60% Police Inquiry Rate 2010 (%)	(9) Between Average +/- SD Police Inquiry Rate 2010 (%)
GINI Index	***0.00856***	***0.00755***
	(0.00367)	*(0.00316)*
PSDB	***0.00123****	***0.000790***
	(0.000455)	*(0.000376)*
Robbery Rate 2009 (per 100,000 inh.)	***-3.64e-07***	***-8.53e-07***
	(1.35e-06)	*(1.13e-06)*
Logged GDP per capita	***-0.000427***	***-0.000356***
	(0.000693)	*(0.000562)*
Urbanization Rate (%)	***2.55e-05****	***3.02e-05***
	(1.36e-05)	*(1.28e-05)*
High School Dropout Rate (%)	***0.000186****	***0.000148****
	(5.65e-05)	*(5.04e-05)*
Employment Rate (%)	***0.00219***	***0.000894***
	(0.00197)	*(0.00160)*
Constant	***0.00580***	***0.00577***
	(0.00656)	*(0.00534)*
Observations	*350*	*473*
F-Statistic	*3.78****	*3.33***
R-squared	*0.088*	*0.056*

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Columns 8 and 9 are the first stages of our two stage least square approach in both of our subsamples. One can clearly see that our instrument remains significant in both cut-outs, yet in the smaller and more restrictive sample, the instrument's p-value and the regressions F-statistic are both kept below the 1% threshold instead of the 5% in column 9. Also, the R^2 of regression in column 8 is also bigger, 8.8% as opposed to 5.6% from regression in 9, which will facilitate some of the inference work in the second stage. Columns 10 and 11, below, are OLS estimations in our subsamples excluding policing (as done in the regression in column 7), while columns 12 and 13 are our second stage estimates for our samples.

VARIABLES	(10) Between 40% - 60% (OLS) Robbery Rate 2010 (per 100,000 inh.)	(11) Between Average +/- SD (OLS) Robbery Rate 2010 (per 100,000 inh.)	(12) Between 40% - 60% (IV) Robbery Rate 2010 (per 100,000 inh.)	(13) Between Average +/- SD (IV) Robbery Rate 2010 (per 100,000 inh.)
Predicted Police Inquiry Rate 2010 (%)			***-9,116***	***-10,245***
			(6,635)	*(9,115)*
GINI Index	***67.43***	***66.87***	***147.6****	***144.9***
	(61.84)	*(49.06)*	*(89.61)*	*(89.49)*
Robbery Rate 2009 (per 100,000 inh.)	***0.844*****	***0.851*****	***0.842*****	***0.843*****
	(0.0287)	*(0.0268)*	*(0.0307)*	*(0.0306)*
Logged GDP per capita	***3.087***	***-1.832***	***-1.247***	***-5.620***
	-8.170	*-7.366*	*(8611)*	*(8501)*
Urbanization Rate (%)	***0.416***	***0.468*****	***0.638*****	***0.771*****
	(0.256)	*(0.213)*	*(0.273)*	*(0.311)*
High School Dropout Rate (%)	***0.158***	***0.106***	***1.860***	***1.626***
	(0.926)	*(0.787)*	*(1538)*	*(1578)*
Formal Employment Rate (%)	***-24.06***	***-16.66***	***-5.717***	***-8.549***
	(29.84)	*(25.14)*	*(24.39)*	*(21.24)*
Constant	***-85.63***	***-46.56***	***-25.02***	***16.79***
	(73.98)	*(67.11)*	*(88.29)*	*(93.35)*
Observations	*350*	*473*	*350*	*473*
F-Statistic	*181.59****	*212.74****		
Wald chi-squared Statistic			*1077.54****	*1110.11****
R-squared	*0.888*	*0.882*	*0.857*	*0.835*

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Here there are some interesting points to be made. First, we turn our attention to column 12, where the GINI index continues significant to a 10% level of significance even in a considerably smaller sample and with a coefficient of 147.6 is very close to our estimate of 172 for the full sample. The weaker instrument in the regression in column 9 has caused a loss of significance in column 13, albeit very small, since the p-value was 10.5%. In terms of magnitude, however, our coefficient continues sufficiently close to our full-sample estimate. In columns 10 and 11 we have repeated our exercise of removing policing and running our regression with OLS.

When this was done for the full sample, the effect of inequality went from 90.37 to 172, nearly doubling; restricting the sample we observed a staggering similar effect in terms of magnitude, from columns 10 to 12 the coefficient grew from 67.43 to 147.6 whereas from columns 11 to 13 it grew from 66.87 to 144.9 (nearly doubling in both cases).

Bottom line is that the exercise of restricting our sample to attenuate the endogenous effect caused in the event of our identification hypothesis failing produced similar results to that observed in the full-sample, admittedly with a loss of significance. This makes us more confident in using our instrument and somehow corroborates our argument that the use a lagged depended variable suffices to control for any unobservable variable that could cause our instrument to be endogenous.

6. Conclusion

Initially we set out to isolate the effect of inequality on crime, but since we could not disregard the interaction of policing and crime, we were forced to include law enforcement in our model, causing a problem of jointly determined variables in our linear regression. To solve such problem, we applied an instrumental variable to cleanse some of the endogenous effect caused by the simultaneity in law enforcement and crime in an attempt to isolate the effect of inequality on crime and to make OLS loosely more consistent. Here, our identification hypothesis could be undermined by a set of unobservable variables such as ideology, however, we argue that the use of our lagged depended variable serves to capture some of those variables, in addition to also capturing the inertial properties of crime as suggested in Glaeser et al (1996) and helping to mitigate our endogeneity problem. To test the robustness of our results, we took subsamples in which we believe the effect of any unobservable variable correlated with PSDB would be less significant and our identification hypothesis could be undermined, and we observed consistency in our method. Yet, our instrument has not been sufficiently strong to compellingly show that policing is negatively correlated with crime, but we believe it was effective enough to better isolate the effect of inequality on crime, our initial goal.

In terms of our results, we were able to show that inequality positively impacts robbery rates, as the literature on the subject anticipated. In terms of magnitude, a coefficient of 172 for the GINI index implies that a 0.01 variation of the index (contained within 0 and 1) would cause an increase of 1.72 robberies per 100 000 inhabitants. Put differently, a city A more unequal by one standard deviation, 0.05, from a city B should expect an additional 8.6 robberies per 100 000 people or approximately 1/30 standard deviations of the average robbery rate. Thus, although the effect of inequality exists and is statistically significant, it is not the major determinant of crime independently; meaning that the effect only of inequality, fixing all other explanatory variables, is not large; which could imply that public policy aiming to reduce criminality should look elsewhere for an efficient solution on the short-run. However, a more equal society could observe larger levels of education and stronger institutions in the long-run both of which could translate into better economic output and employment rates, subsequently mitigating crime through various channels.

References

ARELLANO, Manuel and BOND, Stephen. ***"Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations".*** *The Review of Economic Studies, 1991*

ARELLANO, Manuel and BOVER, Olympia. ***"Another Look at the Instrumental Variable Estimation of Error-Component Models".*** *Journal of Econometrics, 1995*

BECKER, Gary S. ***"Crime and Punishment: An Economic Approach".*** *Journal of Political Economy, 1968*

BOURGUIGNON, François. ***"Crime, Violence and Inequitable development".*** *Paper prepared for the Annual World Bank Conference on Development Economics, Washington, D.C., April 28-30, 1999*

CORMAN, Hope and MOCAN, H. Naci. ***"A Time-Series Analysis of Crime, Deterrence, and Drug Abuse in New York City".*** *American Economic Review, 2000*

DI TELLA, Rafael and SCHARGRODSKY, Ernesto. ***"Do Police Reduce Crime? Estimates Using the Allocation of Police Forces After a Terrorist Attack"****. American Economic Review, 2004*

FAJNZYLBER, Pablo, LEDERMAN, Daniel and LOAYZA, Norman. ***"Inequality and Violent Crime".*** *Journal of Law and Economics, 2002*

GLAESER, Edward L., SACERDOTE Bruce and SCHEINKMAN, José A. ***"Crime and Social Interaction".*** *The Quarterly Journal of Economics, 1996*

LEVITT, Steven D. ***"Using Electoral Cycles In Police Hiring To Estimate The Effects Of Police On Crime".*** *American Economic Review, 1997*

LEVITT, Steven D. ***"Using Electoral Cycles In Police Hiring To Estimate The Effects Of Police On Crime: Reply".*** *American Economic Review, 2002*

McCRARY, Justin. ***"Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime: Comment".*** *The American Economic Review, 2002*

"Dados Estatísticos do Estado de São Paulo". *Secretaria de Segurança Pública, http://www.ssp.sp.gov.br/Estatistica/Pesquisa.aspx. Accessed 12 may 2017.*

"Índice de GINI da Renda Domiciliar per capita". *Instituto Brasileiro de Geografia e Estatística, http://tabnet.datasus.gov.br/cgi/ibge/censo/cnv/ginibr.def. Accessed 7 may 2017.*

"Informação Sobre Municípios Paulistas". *Fundação Sistema Estadual de Análise de Dados, http://www.imp.seade.gov.br/frontend/#/tabelas. Accessed 3 may 2017.*