Crime and Social Expenditure: A Political Economic Approach

Carlos Bethencourt
Universidad de La Laguna*

December 2014

Abstract

High-income and less unequal societies are associated with both lower rates of economic crimes and larger public programs to deter crime. This paper suggests that redistributive effects generated by a crime-control program contribute to explain these facts. Retirees are beneficiaries of the system since they are mostly victims of crimes. Rich young agents, in spite of paying more taxes, become net receivers since they devote relatively less time to criminal activities than poorer ones; thus, a regressive intra-generational redistributive effect arises. All in all, a crime-control program is politically supported by a coalition of high-income young agents and retirees.

Keywords: crime, overlapping generations model, inequality, voting.
JEL Classification: H53, D72, K42.

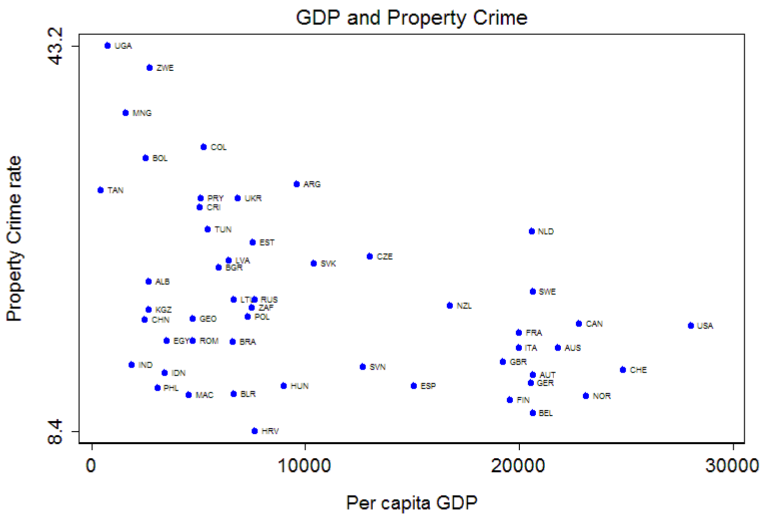
*Carlos Bethencourt, CAERP, Universidad de La Laguna, Campus de Guajara s/n, 38071, La Laguna, Tenerife, tel: +34-922-317954, fax: +34-922-317204, cbethenc@ull.es. I thankfully acknowledge CESifo's financial sponsorship for participation at the Conference on Law and Economics. I thank participants in the Conference and Law and Economics for useful comments. I also thank participants in PET 12 for valuable discussions. Finally, I am grateful to Costas Azariadis, José-Víctor Ríos-Rull, Alessandro Cigno, Lars Kunze, Jochen Mierau, Virginia Sánchez-Marcos and Fernando Perera-Tallo for helpful comments.

1. Introduction

Crime is one of the most deeply entrenched afflictions for societies. The lack of security affects individuals' decisions and has considerable welfare implications. Abundant literature documents the negative effect of criminality on quantity and quality of life (Soares, 2010); investment in physical and human capital, and so in growth (Murphy and Topel, 2003 and Lorentzen et al., 2007); consumption (Field, 1998 and Deadman, 2003); employment (Phillips et al., 1972 and Yamada, 1985); etc. Thus, it is not hard to understand why countries have been concerned with implementing crime-control policies, which have become a basic part of the package of social policies. This paper proposes a theory to explain how the size of the crime sector and the size of the crime-control policy are both determined simultaneously. We study the main factors which explain the large differences that exist among countries in the strength of their law enforcement and the level of their crime rate.

This paper is in the tradition of papers by Becker (1968), Ehrich (1973) and, more recently, İmrohoroğlu, Merlo and Rupert (2006), which have focused on the economic reasons for criminal behavior as opposed to the psychological aspects that are specific to criminals. As a consequence, in our analysis we disregard violent crimes and focus on economic crimes, that is, non-violent crimes motivated by the prospect of direct pecuniary gain. In particular, we consider as criminal activities all the activities which imply the appropriation of uncompensated value from others improperly (without permission, with intentional deceptions, etc.), for example, property crime, fraud, rent-seeking activities, etc.

Contrary to criminology literature which usually cites a positive link between crime and development, recent empirical studies have found systematic negative correlations between various measures of income and crime rates (Bourguignon, 1999 and Soares, 2004). For example, calculations from Soares and Naritomi (2010) show that regions with higher GDP per capita, such as North America and Western Europe also display lower burglary and theft rates. We analyze in detail the pattern of the property crime rate of a large sample of countries. Figure 1.1 shows the result. Clearly, we observe that a negative relationship between property crime and per capita GDP exists with the poorest countries reporting the highest levels of property crime.



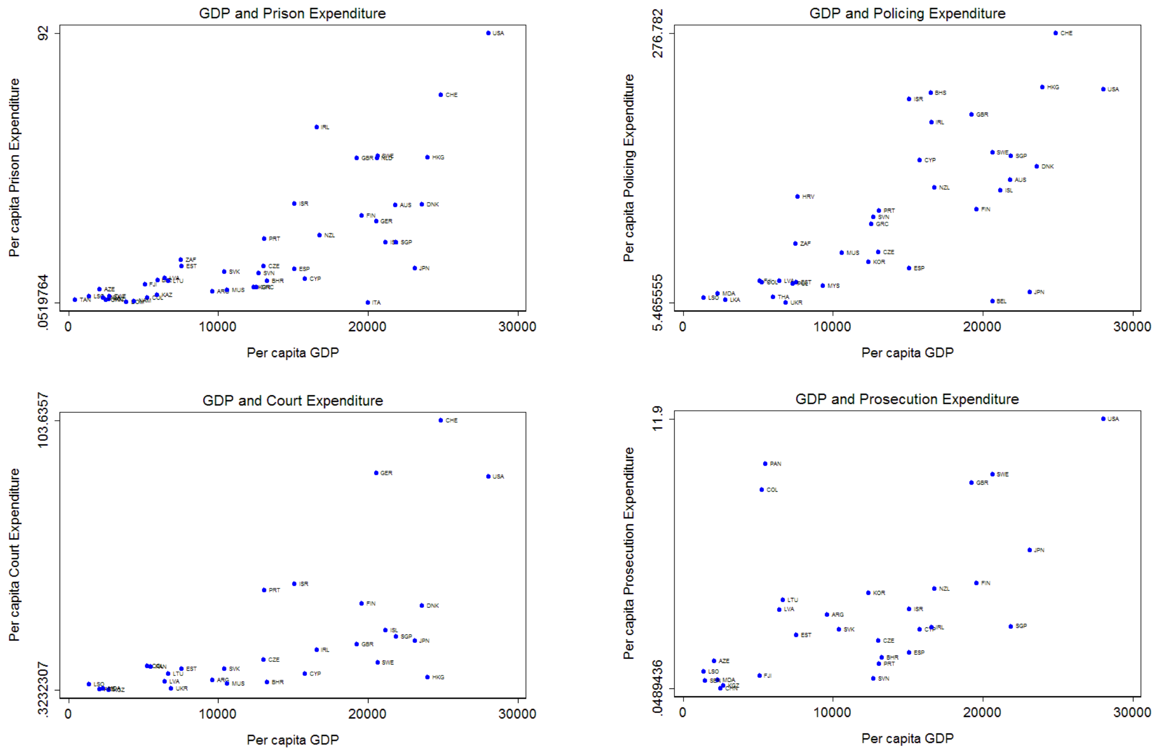
Source: Property crime rate is measured as the combination of the burglary rate and the theft rate (thefts of bicycles, motorcycles and other personal thefts). Data source of crime rates is the International Crime Victim Survey (ICVS), years 1989, 1992 and 1996/7 (see Soares and Naritomi, 2007, for a detailed discussion about the ICVS dataset). Per capita GDP for each country (1996 equivalent dollars) is calculated as the time series average in the range of the years available for the ICVS (data are from the Penn World Table 6.1).

Figure 1.1: **Per capita GDP and Property Crime rate**

Similarly, there have been many studies in the literature which evaluate the strength of the relationship between the crime rate and the effectiveness of the crime control policies. Policies designed by governments to deter crime include incarceration of offenders, effective policing and judicial systems, etc. Empirical evidence indicates that there exists a strong inverse relationship between both the levels of policing and incarceration and the level of criminality (for example, Levitt, 1996, 1997, and 2002, Di Tella and Schargrodsky, 2004, Soares and Naritomi, 2010 and Draca, Machin and Witt, 2011). İmrohoroğlu, Merlo and Rupert (2004) conclude that a higher probability of arrest is one of the most important factors that account for the observed drop in property crime in the US between 1980 and 1996. In a subsequent analysis, İmrohoroğlu, Merlo and Rupert (2006) find that small differences in the probability of arrest can generate quantitatively significant differences in property crime rates across similar environments.

Then, we considered the pattern of expenditure on criminal policies among countries. Previous research has identified a strong relationship between a country's economic welfare and both its expenditure on criminal justice (Newman and Howard, 1999) and policing

(Farrell et al., 2001). In a more recent contribution, Farrell and Clark (2004) find that criminal justice expenditure levels are significantly tied to levels of national income. They provide substantial statistical evidence of a strong relationship between the level of available public money and several measures of crime control policies. Figure 1.2 shows the relationship between the per capita GDP and expenditure on public policing, courts, prosecution and prisons for a large sample of countries. We confirm that the fiscal effort on crime-control is different among countries. The graphs show that higher levels of per capita GDP are associated with higher levels of expenditure for all types of crime-control spending considered.



Source: Per capita GDP for each country (1996 equivalent dollars) is calculated as the time series average for the range of years available for the ICVS (data are from the Penn World Table 6.1). The primary data source on expenditure (national currency) on public policing, courts, prosecution and prisons are obtained from the last release of the Sixth United Nations Survey of Crime Trends and Criminal Justice Systems, 1997. We obtain per capita ($1996) expenditures using the International Monetary Fund (IMF, International Financial Statistics, 2006) exchange rates and the total population for each country from the World Bank database (World Development Indicators, 2009).

Figure 1.2: **Per capita GDP and expenditure on the crime-control program**

Finally, we look at social characteristics associated with crime. Current empirical literature points out the demographical structure of population as one of the most relevant variables in explaining crime patterns. First, it is observed that most crimes are committed by young individuals (see for instance Freeman, 1996, Grogger, 1998, Levitt, 1998 and Mello and Schneider, 2007). Crimes committed by elderly are very rare, according to Plass (2007) usually less than 1 percent of arrestees are aged 65 or older. Second, economic crimes imply the improper taking of objects of economic value from others, mainly properties or accumulated resources [1]. And third, the action of accumulating resources is a lifetime process, so older individuals are usually wealthier than younger, which means that the elderly become natural victims of economic crimes [2].

This paper explains the differences in crime rates and levels of expenditure on crime-control policies observed among countries through the socioeconomic characteristics of these countries. We propose a novel approach which simultaneously considers economic and demographic dimensions of crime. All the empirical facts commented above are used to build a stylized model which allow us to understand the relationship between the level of the crime rate and the amount of crime-control expenditure when both variables are determined endogenously.

We consider a dynamically efficient overlapping generation economy with storage technology. The economy is populated by a large number of heterogenous agents who differ with respect to their labor market productivity. Each young agent decides the amount of time involved in legitimate market activities and in illegitimate activities. Illegitimate (criminal) activities are targeted at the appropriation of accumulated resources from others. Old individuals are retired. Agents value consumption in young and old age. The welfare state collects a proportional income tax on the young labor income, which finances a crime-control program to deter criminality. The crime-control program expenditure is determined in a majority voting game by all agents alive at every election. Voters rationally anticipate that taxes affect the allocation of time devoted to legitimate and il-

[1] All definitions of property crimes (which includes, among others, burglary, larceny, motor vehicle theft, etc.) and white-collar crimes (like embezzlement, bank fraud, etc.) involve the act of appropriating assets or properties, i.e., real or financial wealth. Thus, considering crimes as damaging wealth instead of income is a more suitable assumption.

[2] Tabulations of the U.S. census bureau show that the typical U.S. household headed by a person age 65 or older had a wealth (measured as the value of financial assets) 8 times greater than a household headed by someone under 35 in 2010.

legitimate activities, partly through their impact on criminals' profits. As a consequence, the crime-control program is treated as an intergenerational program. We consider there is a commitment device to future policies so that young voters have an incentive to support the crime-control program.

We show that the level of the expenditure on crime-control depends on both age and income dimensions. First, as old agents do not pay taxes and are the only ones who accumulate resources, they become net receivers of the crime-control program. Second, whereas highly-productive workers pay more taxes than the less productive workers, they devote relatively less time to criminal activities, which implies that their incomes are less damaged by the crime-control program. This implies that highly-productive workers also become net receivers of the program. Consequently, a regressive intra-generational redistributive effect arises and a winning coalition in the majority voting game of highly-productive young workers and old individuals is constituted. We find that the greater the number of agents who have accumulated resources (old agents) and the higher the productivity of young workers in the top tail of the productivity distribution function, the higher the tax rate is to finance the crime-control program. We show that, if there is a sufficiently large proportion of elderly in the population, then a highly-productive young worker median voter arises. This is an interesting result which contrasts with the political economic literature of intergenerational programs, in which the typical result is a poor young median voter and where an increase in proportion of elderly just impoverishes the median voter [3].

Additionally, we find that the tax rate to finance crime-control expenditure increases with the mean of the wage distribution, while it decreases with wage inequality. The crime rate, however, is lower when the mean of the wage distribution is greater and the wage inequality is lower. Finally, we show that an improvement in crime control technology that increases the effectiveness of public crime control in deterring criminals increases the crime-control tax rate and reduces the crime rate.

Since the seminal contribution of Becker (1968), economic literature on crime has been

[3]See for example Tabellini (2000) for social security, Bossi and Gumus (2011) for social security and income mobility, Arawatari and Ono (2011) for social security and public goods, Kinai (2011) for social security and public investment, Bethencourt and Galasso (2008) for social security and health care and Conde-Ruiz and Galasso (2003) for social security and early retirement. For models of three or more generations, these results could change (see for example Kthenbrger, Poutvaara and Profeta, 2007).

focused on analyzing the main economic determinants of crime and thus, the effectiveness of different policies in deterring crime (see among others Ehrlich, 1973, and Polinsky and Shavell, 1984, Ehrlich, 1981, Benoit and Osborne, 1995, Persson and Siven, 2006, and İmrohoroğlu, Merlo and Rupert 2004, 2006). Nevertheless, there have been fewer attempts to explicitly model the simultaneous determination of both the crime rate and the size of the crime control system: Demougin and Schwager (2003) find that if the government only maximizes the utility of high income individuals, then there is negative trade-off between law enforcement expenditures and criminality; whereas, if the government considers the whole population, then the correlation between the same variables may appear positive, which is in contrast to the empirical evidence shown above. İmrohoroğlu, Merlo and Rupert (2000) calibrate a general equilibrium model to assess the relative efficacy of alternative crime-control policies in reducing the U.S. property crime rate. They find a counterintuitive result that an increase in inequality implies an increase in the police expenditure in a median voter setting. Moreover, none of these papers have considered the key role of the age variable in shaping the crime patterns and so, the intertemporal treatment of crime.

This paper proceeds as follows: section 2 describes the model. Section 3 discusses the voting game, and the equilibrium concept, while section 4 characterizes the politico-economic equilibria. Section 5 analyzes the effects on the crime rate and the crime-control tax of an improvement in the mean productivity of agents. Similarly, section 6 analyzes the effects on the crime rate and the crime-control tax of an improvement in the productivity of crime-control program, while section 7 studies the effect of an increase in the income inequality. Section 8 concludes. All proofs are in the appendix.

2. The Economic Model

We propose an overlapping generation model with storage technology. Every period, two generations of non-altruistic agents are alive: young and old. Every young agent survives until the second period. Population is assumed to grow at a constant rate $\eta > 0$.

A simple storage technology allows the transfer of the consumption good from one period to another: a unit of the consumption good today is transformed into $(1+R)$ units of tomorrow's consumption good. R is the exogenously given rate of return on capital.

All private transfers of resources take place through this storage technology. Additionally, we assume that $R > \eta$, and thus the economy is dynamically efficient.

Individuals are endowed with one unit of labor time in their young age, and retire in their old age. Young individuals can spend time working in the market (legitimate sector), l, and/or engaging in criminal activities (illegitimate sector), z [4]. Thus,

$$1 = l + z.$$

For criminal activities, we consider all the activities which imply the appropriation of uncompensated value from others improperly, that is, without permission, without making any contribution to productivity or with intentional deceptions. We include property crimes (burglaries, thefts, etc.), fraud (health fraud, false billing, etc.) and rent-seeking activities.

Agents are assumed to be heterogeneous in their market productivity, w, which is distributed on the support $[\underline{w}, \overline{w}] \subset \Re_{+}$, according to the cumulative distribution function $G(.)$. Nevertheless, they are assumed to be homogeneous in their ability to commit a crime [5]. Thus, an individual born at time t is characterized by a productivity level, and will therefore be denoted by $w_t \in [\underline{w}, \overline{w}]$ where $\underline{w} > 1$. The distribution of abilities is assumed to have mean $\widetilde{w}$, and to be skewed

$$\int_{\underline{w}}^{\overline{w}} w dG(w) = \widetilde{w}, \quad G(\widetilde{w}) > \frac{1}{2}.$$

The labor income of an agent born at time t who works in the market is denoted by wl_t, where w denotes the wage rate (equal to the agent's productivity level), and l_t denotes the number of hours spent working [6]. The non-market (illegitimately obtained) income of an agent born at time t who is engaged in criminal activities is denoted by $b_t(1 - \alpha g_t) z_t^{1/2}$, where b_t is the average level of wealth in the economy owned by the generation born

[4] In our framework, old agents retire not only from labor legitimate activities (as the standard 2-OLG models) but they also retire from engaging in criminal activities. In fact, it is well documented that most crimes are committed by young individuals (see Freeman, 1996, Grogger, 1998, Levitt, 1998 and Mello and Schneider, 2007). For instance, İmrohoroğlu, Merlo and Rupert (2004) find that in 1980, 47.7 percent of all people arrested for property offenses were 18 years of age or younger.

[5] There is no empirical evidence of a clear positive correlation between the market productivity and the skills to commit a crime, that is, individuals with high productivity are not more productive in the crime sector. An important reason for this is the huge variety of activities included in the crime sector, for example, activities ranging from tax fraud to car theft (see for example Lochner, 2004 and Lochner and Moretti, 2004).

[6] If $l = 1, \forall w$, the aggregate labor income in the economy will be equal to $\widetilde{w}$. Otherwise, the labor income in the market will be less than $\widetilde{w}$.

at time $t-1$. For technical simplicity, we have considered the standard assumption in the literature which implies that each criminal steals a fraction $(1-\alpha g_t)\, z_t^{1/2}$ of average wealth in the economy [7]. We call this the "theft fraction". This fraction depends on g_t, the amount of public expenditure devoted to control crime (crime-control program) [8], $\alpha > 1$ which is a measure of the crime-control program's productivity, and z_t denotes the number of the agent's hours spent on performing and producing the offenses [9]. We are assuming that there are diminishing marginal returns on these hours, which is a very reasonable assumption because of the given level of wealth [10].

We assume that each individual is equally sensitive to being the victim of a crime [11]. Thus, the income of a retired agent born at time t is composed of the capitalized wealth, $b_{t+1}(1+R)$, minus the fraction of wealth that is stolen by criminal offenders, which is $b_{t+1}(1-\alpha g_{t+1})\, \widetilde{z}_{t+1}^{1/2}$, where $\widetilde{z}_{t+1}$ denotes the aggregate number of hours in the economy in period $t+1$ spent on producing offenses. Thus, the share of wealth appropriated by a criminal offender depends just on the time that she/he devotes to commit offenses, whereas that the share of wealth lost by a victim depends on the aggregated criminal activity. The larger is the number of criminal offenders, the larger will be the loss that victims face due to the criminal activity. We also consider that there exists diminishing marginal returns on time in the aggregated criminal hours as it happens at the individual level.

Agents are assumed to value consumption in young and old age. Preferences are additively time separable. Consumption in the first period of life enters in a linear utility function and in a logarithmic function as agents get older. The linearity of the utility function in young-age consumption is imposed for convenience. This does simplify the

[7] See İmrohoroğlu, Merlo and Rupert (2004, 2006).

[8] The crime-control program includes all the programs that contribute to deterring the level of criminality. Some of them have a direct impact on reducing the level of crime (for example, police protection), while others have an indirect effect by reducing the incentives to commit a crime (for example, the strength of the legal institutions and courts, correctional facilities, etc.).

[9] A sufficient condition to guarantee that the theft fraction is less or equal to one is $1 > \widetilde{w}\alpha$. Given that the aggregate labor income in the market is less than $\widetilde{w}$, public revenue would be always less than this amount and so would the public expenditure on fighting against the crime.

[10] If there are many gamblers who get in each other's way, then the marginal product of each unit of input will decline as the amount of that input increases.

[11] There exists a significant lack of data about crime incidence by income status of the victims. Thus, we assume that they have the same probability of being a victim of a crime. Our consideration takes into account that, prima facie high income individuals are the most appealing preys. Nevertheless, due to the use of measures for avoiding crimes is widespread among high income individuals (living in a safe neighborhoods, technologies which shield them from crimes, more information about tax procedures, etc.), a reduction in the probability of being a victim takes places. This could overcome the starting high level and hence, the average number of offenses could be equalized among different types of agents at the end.

computations by allowing for all income effects to be assimilated by young-age consumption. Thus, the lifetime utility function is

$$U\left(c_t^t, c_{t+1}^t\right) = c_t^t + \rho \log\left(c_{t+1}^t\right) \tag{2.1}$$

where c is consumption, and $\rho \in (0, 1)$ is the subjective discount factor. Subscripts indicate the calendar time and superscripts indicate the period when the agent was born.

Young agents allocate their disposable income between current consumption and savings. Given these assumptions, the budget constraint at time t of a type-e agent born at time t is

$$c_t^t + b_{t+1} = wl_t(1 - \tau_t) + b_t\left(1 - \alpha g_t\right) z_t^{1/2} \tag{2.2}$$

where τ_t is the tax rate at time t on the young legitimate income to finance the crime-control program.

Similarly, the budget constraint at time $t+1$ of a type-e agent born at time t is

$$c_{t+1}^t = b_{t+1}(1 + R) - b_{t+1}\left(1 - \alpha g_{t+1}\right) \tilde{z}_{t+1}^{1/2}. \tag{2.3}$$

The young agent's problem at time t is choosing the optimum labor supply (in each sector in the economy) and consumption to maximize utility. Formally, the economic problem for the young type-w is:

$$\max_{c_t^t, c_{t+1}^t, z_t} \quad c_t^t + \rho \log\left(c_{t+1}^t\right)$$

$$\begin{aligned} \text{s.t.} \quad c_t^t + b_{t+1} &= w(1 - z_t)(1 - \tau_t) + b_t\left(1 - \alpha g_t\right) z_t^{1/2} \\ c_{t+1}^t &= b_{t+1}(1 + R) - b_{t+1}\left(1 - \alpha g_{t+1}\right) \tilde{z}_{t+1}^{1/2} \end{aligned}$$

It is easy to see that the unique solution to the agent's problem is:

$$z_{w,t}^* = \left[\frac{b_t\left(1 - \alpha g_t\right)}{2w(1 - \tau_t)}\right]^2 \tag{2.4}$$

$$c_{w,t}^{*t} = w(1 - z_t)(1 - \tau_t) + b_t\left(1 - \alpha g_t\right) z_t^{1/2} - \rho \tag{2.5}$$

$$c_{w,t+1}^{*t} = \rho\left(R + 1 - \left(1 - \alpha g_{t+1}\right) \tilde{z}_{t+1}^{1/2}\right). \tag{2.6}$$

Moreover, the optimal demand for savings is $b_{w,t+1}^* = \rho$ [12]. Notice that the supply of

[12] The utility function defined in eq. 2.1 implies that savings are independent of the young income, which could be relevant for the political sustainability of the crime-control program. However, in our framework, high income individuals are political supporters of the crime-control program which benefits them. Thus, a more realistic scenario, with young high income agents displaying higher savings, it just would reinforce the political support of high income individuals. As a consequence, the results of this paper would remain unchanged.

criminal hours is decreasing in the individual market productivity and in the crime-control public spending, while it is increasing in the average wealth in the economy at time t.

We can calculate the aggregate fraction of population involved in the criminal sector, $\widetilde{z}_t$, as the average time that young agents spend on criminal activities, this is,

$$\widetilde{z}_t = \int_{\underline{w}}^{\overline{w}} z^*_{w,t}\, dG(w) = \phi_1 \left[\frac{b_t (1-\alpha g_t)}{2(1-\tau_t)}\right]^2, \qquad \phi_1 = \int_{\underline{w}}^{\overline{w}} \frac{1}{w^2}\, dG(w). \tag{2.7}$$

where the pattern of the crime sector is similar to the individual criminal behavior: it is decreasing in the crime-control public spending and increasing in the average wealth in the economy.

2.1. The Crime-Control Program

At time t, every young individual contributes a proportion, τ_t, of her labor income to the welfare state, and the total revenue is used to finance crime-control expenditure, g_t, which is employed to protect the old agents' wealth. In this sense, the crime-control program works like an intergenerational redistributive system.

In this model, the crime sector encompasses a variety of different activities targeted at the appropriation of uncompensated value from others with intentional deceptions or without people's freely-given consent. These activities are categorized into property crime, fraud and rent-seeking activities, etc.

The crime-control program refers to all the characteristics which shape law enforcement in a society. We include mechanisms that reduce the probability of carrying out a crime successfully, such as police control, as well as, the institutional elements that undermine the incentives to commit a crime, such as the severity of penalties and punishments, the duration of judicial process, the strictness of correctional facilities, etc. The technology of the law enforcement depends on the expenditure on it. In the model, if public crime-control expenditure increases, then the probability of success of a crime, measured as the share of stolen resources, decreases [13].

[13] The arrest technology is not explicitly modeled here, and we simply assume that the higher the expenditure on the crime-control program, the lower the number of successful criminals are and so, the lower the share of stolen wealth is. Likewise we disregard any consideration of the punishment of arrested criminals. Articles by İmrohoroğlu, Merlo and Rupert (2000) and Pyle (1983) deal with specifications of the apprehension technology. Articles by Becker (1968), Eehrlich (1973) and Polinsky and Shavell (1984) among others, model explicitly the punishment of convicted criminals.

The welfare state is assumed to be balanced every period, so that its total expenditure has to be equal to the amount of taxes collected, this is,

$$g_t = \tau_t \int_{\underline{w}}^{\overline{w}} w \left(1 - z_{w,t}^*\right) \, dG(w) = \tau_t \left[\widetilde{w} - \widetilde{z}_t \left(\frac{\phi_2}{\phi_1}\right)\right] \tag{2.8}$$

where $\phi_2 = \int_{\underline{w}}^{\overline{w}} \frac{1}{w} \, dG(w)$. The total expenditure on the crime-control program is decreasing in the share of the economy's time devoted to the illegitimate sector, z_t: if the number of agents involved in criminal activities increases, then the total amount of taxable resources is reduced, and so are the taxes collected[14].

An interesting question to answer is the effect of the tax rate on collected taxes. In this environment, the Laffer curve might not have the typical inverted U-shaped behavior. The reason is that the distortional effect of the taxes could be overcome by an increase in the incentives to participate in the legitimate sector. An increase in the tax rate produces three effects: First, the taxes collected increase, thus increasing the expenditure on the crime-control program and the number of hours devoted to the legitimate sector. Second, the disposable income of the legitimate sector decreases and, thus, the number of hours that agents devote to undertake criminal activities increases. And third, if the number of hours in the labor market decreases, then the total amount of taxable resources falls, thus reducing the expenditure on the crime-control program. Therefore, the return on criminal activities increases and so do the number of hours that agents spend on the illegitimate sector. If the crime-control program is strong enough, then the first effect prevails and thus, the fraction of population involved in the illegitimate sector decreases, and the crime-control expenditure would be monotone increasing in the tax rate.

Lemma 2.1. *If $1/\alpha(\widetilde{w} + \frac{\phi_2}{\phi_1}) \geq \tau$ then, the size of the criminal sector is decreasing in the tax rate, $\frac{\partial \widetilde{z}_t}{\partial \tau_t} < 0$, and the crime-control public expenditure is increasing in the tax rate, $\frac{\partial g_t}{\partial \tau_t} > 0$.*

This lemma guarantees that if the (legitimate) income tax rate, τ, is small enough, then the tax rate has a positive effect on the fraction of population working in the legitimate

[14] Once we have defined and calculated the aggregate variables we can guarantee the existence of an interior solution for the agent's problem, i.e., $z_{w,t}^* \in (0,1), \ \ \forall w$ (defined by eq. 2.4). We know that $b_{w,t+1}^* = b_t = \rho$ and $\rho/2w < 1$. If we assume that $\alpha \left[\widetilde{w} - \frac{\phi_2}{\phi_1}\right] \geq 1$ then, $(1 - \alpha g_t) \leq (1 - \tau_t)$ and thus, $z_{w,t}^* \in (0,1), \ \ \forall w$.

sector, as well as on crime-control public expenditure[15]. We therefore restrict the support of the tax rate according to this lemma.

2.2. The Economic Equilibrium

We can now define the economic equilibrium as follows:

Definition 2.2. *For a given sequence of tax rates, and real interest rates,* $\{\tau_t, R\}_{t=0}^{\infty}$, *and a distribution function* $G(.)$, *an economic equilibrium is a sequence of allocations,* $\left\{c_t^t(w), c_{t+1}^t(w), z_t(w), l_t(w), b_{t+1}(w)\right\}_{w\in[\underline{w},\overline{w}]}^{t=0,\ldots,\infty}$, *such that:*

- in every period, agents maximize their utility function at eq. 2.1 , with respect to $c_t^t(w), c_{t+1}^t(w), z_t(w), l_t^t(w)$ and $b_{t+1}(w)$, subject to the budget constraints at eqs. 2.2 and 2.3;
- the welfare budget constraint is balanced every period, and thus equation 2.8 is satisfied; and
- the goods market clears every period:

$$\int_{\underline{w}}^{\overline{w}} c_t^{t-1} dG(w) + (1+\eta)\int_{\underline{w}}^{\overline{w}} c_t^t dG(w) = (1+\eta)\int_{\underline{w}}^{\overline{w}} w(1-z_t)\, dG(e)$$
$$+(1+\eta)\int_{\underline{w}}^{\overline{w}} b_t\left(1-\alpha\tau_t\left[\widetilde{w}-\widetilde{z}_t\left(\tfrac{\phi_2}{\phi_1}\right)\right]\right) z_t^{1/2} dG(e) \qquad \forall t.$$

The utility level obtained by the agents in an economic equilibrium can be represented by their indirect utility functions. For this purpose, we can use the welfare state budget constraints to obtain an indirect utility function for a type-w young individual that depends on current and future tax rates:

$$\begin{gathered} v_{t,w}^t(\tau_t, \tau_{t+1}, w) = w(1-z_t^*(\tau_t))(1-\tau_t) + \\ b_t\left(1-\alpha\tau_t\left[\widetilde{w}-\widetilde{z}_t(\tau_t)\left(\frac{\phi_2}{\phi_1}\right)\right] z_t^{*1/2}(\tau_t) - \rho\right) + \\ \rho\log\left(\rho\left[R+1-\left(1-\alpha\tau_{t+1}\left[\widetilde{w}-\widetilde{z}_{t+1}(\tau_{t+1})\left(\frac{\phi_2}{\phi_1}\right)\right]\right)\widetilde{z}_{t+1}^{1/2}(\tau_{t+1})\right]\right) \end{gathered} \tag{2.9}$$

For a type-w old individual at time t the indirect utility function is:

$$v_{t,w}^{t-1}(\tau_t, w) = \rho\log\left(\rho\left[R+1-\left(1-\alpha\tau_t\left[\widetilde{w}-\widetilde{z}_t(\tau_t)\left(\frac{\phi_2}{\phi_1}\right)\right]\right)\widetilde{z}_t^{1/2}(\tau_t)\right]\right) \tag{2.10}$$

[15]Note that if the tax rate was high enough, then the direct distortive effect of an increase in the tax rate could overcome the incentives to devote less time to criminal activities (due a large crime-control program) and thus, it would increase the share of time involved in the crime sector, i.e., $\frac{\partial\widetilde{z}_t}{\partial\tau_t} > 0$.

3. The Voting Game

The size of the welfare state is decided by agents through a political system of majoritarian voting. Elections take place every period, and everyone, young and old, cast a ballot on the income tax rate, τ. Individual preferences on this issue are represented by the indirect utility functions in equations 2.9 and 2.10, for the young and the old, respectively. Notice that every agent has zero mass, and thus no individual vote could change the outcome of the election. We, thus, assume that individuals vote sincerely. In other words, voters choose their ideal policies according to their true preferences, independently of any strategic consideration.

Notice that this majoritarian voting game is intrinsically dynamic, since it describes the interaction between successive generations of workers and retirees. Thus, in absence of a commitment device to future policies, young voters have no incentive to support any intergenerational transfer scheme. However, they may expect their current voting behavior to affect future voters' decisions. Bethencourt and Galasso (2008), Conde-Ruiz and Galasso (2003) and Conde-Ruiz and Galasso (2005), among others, show that for political economy models of social security and public health care, a system of punishment and rewards exists, which generates the equilibrium outcome of the game with commitment as a subgame perfect equilibrium outcome of the game without commitment.

To characterize the equilibria of the voting game, we analyze, for simplicity, the case of full commitment, in which voters determine the constant sequence of the parameter of the welfare state (τ). Hence, we drop the subscripts t for ease of exposition, unless it is considered necessary. Current voters, in this case, are assumed to commit to future policies in a once-and-for-all voting game.[16] In the absence of a state variable, this voting game is static, and thus the median voter theorem can be applied to obtain sufficient conditions for a equilibrium to exist. In particular, if preferences are single-peaked throughout the issue space, a sufficient condition for τ^* to be an equilibrium of the voting game with full commitment is that τ^* represents the outcome of majority voting.[17]

Thus, to apply the median voter theorem to our environment, we need to ensure that individuals' preferences are single peaked. The following lemma describes a set of sufficient

[16] See Boadway and Wildasin (1989) and Bénabou and Ok (2001) for details.

[17] See Persson and Tabellini (2000) for a simple explanation of how to apply the median voter theorem.

conditions.

Lemma 3.1. *If* $\frac{\underline{w}}{\alpha\widetilde{w}} / \left(1 - \frac{1}{\overline{w}\phi_1^{1/2}}\right) \geq \rho$, *individuals' preferences are single-peaked over* τ.

We therefore restrict the support of the subjective discount factor according to the lemma 3.1.

4. Voting on the Size of the Welfare State

We now proceed to examine individual votes on τ. Voters cast a ballot on a constant sequence of τ. Votes are then ordered to identify the median vote, which represents the political equilibrium outcome of the voting game with commitment.

Regardless of the composition of the welfare state, the elderly are net recipients of the system. Therefore, they will choose the tax rate that maximizes its size.

Lemma 4.1. *The most preferred tax rate by any old individual,* τ_E^O, *is equal to* $\frac{1}{\alpha(\widetilde{w}+\frac{\phi_2}{\phi_1})}$.

By setting the tax rate to the maximum possible, old retirees are just maximizing their consumption since: first, they do not work and so, do not contribute to financing the crime-control program; second, they do not commit illegitimate activities and so, they are not damaged by the crime-control program, which reduces the gains of committing a crime. Finally, a larger crime-control program increases their (current level of) disposable wealth. Thus, all of them vote for the maximum τ.

Since voters only determine current policies, which may be changed at no cost in the future, young individuals may not be willing to support a crime-control program that is established to protect old individuals' wealth. Young individuals may be willing to vote in favor of the welfare state, and thus to bear the cost of a current program, if their vote will also determine its future size, and thus their future level of protection against crime. In the game with commitment, a type-w young individual chooses her vote, τ_w^Y, by maximizing her indirect utility function at eq. 2.9 with respect to a constant sequence of tax rates, $\tau_t = \tau_{t+1} = \tau_w^Y$. The next lemma characterizes the vote of the young.

Lemma 4.2. *The most preferred tax rate by any type-w young individual is positive,* $\tau_w^Y > 0$, *and increasing in* w, $\frac{\partial \tau_w^Y}{\partial w} > 0$, *if* $\frac{2\left(\frac{\overline{w}}{\rho}\right)^2 + \frac{1}{2}}{\left[\widetilde{w} - \phi_2\left(\frac{\rho}{2}\right)^2\right]} < \alpha$ *and* $R+1 < \frac{\rho\phi_1^{1/2}}{2}\left[1 + \frac{1}{\frac{1}{\alpha\rho}\frac{\epsilon(\underline{w})}{\epsilon(\widetilde{w})} + \frac{\rho}{2\underline{w}}}\right]$, *where* $\epsilon(\underline{w}) = \underline{w}(1 - z_t^*(0))$ *and* $\epsilon(\widetilde{w}) = \widetilde{w} - \left(\frac{\phi_2}{\phi_1}\right)\widetilde{z}_t(0)$, *the legitimate income of the poorest and the average legitimate income evaluated at* $\tau = 0$ *respectively.*

Lemma 4.2 suggests that political support for the welfare state relies heavily on its ability to deter criminal activities. High income individuals, $w > \widetilde{w}$, pay more taxes than average ones, but they enjoy the same amount of crime-control services as everybody else. [18]. However, the crime-control program reduces the gains of committing a crime in the first age. Given the fact that high income individuals are specialized in the legitimate sector, the crime-control program hurts them less than poor individuals. If the crime-control program is strong enough then this positive effect of deterring criminal activities would prevail on the contributive effect and, thus, high income individuals would become net receivers of the crime-control program. On the contrary, low-income types, $w < \widetilde{w}$, while contributing less to finance the crime-control program could become worse off because of the significant reduction in gains coming from the criminal sector. Therefore, the higher the legitimate market productivity, the higher the gains from the crime-control program and the higher the demand for tax rates to finance the crime-control program. In other words, a regressive intra-generational redistributive effect of the crime-control program arises. This is a significant result that contrasts with the typical findings in the literature, where the redistributive effect, if it exists, always prevails and poor individuals are the high demanding agents for tax rates [19].

4.1. Characterization of Equilibria

In the previous section, we analyzed the voting behavior of all individuals on the size of the welfare state. Since preferences are single peaked, we can now apply the median voting theorem, and characterize the equilibria of the game.

[18] If we consider only these two effects, then high income individuals would become net contributors to the welfare state and they would always vote $\tau_w^Y = 0$. This is the typical intragenerationanl redistributive effect that appears in PAYG Social Security modelling which explains why high income individuals vote zero social security when the economy is dynamically efficient. See Conde-Ruiz and Galasso (2005) and Persson and Tabellini (2000) for a detailed explanation.

[19] In our framework, crime is considered as taking of wealth. Notice that if we allow criminals affecting income through losses in labor income as well, then results would not change. In that case, high income workers would have more incentive to support the crime-control thereby, criminals are appropriating labor income and high income workers are obtaining more labor income than low income workers. They now would obtain relatively more benefits and would vote for higher tax rates. The regressive intra-generational redistributive effect of the crime-control program would increase.

Proposition 4.3. *There exists an equilibrium, τ^*, of the voting game with commitment, such that: $\tau^* = \tau_{w_m} \in (0,1)$, i.e., the most preferred tax rate of the median voter.*

Proof: It is now straightforward to order every agent's vote on the size of the welfare state, and to identify the median voter's type. Agents can be ranked according to their age and type, as shown in Figure 4.1, with the elderly and then the high-income young voting for larger welfare state sizes. The median voter is the type-w_m young agent who divides the electorate in halves: $G(w_m) = (2+\eta)/2(1+\eta)$. We identify her most preferred tax rate as τ_{w_m}. ■

An interesting feature of this equilibrium is that old retirees form a coalition with high-income young individuals. This is an uncommon result in the voting literature of intergenerational programs in 2-OLG models where the typical result is a dominant coalition made up of the oldest and poor young individuals [20]. The significance of this outcome is reflected at the time of determining the median voter. Given that the identity of the median voter depends on the proportion of old individuals in the population: a large proportion could imply that the median voter was a high-income young individual, i.e., the median voter could have a type, w, higher than the average, $\widetilde{w}$ (in this case, Figure 5.1 should be modified to display $\widetilde{w} < w_m$).

5. Productivity Increase in the Economy

This section considers an overall increase in workers' productivity and analyzes how the median voter decision would change. Productivity increases are reflected in the increase in the parameter $\widetilde{w}$ of the mean of abilities. We carry out this exercise keeping the type of the median voter unaffected.

Definition 5.1. *Let $F(.)$ the cumulative distribution of labor productivity after an increase in the mean has occurred. A median voter preserving identity implies:*

- $\int_{\underline{w}}^{\overline{w}} wdF(w) > \int_{\underline{w}}^{\overline{w}} wdG(w) = \widetilde{w}, \quad G(\widetilde{w}) > \frac{1}{2}, \quad F(\widetilde{w}) > \frac{1}{2}.$
- $G(w_m) = F(w_m) = (2+\eta)/2(1+\eta)\,.$

[20]See for example Bethencourt and Galasso (2008), Conde-Ruiz and Galasso (2005) and the seminal paper of Browning (1975). For three or more generations results could change (see for example Kthenbrger, Poutvaara and Profeta, 2007).

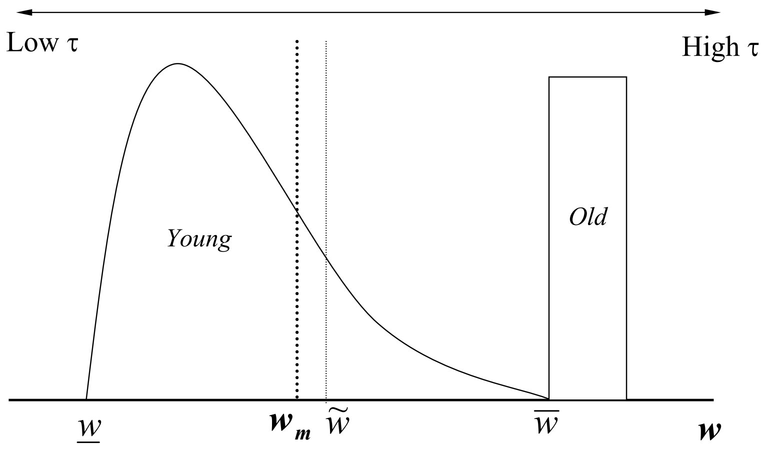


Figure 4.1: **Voting over τ**

The cumulative distribution $F(.)$ has a higher mean than the original distribution, $G(.)$. However, $F(.)$ has the same type-w young agent who divides the electorate into two identical halves as the original distribution, $G(.)$. Thus, the identity of the median voter remains unaffected.

An increase in the mean labor productivity, according to the above definition, implies an increase in public revenues of the welfare state and likewise, in crime-control spending. If the size of the crime-control program grows, then the theft fraction is reduced and thus, the returns from this program increase. As the previous exercise, this direct effect has an impact on young workers, who have more incentives to devote their time to working in the market instead of committing offenses in the criminal sector. This reallocation of labor in the economy reduces the criminal sector and increases the legitimate sector, implying an increase in the crime-control spending due to the larger income in the productive sector (and so, in public revenues). Therefore, these two positive effects on crime-control spending make this program more effective by deterring criminal activities and, therefore, more appealing to the median voter who will vote to increase its size. The next proposition qualifies this finding.

Proposition 5.2. *Consider the equilibrium τ^* in proposition 4.3. An increase in the average productivity of workers, $\widetilde{w}$, leads to an equilibrium τ^{**}, with a larger welfare* *state, $\tau^{**} > \tau^*$, and a smaller size of the criminal sector, $\widetilde{z}^{**} < \widetilde{z}^*$.*

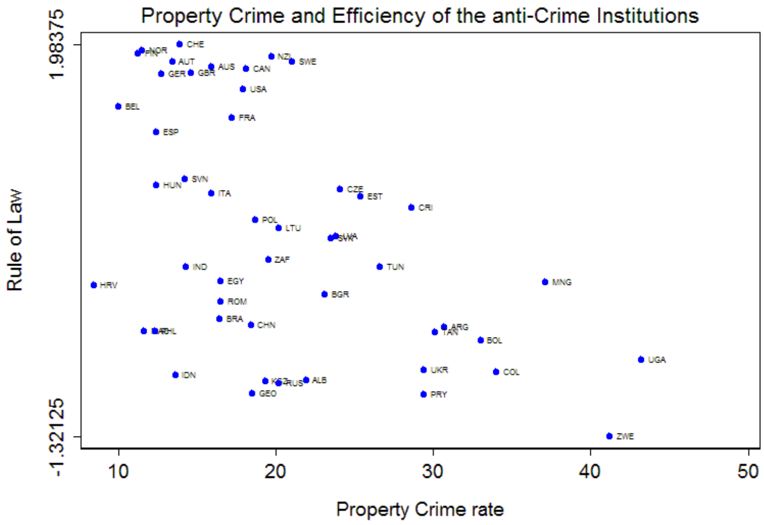
An implication of this proposition is that countries with higher per capita GDP should also be the countries that show larger expenditures on fighting crime and smaller criminal sectors. In this respect, recent studies in the literature find empirical evidence which widely support this result (see for example Graham and Farrell, 2004, Newman and Howard, 1999 and Soares and Naritomi, 2010 for a review). İmrohoroğlu, Merlo and Rupert (2004) identify the increase in the mean earnings distribution as one of the most important factors that account for the observed decline in property crime in US from 1980 to 1996. Also Figures 1.1 and 1.2 in section 1 showed that richer countries have larger crime-control programs and smaller criminal sectors.

6. Productivity Increase in the Crime-Control Program

This section analyzes the impact of an improvement in crime-control technology on the political decisions on the size of the program. Productivity increase is identified by an increase in the parameter α of theft fraction (see section 2), which determines the effectiveness of the crime-control program in deterring crime.

As crime-control spending becomes more efficient in deterring crime and, thus, in reducing the theft fraction, the returns from this program increase. At the same time, young individuals have more incentives to devote time to the legitimate market activities and less time to committing criminal activities. This reallocation of labor in the economy downsizes the criminal sector and increases the legitimate productive sector. As a consequence, the average taxable income in the economy increases and so do the public revenues for financing the crime-control program. Thus, the direct improvement in crime-control productivity added to the increase in public revenue make crime-control spending more effective in fighting against crime and, therefore, more appealing to the median voter, who will vote for increase its size. Productivity increases in the crime-control system will hence lead to more welfare spending. The next proposition qualifies this finding.

Proposition 6.1. *Consider the equilibrium τ^* in proposition 4.3. An increase in the productivity parameter of the crime-control program, α, leads to an equilibrium τ^{**}, with a larger welfare state, $\tau^{**} > \tau^*$, and a smaller criminal sector, $\widetilde{z}^{**} < \widetilde{z}^*$.*



Source: Property crime rate is measured as the combination of the burglary rate and the theft rate (thefts are bicycle, motorcycle and other personal thefts). Data source of crime rates is the International Crime Victim Survey (ICVS), years 1989, 1992 and 1996/7. Rule of Law statistics are the time series average for each country in the years available from 1996 to 2006 (data are from Kaufmann, Kraay and Mastruzzi, 2007).

Figure 6.1: **Property Crime rate and Productivity of the crime-control program**

This proposition has clear empirical implications: countries with reinforced crime-control systems should also be the countries that show larger expenditures on fighting crime and smaller criminal sector. In this respect, we have just observed in Figures 1.1 and 1.2 that countries with larger crime-control programs are also countries with smaller criminal sectors. On the other hand, Figure 6.1 shows the relationship between the productivity of the crime-control program and the level of criminality, for a selected sample of countries. Productivity of the crime-control program is measured using the variable *rule of law*[21]. The level of criminality is approximated by the property crime rate. Clearly, a negative correlation between the productivity of the crime-control policy and the crime rate is well defined. Due to the existent negative relationship between the crime rate and the

[21]The concept of *rule of law* is taken to refer to the manner in which public institutions treat and protect individuals. We use the Kaufmann, Kraay and Mastruzzi (2007) estimations. They construct the *Rule of law* using cross-country surveys of firms and individual and, expert assessments from NGOs, risk rating agencies and governments. Their measure captures in particular the quality of contract enforcement, property rights, the police, and the courts. The index is a standardized measure with range between -2.5 (weakest institutions) and 2.5 (strongest institutions).

expenditure on crime-control (Figures 1.1 and 1.2), we can conclude that, the higher the performance of the crime-control program is, the higher its expenditure is.

There also exists abundant literature that documents the positive effect of strong and effective policies on reducing the crime level. For example Levitt (1996), (1997) and (2002), and Di Tella and Schargrodsky (2004) have documented the crime reducing role of police presence and incarceration of criminals. Soares and Naritomi (2010) have documented the important quantitative role of police and incarceration rates in explaining the differences of reported crime rates among a selected sample of countries. More recently, Draca, Machin and Witt (2011) show that police patrols are a highly effective tool for cutting crime. They do not argue that increased police numbers should be the sole focus on crime-control but they suggest that if the police are resourced properly, the effects can be powerful.

7. A decrease in the earnings inequality

This section studies the effect of a decrease in earnings inequality and analyzes how the median voter decision changes. We analyze the decrease in earnings inequality through a decrease in labor market productivity inequality. We carry out this exercise considering a mean preserving spread.

Definition 7.1. *Let $F(.)$ the cumulative distribution of labor productivity after a decrease in inequality has occurred. A mean preserving spread implies:*

- $\int_{\underline{w}}^{\overline{w}} wdG\left(w\right) = \int_{\underline{w}}^{\overline{w}} wdF\left(w\right) = \widetilde{w}, \quad G\left(\widetilde{w}\right) > \frac{1}{2}, \quad F\left(\widetilde{w}\right) > \frac{1}{2}.$
- $G\left(w^{Gm}\right) = F\left(w^{Fm}\right) = \frac{1}{2}, \quad w^{Gm} < w^{Fm}$ *which implies* $G\left(w^{Fm}\right) > F\left(w^{Fm}\right).$

The cumulative distribution $F(.)$ has the same mean, $\widetilde{w}$, as the original distribution, $G(.)$. Moreover, $F(.)$ is skewed as $G(.)$ but deeper. Thus, the median of cumulative distribution $F(.)$ is greater than the median of distribution $G(.)$. In other words, the mass of low productivity individuals has decreased, while the mean productivity has remained constant.

Proposition 7.2. *Consider the equilibrium τ^* in proposition 4.3. A decrease in the labor productivity inequality which changes the cumulative distribution of types to $F(.)$, leads to an equilibrium τ^{**}, with a larger welfare state, $\tau^{**} > \tau^*$, and a smaller criminal sector, $\widetilde{z}^{**} < \widetilde{z}^*$.*

Proof: The proof is straightforward. The mean labor income in the economy does not change. The decrease in the inequality, due to a decrease in the mass of poor individuals, does change the identity of the median voter. The young median voter becomes richer, with a higher type-w. Due to the fact that the most preferred tax rate by any type-w young agent is positive (proposition 4.2), the richer median voter determines a higher tax rate in the economy. Finally, as we proved in lemma 2.1, an increase in the tax rate implies a decrease in the fraction of population involved in the criminal sector. ∎

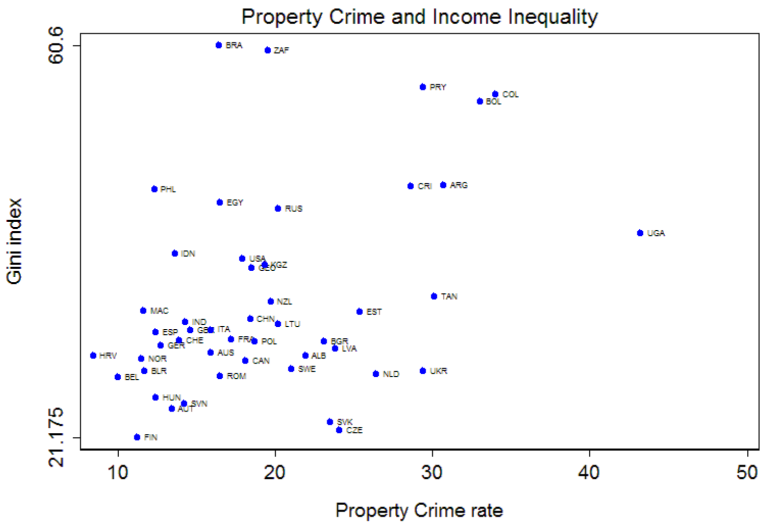
This proposition implies that countries with reduced (legitimate) income inequality should also be the countries that show larger expenditures on fighting against crime and smaller criminal sectors. The existent empirical cross-country literature on crime rates supports these findings, identifying inequality as one of the main economic determinant of crime [22]. For example, İmrohoroğlu, Merlo and Rupert (2006) find that income inequality is one of the most relevant variables explaining property crime rates. Furthermore, in a recent article Soares and Naritomi (2010) document the significant quantitative role of inequality in explaining crime rates among a selected sample of countries. We also document this fact. First, Figures 1.1 and 1.2 show that countries with larger crime-control programs are also the countries with smaller criminal sectors. Second, Figure 7.1 shows the relationship between income inequality and the size of the crime sector, for a selected sample of countries. Income inequality is measured through the Gini index for household income. We observe a positive relationship between income inequality and the crime level.

8. Concluding Remarks

This paper provides a theory to explain the differences in economic crime rates and levels of expenditure of crime-control policies observed among countries through the socioeconomic characteristics of these countries.

Despite of its simplicity, the model delivers a rich set of predictions which are consistent with empirical observations. In particular, we find that countries with both low rates of economic crimes and larger public expenditures devoted to deterring crime are associated

[22] See Bourguignon (1999), Fajnzylber, Lederman and Loayza (2002a) and (2002b), Bourguignon, Nuñez and Sanchez (2003), İmrohoroğlu, Merlo and Rupert (2004), Soares (2004) and more recently Soares and Naritomi (2010).



Source: Property crime rate is measured as the combination of the burglary rate and the theft rate (thefts are bicycle, motorcycle and other personal thefts). Data source of crime rates is the International Crime Victim Survey (ICVS), years 1989, 1992 and 1996/7. Gini index for each country is calculated as the time series average in the range of the years available for the ICVS. Gini indexes are referred to income (or, in some cases, earnings and consumption expenditure) among households. Data are obtained from the World Income Inequality Database (UN) which combines different datasets coming from government statistical agencies, World Bank, Luxembourg Income Study and authors' calculations.

Figure 7.1: **Property Crime rate and Inequality**

with high income levels, low income inequality and greatly strengthened crime-control programs.

A distinctive feature of our analysis is the intertemporal treatment of crime. We present a novel approach which simultaneously considers economic and age dimensions of crime. Formally, a stylized model in which younger agents devote time to criminal activities and older agents become victims of such crimes is presented. Results allow us to conclude that both income and age are key determinants.

Another interesting feature of the analysis is the regressive intra-generational redistributive effect that the crime-control program generates. Richer young agents pay more taxes than average, while enjoying the same amount of crime-control services as everybody else. If we only consider these two effects, then high income workers would become net contributors to the welfare state and would vote for zero taxes. Therefore, we would obtain the typical progressive intra-generational redistributive effect that appears in the literature of PAYG Social Security. However, the crime-control program also reduces the gains of committing a crime. As high income workers devote a relatively minor portion of time to criminal activities, such a program hams them relatively less than poor workers. Thus, if the crime-control program is strong enough, the positive effect of deterring criminal activities would prevail over the contributive effect and, thus, high income individuals would become net receivers of the crime-control program.

We show that public programs to deter crime may be supported by a voting coalition of of high income young individuals and retirees. Two features are crucial to this result. First, the regressive intra-generational redistribution component of the crime-control program benefits high income workers. The second is the political power of the old: unlike the young, the elderly constitute a fairly homogeneous group, they do not contribute to financing the crime-control program and they do not commit illegitimate activities. A larger crime-control program increases their (current level of) disposable wealth, which implies they constitute an uniform electoral block when voting on redistribution issues: the elderly all like the crime-control program and vote for it.

Observed behaviors of individuals are consistent with voting patterns described. High income individuals invest more in crime-control mechanisms such as alarms, private security or living in relatively good neighborhoods. Similarly, there is a wide literature which documents the high level of fear that the elderly suffer regarding being a victim of a crime, justifying why the elderly take more effort to protect themselves. In this sense, a recent contribution by Scarborough et al. (2010) shows that age and income status are positively related to fear of crime.

We view this paper as initial foray into developing a rich setting to assess the impact of different factors on crime rates. Intertemporal issues related to recidivism, life-cycle effects, stigma, long-term consequences of conviction, etc. should also be considered to develop a framework for evaluating dynamically optimal policies. Empirical efforts in obtaining accurate measures of different types of economic crimes are essential to carry out reliable quantitative analysis. We hope that more work will be done along these lines in the future.

Appendix

Proof of Lemma 2.1:

The fraction of hours devoted to the crime sector, $\widetilde{z}_t$, defined by eq. 2.7 depends on the crime-control expenditure, g_t, defined by eq. 2.8. We can the define the following implicit function $F = 2\left(1-\tau_t\right)\frac{\widetilde{z}_t^{1/2}}{\phi_t^{1/2}} - \rho\left[1-\alpha\left(\widetilde{w}-\left(\phi_2/\phi_1\right)\widetilde{z}_t\right)\tau_t\right]$ when we substitute g_t on $\widetilde{z}_t$. Then we obtain $\frac{\partial \widetilde{z}_t}{\partial \tau_t} = -\frac{\frac{\partial F}{\partial \tau_t}}{\frac{\partial F}{\partial \widetilde{z}_t}} < 0$, since $\frac{\partial F}{\partial \tau_t} = -\frac{\rho}{1-\tau_t}\left(1-\alpha\left[\widetilde{w}-\left(\frac{\phi_2}{\phi_1}\right)\widetilde{z}_t\right]\right) > 0$ (see footnote 14) and $\frac{\partial F}{\partial \widetilde{z}_t} = \frac{\rho}{2\widetilde{z}_t}\left[1-\alpha\left(\widetilde{w}+\left(\phi_2/\phi_1\right)\widetilde{z}_t\right)\tau_t\right] > 0$ if $1/\alpha(\widetilde{w}+\frac{\phi_2}{\phi_1}) \geq \tau_t$.

Then, it is easy to see that $\frac{\partial g_t}{\partial \tau_t} = \frac{g_t}{\tau_t} - \tau_t\left(\phi_2/\phi_1\right)\frac{\partial \widetilde{z}_t}{\partial \tau_t} > 0$, since $\frac{\partial \widetilde{z}_t}{\partial \tau_t} < 0$. ■

Proof of Lemma 3.1:

For a type-E old agent, since the first derivative w.r.t. τ is always positive, these agents simply prefer higher τ to lower τ, and preferences are still single peaked, with a maximum in $\tau = 1/\alpha(\widetilde{w}+\frac{\phi_2}{\phi_1})$.

For a type-E young agent, the second derivative w.r.t. τ of her indirect utility function, eq. 2.9, is

$$SOC_w^Y = \underbrace{\alpha\left[\widetilde{w}-\left(\frac{\phi_2}{\phi_1}\right)\widetilde{z}_t\right]\frac{\partial\gamma(w,\tau)}{\partial\tau}}_{1st.term} + \underbrace{\left[w-\alpha\widetilde{w}\gamma(w,\tau)\right]\left(\frac{1}{z_t^*}\right)\frac{\partial z_t^*}{\partial\tau}}_{2nd.term}$$

where z_t^* is defined by eq 2.4, $\widetilde{z}_t$ is defined by eq. 2.7 and $\gamma(w,\tau) = \left(\frac{\rho^2}{c_{t+1}^*} - \frac{1}{w\phi_1^{1/2}}\right)\rho\widetilde{z}_t^{-1/2}$; where c_{t+1}^* is defined by eq 2.6.

We study separately the signs of the two terms in the SOC_w^Y. It is easy to see that the first term is always negative, since the average labor income is always positive, $\left[\widetilde{w}-\widetilde{z}_t\left(\frac{\phi_2}{\phi_1}\right)\right] > 0$ and, since $\frac{\partial\gamma(w,\tau)}{\partial\tau} = \frac{\rho}{2}\left(\gamma(\tau)\widetilde{z}_t^{-1/2} + \frac{\rho^2}{c_{t+1}^{*2}}\left(1-\alpha g_t\right)\right)\frac{\partial\widetilde{z}_t}{\partial\tau} < 0$ by lemma 2.1. We now turn to the second term. From eqs. 2.4 and 2.7 we get that $\frac{\partial z_t^*}{\partial\tau} = \left(\frac{1}{\phi_1 w^2}\right)\frac{\partial\widetilde{z}_t}{\partial\tau} < 0$ by lemma 2.1. Similarly, $\left[w-\alpha\widetilde{w}\gamma(w,\tau)\right] > 0$ for $\frac{\left(\frac{w}{\alpha\widetilde{w}}\right)}{1-\frac{1}{\overline{w}\phi_1^{1/2}}} > \rho$. Thus, the second term is always negative and SOC_w^Y is always negative. ■

Proof of Lemma 4.1:

Trivial. The indirect utility function at eq. 2.10 is always positive, and thus, agents always vote for the maximum τ, this is, $\tau = \frac{1}{\alpha(\widetilde{w}+\frac{\phi_2}{\phi_1})}$. ■

Proof of Lemma 4.2:

Notice that the first order condition w.r.t. τ in the optimization problem of a type-w young voter is equal to the first order condition of a type-w old voter decreased by the distortive effects of the tax, i.e., $FOC_w^Y(\tau) = FOC_w^O(\tau) - w(1 - z_t^*) - \alpha \left[\widetilde{w} - \widetilde{z}_t \left(\frac{\phi_2}{\phi_1}\right)\right] \rho z_t^{*1/2}$.

From Lemma 3.1, since $\frac{w}{\alpha \widetilde{w}} / \left(1 - \frac{1}{\overline{w} \phi_1^{1/2}}\right) \geq \rho$, the indirect utility function is concave over τ, and thus a sufficient condition for richer young agents to vote for higher τ, i.e., $\frac{\partial \tau_w^Y}{\partial w} < 0$, is that the first order condition, evaluated at $\tau = 0$, is decreasing in w: $\frac{\partial FOC_w^Y(\tau=0)}{\partial w} < 0$. It is easy to see that the negative term at $FOC_w^Y(\tau = 0)$ (the distortive effects of the tax) is decreasing in w for $\frac{2\left(\frac{\overline{w}}{\rho}\right)^2 + \frac{1}{2}}{\left[\widetilde{w} - \left(\frac{\rho}{2}\right)^2 \phi_2\right]} < \alpha$ and, thus, $\frac{\partial FOC_w^Y(\tau=0)}{\partial w} < 0$, given that the first term, this is, $FOC_w^O(\tau = 0)$, does not depend on w.

The next step is to guarantee an interior solution for $\frac{w}{\alpha \widetilde{w}} / \left(1 - \frac{1}{\overline{w} \phi_1^{1/2}}\right) \geq \rho$ and a sufficient condition for a type-w young voter to maximize her indirect utility function in an interior, i.e., for $\tau > 0$, is that the first order condition, evaluated at $\tau = 0$, is strictly positive, $FOC_w^Y(\tau = 0) > 0$. Note that $FOC_{\underline{w}}^Y(\tau = 0) > 0 \iff \frac{\widetilde{w}\left(1 - \left(\frac{\rho}{2\widetilde{w}}\right)^2\right)}{\underline{w}\left(1 - \left(\frac{\rho}{2\underline{w}}\right)^2\right)} \alpha \gamma(\underline{w}, \tau = 0) > 1$, where $\gamma(\underline{w}, \tau = 0) = \frac{\rho^2}{2} \left(\frac{\phi_1^{1/2}}{R + 1 - \left(\frac{\rho}{2}\right)\phi_1^{1/2}} - \frac{1}{\underline{w}}\right)$. We rewrite the earnings rate between the average and the poorest as, $\frac{\widetilde{w}\left(1 - \left(\frac{\rho}{2\widetilde{w}}\right)^2\right)}{\underline{w}\left(1 - \left(\frac{\rho}{2\underline{w}}\right)^2\right)} = \frac{\epsilon(\widetilde{w})}{\epsilon(\underline{w})}$. It is easy to see $FOC_{\underline{w}}^Y(\tau = 0) > 0$ for $\frac{\rho \phi_1^{1/2}}{2} \left[1 + \frac{1}{\left(\frac{1}{\alpha \rho}\right)\left(\frac{\epsilon(\underline{w})}{\epsilon(\widetilde{w})}\right) + \frac{\rho}{2\underline{w}}}\right] > R + 1$. Thus, if $FOC_{\underline{w}}^Y(\tau = 0) > 0$ then, $FOC_w^Y(\tau = 0) > 0 \ \forall w$, given that we have guaranteed that $\frac{\partial FOC_w^Y(\tau=0)}{\partial w} < 0$ in the first part of the proof. ■

Proof of Proposition 5.2:

Let τ^* defined in proposition 4.3. We prove that $\frac{\partial \tau^*}{\partial \widetilde{w}} = \frac{\partial \tau^*}{\partial \widetilde{z}^*} \frac{\partial \widetilde{z}^*}{\partial \widetilde{w}} > 0$. First, it is easy to see $\frac{\partial \tau^*}{\partial \widetilde{z}^*} = 1 / \frac{\partial \widetilde{z}^*}{\partial \tau^*} < 0$ for lemma 2.1. We then prove $\frac{\partial \widetilde{z}^*}{\partial \widetilde{w}} < 0$. For that, se start from the implicit function F defined in proof of lemma 2.1. We obtain $\frac{\partial \widetilde{z}^*}{\partial \widetilde{w}} = -\frac{\frac{\partial F}{\partial \widetilde{w}}}{\frac{\partial F}{\partial \widetilde{z}^*}} < 0$, since $\frac{\partial F}{\partial \widetilde{w}} = \alpha \rho \tau_t > 0$ and $\frac{\partial F}{\partial \widetilde{z}^*} = \frac{\rho}{2\widetilde{z}_t}\left[1 - \alpha\left(\widetilde{w} + (\phi_2/\phi_1)\,\widetilde{z}_t\right)\tau_t\right] > 0$ if $1/\alpha(\widetilde{w} + \frac{\phi_2}{\phi_1}) \geq \tau_t$ (see proof 2.1.). ■

Proof of Proposition 6.1:

Let τ^* defined in proposition 4.3. We prove that $\frac{\partial \tau^*}{\partial \alpha} = \frac{\partial \tau^*}{\partial \widetilde{z}^*} \frac{\partial \widetilde{z}^*}{\partial \alpha} > 0$. First, it is easy to see $\frac{\partial \tau^*}{\partial \widetilde{z}^*} = 1/\frac{\partial \widetilde{z}^*}{\partial \tau^*} < 0$ for lemma 2.1. We then prove $\frac{\partial \widetilde{z}^*}{\partial \alpha} < 0$. For that, we start from the implicit function F defined in the proof of lemma 2.1. We obtain $\frac{\partial \widetilde{z}^*}{\partial \alpha} = -\frac{\frac{\partial F}{\partial \alpha}}{\frac{\partial F}{\partial \widetilde{z}^*}} < 0$, since $\frac{\partial F}{\partial \alpha} = \rho g_t > 0$ and $\frac{\partial F}{\partial \widetilde{z}^*} = \frac{\rho}{2\widetilde{z}_t} \left[1 - \alpha \left(\widetilde{w} + \left(\phi_2/\phi_1\right) \widetilde{z}_t\right) \tau_t\right] > 0$ if $1/\alpha(\widetilde{w} + \frac{\phi_2}{\phi_1}) \geq \tau_t$ (see proof 2.1.). ■

References

[1] Arawatari, R. and T. Ono (2011) “Old-age Social Security vs. Forward Intergenerational Public Goods Provision”, Discussion Papers in Economics and Business WP No. 11-26.

[2] Becker, G. S. (1968) “Crime and Punishment: An Economic Approach”, *Journal of Political Economy*, 76 (2), pp. 169-217.

[3] Bénabou, R., and E. A. Ok (2001) “Social Mobility and the Demand for Redistribution: The Poum Hypothesis”, *Quarterly Journal of Economics*, 116 (2), pp. 447-487.

[4] Benoit, J.P. and M.J. Osborne (1995) “Crime, Punishment, and Social Expenditure”, *Journal of Institutional and Theoretical Economics*, 151, pp. 326-347.

[5] Bethencourt, C. and V. Galasso (2008) “Political complements in the welfare state: health care and social security ”, *Journal of Public Economics*, 92 (3-4), pp. 609-632.

[6] Boadway, R. and D. Wildasin (1989) “A Median Voter Model of Social Security”, *International Economic Review*, 30 (2), pp. 307-328.

[7] Bossi, L and G. Gumus (2011) “Income Inequality, Mobility, and the Welfare State”, IZA Discussion Papers WP No. 5009.

[8] Bourguignon, F. (1999) “Crime, violence, and inequitable development”. In: Boris Pleskovic and Joseph Stiglitz (editors). Annual World Bank Conference on Development Economics 1999/2000, Washington DC, World Bank, 2000, pp. 199-220.

[9] Bourguignon, F., Nuñez, J. and F. Sanchez (2003) “A Structural Model of Crime and Inequality in Colombia”, *Journal of the European Economic Association*, 1 (2-3), pp. 440-449.

[10] Conde-Ruiz, J.I. and V. Galasso (2003) “Early retirement ”, *Review of Economic Dynamics*, 6 (1), pp. 12-36.

[11] Conde-Ruiz, J.I. and V. Galasso (2005) “Positive Arithmetic of the Welfare State”, *Journal of Public Economics*, 89 (5-6), pp. 933-955.

[12] Deadman, D. (2003) "Forecasting residential burglary", *International Journal of Forecasting*, l19 (4), pp. 567-578.

[13] Demougin, D. and R. Schwager (2003) "Law enforcement and crimminality: Europe vs. USA", *International Review of Law and Economics*, 23 (2), pp. 217-225.

[14] Di Tella, R. and E. Schargrodsky (2004) "Do Police Reduce Crime? Estimates Using the Allocation of Police Forces after a Terrorist Attack", *American Economic Review*, 94 (1), pp. 115-133.

[15] Draca, M., Machi, S. and R. Witt (2011) "Panic on the Streets of London: Police, Crime and the July 2005 Terror Attacks", *American Economic Review*, 101 (5), pp. 2157-2181.

[16] Ehrlich, I. (1973) "Participation in Illegitimate Activities: A Theoretical and Empirical Investigation", *Journal of Political Economy*, 81 (3), pp. 521-565.

[17] Ehrlich, I (1981) "On the Usefulness of Controlling Individuals: An Economic Analysis of Rehabilitation, Incapacitation, and Deterrence", *American Economic Review*, 71 (3), pp. 307-322.

[18] Fajnzylber, P., Lederman, D. and N. Loayza (2002a) "Inequality and violent crime", *Journal of Law and Economics*, 45 (1), pp. 1-40.

[19] Fajnzylber, P., Lederman, D. and N. Loayza (2002b) "What causes violent crime?", *European Economic Review*, 46 (7), pp. 1323-1357.

[20] Farrell, G. and K. Clark (2005) "What Does the World Spend on Criminal Justice?", HEUNI Paper Series No. 20. Helsinki, Finland: European Institute for Crime Prevention and Control.

[21] Farrell, G., Lane, E., Clark, K. and A. Tseloni (2001) "What does the world spend on policing?", *International Journal of Comparative Sociology*, 42 (1-2), pp. 59-73.

[22] Field, S. (1998) "Trends in crime revisited", Home Office Research Study. London: HMSO.

[23] Freeman, R.B., (1996) "Why Do So Many Young American Men Commit Crimes and What Might We Do About It?", *Journal of Economic Perspectives*, 10 (1), pp. 25-42.

[24] Grogger, J., (1998) "Market Wages and Youth Crime", *Journal of Labor Economics*, 16 (4), pp. 756-791.

[25] Kaufmann, D., Kraay, A. and M. Mastruzzi (2007) "Governance Matters VI: Governance Indicators for 1996-2006", World Bank Policy Research WP No. 4280.

[26] Kinai, Y. (2011) "Growth or Welfare State? Optimal Composition of the Government Expenditure", Osaka University, mimeo.

[27] Koethenbuerger, M., P. Poutvaara, and P. Profeta (2007) "Why are more redistributive social security systems smaller? A median voter approach", *Oxford Economic Papers*, 60 (2), pp. 275-292.

[28] İmrohoroğlu, A., Merlo, A. and M. Rupert (2000) "On the Political Economy of Income Redistribution and Crime", *International Economic Review*, 41 (3), pp. 1-25.

[29] İmrohoroğlu, A., Merlo, A. and M. Rupert (2004) "What accounts for the decline in crime?", *International Economic Review*, 45 (3), pp. 707-729.

[30] İmrohoroğlu, A., Merlo, A. and M. Rupert (2006) "Understanding The Determinants of Crime", *Journal of Economics and Finance*, 30 (2), pp. 270- 284.

[31] Levitt, S. D. (1997) "Using electoral cycles in police hiring to estimate the effect of police on crime", *American Economic Review*, 87 (3), pp. 270-290.

[32] Levitt, S. D. (1998) "Juvenile Crime and Punishment", *Journal of Political Economy*, 106 (6), pp. 1156-1185.

[33] Levitt, S. D. (1999) "The limited role of changing age structure in explaining aggregate crime rates", *Criminology*, 37 (3), pp. 581-598.

[34] Levitt, S. D. (2002) "Using electoral cycles in police hiring to estimate the effect of police on crime: Reply", *American Economic Review*, 92 (4), pp. 1244-1250.

[35] Lochner, L. (2011) "Education Policy and Crime: A Human Capital Approach", in *Controlling Crime: Strategies and Tradeoffs*, Cook, P.J., Ludwig, J. and J. McCrary eds. University Chicago Press, pp. 465-515.

[36] Lochner, L. (2004) "Education, Work And Crime: A Human Capital Approach", *International Economic Review*, 45 (3), pp. 811-843.

[37] Lochner, L. and E. Moretti (2004) "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports", *American Economic Review*, 94 (1), pp. 155-189.

[38] Lorentzen, P., McMillan, J. and R. Wacziarg (2007) "Death and Development", *Journal of Economic Growth*, 13 (2), pp. 81-124.

[39] Mello, J. and A. Schneider (2007) "Demographic Change and Homicides: The São Paulo Case", unpublished manuscript, PUC-Rio.

[40] Murphy, K. and R. Topel (2003) "The Economic Value of Medical Research", in: Kevin M. Murphy and Robert Topel (eds). Measuring the Gains from Medical Research: An Economic Approach. Chicago, University of Chicago Press, pp. 41-73.

[41] Newman, G. and G. Howard (1999) "Introduction: data sources and their use", in: Graeme Newman (editor). Global Report on Crime and Justice, published for the United Nations Center for International Crime Prevention. Oxford and New York: Oxford University Press.

[42] Persson, M. and C. Siven (2006) "Incentive and incarceration effects in a general equilibrium model of crime", *Journal of Economic Behavior and Organization*, 59 (2), pp. 214-229.

[43] Persson, T. and G. Tabellini (2000) "Political Economics. Explaining Economic Policy". The MIT Press.

[44] Phillips, L., Votey, H.L., and D. Maxwell (1972) "Crime, Youth and The Labor Market", *Journal of Political Economy*, 80 (3), pp. 491-504.

[45] Plass, P.S. (2007) "Age and Crime", in: Ritzer, G. (editor). Blackwell Encyclopedia of Sociology, Blackwell Publishing.

[46] Polinsky, A. M. and S. Shavell (1984) "The Optimal Use of Fines and Imprisonment", *Journal of Public Economics*, 24 (1), pp. 89-99.

[47] Pyle, D. J. (1983) "The Economics of Crime and Law Enforcement". New York: St. Martin Press.

[48] Scarborough, B. K., Like-Haislip, T. Z., Novak, K. J., Lucas, W. L. and L. F. Alarid (2010) "Assessing the relationship between individual characteristics, neighborhood context, and fear of crime ", *Journal of Criminal Justice*, 38 (4), pp. 819-826.

[49] Soares, R. (2004) "Development, crime, and punishment: Accounting for the international differences in crime rates", *Journal of Development Economics*, 73 (1), pp. 155-184.

[50] Soares, R. (2010) "Welfare Costs of Crime and Common Violence: A Critical Review", unpublished manuscript, PUC-Rio.

[51] Soares, R. and J. Naritomi (2010) "Understanding High Crime Rates in Latin America: The Role of Social and Policy Factors", in: Rafael Di Tella, Sebastian Edwards, and Ernesto Schargrodsky (editors). The Economics of Crime: Lessons for and from Latin America, University of Chicago Press, pp. 19-55.

[52] Yamada, T. (1985) "The Crime Rate and the Condition of the Labor Market: A Vector Autoregressive Model", NBER WP No. 1782.