INEQUALITY AND VIOLENT CRIME*

PABLO FAJNZYLBER,
University of Minas Gerais

DANIEL LEDERMAN,
World Bank

and

NORMAN LOAYZA
World Bank

Abstract

We investigate the robustness and causality of the link between income inequality and violent crime across countries. First, we study the correlation between the Gini index and homicide and robbery rates within and between countries. Second, we examine the partial correlation by considering other crime determinants. Third, we control for the endogeneity of inequality by isolating its exogenous impact on these crime rates. Fourth, we control for measurement error in crime rates by modeling it as both unobserved country effects and random noise. Finally, we examine the robustness of this partial correlation to alternative measures of inequality. The panel data consist of nonoverlapping 5-year averages for 39 countries during 1965–95 for homicides and 37 countries during 1970–94 for robberies. Crime rates and inequality are positively correlated within countries and, particularly, between countries, and this correlation reflects causation from inequality to crime rates, even after controlling for other crime determinants.

I. Introduction

The relationship between income inequality and the incidence of crime has been an important subject of study since the early stages of the economics literature on crime. According to Gary Becker's analytical framework, crime rates depend on the risks and penalties associated with apprehension and also on the difference between the potential gains from crime and the associated

* We are grateful for comments and suggestions from François Bourguignon, Dante Contreras, Francisco Ferreira, Edward Glaeser, Sam Peltzman, Debraj Ray, Luis Servén, and an anonymous referee. Norman Loayza worked at the research group of the Central Bank of Chile during the preparation of the paper. This study was sponsored by the Latin American Regional Studies Program of the World Bank. The opinions and conclusions expressed here are those of the authors and do not necessarily represent the views of the institutions with which they are affiliated.

[*Journal of Law and Economics,* vol. XLV (April 2002)]
© 2002 by The University of Chicago. All rights reserved. 0022-2186/2002/4501-0001$01.50

opportunity cost.[1] These net gains have been represented theoretically by the wealth differences between the rich and poor, as shown by the work of François Bourguignon,[2] or by the income differences among complex heterogeneous agents, as shown by the work of Ayse Imrohoroglu, Antonio Merlo, and Peter Rupert.[3] Similarly, in their empirical work, Belton Fleisher,[4] Isaac Ehrlich,[5] and, more recently, Morgan Kelly[6] have interpreted measures of income inequality as indicators of the distance between the gains from crime and its opportunity costs.

The relationship between inequality and crime has also been the subject of sociological theories on crime. Broadly speaking, these have developed as interpretations of the observation that "with a degree of consistency which is unusual in social sciences, lower-class people, and people living in lower-class areas, have higher official crime rates than other groups."[7] One of the leading sociological paradigms on crime, the theory of "relative deprivation," states that inequality breeds social tensions as the less well-off feel dispossessed when compared with wealthier people (see the work by Steven Stack for a critical view).[8] The feeling of disadvantage and unfairness leads the poor to seek compensation and satisfaction by all means, including committing crimes against both poor and rich.

It is difficult to distinguish empirically between the economic and sociological explanations for the observed correlation between inequality and crime. The observation that most crimes are inflicted by the poor on the poor does not necessarily imply that the economic theory is invalid given that the characteristics of victims depend not only on their relative wealth but also on the distribution of security services across communities and social classes. In fact, crime may be more prevalent in poor communities because the distribution of police services by the state favors rich neighborhoods[9,10] or because poor people demand lower levels of security given that it is a normal

[1] Gary S. Becker, Crime and Punishment: An Economic Approach, 76 J. Pol. Econ. 169 (1968).

[2] François Bourguignon, Crime, Violence, and Inequitable Development, in Annual World Bank Conference on Development Economics 1999, at 199 (Boris Pleskovic & Joseph E. Stiglitz eds. 2000).

[3] A. Imrohoroglu, A. Merlo, & P. Rupert, On the Political Economy of Income Redistribution and Crime, 41 Int'l Econ. Rev. 1 (2000).

[4] Belton Fleisher, The Effect of Income on Delinquency, 56 Am. Econ. Rev. 118 (1966).

[5] Isaac Ehrlich, Participation in Illegitimate Activities: A Theoretical and Empirical Investigation, 81 J. Pol. Econ. 521 (1973).

[6] Morgan Kelly, Inequality and Crime, 82 Rev. Econ. Stat. 530 (2000).

[7] John Braithwaite, Inequality, Crime, and Public Policy (1979).

[8] Steven Stack, Income Inequality and Property Crime: A Cross-National Analysis of Relative Deprivation Theory, 22 Criminology 229 (1984).

[9] Jere R. Behrman & Steven G. Craig, The Distribution of Public Services: An Exploration of Local Government Preferences, 77 Am. Econ. Rev. 37 (1987).

[10] Bourguignon, *supra* note 2.

good.[11] Similarly, contrasting or consistent evidence on the effect of inequality on different types of crime cannot be used to conclusively reject one theory in favor of the other. For example, if income inequality leads to higher theft and robbery rates but not to higher homicide rates (as Kelly finds for the United States),[12] the economic model could still be valid given that, first, homicides are also committed for profit-seeking motives and, second, homicide data are more reliable and produce more precise regression estimates than property crime data. By the same token, if income inequality leads to both higher robbery and higher homicide rates (as we find in this cross-country paper), we cannot conclude that the sociological model is incorrect because social deprivation can have both nonpecuniary and pecuniary manifestations. At any rate, the objective of this paper is not to distinguish between various theories of the link between inequality and crime; rather, we attempt to provide a set of stylized facts on this relationship from a cross-country perspective. This initial evidence could then be used in further, more analytically oriented, research to discriminate among competing theories.

As the preceding remarks try to convey, the correlation between income inequality and crime is a topic that has intrigued social scientists from various disciplines. Most economic studies on the determinants of crime rates have used primarily microeconomic-level data and focused mostly on the United States (see the papers by Ann Dryden Witte;[13] Helen Tauchen, Witte, and Harriet Griesinger;[14] Jeffrey Grogger;[15] and Naci Mocan and Daniel Rees[16]). In the 1990s, the interest in cross-country studies awakened, in part because of the appearance of internationally comparable data sets on national income and production,[17] income inequality,[18] and crime rates.[19] In one of these cross-country studies, Pablo Fajnzylber, Daniel Lederman, and Norman Loayza[20] found that income inequality, as measured by the Gini index, is an important factor that drives violent crime rates across countries and over time. Far from

[11] Menno Pradhan & Martin Ravallion, Demand for Public Safety (Working Paper No. 2043, World Bank 1998).

[12] Kelly, *supra* note 6.

[13] Ann Dryden Witte, Estimating the Economic Model of Crime with Individual Data, 94 Q. J. Econ. 57 (1980).

[14] Helen Tauchen, Ann Dryden Witte, & Harriet Griesinger, Criminal Deterrence: Revisiting the Issue with a Birth Cohort, 76 Rev. Econ. Stat. 399 (1994).

[15] Jeffrey Grogger, Market Wages and Youth Crime, 16 J. Lab. Econ. 756 (1998).

[16] H. Naci Mocan & Daniel I. Rees, Economics Conditions, Deterrence and Juvenile Crime: Evidence from Micro Data (1999).

[17] Robert Summers & Alan Heston, The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950–1988, 106 Q. J. Econ. 327 (1991).

[18] Klaus Deininger & Lyn Squire, A New Data Set Measuring Income Inequality, 10 World Bank Econ. Rev. 565 (1996).

[19] United Nations crime surveys and the World Health Organization data sets.

[20] Pablo Fajnzylber, Daniel Lederman, & Norman Loayza, What Causes Violent Crime? 46 Eur. Econ. Rev. 1323 (2002).

settling the issue, this result opened a variety of questions on the plausible interactions between crime rates, measures of income distribution, and other potential determinants of crime. Some of these questions examine the robustness of the crime-inequality link to changes in the sample of countries, the data dimension (time series or cross-country), the method of estimation, the measures of inequality and crime, and the types of control variables. Other questions put in doubt the direct effect of inequality on crime. For instance, Bourguignon argues that "the significance of inequality as a determinant of crime in a cross-section of countries may be due to unobserved factors affecting simultaneously inequality and crime rather than to some causal relationship between these two variables."[21]

In this paper, we investigate the robustness and causality of the link between inequality and crime rates from an empirical cross-country perspective. Figures 1 and 2 plot the simple correlation between the Gini index and, respectively, the homicide and robbery rates in a panel of cross-country and time-series observations. In both cases, the correlation is positive and significant. In what follows, we go behind this correlation to assess issues of robustness and causality. We present the stylized facts starting from the simplest statistical exercises and moving gradually to a dynamic econometric model of the determinants of crime rates. First, we study the correlation between the Gini index and, separately, homicide and robbery rates along different dimensions of the data, namely, between countries, within countries, and pooled cross-country and time-series data. Second, along the same data dimensions, we examine the link between income inequality and homicide and robbery rates when other potential crime determinants are controlled for. These include the level of development (proxied by real gross national product (GNP) per capita), the average years of education of the adult population, the growth rate of the gross domestic product (GDP), and the level of urbanization. We also include the incidence of crime in the previous period as an additional explanatory variable, which makes the crime model dynamic.

Third, we control for the likely joint endogeneity of income inequality in order to isolate its exogenous impact on the two types of crime under consideration. Fourth, we control for the measurement error in crime rates by modeling it as both an unobserved country-specific effect and random noise. We correct for joint endogeneity and measurement error by applying an instrumental variable estimator for panel data. Fifth, using the same panel estimator, we examine the robustness of the inequality-crime link to alternative measures of inequality such as the ratio of the income share of the poorest to the richest quintile, an index of income polarization (calculated

[21] François Bourguignon, Crime as a Social Cost of Poverty and Inequality: A Review Focusing on Developing Countries (unpublished manuscript, World Bank, Dev. Econ. Res. Group 1998).

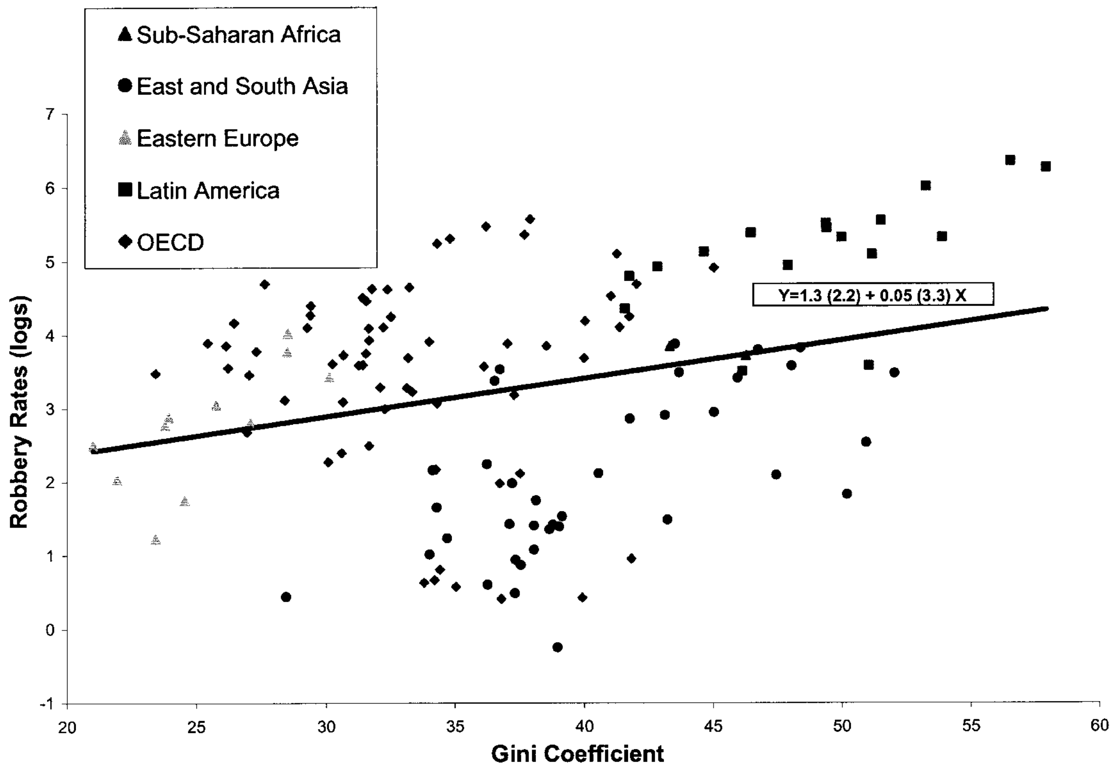

FIGURE 1.—Income distribution and intentional homicide rates, 1965–94 (5-year averages)

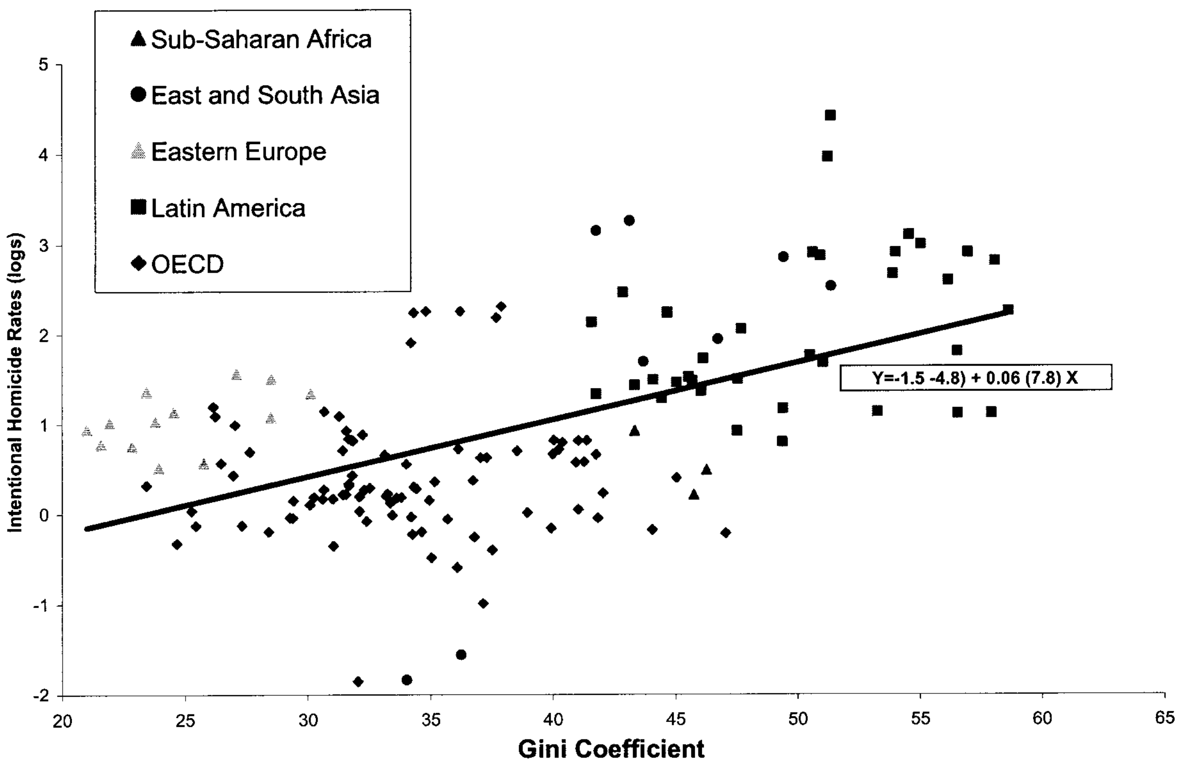

FIGURE 2.—Income distribution and robbery rates, 1970–94 (5-year averages)

following Joan-Maria Esteban and Debraj Ray),[22] and an indicator of educational inequality (taken from José De Gregorio and Jong-Wha Lee).[23] Finally, we test the robustness of this link to the inclusion of additional variables that may be driving both inequality and crime, such as the population's ethnolinguistic fractionalization, the availability of police in the country, a Latin America–specific effect, and the proportion of young males in the national population.

As we said above, this paper adopts a comparative cross-country perspective. Although there are well-known advantages to using microlevel data for crime studies, cross-national comparative research has the following advantage. Using countries as the units of observation to study the link between inequality and crime is arguably appropriate because national borders limit the mobility of potential criminals more than neighborhood, city, or even provincial boundaries do. In this way, every (country) observation contains independently all information on crime rates, inequality measures, and other crime determinants, which thus allows us to avoid the need to account for cross-observation effects.

The main conclusion of this paper is that an increase in income inequality has a significant and robust effect of raising crime rates. In addition, the GDP growth rate has a significant crime-reducing impact. Since the rate of growth and distribution of income jointly determine the rate of poverty reduction, the two aforementioned results imply that the rate of poverty alleviation has a crime-reducing effect. The rest of the paper is organized as follows. Section II presents the data and basic stylized facts. Section III introduces the methodology and presents the results from the generalized method of moments (GMM) estimations, including several robustness checks. Section IV concludes.

II. Data and Stylized Facts

This section reviews the data and presents the basic stylized facts concerning the relationship between violent crime rates and income inequality. Section II*A* presents the sample of observations used in the various econometric exercises in the paper. Sections II*B* and II*C* review the quality and sources of data for the dependent variable (crime rates) and the main explanatory variable (income inequality), respectively. Detailed definitions and sources of all variables used in the paper are presented in Appendix A, Table A1. Section II*D* examines the bivariate correlations between homicide and robbery rates and the Gini coefficient of income inequality. Finally, Section

[22] Joan-Maria Esteban & Debraj Ray, On the Measurement of Polarization, 62 Econometrica 819 (1994).

[23] José De Gregorio & Jong-Wha Lee, Education and Income Distribution: New Evidence from Cross-Country Data (unpublished manuscript, Univ. Chile & Korea Univ. 1999).

II*E* presents ordinary least squares (OLS) estimates of multivariate regression for both types of crime.

A. Sample of Observations

We work with a pooled sample of cross-country and time-series observations. The time-series observations consist of nonoverlapping 5-year averages spanning the period 1965–94 for homicides and 1970–94 for robberies. The pooled sample is unbalanced, with at most six (time-series) periods per country. All countries included in the samples have at least two consecutive 5-year observations. The sample for the homicide regressions contains 20 industrialized countries; 10 countries from Latin America and the Caribbean; four from eastern and central Europe; four from East Asia, South Asia, and the Pacific; and one from Africa. The sample for robberies contains 17 industrialized countries; five countries from Latin America and the Caribbean; four from eastern and central Europe; 10 from East Asia, South Asia, and the Pacific; and one from Africa. Appendix B Tables B1 and B2 show the summary statistics for, respectively, homicide and robbery rates for each country in the sample.

B. National Crime Statistics

We proxy for the incidence of violent crime in a country by its rate of intentional homicide and robbery rates. These rates are taken with respect to the country's population; specifically, they are the number of homicides or robberies per 100,000 people. Cross-country studies of crime face severe data problems. Most official crime data are not comparable across countries because each country suffers from its own degree of underreporting and defines certain crimes in different ways. Underreporting is worse in countries where the police and justice systems are not reliable, where the level of education is low, and perhaps where inequality is high. Country-specific crime classifications, arising from different legal traditions and different cultural perceptions of crime, also hinder cross-country comparisons. The type of crime that suffers the least from underreporting and idiosyncratic classification is homicide. It is also well documented that the incidence of homicide is highly correlated with the incidence of other violent crimes.[24] These reasons make the rate of homicides a good proxy for crime, especially violent crime. To account for likely nonlinearities in the relation between homicide rates and its determinants, we use the homicide rate expressed in natural logs.

The homicide data we use come from the World Health Organization (WHO), which in turn gathers data from national public health records. In the WHO data set, a homicide is defined as a death purposefully inflicted

[24] Pablo Fajnzylber, Daniel Lederman, & Norman Loayza, Crime and Victimization: An Economic Perspective, 1 Economia 219 (2000).

by another person, as determined by an accredited public health official. The other major source of cross-country homicide data is the United Nations' *World Crime Survey*, which reports data from national police and justice records.[25] In this paper, we use the WHO data set because of its larger time coverage for the countries included. Counting with sufficient time coverage is essential for the panel data econometric procedures we implement (see Section III below).

To complement the analysis of the homicide rate, we consider the robbery rate as a second proxy for the incidence of crime. Although data on robberies are less reliable than homicide data for cross-country comparisons, they are likely to be more reliable than data on lesser property crimes such as theft. This is so because robberies are property crimes perpetrated with the use or threat of violence; consequently, their victims have a double incentive to report the crime, namely, the physical and psychological trauma caused by the use of violence and the loss of property. Robbery's close connection with property crimes, to which economic theory is more readily applicable, makes its study a good complement to that of homicide. The robbery data we use come from the United Nations' *World Crime Survey*. The robbery rates are also expressed in natural logs.

C. National Income Inequality Data

Most of the empirical exercises presented below use the Gini coefficient as the proxy for income inequality. In a couple of instances, we also use the ratio of the income share of the poorest to the richest quintile of the population. In addition, we use income quintile shares to construct a measure of income polarization (see Appendix C for details).

Data on the Gini coefficient and the income quintile shares come from the Klaus Deininger and Lyn Squire database.[26] We only use what these authors label "high-quality" data, which they identify through the following three criteria.[27] First, income and expenditure data are obtained only from household or individual surveys. In particular, high-quality Gini index and income quintile shares are not based on estimates generated from national accounts and assumptions about the functional form of the distribution of income taken from other countries. Second, the measures of inequality are derived only from nationally representative surveys. Thus, these data do not suffer from biases stemming from estimates based on subsets of the population in any country. Third, primary income and expenditure data are based on comprehensive coverage of different sources of income and type of ex-

[25] See Pablo Fajnzylber, Daniel Lederman, & Norman Loayza, Determinants of Crime Rates in Latin America and the World (1998), for a description of the United Nations' World Crime Survey statistics.

[26] Deininger & Squire, *supra* note 18.

[27] *Id.* at 568–71.

TABLE 1

PAIRWISE CORRELATIONS BETWEEN THE GINI INDEX AND HOMICIDE AND ROBBERY RATES

	Homicides		Robberies	
	Correlation	N	Correlation	N
Pooled levels	.54 (.00)	148	.28 (.00)	132
Pooled first differences[a]	.26 (.01)	106	.21 (.05)	94
Country averages	.58 (.00)	39	.26 (.12)	37

SOURCES.—Authors' calculations using data from the World Health Organization's mortality statistics; United Nations' world crime surveys; and Klaus Deininger & Lyn Squire, A New Data Set Measuring Income Inequality, 10 World Bank Econ. Rev. 565 (1996).

NOTE.—Crime rates are expressed in natural logs. *P*-values are in parentheses.

[a] Differences are obtained from consecutive country-period observations. Three observations are lost for homicides (one for robberies) for countries for which we have nonconsecutive data.

penditure. Therefore, the high-quality inequality data do not contain biases derived from the exclusion of nonmonetary income.

D. Bivariate Correlations

Table 1 presents the bivariate correlations between both crime rates and the Gini coefficient for three dimensions of the data, namely, pooled levels, pooled first differences, and country averages. The first set contains the correlation estimated for the pooled sample in levels, that is, using both the cross-country and over-time variation of the variables. The second set presents the correlations between the first differences of the crime rates and the first differences of the Gini index. These correlations, therefore, reflect only the over-time relationship between crime rates and inequality, thus controlling for any country characteristics that are fixed over time, such as geographic location or cultural heritage. The third set shows the correlations across countries only, based on the country averages for the whole periods (1965–94 for homicides and 1970–94 for robberies). Consequently, these correlations do not reflect the influence of country characteristics that change over time. All correlations of both crime rates with the Gini coefficient are positive and statistically significant (the largest *P*-value is .12). The smallest, but still positive, correlations are those estimated using the data in first differences. While there is not much disparity between the correlations estimated for the three data dimensions for the robbery rate, in the case of homicides the correlation drops from .54 for the data in pooled levels and .58 for country averages to .26 for first differences. This result suggests that almost half of the correlation between the Gini and homicide rates is due to country characteristics that are persistent over time.

Table 2 presents a second group of bivariate correlations for two cuts of the cross-country sample, namely, within countries and within time periods. The table contains the mean and the median of the correlations between each crime rate and the Gini index, which were obtained using, respectively, all

TABLE 2

WITHIN-COUNTRY AND WITHIN-PERIOD PAIRWISE CORRELATIONS BETWEEN THE GINI INDEX AND HOMICIDE AND ROBBERY RATES (in logs)

	HOMICIDES		ROBBERIES	
	Within Country	Within Period	Within Country	Within Period
Mean correlation	.22	.52	.23	.28
Median correlation	.48	.55	.58	.25
Positive correlations:				
%	62	100	59	100
N	39	6	37	5

SOURCES.—Authors' calculations using data from the World Health Organization's mortality statistics; United Nations' world crime surveys; and Klaus Deininger & Lyn Squire, A New Data Set Measuring Income Inequality, 10 World Bank Econ. Rev. 565 (1996).
NOTE.—Crime rates are expressed in natural logs.

the observations available for each country ("within country") and for each 5-year period ("within period"). In addition, we report the percentage of, respectively, countries and periods for which the correlation between crime rates and inequality is positive. All the estimated mean and median correlations are positive. In fact, for each of the 5-year periods, the cross-country correlation of crime and inequality is positive, while for about 60 percent of the countries, the time-series correlation is also positive. The fact that for both homicides and robberies the median within-country correlation is higher than the mean indicates that there are some outliers having negative correlations that depress the average.

An important problem for the interpretation of these bivariate correlations is that the apparent positive link between crime rates and income inequality might in fact be driven by other variables that are correlated with both of them. To address this issue, the following section studies the relationship between the Gini index and homicide and robbery rates, while controlling for other potential correlates of crime.

E. Multivariate Regression Analysis

On the basis of previous micro- and macrolevel crime studies, we consider the following variables as the basic correlates of homicide and robbery rates in addition to inequality measures: (1) GNP per capita (in logs) as both a measure of average national income and a proxy for overall development.[28] (2) The average number of years of schooling of the adult population as a measure of average educational attainment.[29] (3) The GDP growth rate to

[28] Norman Loayza *et al*., A World Savings Data-base (unpublished manuscript, World Bank, Policy Res. Dep't 1998).

[29] Robert Barro & Jong-Wha Lee, New Measures of Educational Attainment (unpublished manuscript, Harvard Univ. 1996).

proxy for employment and economic opportunities in general.[30] (4) The degree of urbanization of each country, which is measured as the percentage of the population in the country that lives in urban settlements (from World Bank data). Appendix A, Table A1, contains a detailed description of the data sources for these and the other variables used in this paper.[31]

The basic OLS multivariate regression results are shown in Table 3. The homicide and robbery regressions were run on the same data dimensions as in Table 1. The first regression was estimated using the pooled sample in levels; the second uses pooled first differences, thus focusing on the within-country variation; and the third regression uses country averages to isolate the pure cross-country dimension of the data. The results indicate that the Gini index maintains its positive and significant correlation with both crime rates. As expected, the models estimated in first differences present the lowest magnitudes for the coefficient on the Gini index. When the cross-country variation is taken into account, the coefficient on the Gini index increases from .02 to .06 in the case of homicides and from .04 to .11 in the case of robberies. Hence, in both cases, two-thirds of the conditional correlation between crime rates and inequality seems to be due to country characteristics that do not change over time.

Of the additional crime regressors, the most important one seems to be the GDP growth rate. This variable appears consistently with a negative sign, as expected, for both crimes. It is also statistically significant, although only marginally so in the robbery regression using country averages. In contrast, the other crime regressors do not show a consistent sign or are not statistically significant in at least half of the specifications.

The OLS estimates just discussed might be biased for three reasons. First, these regressions do not take into account the possibility that crime tends to persist over time; that is, they ignore yet another potential determinant of crime, which is the crime rate of the previous period. Second, these estimates might be biased because of the possibility that crime rates themselves (our dependent variables) might affect the right-hand-side variables. Third, it is very likely that the crime rates are measured with error, and this error might be correlated with some of the explanatory variables, particularly income inequality. The following section examines alternative specifications that include the lagged crime rate as an explanatory variable, account for certain types of measurement error, and allow for jointly endogenous explanatory variables.

[30] Loayza *et al*., *supra* note 27.

[31] Appendix B, Table B3, part A, contains the matrix of bivariate correlations among the basic set of dependent and explanatory variables. Note that the Gini is indeed significantly correlated with log of income per capita (negatively), educational attainment of the adult population (negatively), and the GDP growth rate (positively).

TABLE 3

Basic Economic Model (Ordinary Least Squares Estimation)

	Homicide Rate			Robbery Rate		
Dependent Variable (in logs)	Pooled Levels (1)	Pooled First Differences (2)	Country Averages (3)	Pooled Levels (4)	Pooled First Differences (5)	Country Averages (6)
Income inequality: Gini coefficient	.064 (6.418)	.023 (3.121)	.067 (2.923)	.105 (7.634)	.039 (2.476)	.111 (4.204)
Growth rate: % annual change in real gross domestic product	−7.959 (−2.785)	−2.032 (−2.184)	−12.026 (−1.668)	−11.963 (−3.371)	−4.963 (−2.294)	−9.751 (−1.251)
Average income: log of gross national product per capita in U.S. dollars	−.343 (−2.966)	.106 (.620)	−.351 (−1.391)	−.053 (−.349)	−.223 (−.624)	−.101 (−.351)
Urbanization: % urban population	.000 (−.050)	.039 (3.068)	.003 (.254)	.026 (3.449)	.015 (.518)	.030 (2.089)
Educational attainment: average years of education for adults	.081 (1.646)	−.023 (−.520)	.044 (.360)	.153 (2.260)	.254 (2.332)	.175 (1.304)
Intercept	1.112 (6.418)		1.165 (.579)	−2.422 (−2.427)		−2.838 (−1.527)
Adjusted R^2	.38	.24	.34	.49	.25	.49
Number of countries	39	39	39	37	37	37
N	148	106	39	132	94	37

Sources.—Authors' calculations. For details on definitions and sources of variables, see Appendix A, Table A1. Homicides data source: World Health Organization's mortality statistics; robbery data source: United Nations' world crime surveys.

Note.—The *t*-statistics are in parentheses.

III. A Dynamic Empirical Model of Crime Rates

A. Econometric Issues

The evidence presented so far suggests that, from a cross-country perspective, there is a robust correlation between the incidence of crimes and the extent of income inequality. However, there are several issues we must confront in order to assure that this correlation is not the result of estimation biases. First, as mentioned, the incidence of violent crime appears to have inertial properties (for example, persistence) that are noted in the theoretical literature and documented in the micro- and macroempirical work (see the work of Edward Glaeser, Bruce Sacerdote, and Jose Scheinkman[32] and Fajnzylber, Lederman, and Loayza[33]). To account for criminal inertia, we need to work with a dynamic, lagged dependent econometric model.

The second issue we must address is that the relationship between violent crime rates and their determinants is often characterized by a two-way causality. Failure to correct for the joint endogeneity of the explanatory variables would lead to inconsistent coefficients, which depending on the sign of the reverse causality would render an over- or underestimation of their effects on violent crime rates. We address the problem of joint endogeneity by employing an instrumental variable procedure applied to dynamic models of panel data. This is the GMM estimator that uses the dynamic properties of the data to generate proper instrumental variables.

The third estimation difficulty is that, despite our use of intentional homicide and robbery rates as the best proxies for the incidence of violent crimes, it is likely that measurement error still afflicts our crime data. Ignoring this problem might also result in biased estimates, especially because crime underreporting is not random measurement error but is strongly correlated with factors affecting crime rates such as inequality, education, the average level of income, and the rate of urbanization. Even if measurement error were random, the coefficient estimates would still be biased given the dynamic nature of our model. To control for measurement error, we model it as either random noise or a combination of an unobserved country-specific effect and random noise.

The econometric methodology is as follows: We implement a GMM estimator applied to dynamic (lag-dependent variable) models that use panel data. This method was developed by Manuel Arellano and Stephen Bond[34]

[32] Edward L. Glaeser, Bruce Sacerdote, & Jose A. Scheinkman, Crime and Social Interactions, 111 Q. J. Econ. 507 (1996).

[33] Fajnzylber, Lederman, & Loayza, *supra* note 24.

[34] Manuel Arellano & Stephen Bond, Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations, 58 Rev. Econ. Stud. 277 (1991).

and Arellano and Olympia Bover.[35] It controls for (weak) endogeneity through the use of instrumental variables that consist of appropriately lagged values of the explanatory variables. When the model does not include an unobserved country-specific effect, the model is estimated in levels, for both the regression equation and the set of instruments. This is called the GMM levels estimator. When the model includes an unobserved country-specific effect (resulting from time-invariant omitted factors such as systematic measurement error), the model is estimated in both differences and levels jointly in a system. This is called the GMM system estimator. For each estimator, the correct specification of the regression equation and its instruments is tested through a Sargan-type test and a serial-correlation test.[36] Appendix D presents the econometric methodology in detail.

B. Basic Results

Table 4 shows GMM estimates for the basic set of determinants of the homicide and the robbery rates. Columns 1 and 2 and columns 4 and 5 present the results obtained for the model that assumes no unobserved country-specific effects and that was estimated using the GMM levels estimator. The difference between columns 1 and 2 and columns 4 and 5 is related to the samples used in each case, which are restricted to, respectively, countries with at least two and three consecutive observations. Columns 3 and 6 report the results obtained for the model that allows unobserved country-specific effects and were estimated using the GMM system estimator. The system estimator uses not only levels but also differences of the variables and requires at least three consecutive observations for each country. Thus, the results in columns 2 and 3 and columns 5 and 6 are obtained from the same samples but are based on different estimators.

In the basic levels specification for homicides, using the largest possible sample (column 1), the lagged homicide rate, the level of income inequality, and the growth rate of the GDP have significant coefficients with the expected signs. The rate of urbanization also appears to be significantly associated with homicide rates but, unexpectedly, in a negative way: countries with a larger fraction of their population in cities would appear to have lower crime rates. Qualitatively similar results are obtained with the smaller sample used in columns 2 and 5, although in these cases the population's average income and educational attainment are significant, both with negative signs. Re-

[35] Manuel Arellano & Olympia Bover, Another Look at the Instrumental Variable Estimation of Error-Component Models, 68 J. Econometrics 29 (1995).

[36] In both tests, the null hypothesis denotes correct specification. For the GMM levels estimator, serial correlation of any order implies misspecification, while for the GMM system estimator, only second- and higher-order serial correlation denotes misspecification (see Appendix D for details).

TABLE 4

BASIC ECONOMIC MODEL (Generalized Method of Moments Estimation)

	HOMICIDE RATE			ROBBERY RATE		
	(1)	(2)	(3)	(4)	(5)	(6)
Regression specification	Levels	Levels[a]	Levels and differences	Levels	Levels[a]	Levels and differences
Lagged dependent variable	.8957 (46.2310)	.9282 (114.5404)	.8137 (25.4593)	.7254 (12.3614)	.7528 (23.1968)	.7222 (50.0311)
Income inequality: Gini coefficient	.0069 (3.9761)	.0032 (2.3130)	.0155 (7.0490)	.0331 (3.5354)	.0223 (4.8849)	.0307 (11.8691)
Growth rate: % annual change in real gross domestic product	−1.9270 (−2.9066)	−3.3952 (−6.5945)	−4.2835 (−4.9471)	−8.4505 (−7.2343)	−7.1754 (−9.3441)	−11.1536 (−18.1176)
Average income: log of gross national product per capita in U.S. dollars	−.0570 (.0187)	−.0396 (−4.0141)	.0151 (.8876)	−.0541 (−.7641)	−.0923 (−2.9946)	−.0287 (−2.0664)
Urbanization: % urban population	−.0032 (−4.9258)	−.0023 (−5.9852)	−.0019 (−.8959)	.0078 (4.0733)	.0106 (6.9098)	.0053 (6.0005)
Educational attainment: average years of education for adults	−.0090 (−1.1584)	−.0153 (−2.7694)	−.0300 (−3.4280)	.0901 (3.9416)	.0634 (4.2398)	.0382 (6.5551)
Intercept	.7935 (5.3834)	.7584 (9.7501)		−.5486 (−.8663)	.0669 (.2277)	
Number of countries	39	27	27	37	29	29
N	106	91	91	94	85	85
Specification tests, *P*-values:						
Sargan test	.651	.581	.958	.531	.314	.430
Serial correlation:						
First order	.683	.873	.048	.035	.103	.082
Second order	.239	.498	.240	.147	.225	.879

SOURCES.—Authors' calculations. For details on definitions and sources of variables, see Appendix A, Table A1. Homicides data source: World Health Organization's mortality statistics; robbery data source: United Nations' world crime surveys.

NOTE.—The *t*-statistics are in parentheses.

[a] The sample is restricted to the countries that have at least three consecutive observations.

gardless of the sample, both the Sargan and the serial-correlation tests validate the results obtained using the levels estimator for homicides.

Columns 3 and 6 show the results using the GMM system estimator. As in the case of the levels estimator, both the Sargan and the serial-correlation tests support the specification of the system estimator. The main results are as follows:

First, homicide rates show a sizeable degree of inertia. The coefficient on the lagged homicide rate is close to unity (although not as large as when country-specific effects are ignored). The size of this coefficient implies that the half-life of a unit shock lasts about 17 years.[37]

Second, income inequality, measured by the Gini index, has a positive and significant effect on homicide rates. By using the corresponding coefficient estimate, we can evaluate the crime-reducing effect of a decline in inequality in a given country. If the Gini index falls permanently by the within-country standard deviation in the sample (about 2.4 percentage points), the intentional homicide rate will decrease by 3.7 percent in the short run and 20 percent in the long run.[38] If the Gini index were to fall by its cross-country standard deviation, the decline in inequality would be much larger; however, a change in inequality by the magnitude of cross-country differences is implausibly large to be attained by a country in a reasonable amount of time. It is noteworthy that the estimated coefficient on the Gini index is much larger with the system estimator than with the levels estimator, although they are both based on a common sample of 27 countries. It is possible that the higher magnitude obtained with the system estimator is due to the fact that this estimator corrects for the positive correlation between inequality and the degree of crime underreporting (for example, the measurement error).[39]

Third, the GDP growth rate has a significantly negative effect on the homicide rate. According to our estimates, the impact of a permanent 1 percentage point increase in the GDP growth rate is associated with a 4.3 percent fall in the homicide rate in the short run and a 23 percent decline in the long run. Fourth, our measure of educational attainment remains negative and significant, but the GNP per capita and the urbanization rate now lack statistical significance. The pattern of significance (or lack thereof) of the basic explanatory variables is quite robust to all the various empirical

[37] The half-life (HL) of a unit shock is obtained as follows: $HL = \ln(.5)/\ln(a)$, where a is the estimated autoregressive coefficient. According to Table 4, column 3, $a = .8137$.

[38] The within-country standard deviation is calculated after applying a "within" transformation to the Gini index, which amounts to subtracting from each observation the average value of that variable for the corresponding country and adding the global mean (based on all observations in the sample).

[39] This finding is interesting because we expected that the magnitude of the effect of the levels estimator would be higher, because the analysis of the bivariate and conditional OLS correlations showed that a large portion of the correlation between the crime rates and the Gini was due to country characteristics that do not change over time and that are lost in the first-differenced data.

exercises of this paper. It is also similar to what we found in our first empirical cross-country study on violent crime rates.[40]

In columns 4–6 of Table 4, we report analogous estimates for the determinants of robbery rates. For robberies, the results are qualitatively similar across samples and specifications. In the cases of the lagged dependent variable, the growth rate, and income inequality, the results for robberies are similar to those for homicides. Indeed, there is evidence that robberies are also subject to a sizeable degree of inertia, although somewhat smaller than in the case of homicides: the half-life of the effects of a permanent shock is between 11 and 12 years, depending on the specification. The coefficients on income inequality are also positive and significant in all specifications. On the basis of the results in column 6, we find that a decline of 1 within-country standard deviation in the Gini coefficient (about 2.1 percent) leads to a 6.5 percent decline of the robbery rate in the short run and a 23.2 percent decline in the long run. Similarly, a permanent 1 percentage point increase in the GDP growth rate produces 11 and 45 percent declines of the robbery rate in the short and long runs, respectively. As in the case of homicides, note that the magnitude of the estimated impact of the Gini index on robbery rates is larger for the system than for the levels estimator (for equal samples, of course).

As for the other variables, their signs and significance vary from homicides to robberies. The average income appears with a negative sign but is significant only for the smaller samples (columns 5 and 6). Educational attainment and urbanization are significant in all specifications, both with a positive sign. The latter result was expected, as robberies appear to be mostly an urban phenomenon. However, the finding that robberies are positively associated with education is puzzling.

With regard to the GMM specification tests for the robbery models, all regressions are supported by the Sargan test on the validity of the instrumental variables. However, in the levels regressions there is evidence that the residuals suffer from first-order serial correlation, especially in the case of the largest sample of 37 countries.

C. Alternative Measures of Inequality

This section studies the crime effect of alternative measures of income inequality and thus checks the robustness of the results obtained with the Gini coefficient. The alternative measures we consider are the ratio of income of the richest to the poorest quintile of the population, an index of income polarization, and the standard deviation of the educational attainment of the

[40] Fajnzylber, Lederman, & Loayza, *supra* note 19.

adult population.[41] Given the fact that the new variables lead to further restrictions in sample size, we choose to maintain our basic levels specification, which allows the largest possible sample in the context of a dynamic model. The results are presented in Table 5.

In columns 1 and 4, the ratio of the income shares of the first to the fifth quintile is substituted for the Gini coefficient in the basic regressions for homicides and robberies, respectively. The results are qualitatively analogous to those reported in Table 4. The new measure of income inequality is positively and significantly associated with both crime rates. A permanent decline of 1 within-country standard deviation in the quintile ratio (about 1.3) leads to a 2 percent decline in the intentional homicide rate in the short run and a 16.2 percent fall in the long run. The corresponding impacts on the robbery rate are 4.7 and 21.5 percent, respectively, in the short and the long runs. In further exercises (not presented in the tables), we examined the significance of the income levels of the poor and rich separately. We found that when the income of the poor was included by itself, its coefficient was not generally significant. When we included the incomes of both the poor and the rich, neither was statistically significant, which can be explained by the fact that they are highly correlated with each other. These results contrast with the significant crime-inducing effect of the difference between the income levels of the rich and poor (or more precisely, the log of the ratio of top to bottom income quintiles of the population). Coupled with the general lack of significance of per capita GNP in our crime regressions, the aforementioned results indicate that it is not the level of income that matters for crime but the income differences among the population.

In columns 2 and 5, we substitute an index of polarization for the Gini index. Some authors argue that a society's degree of polarization may be the cause of rebellions, civil wars, and social tension in general (see, for example, Esteban and Ray[42] and Paul Collier and Anke Hoeffler[43]). Similar arguments can be applied to violent crime. The concept of polarization was formally introduced by Esteban and Ray.[44] Although they are linked to standard measures of income inequality, the polarization indicators proposed by these authors do not consider only the distance between the incomes of various groups but also the degree of homogeneity within these groups. Thus, the social tension that leads to violence and crime would be produced by the heterogeneity of internally strong groups. Following the principles proposed by Esteban and Ray, we constructed a polarization index from data on national

[41] Appendix B, Table B3, part B, shows the bivariate correlations between the Gini index and these three alternative indicators of inequality. As expected, these correlations are statistically significant and high in magnitude, ranging from .62 to .88.

[42] Esteban & Ray, *supra* note 21.

[43] Paul Collier & Anke Hoeffler, On the Economic Causes of Civil War, 50 Oxford Econ. Papers 563 (1998).

[44] Esteban & Ray, *supra* note 21.

TABLE 5

ALTERNATIVE INEQUALITY MEASURES (Generalized Method of Moments Levels Estimation)

DEPENDENT VARIABLE (in logs)	HOMICIDE RATE (1)	(2)	(3)	ROBBERY RATE (4)	(5)	(6)
Lagged dependent variable	.8780 (48.8267)	.8688 (56.5593)	.9268 (58.2480)	.7818 (18.7384)	.7784 (25.1457)	.8698 (18.9108)
Growth rate: % annual change in real gross domestic product	−3.2533 (−3.8806)	−3.1665 (−4.2468)	−.9411 (−1.2659)	−4.1422 (−3.6406)	−4.7806 (−4.1412)	−6.9070 (−2.3018)
Average income: log of gross national product per capita in U.S. dollars	−.0666 (−4.2536)	−.0761 (−4.1829)	−.0590 (−2.8788)	−.0666 (−1.8281)	−.1135 (−3.1697)	.0373 (.6770)
Urbanization: % urban population	−.0019 (−2.1161)	−.0023 (−2.4541)	−.0031 (−2.5615)	.0078 (3.6231)	.0095 (4.3303)	.0028 (.9051)
Educational attainment: average years of education for adults	−.0205 (−3.1268)	−.0151 (−2.5024)	−.0083 (−.7124)	.0568 (2.6936)	.0699 (2.9841)	.0632 (1.9850)
Intercept	1.0556 (8.5890)	.7916 (5.6564)	.9411 (5.1455)	.4020 (1.5376)	.0463 (.1752)	−.7550 (−1.2889)
Ratio of the first to the fifth quintile	.0152 (3.7918)			.0426 (4.0275)		
Income polarization: log of income polarization index		.0019 (4.9967)			.0037 (3.6931)	
Educational inequality: standard deviation of schooling years			.1036 (.8801)			.8976 (2.0666)
Number of countries	39	39	37	36	36	35
N	96	96	103	86	86	91
Specification tests: *P*-values:						
Sargan test	.533	.789	.512	.168	.201	.180
Serial correlation:						
First order	.662	.742	.829	.145	.078	.023
Second order	.272	.272	.174	.283	.136	.359

SOURCES.—Authors' calculations. For details on definitions and sources of variables, see Appendix A, Table A1. Homicides data source: World Health Organization's mortality statistics; robbery data source: United Nations' world crime surveys.

NOTE.—The *t*-statistics are in parentheses.

income shares by quintiles (see Appendix C for details). The results concerning polarization presented in Table 5 are similar to those obtained with the other inequality indicators. The effect of polarization on crime appears to be positive and significant for both homicides and robberies, and the signs and significance of the other core variables are mostly unchanged. As for the size of the polarization coefficient, a permanent reduction of 1 within-country standard deviation (about 7.6 percent) in this variable leads to a decline in the homicide and the robbery rate of, respectively, 3.8 and 3 percent in the short run. In the long run, the corresponding reductions are 28.7 and 19.2 percent for the homicide and robbery rates.

Columns 3 and 6 address the question of whether the underlying inequality of educational attainment has the same impact on crime rates as the Gini index does. We measure the inequality of educational attainment as the standard deviation of schooling years in the adult population, as estimated by De Gregorio and Lee.[45] The basic results discussed above remain essentially unaltered. When we substitute the measure of education inequality for the Gini index, the estimated coefficient of this variable acquires the sign of the Gini index in the benchmark regression but appears significant only in the robbery regression. A decline of 1 within-country standard deviation (about 4 percent) in our measure of educational inequality leads to a reduction in the robbery rate of 3.6 and 27.6 percent in the short and long runs, respectively.

D. Additional Controls

This section focuses on the potential role played by additional control variables in the crime-inducing effect of income inequality.[46] The regression results are presented in Table 6. Columns 1 and 5 show the results for the regression on the basic explanatory variables, with the addition of a measure of ethnic diversity. This measure is the index of ethnolinguistic fractionalization employed by Pablo Mauro[47] and William Easterly and Ross Levine[48] in their cross-country growth studies. Our results indicate that this index is significantly associated with higher homicide rates but its link with robberies is not significant (with a negative point estimate). As to its quantitative effect on homicides, an increase of 1 standard deviation (about 4 percent) in ethnolinguistic fractionalization is associated with an increase in the homicide rate of 3.8 and 31.6 percent in the short and long runs, respectively. Most

[45] De Gregorio & Lee, *supra* note 22.

[46] Appendix B, Table B3, part C, contains the bivariate correlations between these new control variables and the set of basic variables used in the paper. Of the additional control variables, only the share of young males in the national population exhibits a high and significant correlation with the Gini index. This variable is also positively and significantly correlated with both crime rates.

[47] Paolo Mauro, Corruption and Growth, 110 Q. J. Econ. 681 (1995).

[48] William Easterly & Ross Levine, Africa's Growth Tragedy: Policies and Ethnic Divisions, 112 Q. J. Econ. 1203 (1997).

TABLE 6

Additional Control Variables (Generalized Method of Moments Levels Estimation)

	Homicide Rate				Robbery Rate			
Dependent Variable (in logs)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lagged dependent variable	.8792	.8987	.9040	.9106	.8195	.7358	.7414	.7849
	(57.6962)	(62.3507)	(45.7224)	(42.0085)	(15.2561)	(12.9489)	(12.2210)	(16.3984)
Income inequality: Gini coefficient	.0069	.0055	.0049	.0073	.0275	.0307	.0228	.0226
	(2.7692)	(3.4744)	(2.1574)	(3.8160)	(3.3464)	(3.3330)	(2.4784)	(3.8775)
Growth rate: % annual change in real gross domestic product	−1.3845	−1.0327	−1.8185	−.9918	−7.8263	−8.6606	−7.1703	−8.1747
	(−2.8885)	(−1.3381)	(−2.4141)	(−1.1735)	(−6.9305)	(−6.2523)	(−4.9902)	(−6.0847)
Average income: log of gross national product per capita in U.S. dollars	−.0335	−.0689	−.0432	−.0448	.0115	−.0332	−.0477	−.0865
	(−1.2479)	(−6.2284)	(−2.4857)	(−2.9894)	(.1548)	(−.5184)	(−.7148)	(−1.6034)
Urbanization: % urban population	−.0033	−.0029	−.0033	−.0031	.0047	.0065	.0045	.0078
	(−5.6910)	(−4.7985)	(−5.0455)	(−4.9780)	(2.4969)	(2.9605)	(2.3226)	(4.6855)
Educational attainment: average years of education for adults	−.0203	−.0042	−.0103	−.0029	.0453	.0832	.1081	.0707
	(−2.1061)	(−.6974)	(−1.2640)	(−.4580)	(3.0588)	(3.7235)	(4.5435)	(4.6855)
Intercept	.6323	.8846	.7445	.7186	−.6947	−.6327	−.2877	.0029
	(2.5482)	(8.5859)	(4.7312)	(6.5509)	(−1.0287)	(−1.0006)	(−.4495)	(.0089)

Ethnolinguistic fractionalization	.1706				−.1092			
	(2.1826)				(−.7786)			
Police per 100,000 population		−.0001				.0004		
		(−.9153)				(1.5252)		
Latin America: dummy variable			.0493				.3486	
			(.9198)				(2.7248)	
Young male population share: 15–29-year-olds as % of total population				−.0103				.0039
				(−1.5412)				(.1930)
Number of countries	34	35	39	39	31	36	37	37
Number of observations	96	97	106	106	83	92	94	94
Specification tests: *P*-values:								
Sargan test	.674	.533	.471	.536	.554	.575	.349	.539
Serial correlation:								
First order	.486	.702	.732	.782	.060	.025	.044	.043
Second order	.123	.240	.279	.244	.153	.120	.175	.194

SOURCES.—Authors' calculations. For details on definitions and sources of variables, see Appendix A, Table A1. Homicides data source: World Health Organization's mortality statistics; robbery data source: United Nations' world crime surveys.

NOTE.—The *t*-statistics are in parentheses.

important for our purposes, the Gini index keeps its sign, size, and significance in the homicide and robbery regressions even when ethnic diversity as a crime determinant is controlled for.

Columns 2 and 6 address the possibility that the crime-inducing effect of the Gini coefficient in fact reflects an unequal distribution of protection from the police and the judicial system. We do this by adding the number of police per capita to the core explanatory variables.[49] This is an average measure for the whole population and may not represent egalitarian protection by the police and the law. However, it is an appropriate control under the assumption that an unequal distribution of protection is more likely to occur when there is scarcity of police resources. Although for homicides the number of police per capita does present the expected negative sign, for both crimes this variable presents statistically insignificant coefficients. Most important, the sign, size, and statistical significance of the Gini coefficient appear to be unaltered by the inclusion of this proxy for police deterrence.

In columns 3 and 7, we add a Latin American dummy to the basic explanatory variables. We do it to assess whether the apparent effect of inequality on crime is merely driven by a regional effect, given that countries in Latin America have among the highest indices of income inequality in the world and, in many cases, also very high crime rates. We find that in fact the Latin American dummy has a positive coefficient in the regressions for both crimes, although it is statistically significant only in the case of robberies. Quantitatively, the results suggest that in Latin America the rate of robberies is 35 percent higher than what our basic model predicts, given the economic characteristics of the countries in that region. Most important for our purposes, the signs and significance of our basic explanatory variables, especially the Gini index, are not altered by the inclusion of the Latin American dummy.

Finally, columns 4 and 8 report results of regressions in which we introduce the percentage of young males (aged 15–29 years) in the population as an additional explanatory variable. It is well known that the rate of crime participation of individuals is highest at the initial stages of adulthood, so one could expect countries with large populations in that age group to have high crime rates. At the same time, countries with younger adult populations may experience more income inequality through a Kuznets-like effect. The inclusion of the proportion of young males as a determinant of crime allows us to test whether the inequality-crime link is driven by this demographic factor. Our results indicate that after controlling for our basic crime determinants, the share of young males in the population does not have a statis-

[49] We average the available observations of this variable for the 1965–95 period, and then we use the average as a constant observation for all 5-year periods in the regression. We do this to increase the number of usable observations per country and, most important, to minimize the reverse causation of this variable to changes over time in homicide rates (although this does not solve the cross-country dimension of reverse causation).

tically significant effect on either homicide or robbery rates. In fact, for the former crime, the point estimate of that variable is actually negative. As in previous robustness exercises, controlling for the proportion of young males does not lead to any substantial change in the estimated effect of inequality on crime.

E. Poverty Alleviation and Crime

Although the main objective of this paper is to analyze the relationship between income inequality and crime, our empirical findings suggest that there is also an important correlation between the incidence of crime and the rate of poverty alleviation. This relationship exists as a consequence of the joint effects of income inequality and economic growth on crime rates. The level of poverty in a country is measured as the percentage of the population that receives income below a threshold level, which is usually determined by the necessary caloric intake and the local monetary cost of purchasing the corresponding food basket. Simply put, the level of poverty is jointly determined by the national income level and by the pattern of distribution of this income. When a reduction in income inequality is coupled with a rise in economic growth, the rate of poverty alleviation improves.

Through the several econometric exercises performed in the paper, we find that the GDP growth rate and the Gini index are the most robust and significant determinants of both homicide and robbery rates. Consequently, these results also indicate that the rate of change of poverty is also related to the incidence of crime. That is, when poverty falls more rapidly, either because income growth rises or the distribution of income improves, then crimes rates tend to fall. Estimating the precise effect of poverty reduction on violent crime and designing a strategy for crime-reducing poverty alleviation remain important topics for future research.

IV. Conclusions

The main conclusion of this paper is that income inequality, measured by the Gini index, has a significant and positive effect on the incidence of crime. This result is robust to changes in the crime rate when it is used as the dependent variable (whether homicide or robbery), the sample of countries and periods, alternative measures of income inequality, the set of additional variables explaining crime rates (control variables), and the method of econometric estimation. In particular, this result persists when using instrumental variable methods that take advantage of the dynamic properties of our cross-country and time-series data to control for both measurement error in crime data and the joint endogeneity of the explanatory variables.

In the process of arriving at this conclusion, we found some interesting results; the following are among them: First, the incidence of violent crime has a high degree of inertia, which justifies early intervention to prevent crime waves. Second, violent crime rates decrease when economic growth improves. Since violent crime is jointly determined by the pattern of income distribution and by the rate of change of national income, we can conclude that faster poverty reduction leads to a decline in national crime rates. And third, the mean level of income, the average educational attainment of the adult population, and the degree of urbanization in a country are not related to crime rates in a significant, robust, or consistent way.

The main objective of this paper has been to characterize the relationship between inequality and crime from an empirical perspective. We have attempted to provide a set of stylized facts on this relationship. Crime rates and inequality are positively correlated (within each country and, particularly, between countries), and it appears that this correlation reflects causation from inequality to crime rates, even controlling for other crime determinants. If anything, the contribution of this paper is empirical. Analytically, however, this paper has two important shortcomings. First, we have not provided a way to test or distinguish between various theories on the incidence of crime. In particular, our results are consistent with both economic and sociological paradigms. Although our results for robbery (a typical property crime) confirm those for homicide (a personal crime with a variety of motivations), this cannot be used to reject the sociological paradigm in favor of the economic one. The reason is that the satisfaction that the "relatively deprived" people in sociological models seek can lead to both pure manifestations of violence and illicit appropriation of material goods. A more nuanced econometric exercise than what we offer here is required to shed light on the relative validity of various theories on the inequality-crime link.

The first shortcoming of the paper leads to the second, which is that we have not identified the mechanisms through which more pronounced inequality leads to more crime. Uncertainty about these mechanisms raises a variety of questions with important policy implications. For instance, should police and justice protection be redirected to the poorest segments of society? How important for crime prevention are income-transfer programs in times of economic recession? To what extent should public authorities be concerned with income and ethnic polarization? Do policies that promote the participation in communal organizations and help develop "social capital" among the poor also reduce crime? Hopefully, this paper will help stir an interest in these and related questions on the prevention of crime and violence.

Appendix A

Data Definitions and Sources

TABLE A1

Description and Sources of the Variables

Variable	Description	Source
Intentional homicide rate	Number of deaths purposely inflicted by another person, per 100,000 population	Constructed from mortality data from the World Health Organization (WHO). Most of these data are available by FTP from the WHO server (WHO-HQ-STATS01.WHO.CH) in the directory \FTP\MORTALIT. Additional data were extracted from the WHO publication "World Health Statistics Annual." The data on population were taken from the World Bank's International Economic Department database
GNP per capita	Gross national product expressed in constant 1987 U.S. prices and converted to U.S. dollars on the basis of the "notional exchange rate" proposed by Norman Loayza *et al*., A World Savings Data-base (unpublished manuscript, World Bank, Policy Res. Dep't 1998)	Most data were taken from Loayza *et al*., *supra*. For some countries, the variable was constructed on the basis of the same methodology using data from the World Bank's International Economic Department database
Gini index	Income-based Gini coefficient. Constructed by adding 6.6 to expenditure-based indexes to make them comparable to income-based indexes. Data of "high quality" were used when available. Otherwise, an average of the available data was used	Klaus Deininger & Lyn Squire, A New Data Set Measuring Income Inequality, 10 World Bank Econ. Rev. 565 (1996). The data set is available on the Internet from the World Bank's server, at http://www.worldbank.org/html/prdmg/grthweb/datasets.htm

Educational attainment	Average years of schooling of the population over 15	Robert Barro & Jong-Wha-Lee, New Measures of Educational Attainment (unpublished manuscript, Harvard Univ. 1996). The data set is available on the Internet from the World Bank's server, at http://www.worldbank.org/html/prdmg/grthweb/datasets.htm
GDP growth	Growth in the gross domestic product constructed as the log-difference of GDP at constant local 1987 market prices	Loayza *et al*., *supra*
Standard deviation of educational attainment	Standard deviation of the distribution of education for the total population over age 15. The population is distributed in seven categories: no formal education, incomplete primary, complete primary, first cycle of secondary, second cycle of secondary, incomplete higher, and complete higher. Each person is assumed to have an educational attainment of $\log(1 +$ years of education$)$	José De Gregori & Jong-Wha Lee, Education and Income Distribution: New Evidence from Cross-Country Data (unpublished manuscript, Univ. Chile & Korea Univ. 1999)
Ethnolinguistic fractionalization	Probability that two randomly selected people from a given country will not belong to the same ethnolinguistic group (1960)	William Easterly & Ross Levine, Africa's Growth Tragedy: Policies and Ethnic Divisions, 112 Q. J. Econ. 1203 (1997). The data set is available on the Internet from the World Bank's server, at http://www.worldbank.org/html/prdmg/grthweb/datasets.htm
Police per 100,000	Number of police personnel per 100,000 population	Constructed from the United Nations' *World Crime Surveys of Crime Trends and Operations of Criminal Justice Systems*, various issues. The data are available on the Internet at http://www.ifs.univie.ac.at/~uncjin/wcs.html#wcs123
Young male population share	Male population 15–29 years of age as a share of the total population	World Bank data
Income of the fifth quintile relative to the first quintile	Income of the population in the fifth quintile of the distribution of income divided by the income of the first quintile	World Bank data

APPENDIX B

Summary Statistics

TABLE B1

Summary Statistics, Homicide Rates (Number of Homicides per 100,000 Population)

Country	N	Mean	Standard Deviation	Minimum	Maximum
Australia	4	1.91	.06	1.84	1.98
Belgium	2	1.62	.15	1.51	1.73
Brazil	3	12.97	3.49	9.39	16.36
Bulgaria	3	3.41	.92	2.71	4.45
Canada	6	2.08	.34	1.51	2.48
Chile	5	3.45	.61	3.00	4.33
China	2	.18	.03	.16	.21
Colombia	4	42.80	29.43	17.28	80.61
Costa Rica	4	4.21	.43	3.56	4.51
Denmark	4	1.07	.25	.69	1.23
Dominica	3	4.70	.89	4.12	5.72
Finland	6	2.84	.36	2.22	3.24
France	4	.94	.16	.79	1.14
Germany	4	1.22	.05	1.18	1.29
Greece	2	.89	.07	.84	.94
Hong Kong	5	1.68	.35	1.24	2.18
Hungary	5	2.65	.68	2.08	3.75
Ireland	2	.87	.09	.81	.93
Italy	5	1.69	.50	1.03	2.38
Japan	6	1.04	.31	.61	1.41
Mauritius	3	1.76	.65	1.21	2.48
Mexico	4	18.37	.83	17.92	19.62
Netherlands	4	.94	.14	.81	1.14
New Zealand	5	1.57	.43	1.09	2.02
Norway	6	.92	.30	.54	1.28
Panama	2	4.23	2.49	2.46	5.99
Peru	2	2.68	.70	2.19	3.17
Philippines	2	14.65	3.29	12.32	16.98
Poland	3	2.09	.69	1.64	2.89
Romania	2	4.28	.59	3.86	4.70
Singapore	4	2.13	.19	1.84	2.23
Spain	4	.63	.43	.15	1.01
Sri Lanka	2	6.10	1.07	5.35	6.85
Sweden	4	1.27	.08	1.18	1.35
Thailand	2	24.22	1.88	22.89	25.55
Trinidad and Tobago	3	4.85	.97	3.74	5.52
United Kingdom	6	.97	.21	.71	1.35
United States	6	8.89	1.17	6.62	9.91
Venezuela	5	10.20	2.68	7.68	14.19

Sources.—Homicide data are from the World Health Organization; population data are from the World Bank.

TABLE B2

SUMMARY STATISTICS, ROBBERY RATES (Number of Robberies per 100,000 Population)

Country	Observations	Mean	Standard Deviation	Minimum	Maximum
Australia	4	44.34	18.70	23.61	68.60
Bangladesh	3	3.08	1.78	1.79	5.11
Belgium	2	38.83	34.67	14.31	63.35
Bulgaria	3	22.70	27.93	5.65	54.94
Canada	5	88.09	18.74	58.70	107.58
Chile	3	490.43	82.55	398.70	558.74
China	3	4.48	4.14	1.52	9.22
Denmark	2	70.72	44.52	39.24	102.21
Finland	5	37.61	5.97	31.18	46.33
Germany	3	27.84	7.17	21.58	35.66
Greece	2	2.03	.74	1.50	2.55
Hong Kong	3	133.56	26.99	106.72	160.70
Hungary	3	19.22	9.64	11.91	30.14
India	5	4.12	.98	2.87	5.63
Indonesia	5	4.54	2.12	2.51	8.14
Italy	4	35.42	23.46	7.09	59.51
Jamaica	4	177.73	35.63	137.08	213.61
Japan	5	1.83	.26	1.47	2.19
Korea	4	5.88	2.35	3.37	8.51
Malaysia	4	31.02	13.61	12.36	44.80
Mauritius	2	43.28	3.86	40.55	46.01
Netherlands	4	53.79	26.28	22.11	79.89
New Zealand	4	19.33	11.49	9.50	34.85
Norway	4	14.34	7.40	8.11	24.68
Pakistan	2	1.18	.58	.76	1.59
Peru	3	240.25	12.04	227.64	251.63
Philippines	2	24.26	7.90	18.66	29.85
Poland	3	26.93	13.61	17.52	42.54
Romania	2	9.79	9.11	3.34	16.23
Singapore	4	65.63	18.25	47.59	90.82
Sri Lanka	5	37.10	8.05	28.66	47.40
Sweden	4	49.13	14.24	36.22	68.76
Thailand	4	12.29	6.14	6.11	17.99
Trinidad	3	62.27	49.05	32.70	118.89
United Kingdom	4	59.75	28.91	31.82	99.66
United States	5	216.00	28.83	184.85	256.83
Venezuela	4	144.26	52.39	76.62	200.40

SOURCES.—Robbery data are from the United Nations' *World Crime Surveys*; population data are from the World Bank.

TABLE B3

Bivariate Correlations of Variables Included Simultaneously in Multivariate Regressions

	Log of Homicide Rate	Log of Robbery Rate	Gini Index	Log of Gross National Product per Capita	Educational Attainment	Gross Domestic Product Growth	Urbanization
A. Variables used in basic regressions:							
Log of homicide rate	1.00						
Log of robbery rate	.46 (96)	1.00					
Gini index	.54 (148)	.28 (132)	1.00				
Log of gross national product per capita	−.44 (148)	.40 (132)	−.46 (148)	1.00			
Educational attainment	−.31 (148)	.35 (132)	−.63 (148)	.64 (148)	1.00		
Gross domestic product growth	−.07* (148)	−.28 (132)	.26 (148)	−.15* (148)	−.37 (148)	1.00	
Urbanization	−.24 (148)	.51 (132)	−.19 (148)	.68 (148)	.41 (148)	−.10* (148)	1.00
B. Alternative indicators of inequality used instead of the Gini index in GMM regressions:							
Ratio of income share of fifth to first quintile	.61 (96)	.31 (86)	.88 (96)	−.45 (96)	−.52 (96)	.15* (96)	−.23 (96)
Log of income polarization	.53 (96)	.35 (86)	.88 (96)	−.28 (96)	−.45 (96)	.23 (96)	−.12* (96)
Standard deviation of the log (1 + years of education)	.31 (103)	−.25 (91)	.62 (103)	−.53 (103)	−.74 (103)	.41 (103)	−.07* (103)
C. Additional control variables used in GMM regressions:							
Police per 100,000 population	−.18* (97)	.38 (92)	.03* (97)	.06* (97)	−.00* (97)	.31 (97)	.25 (97)
Young male (15–29-year-olds) population as share of total	.36 (106)	.06 (94)	.58 (106)	−.36 (106)	−.43 (106)	.48 (106)	−.04* (106)
Ethnolinguistic fraction in 1960	.20* (96)	−.12* (83)	−.03* (96)	−.07* (96)	.19* (96)	.03* (96)	−.14* (96)
Latin America dummy	.59 (148)	.49 (132)	.74 (148)	−.49 (148)	−.58 (148)	−.01* (148)	−.13* (148)

Note.—GMM: generalized method of moments. Numbers of 5-year-average observations are in parentheses.
* Not significant at the .05 level.

APPENDIX C

On the Empirical Implementation of Esteban and Ray's Measure of Polarization

In this Appendix, we briefly describe a possible empirical implementation of the measure of polarization proposed by Esteban and Ray[50] (hereafter ER). More precisely, we propose an implementation of ER's equation (3), extended to incorporate the possibility of identification between individuals belonging to different income groups.

We use data on the percentages of total income held by different quintiles of the distribution of income within a given country. We thus consider a population that is initially subdivided into five groups (the quintiles). Since we do not have information on the degree of income heterogeneity within each quintile, we assume that they are equally homogeneous and thus treat each quintile as having the same degree of "identification" (as defined by ER).

Following the suggestion contained in Section 4 of ER, we also permit "identification across income groups that are 'sufficiently' close."[51] We implement this idea by assuming that two or more quintiles may group themselves into a new unit if their incomes are sufficiently similar. As emphasized by ER, the definition of the "domain over which a sense of identification prevails"[52] cannot be specified a priori. Thus, we test with different values of the minimum logarithmic difference (D) that gives rise to the merger of two quintiles into a new group. In our empirical exercise, this minimum (percentage) distance is allowed to vary between 10 and 100 percent.

We also assume that individuals act as "social climbers": when a given quintile is within the range of identification with both a quintile with higher income and a quintile with lower income, the merger takes place first between the two "superior" quintiles. Moreover, once two (or more) quintiles have merged, the decision to form a larger group with another quintile rests on the quintile with the highest income within the (pre-) existing grouping. That is, the new merger takes place only if the new "candidate" is within the range of identification of the highest quintile within the previously existing group.

In practice, given our assumptions, there are 16 (or 2 to the fourth power) possible structures of groups, each formed by one or more quintiles: either the highest quintile merges with the fourth or not, either the fourth quintile (with or without the fifth) merges or with the third or not, and so on. On the basis of the Deininger and Squire international data set on income inequality,[53] we apply a simple algorithm that implements our assumptions and determines, for each country and for each value of the parameter D, the types of groups that are expected to emerge.

Once the structure of groups in a given country is defined, we calculate, for each year and country, the value of a measure of polarization Q, using a modified version of ER's equation (3). We assume that the degree of identification of a group depends positively on its size and negatively on the log difference between the average income

[50] Esteban & Ray, *supra* note 21, at 834.

[51] *Id*. at 846.

[52] *Id*.

[53] Deininger & Squire, *supra* note 18.

of the two quintiles that, within the group, are situated farthest away from each other:

$$Q(\pi, y) = \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\frac{\pi_i}{1 + |y_i^{\max} - y_i^{\min}|} \right)^{1+\alpha} \pi_j |y_j - y_j|,$$

where y_i is the log of the average income of group *i* (formed by one or more quintiles), $y_i^{\max}$ is the average income of the highest quintile within group *i*, $y_i^{\min}$ is the average income of the lowest quintile within group *i*, and π_i is twice the number of quintiles that form group *i* (the number of deciles that were merged to create group *i*). Following the analysis in ER, we allow the parameter α to vary between 1.0 and 1.6.

Our preliminary exercises show that the measure of polarization increases in *D* for sufficiently low values of this parameter and then decreases in *D*. The value of *D* after which *Q* starts to decrease with *D*, in turn, increases in α. As expected, the correlation between *Q* and the Gini index decreases with the parameter α and varies from .74 for α equal to 1 to .58 for α equal to 1.6.

APPENDIX D

Dynamic Panel Generalized Method of Moments Methodology

A. Assuming No Unobserved Country-Specific Effects: Moment Conditions

We use a dynamic model to explain the homicide and robbery rates. The basic model is given by

$$y_{i,t}^{*} = \alpha y_{i,t-1}^{*} + \beta' X_{i,t} + \xi_{i,t}, \tag{D1}$$

where y^* is the "true" crime (homicide or robbery) rate, *X* is the set of explanatory variables, and ξ is the unexplained residual. The subscripts *i* and *t* denote country and time period, respectively.

Available crime data suffer from measurement error. For this section, let us assume that measurement error is only standard random noise (we relax this assumption below). Then

$$y_{i,t} = y_{i,t}^{*} + \nu_{i,t} \quad \text{and} \quad \nu_{i,t} \text{ is i.i.d.}, \tag{D2}$$

where *y* represents the measured crime rate. Substituting equation (D2) into (D1),

$$y_{i,t} = \alpha y_{i,t-1} + \beta' X_{i,t} + \varepsilon_{i,t} \quad \text{where} \quad \varepsilon_{i,t} = \xi_{i,t} + \nu_{i,t} - \alpha \nu_{i,t-1}. \tag{D3}$$

Equation (D3) is our basic regression model. Estimation via OLS would lead to inconsistent parameter estimates because the explanatory variables are not independent with respect to the error term: $y_{i,t-1}$ is correlated by construction with $\nu_{i,t-1}$, and *Xi,t* is potentially correlated with *ξi,t*. Consistent estimation requires the use of instrumental variables. Specifically, we use the GMM estimators developed for dynamic models of panel data that were introduced by Douglas Holtz-Eakin, Whitney Newey, and Harvey S. Rosen,[54] Arellano and Bond,[55] and Arellano and Bover.[56] Given that for this section we assume that there is no country-specific effect, we base our estimates on the so-called levels GMM estimator. The use of instruments is required

[54] Douglas Holtz-Eakin, Whitney Newey, & Harvey S. Rosen, Estimating Vector Autoregressions with Panel Data, 56 Econometrica 1371 (1988).

[55] Arellano & Bond, *supra* note 33.

[56] Arellano & Bover, *supra* note 34.

to deal with both the random noise measurement error in the lagged dependent variable and the likely endogeneity of the remaining explanatory variables, X, which may be affected by crime rates (reverse causation) and/or jointly caused by other variables (simultaneity). Instead of assuming strict exogeneity of X (for example, that the explanatory variables are uncorrelated with the error term at all leads and lags), we allow for a limited form of simultaneity and reverse causation. Specifically, we adopt the more flexible assumption of weak exogeneity, according to which current explanatory variables may be affected by past and current realizations of the dependent variable (the homicide or the robbery rate) but not by its future innovations. Under the assumptions that (1) the error term, ε, is not serially correlated and (2) the explanatory variables are weakly exogenous, the following moment conditions apply:

$$E(y_{i,t-s} \times \varepsilon_{i,t}) = 0 \quad \text{for } s \geq 2 \tag{D4}$$

and

$$E(X_{i,t-s} \times \varepsilon_{i,t}) = 0 \quad \text{for } s \geq 1. \tag{D5}$$

B. Allowing and Controlling for Unobserved Country-Specific Effects: Moment Conditions

Our second specification allows for the existence of persistent country-specific measurement error. This alternative model is given by

$$y^*_{i,t} = \alpha y^*_{i,t-1} + \beta' X_{i,t} + \eta_i + \xi_{i,t}, \tag{D6}$$

where y^* is the "true" crime rate and ηi is a country-specific unobserved factor, which may be correlated with the explanatory variables. We now assume that the mismeasurement in crime rates is driven not only by random errors but most important by specific and persistent characteristics of each country. These characteristics can be related to the variables that explain crime rates, such as the average level of income, educational attainment, and income inequality. Then we model measurement error as the sum of random noise and a country-specific effect:

$$y_{i,t} = y^*_{i,t} + \nu_{i,t} + \psi_i, \tag{D7}$$

where ν is i.i.d. and y is a country-specific effect. Substituting equation (D7) into (D6)

$$y_{i,t} = \alpha y_{i,t-1} + \beta' X_{i,t} + \mu_i + \varepsilon_{i,t}, \tag{D8}$$

where

$$\mu_{i,t} = \eta_i + (1-\alpha)\psi_i \qquad \text{and} \qquad \varepsilon_{i,t} = \nu_{i,t} - \alpha\nu_{i,t-1} + \xi_{i,t}.$$

Thus, the measurement error in crime rates is subsumed into both the unobserved country-specific effect and the time-varying residual. Equation (D8) is our second regression model. To estimate it we use the so-called system GMM estimator, which joins in a single system the regression equation in both differences and levels, each with its specific set of instrumental variables.

For ease of exposition, we discuss each section of the system separately, although the actual estimation is performed using the whole system jointly. Specifying the regression equation in differences allows direct elimination of the country-specific effect. First-differencing equation (D8) yields

$$y_{i,t} - y_{i,t-1} = \alpha(y_{i,t-1} - y_{i,t-2}) + \beta'(X_{i,t} - X_{i,t-1}) + (\varepsilon_{i,t} - \varepsilon_{i,t-1}). \tag{D9}$$

In addition to the likely endogeneity of the explanatory variables, X, and the random measurement error of the lagged crime rate, the use of instruments is here required to deal with the correlation that, by construction, is generated between the new error term, $(e_{i,t} - e_{i,t-1})$, and the differenced lagged dependent variable, $(y_{i,t-1} - y_{i,t-2})$. Once again, we adopt the assumption of weak exogeneity, which together with the assumption of no serial correlation in the error term yields the following moment conditions:

$$E[y_{i,t-s} \times (\varepsilon_{i,t} - \varepsilon_{i,t-1})] = 0 \quad \text{for } s \geq 3 \tag{D10}$$

and

$$E[X_{i,t-s} \times (\varepsilon_{i,t} - \varepsilon_{i,t-1})] = 0 \quad \text{for } s \geq 2. \tag{D11}$$

The GMM estimator simply based on the moment conditions in equations (D10) and (D11) is known as the differences estimator. Although asymptotically consistent, this estimator has low asymptotic precision and large biases in small samples, which leads to the need to complement it with the regression equation in levels.[57]

For the regression in levels, the country-specific effect is not directly eliminated but must be controlled for by the use of instrumental variables. The appropriate instruments for the regression in levels are the lagged differences of the corresponding variables if the following assumption holds. Although there may be correlation between the levels of the right-hand-side variables and the country-specific effect, there is no correlation between the differences of these variables and the country-specific effect. This assumption results from the following stationarity property:

$$E(y_{i,t+p} \times \eta_i) = E(y_{i,t+q} \times \eta_i) \quad \text{and} \quad E(X_{i,t+p} \times \eta_i) = E(X_{i,t+q} \times \eta_i)$$

$$\text{for all } p \text{ and } q.$$

Therefore, the additional moment conditions for the second part of the system (the regression in levels) are given by the following equations:[58]

$$E[(y_{i,t-s} - y_{i,t-s-1}) \times (\eta_i + \varepsilon_{i,t})] = 0 \quad \text{for } s = 2 \tag{D12}$$

[57] C. Alonso-Borrego & M. Arellano, Symmetrically Normalized Instrumental Variable Estimation Using Panel Data (CEMFI Working Paper No. 9612, 1996), and R. Blundell & S. Bond, Initial Conditions and Moment Restrictions in Dynamic Panel Data Models, 87 J. Econ. 115 (1998), show that when the lagged dependent and the explanatory variables are persistent over time, lagged levels of these variables are weak instruments for the regression equation in differences. This weakness has repercussions on both the asymptotic and small-sample performance of the differences estimator. As persistence increases, the asymptotic variance of the coefficients obtained with the differences estimator rises (for example, deteriorating its asymptotic precision). Furthermore, Monte Carlo experiments show that the weakness of the instruments produces biased coefficients in small samples. This is exacerbated with the variables' over-time persistence, the importance of the country-specific effect, and the smallness of the time-series dimension. An additional problem with the simple differences estimator relates to measurement error: differencing may exacerbate the bias because of errors in variables by decreasing the signal-to-noise ratio (Zvi Griliches & J. Hausman, Errors in Variables in Panel Data, 31 J. Econ. 93 (1986)). Blundell & Bond, *supra*, suggests the use of the system estimator as put forth in Arellano & Bover, *supra* note 34, which reduces the potential biases and imprecision associated with the traditional differences estimator.

[58] Given that lagged levels are used as instruments in the differences specification, only the most recent difference is used as an instrument in the levels specification. Other lagged differences would result in redundant moment conditions (Arellano & Bover, *supra* note 34).

and

$$E[(X_{i,t-s} - X_{i,t-s-1}) \times (\eta_i + \varepsilon_{i,t})] = 0 \quad \text{for } s = 1. \tag{D13}$$

C. Estimation

Using the moment conditions presented in equations (D4) and (D5) and, alternatively, (D10)–(D12), and following Arellano and Bond[59] and Arellano and Bover,[60] we employ a GMM procedure to generate consistent estimates of the parameters of interest and their asymptotic variance-covariance. These are given by the following formulas:

$$\hat{\theta} = (\bar{X}'Z\hat{\Omega}^{-1}Z'\bar{X})^{-1}\bar{X}'Z\hat{\Omega}^{-1}Z'\bar{y} \tag{D14}$$

and

$$\text{AVAR}(\hat{\theta}) = (\bar{X}'Z\hat{\Omega}^{-1}Z'\bar{X})^{-1}, \tag{D15}$$

where θ is the vector of parameters of interest (α, β), $\bar{y}$ is the dependent variable (stacked first in differences and then in levels in the case of the system estimator), X is the explanatory variable matrix including the lagged dependent variable (y_{t-1}, X) (also stacked first in differences and then in levels for the system estimator), Z is the matrix of instruments derived from the moment conditions, and Ω is a consistent estimate of the variance-covariance matrix of the moment conditions.[61]

D. Specification Tests

The consistency of the GMM estimators depends on whether lagged values of the explanatory variables are valid instruments in the crime rate regression. We address this issue by considering two specification tests suggested by Arellano and Bond[62] and Arellano and Bover.[63] The first is a Sargan test of overidentifying restrictions, which tests the overall validity of the instruments by analyzing the sample analog of the moment conditions used in the estimation process. Failure to reject the null hypothesis gives support to the model. The second test examines the null hypothesis that the error term $\varepsilon i,t$ is not serially correlated. As in the case of the Sargan test, the model specification is supported when the null of no serial correlation is not rejected. In our levels (basic) specification, we test whether the error term is first-order serially correlated. In our system (alternative) specification, we test whether the differenced error term (that is, the residual of the regression in differences) is second-order serially correlated. First-order serial correlation of the differenced error term is expected even if the original error term (in levels) is uncorrelated, unless the

[59] Arellano & Bond, *supra* note 33.

[60] Arellano & Bover, *supra* note 34.

[61] In practice, Arellano & Bond, *supra* note 33, suggests the following two-step procedure to obtain consistent and efficient GMM estimates. First, assume that the residuals, *e*, *i*, and *t*, are independent and homoskedastic both across countries and over time. This assumption corresponds to a specific weighting matrix that is used to produce first-step coefficient estimates. Then construct a consistent estimate of the variance-covariance matrix of the moment conditions with the residuals obtained in the first step and use this matrix to reestimate the parameters of interest (for example, second-step estimates). Asymptotically, the second-step estimates are superior to the first-step ones insofar as efficiency is concerned.

[62] Arellano & Bond, *supra* note 33.

[63] Arellano & Bover, *supra* note 34.

latter follows a random walk. Second-order serial correlation of the differenced residual indicates that the original error term is serially correlated and follows a moving average process at least of order 1. This would reject the appropriateness of the proposed instruments (and would call for higher-order lags to be used as instruments).

Bibliography

Alonso-Borrego, C., and Arellano, M. "Symmetrically Normalised Instrumental Variable Estimation Using Panel Data." Working Paper No. 9612. Madrid: CEMFI, 1996.

Arellano, Manuel, and Bond, Stephen. "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." *Review of Economic Studies* 58 (1991): 277–97.

Arellano, Manuel, and Bover, Olympia. "Another Look at the Instrumental Variable Estimation of Error-Components Models." *Journal of Econometrics* 68 (1995): 29–51.

Barro, Robert, and Lee, Jong-Wha. "New Measures of Educational Attainment." Unpublished manuscript. Cambridge, Mass.: Harvard University, 1996.

Becker, Gary S. "Crime and Punishment: An Economic Approach." *Journal of Political Economy* 76 (1968): 169–217.

Behrman, Jere R., and Craig, Steven G. "The Distribution of Public Services: An Exploration of Local Government Preferences." *American Economic Review* 77 (1987): 37–49.

Blundell, R., and Bond, S. "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models." *Journal of Econometrics* 87 (1998): 115–43.

Bourguignon, François. "Crime as a Social Cost of Poverty and Inequality: A Review Focusing on Developing Countries." Unpublished manuscript. Washington, D.C.: World Bank, Development Economics Research Group, 1998.

Bourguignon, François. "Crime, Violence, and Inequitable Development." In *Annual World Bank Conference on Development Economics 1999*, edited by Boris Pleskovic and Joseph E. Stiglitz, pp. 199–220. Washington, D.C.: World Bank, 2000.

Braithwaite, John. *Inequality, Crime, and Public Policy*. London and Boston: Routledge & Kegan Paul, 1979.

Collier, Paul, and Hoeffler, Anke. "On the Economic Causes of Civil War." *Oxford Economic Papers* 50 (1998): 563–73.

De Gregorio, José, and Lee, Jong-Wha. "Education and Income Distribution: New Evidence from Cross-Country Data." Unpublished manuscript. Santiago: Universidad de Chile; and Seoul: Korea University, 1998.

Deininger, Klaus, and Squire, Lyn. "A New Data Set Measuring Income Inequality." *World Bank Economic Review* 10 (1996): 565–91.

Easterly, William, and Levine, Ross. "Africa's Growth Tragedy: Policies and Ethnic Divisions." *Quarterly Journal of Economics* 112 (1997): 1203–50.

Ehrlich, Isaac. "Participation in Illegitimate Activities: A Theoretical and Empirical Investigation." *Journal of Political Economy* 81 (1973): 521–65.

Esteban, Joan-Maria, and Ray, Debraj. "On the Measurement of Polarization." *Econometrica* 62 (1994): 819–51.

Fajnzylber, Pablo; Lederman, Daniel; and Loayza, Norman. *Determinants of Crime Rates in Latin America and the World*. Washington, D.C.: World Bank, 1998.

Fajnzylber, Pablo; Lederman, Daniel; and Loayza, Norman. "Crime and Victimization: An Economic Perspective." *Economia* 1 (2000): 219–302.

Fajnzylber, Pablo; Lederman, Daniel; and Loayza, Norman. "What Causes Violent Crime?" *European Economic Review* 46 (2002):1323–57.

Fleisher, Belton M. "The Effect of Income on Delinquency." *American Economic Review* 56 (1966): 118–37.

Glaeser, Edward L.; Sacerdote, Bruce; and Scheinkman, Jose A. "Crime and Social Interactions." *Quarterly Journal of Economics* 111 (1996): 507–48.

Griliches, Zvi, and Hausman, J. "Errors in Variables in Panel Data." *Journal of Econometrics* 31 (1986): 93–118.

Grogger, Jeffrey. "Market Wages and Youth Crime." *Journal of Labor Economics* 16 (1998): 756–91.

Holtz-Eakin, Douglas; Newey, Whitney; and Rosen, Harvey S. "Estimating Vector Autoregressions with Panel Data." *Econometrica* 56 (1988): 1371–95.

Imrohoroglu, A.; Merlo, A.; and Rupert, P. "On the Political Economy of Income Redistribution and Crime." *International Economic Review* 41 (2000): 1–25.

Kelly, Morgan. "Inequality and Crime." *Review of Economics and Statistics* 82 (2000): 530–39.

Loayza, Norman, et al. "A World Savings Data-base." Unpublished manuscript. Washington, D.C.: World Bank, Policy Research Department, 1998.

Mauro, Paolo. "Corruption and Growth." *Quarterly Journal of Economics* 110 (1995): 681–712.

Mocan, H. Naci, and Rees, Daniel I. "Economic Conditions, Deterrence and Juvenile Crime: Evidence from Micro Data." Working Paper No. 7405. Cambridge, Mass.: National Bureau of Economic Research, 1999.

Pradhan, Menno, and Ravallion, Martin. "Demand for Public Safety." World Bank Policy Research Working Paper No. 2043. Washington, D.C.: World Bank, 1999.

Sen, Amartya. *On Economic Inequality*. Oxford: Clarendon Press, 1973.

Stack, Steven. "Income Inequality and Property Crime: A Cross-National Analysis of Relative Deprivation Theory." *Criminology* 22 (1984): 229–57.

Summers, Robert, and Heston, Alan. "The Penn World Table (Mark 5): An

Expanded Set of International Comparisons, 1950–1988." *Quarterly Journal of Economics* 106 (1991): 327–68.

Tauchen, Helen; Witte, Ann Dryden; and Griesinger, Harriet. "Criminal Deterrence: Revisiting the Issue with a Birth Cohort." *Review of Economics and Statistics* 76 (1994): 399–412.

Witte, Ann Dryden. "Estimating the Economic Model of Crime with Individual Data." *Quarterly Journal of Economics* 94 (1980): 57–84.