Essays on Crime, Hysteresis, Poverty and Conditional Cash Transfers

André Oliveira Ferreira Loureiro

Doctor of Philosophy
The University of Edinburgh
2013

Abstract

This thesis encompasses three essays around criminal behaviour with the first one analysing the impact of programmes aimed at poverty reduction, the second one developing a theoretical model of hysteresis in crime, and the third one empirically investigating the hysteresis hypothesis in crime rates. In the first chapter I investigate the impact of conditional cash transfers (CCT) on crime rates by analysing the Brazilian Bolsa Familia, the largest CCT programme in the world, in a panel data between 2001 and 2008. The related existing economic literature analysing general welfare programmes usually ignores the crucial endogeneity involved in the relationship between crime rates and social welfare policies through poverty, since poorer regions are focused in the distribution of resources. I use the existing temporal heterogeneity in the implementation of the programme across the states to identify the causal impact of CCT programmes on poverty and criminality. The guidelines of the Brazilian programme established that the amount of resources available for each state should be based on the poverty levels in the 2000 Census. However, due to reasons unrelated to poverty levels and crime rates, some states were able to implement the programme to a greater extent more quickly than others. States that reached the level of cash transfer expenditures proposed by the guidelines of the programme more promptly had a more significant reduction in poverty rates. Similar but less robust results are found for crime rates as robbery, theft and kidnapping, while no significant effects were found for homicide and murder, indicating a weak or non-existent relationship between conditional cash transfers and crime. I also develop, to my knowledge, the first theoretical model to explicitly account for hysteresis - a situation where positive exogenous variations in the relevant economic variables have a different effect from negative variations - in both criminal behaviour and crime rates in order to fill the gap between the theoretical predictions and the empirical evidence about the efficiency of policies in reducing crime rates. The majority of the theoretical analyses predict a sharp decrease in crime rates when there are significant improvements in the economic conditions or an increase in the probability of punishment. However, the existing empirical studies have found lower than expected

effects on crime rates from variations in variables related to those factors. One important consequence of hysteresis is that the effect on an outcome variable from positive exogenous variations in the determining variables has a different magnitude from negative variations. For example, if hysteresis is present in the criminal behaviour and part of the police force in a city are dismissed in a given year, resulting in an escalation in crime, a reversal of the policy in the following year by readmitting all sacked police officers in an attempt to restore the original crime levels will result in lower crime rates, but higher than the original ones, yielding an asymmetric relationship between police and crime. Hysteresis is considered in a simple framework to model illicit behaviour. At the individual level, if criminal activity is associated with intrinsic sunk costs and learning, then the cost of leaving a criminal career is higher than entering it. At the aggregate level with homogeneous agents, this is translated into a hysteresis effect that will only occur if a specific threshold is surpassed. With heterogeneous agents, this phenomenon is reinforced generating a hysteresis effect that exists for all possible values of the variable affecting the crime decision. There are multiple equilibria at both levels. In the last chapter I empirically investigate the existence of hysteresis in crime rates. To my knowledge, this is the first empirical study to consider the existence of asymmetric effects on crime from variations in the probability of punishment and in the opportunity cost of crime. More specifically, I investigate whether positive variations on variables associated to those factors, respectively police officers and average level of income, are statistically different from negative variations. Using US crime data at the state level between 1977 and 2010, I find that police force size and real average income of unskilled workers have asymmetric effects on most types of crimes. The absolute value of the average impact of positive variations in those variables on property and violent crime rates are statistically smaller than the absolute value of the average effect of negative variations. These effects are robust under several specifications. A closer inspection of the data reveals a relatively monotonic negative relationship between wages and property crime rates, as well as negative variations in police and most crime rates. However, the relationships between positive variations in law enforcement size and most crime rates are non-linear. The magnitude of the observed asymmetries supports the hypothesis of hysteresis in crime, and suggests that no theoretical or empirical analysis would be complete without careful consideration of that important feature in the relationships between crime, police and legal income. These results corroborate the argument that policy makers should be more inclined to set pre-emptive policies rather than mitigating measures.

For Cintia.

Acknowledgements

I would like to thank my two supervisors, Professor Maia Güell and Dr. Dimitra Politi, for all their reassuring support and invaluable insights. I am also obliged to Professor Ed Hopkins and Professor Michèle Belot, PhD directors during my years in Edinburgh, that also provided a continuous source of encouragement and motivation. I also would like to gratefully acknowledge the support from *Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico* - FUNCAP. I made many friends in the School of Economics that I would like to thank. Any list would inevitably leave important names out, but I would like to especially thank David Comerford, James Best and Nicholas Myers for their support and friendship since the first year in the PhD and Vasileios Vlaseros for our long discussions about hysteresis and crime. Finally, I would like to thank my parents, Rubens and Nagmar, for their love and support, essential to reach this important milestone in my life.

Declaration

I hereby declare that this thesis has been composed by myself and is the result of my own work under the guidance of my supervisors as stated in the acknowledgements. None of the work contained in this thesis has been submitted for any other degree or professional qualification.

André O. F. Loureiro
Edinburgh, 2013

Contents

List of Figures

List of Tables

Introduction

"*...the hope of evading such taxes by smuggling gives frequent occasion to forfeitures and other penalties which entirely ruin the smuggler; a person who, though no doubt highly blamable for violating the laws of his country, is frequently incapable of violating those of natural justice, and would have been, in every respect, an excellent citizen had not the laws of his country made that a crime which nature never meant to be so.*"

- Adam Smith, The Wealth of Nations, 1776.

Crime and Economics

Crime is much more pervasive in all societies than we initially realise, as the word is often associated only with the more serious infractions. Most countries have legislation which deems a myriad of acts and activities illegal, beyond the smaller number of "natural crimes", such as theft and murder. Smuggling, tax evasion, cartels, illegal immigration, industrial pollution and copyright infringement are only a few examples of acts that contravene the established societal norms and are classified as crimes.

The classification of crimes is varied, but one important element of their taxonomy regards the degree of damage the perpetrator causes to the victim or the society and consequently, the severity of the punishment. People who are convicted of lesser and petty crimes are usually punished with probation, community service or fines, rather than a loss of freedom through imprisonment.[1] Indeed, the penalties imposed to the majority of crimes do not involve prison in most countries. The fraction of sentences involving confinement in the UK and the US, for example, is approximately 5%.[2]

[1] The lexicon for the different levels of seriousness of crime varies, even across countries with a shared language. In the English language, the legal term for serious, lesser and petty crimes are respectively felony, misdemeanour and infraction in the US, and indictable, triable either way and summary offence in the UK and Commonwealth countries.

[2] See Sentencing Statistics - UK Ministry of Justice and US Sentencing Commission.

Criminality is not only a relevant social problem, but also an important economic matter because it imposes a huge financial burden on society, as well as government. In order to prevent crime, considerable amounts of resources are allocated to private and public security. Another potential economic loss due to the presence of crime that is more difficult to measure is the economic activity that emigrates or is undertaken in another area or country because of high levels of crime. A cost which is even more difficult to measure is that associated with violent crime which implies loss of life and deep emotional harm to those involved.

Even though property crime is effectively a transfer of resources between agents, a significant fraction of the subtracted goods are destroyed. This type of crime can also have a harmful impact on incentives as it threatens the very principle of private property. Even petty crime or diversions from social norms reduce the efficiency of the economy: corruption and costly bureaucracy that is imposed to reduce illegal acts or process those perpetrators who are already in the legal system for an infraction.

Due to the unquestionable relevance of crime, all social sciences make an effort to contribute to the base of knowledge of criminal behaviour, which is enriched by their different perspectives. Nevertheless the Economics of Crime, established by the seminal article by Gary Becker published in the Journal of Political Economy in 1968, stands out with two key contributions: 1) It is the only science that provides a formal theoretical framework to criminal behaviour. 2) It provided some of the first empirical studies on crime, and even though criminology emerged as a frequently quantitative character and boasts an increasing quality of empirical analysis, economics remains the leading social science at the forefront of quantitative methods. This is especially true when one examines the relationship between crime and socioeconomic variables, such as those examined in this manuscript, including income, poverty and welfare benefits.

The economics of crime has almost entirely focused on serious crimes and for that reason the theoretical models that followed have generally used prison as the punishment of unlawful acts. That choice is certainly based on the relevance and prominence of more serious crimes, but also frequently due to the greater tractability of the models and the fact that virtually all empirical studies that use those models as a framework analyse more serious crimes, like murder and theft, that are officially registered and readily available. However, that focus may thwart the depiction of the aspects of more general and common unlawful activities not involving jail, arguably more economically motivated than the more serious crimes.

Certainly all violations of the social norms, in a broader definition, share key aspects from the economic perspective. Several decisions, from apparently inoffensive cheating like illegal parking, on to drug use and cybercrime, all the way up to capital crime share common elements: a decision between the benefits of a certainty equivalent (no crime) and a lottery with the expected pecuniary and/or nonpecuniary gain from the criminal act versus the associated probability of punishment and penalty. As in any lottery, risk preferences play a relevant role in the criminal decision. Additionally, two people facing exactly the same levels of expected gain and probability and severity of punishment can make different decisions, and one should also examine the moral cost and the opportunity cost of taking an action that breaks the norms.

Nevertheless, because the punishment for more serious crimes involves jail, that deeply affect the individual's life and crucially his/hers knowledge about crime technology, whereas lesser and petty crimes do not, theoretical and empirical analyses should be careful about the distinction of crimes by the usual type of punishment.

The economic approach also allows one to examine the criminal choice from a dynamic perspective as in the other mainstream economics models. Individuals would calculate the expected present value of the net benefit of breaking the law. In that sense, crime could also vary with the level of time preference rate. However, as Akerlof (1991) highlights, any analysis that considers potential criminals evaluating all periods in the future is based on an unreasonable assumption. The individual at the margin would be likely, in most cases, to be very myopic.

Another potentially relevant economic variable to explain crime is unemployment, which still is an open question in the empirical literature. Most studies examining the relationship between labour market and crime have focused on this matter, even though, until recently unemployment was not at the heart of the theoretical models in economics of crime.[3] Some studies find a positive relationship between unemployment and crime, but many others find no significant or negative effects of unemployment on crime.[4] One explanation for that ambiguity is that are the flows, rather than the levels of unemployment and nonparticipation in the labour market, that have a more important effect on crime.[5] The next step of this literature should focus at the analysis of the transition intensities among employment/participation states, that would allow to disentangle the impact of variations in the job finding and firing/quitting probabilities

[3]That was changed by Burdett, Lagos, and Wright (2003).

[4]Mustard (2010) provides an literature review on the empirical evidence on the impact of labour markets on crime.

[5]As Shimer (2012) shows, there was a great variability of the sources of variation in unemployment levels in the US economy in the last decades, particularly the job exit and job finding probabilities.

on crime rates.

Serious crime rates are at lower level in the developed world than they were three decades ago, but there are still important hurdles to be overcome by developing countries regarding criminality. And less serious crimes permeates most aspects of the socioeconomic life. For those reasons, a better understanding of the mechanisms that drive criminal behaviour and crime rates is fundamentally important to the implementation and efficacy of policies that seek to reduce crime in all its different forms.

Overview

The three chapters that follow investigate important aspects of crime, providing a contribution to the knowledge about the relationship between crime rates, poverty and conditional cash transfers (CCT) in developing countries, as well as a theoretical and empirical contribution to a better understanding of the phenomenon of hysteresis and asymmetric effects in crime.

The first chapter contains, to my knowledge, the first analysis of the impact of CCT programmes on crime, and solves the inherent problem of endogeneity in the relationship between welfare programmes, poverty and crime rates by exploiting the heterogeneous implementation of the Brazilian CCT programme. The guidelines of the Brazilian programme established that the amount of resources available for each state should be based on the poverty levels in the 2000 Census. However, by reasons unrelated to poverty levels and crime rates, some states were able to implement the programme to a greater extent more quickly than others. States that reached the level of cash transfers expenditures proposed by the guidelines of the programme more promptly had a more significant reduction in poverty rates. Similar, albeit less robust results are found for the crime rate of robbery, theft and kidnapping, while no significant effects were found for homicide and murder, indicating a weak or non-existent relationship between conditional cash transfers and crime. It could be argued that because individuals engaged in the criminal career have lower costs to commit a crime, they are not affected by variations in their legal income and a CCT programme would only influence the decision of individuals at the margin of the threshold between crime and no crime.

In the second chapter, I develop, to my knowledge, the first theoretical model to explicitly account for hysteresis in both criminal behaviour and crime rates. It also provides a clear and simple transition from the individual decision to the aggregate crime rates, which constitutes a useful development in the theoretical crime literature, as it allows to observe the relevant role of heterogeneity of individuals in the composition of the crime rates. Intrinsic sunk costs and learning in criminal activity generate hysteresis, and as a consequence, the effect on crime from positive exogenous variations in variables such as law enforcement has a different magnitude from negative variations. Because the sources of hysteresis are intrinsic to the individual, they take effect even if the perpetrators are not caught and punished. If hysteresis is present in the criminal behaviour, crime reduction policies will have a diminished impact when compared to the expected impact where individuals with a criminal past will behave similarly to individuals without a criminal past. Such asymmetric effects will be very clear in a situation in which the crime reduction policy in a given period is simply a reversal of a deterioration of one determinant of crime. A concrete example would be a situation where part of the police officers in a city are dismissed in a given year, resulting in a escalation in crime. If all sacked police officers are readmitted in the following year in an attempt to restore the original crime levels and hysteresis is present in the criminal behavior, that policy will result in a lower crime rate, but higher than the original one.

This prediction is subsequently tested in the third chapter with the focus on the asymmetric effects on crime from variations in police and income. Using US crime data at the state level between 1977 and 2010, I find that two of the main factors that affect crime - police force size[6] and real average income of unskilled workers - have asymmetric effects on most types of crimes. The absolute value of the average impact of positive variations in those variables on property and violent crime rates, are statistically smaller than the absolute value of the average effect of negative variations. A closer inspection with a semiparametric estimation reveals that the slopes are positive only for lower levels of police officers and they are negative for higher levels, except to murder, but the slopes are still significantly larger in absolute value when compared to negative variations in police, respectively to each type of crime. As will be discussed, no theoretical or empirical analysis would be complete without careful consideration of the various aspects raised in the following chapters. These results are relevant for any empirical analysis of policies at crime reduction, but they are particularly important for evaluations of policies based on increases of police force size.

[6] After correcting for endogeneity, as it will be discussed in chapter 3.

References

AKERLOF, G. A. (1991): “Procrastination and Obedience,” *The American Economic Review*, 81(2), pp. 1–19.

BURDETT, K., R. LAGOS, AND R. WRIGHT (2003): “Crime, Inequality, and Unemployment,” *The American Economic Review*, 93(5), 1764–1777.

MUSTARD, D. B. (2010): “How Do Labor Markets Affect Crime? New Evidence on an Old Puzzle,” IZA Discussion Papers 4856, Institute for the Study of Labor (IZA).

SHIMER, R. (2012): “Reassessing the ins and outs of unemployment,” *Review of Economic Dynamics*, 15(2), 127–148.

Chapter 1

Can Conditional Cash Transfers Reduce Poverty and Crime? Evidence from Brazil

1.1 Introduction

Crime is a central theme in the discussions of the main public issues of developing countries. Many of these countries have recently implemented policies based on conditional behaviour that are aimed at vulnerable people: Conditional Cash Transfers (CCT) Programmes, where the recipients receive a monthly benefit that represents a significant increase in their initial income. These policies have been the main device used by governments of countries such as Brazil (*Bolsa Família*), Mexico (*Oportunidades*) and Chile (*Chile Solidário*) to reduce poverty.[1] It is a stylised fact that higher levels of income inequality and poverty are associated with increased criminal behaviour, and these welfare programmes are deemed to be effective to mitigate these social problems.[2] A natural question would be whether those policies have also affected crime rate levels and how effective this kind of policy is when compared to increases in law enforcement.

This chapter uses a panel data set to analyse the impact of CCT programmes on crime rates in the Brazilian states, a relationship that still has not been investigated by formal theoretical or empirical analyses. As the vast majority of the empirical literature in economics of crime, the present analysis is focused on more serious crimes, since

[1] Other countries that adopted CCT programmes in the recent years are: India, Indonesia, Nigeria, Turkey and most Latin American countries. For further discussion on CCT programmes, see Medeiros, Brito, and Soares (2008) and Fiszbein and Schady (2009).

[2] As show Rawlings and Rubio (2003), Skoufias and Maro (2008) and Resende and Oliveira (2008).

the availability and reliability of data on less serious crime and petty crime is very limited.

Unlike the related literature that studies the effect of general social welfare, the channels through which this relationship occurs are analysed in some detail.[3] Here the analysis is deepened by verifying the specific impact of conditional cash transfer programmes on crime levels and the potential sources of endogeneity. By accomplishing this task, it will be possible to shed some light on the effects on crime rates of this kind of policy originally aimed at reducing poverty and income inequality through minimum income policy and a conditional component.

Another pivotal difference is that the existing literature on unconditional programmes generally ignores the crucial endogeneity involved in the relationships among poverty, income inequality and social welfare programmes.[4] If it is present and not taken into account, the estimates are biased and inconsistent. Temporal heterogeneity in the implementation of the programme across the states is used in this paper to identify the causal impact of CCT programmes on poverty and criminality.

CCT programmes have also the advantage of providing significant changes in social expenditures. The Brazilian CCT programme started in all 27 States in December 2003, adding up to the existing unconditional programme Continuous Cash Benefit (CCB), aimed at elderly and disabled poor people that started to operate in 1995. Figure 1.1 shows the trend in homicide rates per 100,000 inhabitants, cash transfers and law enforcement real expenditures[5] per capita between 2001 and 2008.[6] It is possible to observe the relevance of the change in the cash transfers in 2004 with the beginning of the Bolsa Familia programme.[7] A sharp decrease in the homicide rate in the same period is also observed.

The discussion of effect of social policies on crime is present in the texts of the first economists, but since the seminal paper of Becker (1968), economic studies consider formally the possible effects of socioeconomic variables on criminal

[3] All papers in this area focus on general social welfare or specific unconditional programmes. More details in the literature review (section 1.2.2).

[4] The exceptions are given by Pratt and Godsey (2002) (Homicide only), Johnson, Kantor, and Fishback (2007) and Worrall (2009) (Homicide only) that use instrumental variables to account for the simultaneity between welfare expenditure and crime. However, the instruments are either arguably weak (people immunized for measles, % of voting for Democratic Party and months of extreme wetness) or do not satisfy the required exclusion restriction for an instrument (lagged welfare spending).

[5] This rubric refers to all expenditures on policing, with the vast majority being spent on wages and equipment.

[6] All data sources and definitions of the variables are given in the appendix.

[7] It can also be seen that CCB expenditures were also increasing, even after inflation is taken into account. This shows the effect of the approach of the new government that proposed to reach more vulnerable people.

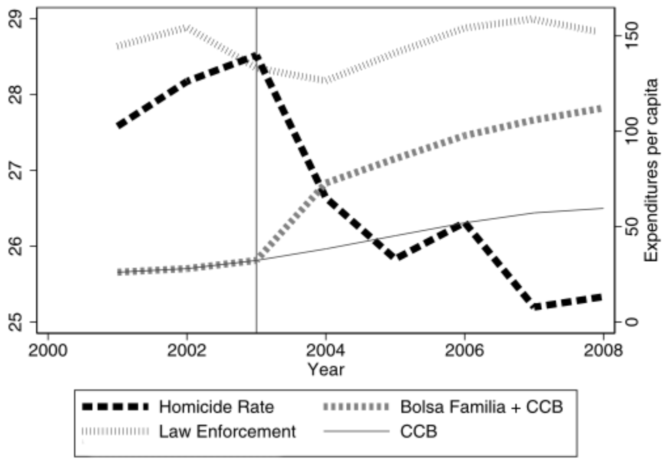


Figure 1.1: Per Capita Cash Transfers and Law Enforcement Expenditures and Homicide Rate - Brazil - 2001-2008
Data Sources: STN, MDS and SENASP.

behaviour. However, just recently the specific discussion of the effect of welfare programmes on criminal behaviour has been formalised. The literature that formally models the specific effect of social welfare programmes on crime is restricted to three papers: Benoit and Osborne (1995), Zhang (1997) and Imrohoroglu, Merlo, and Rupert (2000).

Empirical evidence, usually restricted to homicide rates, is provided by Chamlin, Cochran, and Lowenkamp (2002), Pratt and Godsey (2002), Burek (2005), Johnson, Kantor, and Fishback (2007), Savage, Bennett, and Danner (2008) and Worrall (2005, 2009). Ambiguous effects are found. Results are frequently plagued by endogeneity issues and not always properly addressed.

The effect of CCT programmes on legal income have some features that may intensify their effects on crime when compared with general welfare programmes. Firstly, this type of policy is more extensive than other welfare programmes (in Brazil it reaches about 1/4 of population, which represent the vast majority of the poor population), which makes one more likely to be affected by the programme, directly or indirectly; Secondly, the conditional aspect (incentivise children's education & health care) would affect the expected income and the decision to engage in the illegal market; Another distinguishing aspect would be a social altruism effect, where recipients of the benefit increase their sensation of social protection, making them more inclined to tolerate higher levels of poverty and income inequality.

Overall, because of its conditional component and extensive coverage, policies of this kind should also affect the expected income of individuals on the edge to commit a crime as well as their attitude towards society and government.

In the following section I discuss the links between CCT programmes and criminal behaviour. The data used is described in section 1.3. Sections 1.4 and 1.5 analyse the relationships among CCT, Poverty and Crime. The effect of CCT on crime in provided in section 1.6 and section 1.7 concludes.

1.2 Context

The sign and the magnitude of the effect of CCT programmes on crime is a relevant question for the policy maker. Welfare policies in general and specifically CCT programmes are not explicitly aimed at crime reduction, however, its likely effect on income distribution, especially on poverty could also affect the decisions involved in the criminal behaviour.

1.2.1 Conceptual Framework

The decision of an individual to commit a crime results from a expected utility maximization process, in which the individual at the margin would face, on one hand, the potential net gains arising from the criminal action, the value of punishment and the likelihood of arrest and conviction and, on the other hand, the opportunity cost of committing a crime represented by the legal income obtained in the legal labour market and/or from welfare benefits.[8]

Theoretical models and empirical findings in economics of crime point out poverty and income inequality to be major causes of crime.[9] In the economic theory of crime, areas with higher inequality place poor individuals who have low returns from market activity next to high-income individuals who have goods worth taking, thereby increasing the returns to time allocated to criminal activity. Faced with the relative success of others around them unsuccessful individuals feel frustration at their situation. A rise in inequality may also have a crime-inducing effect by reducing the individual's risk aversion and moral threshold to commit a crime (envy effect).

CCT programmes, as some other welfare policies, provide a boost in the income of the poorest individuals, generally deemed to be more likely to get involved in criminal activity, because of their lower opportunity costs. These programmes help individuals to reach a more acceptable subsistence income and also promotes income

[8]For an extensive critical literature review, see Dills, Miron, and Summers (2008).

[9]See Bourguignon (1998) for discussion. Huang, Laing, and Wang (2004) propose a search model analysis on the relation between poverty and crime.

redistribution, reducing inequalities and the associated "envy effect". These effects raise the opportunity cost of crime for the recipients of these programmes, which affects the propensity of an individual committing a crime. Because the involvement in illegal activities can result in the loss of the benefit if arrested, recipients have an additional increase in the opportunity cost.

Therefore, the major effect of CCT programmes on criminal behaviour would come from the effect on poverty levels and income inequality. However, this type of policy generally follows a natural rule to provide more resources in regions with higher levels of poverty. This is especially true in the Bolsa Familia programme, which has a formal law to determine the amount spent in each state. The amount of resources that should be spent in each state was based on the levels of poverty obtained by the Census carried out in 2000. This rule creates a relevant source of endogeneity in the estimation of the effect of CCT on poverty.

Unlike most welfare programmes, CCT programmes make the payment of the benefit conditional on health and educational attainments, which implies not only a higher short term income, but also a higher expected income for the family, given the programme would provide opportunities for higher levels of education.[10] Because CCT programmes generally have a considerable magnitude, a possible additional effect is a situation of social altruism, in which the individuals that receive the benefit have a real perception of the government action, with the approval of the political elite, creating an environment/sensation of social protection. This effect would raise the "moral" opportunity costs of committing an illicit act for some individuals.[11]

1.2.2 Welfare Payments and Crime Literature

The number of papers that investigate the relationship between social welfare spending and crime is more restricted and more recent than those that analyse similar issues like the effects of law enforcement on criminality. Moreover, there is no empirical evidence on this issue for developing countries, where the different economic environment could lead to contrasting results.

The existing theoretical literature that formally model this specific issue is restricted to three articles. In general, they suggest that expenditures on welfare programmes have a negative effect on crime rates, as discussed by Benoit and

[10] Lochner and Moretti (2004) and Stephen Machin and Vujić (2010) provide empirical evidence in this direction.

[11] Sickles and Williams (2008) propose and estimate a dynamic theoretical model where "social capital" is a relevant aspect to influence criminal behaviour while Buonanno, Montolio, and Vanin (2009) estimate the effect of several measures of social capital on crime rates.

Osborne (1995), Zhang (1997) and Imrohoroglu, Merlo, and Rupert (2000).[12] The idea behind this negative effect is that the welfare spending would impact the model with a reduction in incentives to commit a crime by raising the opportunity costs of the potential criminal.

In Benoit and Osborne (1995), under a theoretical setting, a formal model is developed in order to integrate spending on social assistance in the economic model of crime. Consideration is given to how individuals in the society behave in order to decide the optimal amount of investment in police and welfare payments needed to reduce crime. Under some specific assumptions, income redistribution reduce crime rates.

Zhang (1997) establishes a simple economic model in which, under some assumptions, criminal behaviour is reduced when policies aimed at redistribution are emphasised. This prediction is confirmed by his empirical findings for the US economy (Cross-sectional data for States). Reverse causality between Welfare Payment and crime is tested by using the amount of federal aid to the state governments and the percentage of the elderly in the population as instruments. No endogeneity problem in this sense is identified with these data.

In Imrohoroglu, Merlo, and Rupert (2000), a general equilibrium model is built in order to explain the relationship between public expenditures and criminality. The amount spent on police and income redistribution are determined endogenously through majority voting. In addition to the theoretical analysis, the authors consider the effects of increases in public spending on social welfare and in police over crime rates through calibration. Based on this structural model, it is found that the effect of the redistribution varies according to the characteristics of each region. However, as the authors emphasise, due to the fact that the model estimated is static, all possible dynamic aspects are ignored in these estimates.

Empirical evidence that find a negative effect of welfare programmes on crime rates is provided by Pratt and Godsey (2002) (Homicide only), Johnson, Kantor, and Fishback (2007), Savage, Bennett, and Danner (2008) and Worrall (2009) (Homicide only).

Pratt and Godsey (2002) use a panel data of 46 countries to estimate the relationship between social support (% of GDP spent on health care and education) and homicide. The percentage of people immunized for measles is used as instrumental variable. A negative effect is found and relatively robust. Alternative policies (law enforcement expenditures) are not considered.

[12] A few other papers, as Burdett, Lagos, and Wright (2003) and Merlo (2003), consider the role of redistribution on crime in general terms.

Johnson, Kantor, and Fishback (2007) explore a panel data set for 81 large American cities in order to estimate the effect of the relief effort in the years following the great depression. By using instrumental variables (mean of percentage voting for the Democratic Party and months of extreme wetness) the authors find a negative effect of public welfare spending on crime rates.

Savage, Bennett, and Danner (2008) analyse a panel data set of 25 countries for 13 years to explore the relationship between crime and social welfare spending. After unobserved heterogeneity and dynamic aspects are considered, a negative and curvilinear (by adding a quadratic term) relationship is found. Short run and long run results vary in sign and statistical significance. Endogeneity issues are not mentioned. Alternative policies (law enforcement expenditures) are not considered.

Worrall (2009) re-estimates the model considered in Worrall (2005) taking the likely endogeneity issues into account. A panel data from California counties is used in order to assess the effect of welfare spending on homicide. The lagged welfare spending is used as instrument. A negative effect is found, but it is not robust to different econometric approaches and specifications.[13]

The other papers on this issue find little or no evidence of the effect of social welfare spending on criminal behaviour, as argued by Chamlin, Cochran, and Lowenkamp (2002) (except homicide), Burek (2005) (less serious crimes) and Worrall (2005). Some of them predict no relationship between those variables or even a positive relationship when endogeneity issues are not taken into account and others do not encounter any effect even when the appropriate econometric models are considered.[14]

As discussed below, the econometric estimation of the relationship between welfare payments and crime is complicated by the likely presence of problems of endogeneity via poverty. It can also be argued that the welfare expenditures are intensified in places and/or in periods of higher economic hardship or poverty. Section 1.5 provides some empirical evidence on this issue. It is therefore not surprising to find that crime in its

[13]Two papers were published after the first drafts of the present paper that also exploit the timing of welfare programmes, but in a very different fashion. The first one is Foley (2011) that exploits a daily database on crime in the US and find that crime rates increase over the course of monthly welfare payment cycles. The second one is Chioda, De Mello, and Soares (2012) (that acknowledges the present paper as the first to examine the relationship between CCT and crime) that also examines the impact of CCT on crime, but only at urban schools areas in one Brazilian city, São Paulo. They exploit the expansion of the Bolsa Familia programme to families with adolescents in 2008 to identify the relationship. They find some evidence of a negative relationship between Bolsa Familia and an overall index of crime, but with very small magnitudes. They do not find any significant impact of Bolsa Familia on theft and vandalism.

[14]Witte and Witt (2001) also provides an extensive discussion on the government role to reduce crime rates.

various forms is positively correlated with the spending on welfare. This problem may be controlled for by using the appropriate econometric methods.

1.2.3 Conditional Cash Transfer Programmes: The *Bolsa Familia* Case

Conditional Cash Transfer (CCT) programmes are policy mechanisms aimed at reducing poverty by making welfare programmes conditional upon the receivers' actions. The government only transfers the money to individuals/families that meet specific criteria. In addition, after these agents have engaged in the programme and in order to keep the benefit they must follow some educational and health requirements.[15]

This kind of policy provides emergency assistance, while the conditionalities (requirement to the families) promote long-term investments in human capital. This kind of policy addresses the problem of underinvestment in human capital not only by compensating individuals in the short-term for the real costs of investing in health, nutrition, and education, but also by adding requirements for households to use public services that have long-term payoff in these areas.[16]

CCT programmes have some important features that distinguish them from other welfare programmes. Firstly, they have eligibility requirements and conditionalities over the recipients' actions. Additionally, the grant is paid in cash, providing a minimum income. Another distinguishing feature is that the responsible individual for the benefit is almost always a woman. In the Bolsa familia programme, the percentage of households where the payment is made to an adult female is around 95%.

In the last few years the Brazilian government has significantly boosted cash transfers policies to poor people. Two major welfare programmes are now being carried out in Brazil: Continuous Cash Benefit (CCB) and *Bolsa Familia* (BF).

CCB is a cash transfer programme implemented in the country in 1995 and aimed at individuals over 65 and/or with severe disabilities. In both cases the income per capita in the family must be below 1/4 of the minimum wage.[17]

[15] For further details on CCT programmes, see Fiszbein and Schady (2009).

[16] Several empirical papers address these issues, as Gertler (2004), Glewwe and Kassouf (2008) and Reis (2010). In general they provide evidence that CCT programmes are effective to boost recipients' health and education.

[17] It should not be confused with a pension or retirement benefit. This confusion is common even among recipients of this benefit.

Bolsa Familia is a conditional cash transfer programme that took place in all 27 Brazilian states at the end of 2003 targeted at families with low income.[18] It is mainly aimed at poor families with children and establishes education and health requirements.[19]

As with most CCT programmes around the world, the payment of the Bolsa Familia transfer is made directly to the recipient. The money comes from the federal government, with no involvement of the state or municipality administrations. Once the family is registered and considered eligible to receive the benefit, someone in the household will receive a bank card that can be used exclusively to withdraw the cash from the programme in any branch of the largest government-owned bank in Brazil.[20] The person withdrawing the money is, in almost all cases, a woman, even if she is not considered the head of the family in other contexts.

However, the transfer would only start taking place at a monthly basis to a family after the local governments had registered all individuals in the household into a central database.[21] The continuation of the monthly payments is conditional on the children having a school attendance rate of least 85% and younger children (up to 7 years old) having their vaccination cards, where all history of vaccines are held, up to date and according to the timetable recommended by the Ministry of Health.

The payments of the *Bolsa Familia* benefit can be suspended or cancelled if any of the educational and health conditions are not met. The family can also stop receiving the benefits if the level of income per capita (excluding the benefit) surpass the minimum threshold (1/4 of the Brazilian minimum wage per capita).

As mentioned above, the *Bolsa Familia* benefit is paid directly to the final recipient by the federal government. However, the municipality government also participates in the process by registering the potential receivers of the benefit. The participation of the state government is restricted to the administrative and technical support to the local governments.

---

[18] Bolsa Familia was a programme that provided additional income to most poor families in Brazil. Nevertheless, there were previous programmes of smaller magnitude that were absorbed into the new programme, like *Bolsa Escola*, *Vale Gas* etc. Data on those programmes were not included in this paper. However, this omission will only underestimate any effect found.

[19] For further discussion on cash transfer programmes in Brazil, see Medeiros, Brito, and Soares (2008).

[20] The bank is *Caixa Econômica Federal*, which is the forth largest bank in Brazil, with branches in all states and in most of the 5570 municipalities in Brazil. The recipients can alternatively withdraw the money in one of the more numerous *Casas Lotéricas*, offical lottery stores registered by *Caixa Econômica Federal*.

[21] This is a relatively famous database in Brazil, named *Cadastro Único*, where all *Bolsa Familia* recipients need to be registered to receive the benefit.

In 2009 Bolsa Familia was the largest conditional cash transfer programme in the world, although the Mexican programme *Oportunidades* was the first nation-wide programme of this kind. I compare the figures related to the cash transfers programmes in Brazil in Table 1.1.

Table 1.1: Bolsa Familia and CCB's Figures - Jan-Dec - 2009

	Bolsa Familia	**CCB**
Amount Spent	US$ 6,610,220,967	US$ 9,917,301,532
Families receiving the benefit	12,472,540	3,166,845
People receiving the benefit	51,636,316	13,110,738
Average Benefit per person per year	US$ 529.98	US$ 3,131.60
Percentage of Federal Social Expenditures	37.26%	57.01%
Percentage of National Social Expenditures	33.53%	51.32%
Percentage of Federal Total Expenditures	0.76%	1.16%
Percentage of GDP	0.33%	0.51%

Sources: Calculated by the author with data from the National Treasury Secretariat and Ministry of Social Development.
Notes: Monetary figures converted from *Real* (R$). People receiving the benefit and derived quantities are estimates.

The average value of the Continuous Cash Benefit (CCB) per person per year is almost 6 times higher than the average value paid for *Bolsa Familia* recipients. Data from the PNAD[22] shows the average annual income per capita in poor households,[23] excluding benefits, was US$ 577.65, [24] which evidences that the *Bolsa Familia* benefit represent almost half of the income in an average household receiving the benefit. The impact on the initial income of the families is even more substantial if the average income is calculated only for families below the line of extreme poverty[25] (US$ 267.57).

Social welfare policies are accomplished by the three levels of government.[26] Nevertheless, unlike other areas, such as spending on police, the percentage that each level of government applies in welfare programmes varies considerably across states

[22]National Survey of Household Sample run by IBGE - Brazilian Institute of Geography and Statistics.
[23]Monthly income per capita in the household below 1/2 of the value of the monthly minimum wage.
[24]Values converted from *Real* (R$) in 31 December 2009.
[25]Households with income per capita in the below 1/4 of the value of the monthly minimum wage.
[26]In Brazil, the governmental attributions are carried out by federal, state and municipality administrations.

and over the years. The amount the Federal Government spends on welfare represents about 90% of the whole expenditure in the states. In 2009 those two programmes represented about 94% of all federal spending on social welfare. Expenditure on *Bolsa Familia* corresponds to approximately 2/3 of the amount spent on Continuous Cash Benefit, although the former covers 4 times more families. Another relevant fact is that such expenditures account for less than 2% of the federal budget. It should be also noticed that, unlike many developing countries, effectively all resources for funding *Bolsa Familia* and CCB expenditures come from federal tax revenue, with insignificant contribution from foreign aid.

1.3 Data

A new data set is created by linking aggregate variables I constructed from the state representative socioeconomic micro data from annual Brazilian house survey of PNAD (National Survey of Household Sample by the IBGE - Brazilian Institute of Geography and Statistics) with the criminal registers from the Public Security national agency, from 2001 to 2008.[27]

1.3.1 Criminal Data

I use data from SENASP - National Secretariat of Public Security, agency of the Ministry of Justice, which compiles information from the State Secretariats of Public Security, and indicators of the incidence of crime in the Brazilian states the following indices: murder rate per 100 thousand inhabitants, total rate of robberies per 100 thousand inhabitants, total rate of theft per 100 thousand inhabitants and extortion through kidnapping rate per 100 thousand inhabitants. The data are to be used annually for all 27 federal units of Brazil and covering the period from 2001 to 2008.[28] Alternative measures of homicide rates are also calculated using data from the Ministry of Health, that provides state level data.[29]

[27] I computed all variables from PNAD, except the variables associated with income, computed by IPEA, as specified in table 1.7 in the appendix.

[28] The national agency decided not to make public the figure for property crimes for some states after 2005. For that reason, this type of crime will be analysed over the period between 2001 and 2005.

[29] Murder are always intentional crimes, whereas homicide comprises both voluntary and involuntary killing.

1.3.2 Expenditures on Cash Transfers and Law Enforcement

As mentioned before, the two main federal welfare programs are direct cash transfers: *Bolsa Familia*, aimed at poor people, especially those with young children and Continuous Cash Benefit, which is primarily aimed at elderly and disabled people. The variable used in this chapter is the sum spent on both programmes per capita and corrected for inflation. Data come from the Ministry of Social Development (MDS).

Information about the law enforcement spending and revenue of the states was obtained from the Bulletin of Public Finance of Brazil, issued by the Secretariat of the National Treasury (STN). Such information relates to all public expenditure made by state governments and the Federal District within the respective units of the federation. The variables used were corrected for inflation using INPC index from IBGE with 2007 as the base year.

Like the other public expenditures in Brazil, law enforcement is run by all the three levels of government: Federal, State and Municipality governments. However, as usually happens for all types of expenditures in Brazil, the main responsibility lies with one of these levels. Law enforcement is a primary duty of state governments. In fact, the total amount spent on police in the Brazilian States in 2007 is about 88% made by the state governments.

1.3.3 Poverty Rates and other Explanatory Variables

Poverty Rates and the other explanatory variables were constructed by aggregating state representative micro data of PNAD (National Survey of Household Sample by the IBGE - Brazilian Institute of Geography and Statistics) from 2001 to 2008. Besides poverty rates and extreme poverty rates, other poverty related measures are calculated. Poverty lines were determined as a function of minimum consumption levels.[30] The other variables considered are income inequality (GINI index), years of schooling, average labour income, unemployment rate, % of one-parent households, percentage of young males, informality degree in the labour market. I correct all the monetary variables for inflation using INPC index from IBGE with 2007 as the base

[30]Poverty is measured here using the standard index (FGT index) suggested by Foster, Greer, and Thorbecke (1984): $P(\alpha) = (\frac{1}{N}) \sum_{i=1}^{q} \left(\frac{z-y_i}{z}\right)^{\alpha}$, where N is the total number of households, y_i is the per capita income of the ith household, z is the poverty line, q is the number of poor individuals. With α=0, the FGT measure becomes the incidence of poverty index P(0) or simply the percentage of the population that is below the poverty line (poverty rate). With α=1 the FGT measure gives the poverty gap P(1), a measure of the average intensity of poverty. With α=2, the FGT index becomes the severity of poverty index.

year. Table 1.7 in the appendix summarizes the description of each variable used in the estimates, and the origin of the data.[31]

1.4 Empirical Framework

In this section I present the empirical relationships of interest. However, the main purpose of the following paragraphs is to discuss the inherent problems of estimating the effect of welfare programmes on crime rates.

The basic assumption is that CCT programmes affect poverty rates, as described by:

$$Pov_{it} = \alpha_i + \pi_t + \mathbf{w}'_{it}\sigma + \rho CCT_{it} + \tau_{it} \tag{1.1}$$

where Pov_{it} represents the poverty rate in each state i in a specific year t, CCT_{it} represents the amount per capita spent in cash transfers, $\mathbf{w}_{it}$ is a vector of covariates, α_i and π_t are respectively state and year fixed effects.

Furthermore, in order to pursuit principles of efficiency and fairness, the policymaker generally establishes a rule that assures more resources to the poorer regions. This fact establishes the following relationship:

$$CCT_{it} = \phi_i + \lambda_t + \mathbf{x}'_{it}\varrho + \theta Pov_{it} + \mu_{it} \tag{1.2}$$

where $\mathbf{x}_{it}$ is a vector of factors that can also affect the policy maker's allocation decisions, ϕ_i and λ_t are respectively state and year fixed effects.

Consider now equation 1.3 that describes the determinants of crime rates,[32] depending on expenditures on law enforcement,[33] poverty rates and a vector with other control variables $\mathbf{z}_{it}$:

$$Crime_{it} = \psi_i + \omega_t + \mathbf{z}'_{it}\beta + \gamma Law_{it-1} + \delta Pov_{it} + \epsilon_{it}. \tag{1.3}$$

$Crime_{it}$ is one of the crime rates mentioned before: Murder, Homicide, Robbery, Theft and Kidnapping (per 100,000 inhabitants). Law_{it-1} is the total expenditure in law enforcement per capita in the previous period, $\mathbf{z}'_{it}$ is a vector of socioeconomic

[31] Figure 1.9 in the appendix depicts the income distribution in Brazil before and after the implementation of the programme.

[32] The theoretical and the empirical literature in the economics of crime converge for a similar specification. However, poverty rate is often omitted.

[33] Law enforcement itself could also be potentially endogenous due to the simultaneity with crime rates. The use of lagged law enforcement would minimise that problem. As argued by Chalfin and McCrary (2013) this endogeneity is less likely to be the case as the response from the policy makers in the short run is very limited.

variables, namely: income inequality (GINI index), years of schooling, average labour income, unemployment rate, % of one-parent households, informality degree in the labour market and ψ_i and ω_t are respectively state and year fixed effects.[34]

As discussed before, the underlying assumption is that individuals are maximisers of their respective expected utility, making rational choices in order to participate in the criminal sector in response to the costs and benefits of illegal activities and in relation to the alternative gain from the legal market. This specification attempts to capture the fact that the participation of an individual in criminal activities depends on the monetary return on these actions in relation to legal activities, the economic conditions under which the individual is living, their cultural and social condition (including the environment that he/she is surrounded) and the degree in which the police system is able to affect the likelihood of imprisonment and punishment.

One assuming exogeneity of the redistribution policy would ignore equation 1.2 and estimate the effect of cash transfers on crime rates after plugging equation 1.1 on equation 1.3. However, as it will be shown in the next section, an endogenous resources allocation should hold in most situations, leading to inconsistent estimates.[35]

Another confounding factor would raise if the cash transfers are assumed to affect crime rates directly (not only through poverty rates). That would account for any social altruism effect on the potential criminals, as the social program would make them perceive a real presence of the government in their lives. The benefit would then have an effect on the individuals' decisions beyond the monetary transfer. If this is the case, the error term in the equation 1.3 would be $\epsilon_{it} = \eta CCT_{it} + \zeta_{it}$. [36]

Therefore, in order to consistently estimate the effect of cash transfers on crime rates, these endogeneity problems must be taken into consideration.

[34] As emphasised by Cornwell and Trumbull (1994), there are at least three reasons to expect the presence of the terms of spatial and time unobserved heterogeneity ψ_i and ω_t. First, it is to be expected that there are unobserved heterogeneity such as cultural characteristics relatively stable over time, what makes some states to have higher crime rates than others. A second reason is to account for the presence of measurement error in the rates of crime. Time-specific effects common to all states, like changes in federal law, justifies the presence of ω_t.

[35] Another possibility would be to use CCT as an IV for Pov. That not only would generate inconsistent estimates as the measure of the effect of CCT on crime would be lost.

[36] An alternative approach, where CCT is explicitly included in the crime equation, is also considered in the appendix. Similar problems of endogeneity would arise.

1.5 Effect of Conditional Cash Transfer Programmes on Poverty

1.5.1 Will CCT always reduce Poverty?

The effect of CCT expenditures on Poverty is not obvious as it initially seems. One important reason is the simultaneity between poverty and this kind of policy. Poorer places often are focused in the distribution of resources, which creates a confounding factor in the analysis.

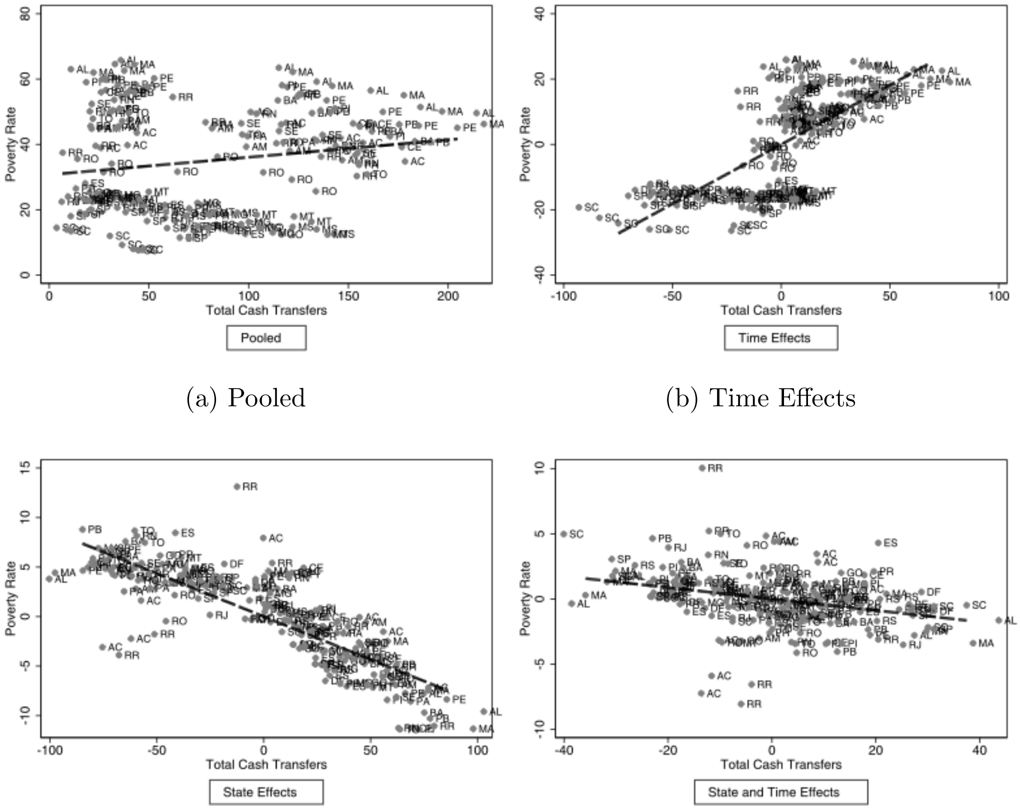
Another important issue is that more cash transfers may affect fertility and family formation decisions, diluting the expected reduction in poverty levels. It may also reduce the incentives to mothers to work, reducing the total income in the family. Although this type of governmental action will alleviate poverty, the additional resource will not necessarily make a poor family "non-poor" if it is very far below the poverty line.

1.5.2 A More Detailed Look on the Data

Figure 1.2 shows the correlations between CCT expenditures per capita and poverty rates across the Brazilian states between 2001 and 2008. By observing the very different correlations in each case, the reason for being careful in estimating the relationship becomes clear.

Figure 1.2a shows a positive correlation between cash transfers and poverty in the pooled data. When time effects are added, the relationship becomes stronger and with a higher slope (figure 1.2b). This suggests that once the time dimension is removed, higher poverty rates imply more money being spent on cash transfers.

However, when alternatively state specific effects are included, the relationship becomes negative, as show figure 1.2c. A similar result is obtained when both time and state effects are taken into account, but with a lower slope, as show figure 1.2d. This could indicate that once the state specific unobserved heterogeneity is removed more resources for cash transfers results in lower levels of poverty.



(a) Pooled

(b) Time Effects

(c) State Effects

(d) State and Time Effects

Figure 1.2: Real Per capita Cash Transfers Expenditures and Poverty Rates
Data Sources: MDS/IPEA/PNAD.

Table 1.2 presents the regression results for the cases illustrated, including average labour income as a control, in order to capture reductions in poverty due to improvements in the labour market. A negative and statistically significant coefficient prevails when time and state effects are considered.

This analysis suggests that the association between cash transfers and poverty is negative and significant. The coefficients of the last equation correspond to elasticities of -0.0691 and -0.3929 for respectively Cash Transfers and Average Labour Income. However, as seen above, endogeneity is present in this relationship: because poorer states receive more cash transfers, this prevents any causal interpretation of the results. In spite of the fact that fixed effects estimation could mitigate concerns about this problem.

Table 1.2: Cash Transfers Expenditures and Poverty Rates: State and Time Effects

	(1) Poverty	(2) Poverty	(3) Poverty	(4) Poverty	(5) Poverty	(6) Poverty
Cash Transfers	0.050**	-0.084***	-0.050***	0.361***	-0.042***	-0.029**
	(0.016)	(0.005)	(0.006)	(0.016)	(0.008)	(0.008)
Avg. Labour Income			-0.022***			-0.018***
			(0.008)			(0.004)
Constant	30.9***	41.9***	56.6***	-23.9***	39.3***	51.4***
	(1.971)	(0.544)	(3.015)	(2.513)	(1.101)	(4.196)
State Effects	no	yes	yes	no	yes	yes
Time Effects	no	no	no	yes	yes	yes
Observations	216	216	216	216	216	216
R^2	0.0251	0.6033	0.6189	0.5292	0.6887	0.7598

Notes: Standard errors robust to heteroscedasticity and clustering on state in parentheses. p-values for the test for unobserved state, time and two-way effects, respectively: 0.0000 ; 0.0000; 0.0000. Robust Hausman test rejects Random Effects: p-value=0.0001. p-value test for heteroscedasticity = 0.0000. p-value of the test for strict exogeneity = 0.6597. Wooldridge test for autocorrelation in panel data does not reject H_0: no first-order autocorrelation: p-value=0.0894
*** p<0.01, ** p<0.05, * p<0.1
Data Sources: SENASP/MDS/IPEA/PNAD.

Therefore, a source of exogenous variation is necessary to identify this effect. As shown in the next section, a specific characteristic in the timing of the distribution of the resources helps to identify the relationship between CCT spending and poverty rates.

Another important concern would be serial correlation. The Wooldridge test for autocorrelation in panel data does not reject the null hypothesis of no first-order autocorrelation. However, as the non-rejection is only marginal (p-value=0.0894), it could be argued that autocorrelation is present and first-difference (FD) estimator should be used. The estimated coefficients with the FD estimator (not reported) are very similar to the FE estimates, but with larger standard errors, suggesting the FD estimator is inefficient in this analysis. [37]

[37] Wooldridge (2002) highlights the fact that if the idiosyncratic errors are uncorrelated, first differencing will introduce serial correlation in the analysis.

1.5.3 Constructing exogenous variation of CCT programmes

As figure 1.3 show, some states had a more intensive change of total cash transfers (after Bolsa Familia Programme in 2004) than others. The guidelines of Bolsa Familia programme established that the amount of resources available for each state should be based on the poverty levels in 2000. However, by reasons unrelated to poverty levels and crime rates, some states were able to implement the programme to a greater extent more quickly than others. As it will be shown, the heterogeneity in the promptness of the implementation of the programme was not associated with poverty or crime rates. It is more likely that such variation is related to the degree of efficiency or bureaucracy of the local governments in each state.

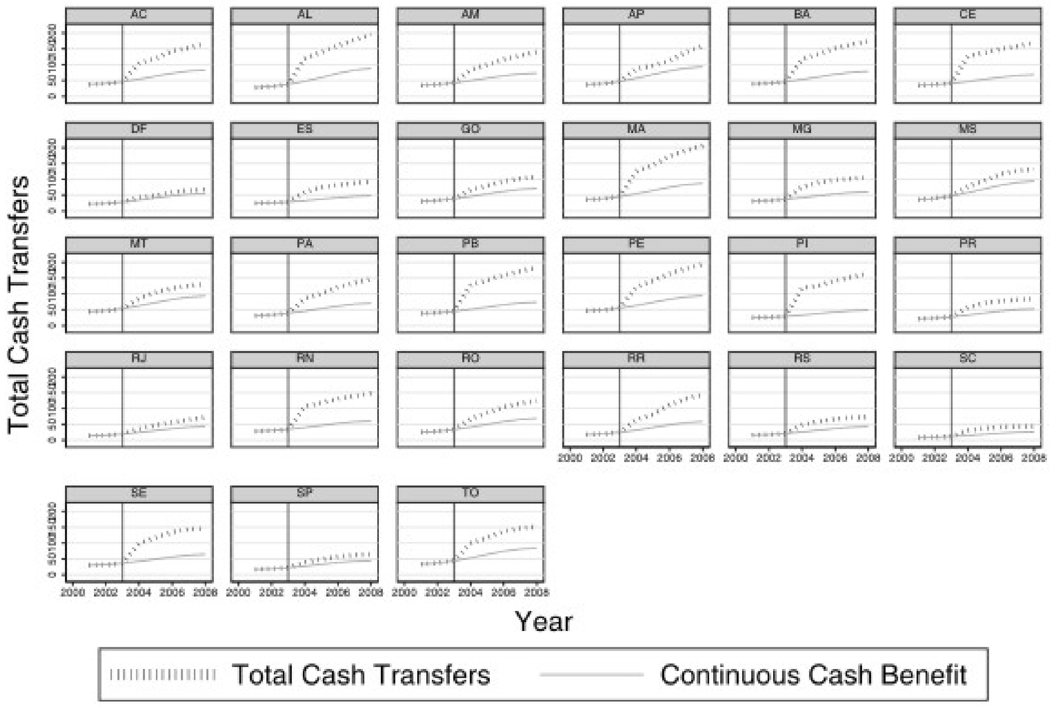
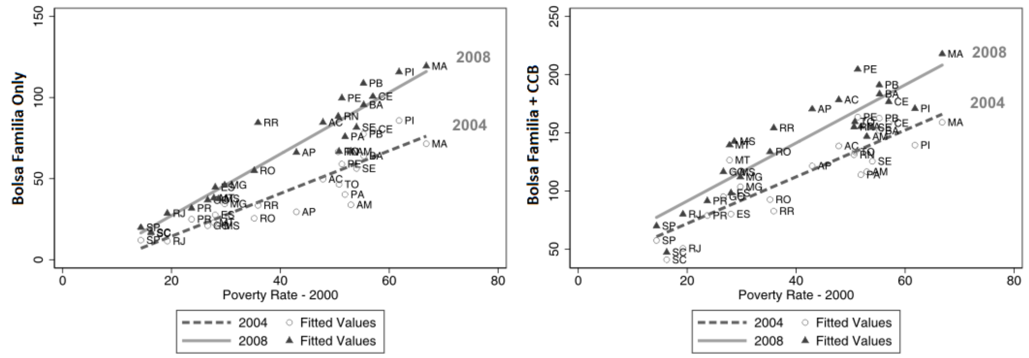


Figure 1.3: Total Cash Transfers over time by State

Data Source: MDS.

Another way to see this is to observe that some states had a significant increase in cash transfer spending in the 2004 with the beginning of *Bolsa Familia* programme. Those states had a smaller variation in cash transfers between 2004 and 2008 than the states that had a "late" start.

Because the guidelines of the programme establish that the amount of resources for each state depends on the poverty levels of the 2000 census, it is interesting to examine the relationship between poverty rates in 2000 and Cash Transfers expenditures in 2004 and 2008. Figure 1.4 displays those relationships for *Bolsa Familia* only and total cash transfers (*Bolsa Familia*+ CCB).



(a) Bolsa Familia (b) Total Cash Transfers

Figure 1.4: Poverty Rates in 2000 X Per Capita Cash Transfers Expenditures (R$ - 2007)
Data Sources: MDS/IPEA/PNAD.

In a situation of perfect implementation of the programme according to its guidelines, this relationship would be a graph proportional to a 45 degrees line.[38] However, in 2004 many states spend less than they were supposed to be spending. The resources were available from the federal government, but because some states and local governments were not able to register all eligible recipients, the money was not fully spent. Note that in 2008 most states reach the full implementation of the programme.

Figure 1.5 presents the distribution of variation in *Bolsa Familia* between 2004 and 2008. The variation in CCT spending between 2004 and 2008 was between 40% and 150% and 5 out of 27 states had variation above 100%.

[38]Strictly speaking, because the amount of money varies with the number and age of children in a family, the present graph would never be a perfect 45 degrees line. However, this is a good approximation to convey the idea of compliance.

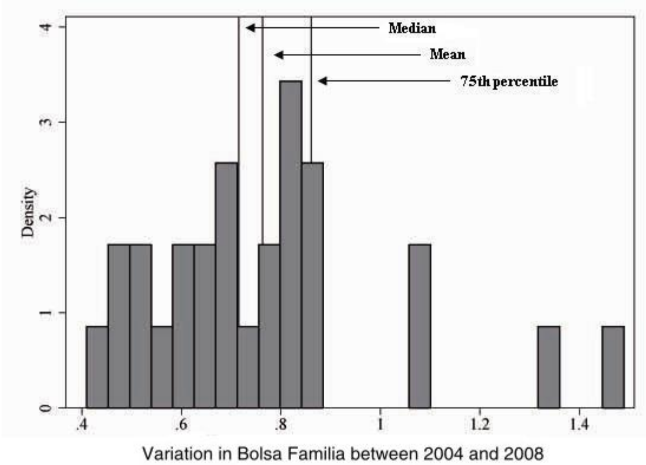


Figure 1.5: Distribution of variation in Real Per Capita Cash Transfers between 2004 and 2008
Data Source: MDS.

As mentioned above, this heterogeneity could stem from different operational capabilities and efforts among states. On one hand, it is possible to argue that states with higher poverty rates would be more determined to implement the programme as soon as possible. On the other hand, it is likely that states with a higher relative number of poor individuals would face more difficulties to identify and register all eligible families. As it is discussed below, poverty rates do not seem to have played a significant role in the promptness of the execution of the programme. The efficiency of the local governments and political issues in the states were more likely to determine the speed of the implementation.

The heterogeneity in the implementation of the *Bolsa Familia* programme allows the definition of an artificial control group, as some states lag considerably behind in a way that they virtually do not put the programme into practical effect when compared with the levels of the programme in the same state a few years later or with the level of implementation observed in other states.

In order to account for the levels of poverty before the rollout of the *Bolsa Familia* programme and the fact that even states that had a very slow start in the implementation of the programme, actually had some families receiving the *Bolsa Familia* benefits, I use the differences-in-differences (DD) approach.

If unobserved factors, such as the efficiency of the local governments are the main drivers to determine the speed of the implementation and they are relatively stable over time, the DD approach will account for such unobserved heterogeneity.

In order to take advantage of this heterogeneous variation in the implementation in CCT to analyse its effect on poverty, define D_i as:

$$D_i = \begin{cases} 1, \text{ if } CCT \text{ gap in state } i \text{ in 2004 is below 75th percentile value (Low Gap)} \\ 0, \text{ otherwise (High Gap)} \end{cases} \tag{1.4}$$

And defining pov_1 to be the poverty rate before the beginning of the programme (2004) and pov_2 poverty rate to after the programme started, DD estimator in this context can be defined as:

$$\hat{\delta} = (pov_2^{lowgap} - pov_1^{lowgap}) - (pov_2^{highgap} - pov_1^{highgap}) \tag{1.5}$$

The DD estimate $\hat{\delta}$ can be obtained by estimating:

$$pov_{it} = \alpha + \beta D_i + \gamma 1(After\ Programme)_t + \delta D_i \cdot 1(After\ Programme)_t + \varepsilon_{it} \tag{1.6}$$

or equivalently:

$$pov_{it} = \alpha + \delta D_i \cdot 1(After\ Programme)_t + \phi_i + \lambda_t + \varepsilon_{it} \tag{1.7}$$

where $1(\cdot)$ is an indicator function and ϕ_i and λ_t denote respectively state and time effects.

A crucial assumption in the DD framework is that λ_t is common across "treated" and "untreated".[39] This assumption can be tested by running the following regression:

$$pov_{it} = \alpha + \delta D_i \cdot 1(After\ Programme)_t + \psi_i t + \phi_i + \lambda_t + \varepsilon_{it} \tag{1.8}$$

In the case of poverty, significant values for ψ_i and insignificant value for δ tell us that the dependent variable would be declining even without treatment. In any case, the pre-treatment data must establish a clear trend that can be extrapolated into the post-treatment period.

Table 1.3 displays the averages for the relevant variables for both high and low variation in CCT expenditures between 2004 and 2008.[40]

Apart from murder and homicide rates, all other variables have no significant differences between the groups, corroborating the assumption that this variation was exogenous in this context. Table 1.10 in the appendix presents the regression of poverty rates in 2001 on the variation in CCT expenditures between 2004 and 2008.

[39] See Cameron and Trivedi (2005) and Angrist and Pischke (2009) for further discussion.

[40] Imbens and Wooldridge (2008) highlight the importance of comparing descriptive statistics in this setting.

No statistical significance is found in the relationship, whereas a similar regression on the variation in the spending between 2003 and 2004 is positive and significant. Figure 1.8 in the appendix presents the spatial distribution of these variables, which also corroborate this idea.

Tables 1.8 and 1.9 in the appendix replicate table 1.3 the threshold in equation 1.4 are replaced by respectively the mean and the median of the distribution of variation in the cash transfers. The conclusions about the significant differences for the outcome variables across the sub-samples are unaltered, except to robbery rates that are now statistically different and homicide rates that are now statistically indistinguishable across the groups.

Because the averages of murder and homicide rates are statistically smaller in states with a swifter implementation of the *Bolsa Familia* programme relative to the averages for states with a slow implementation when the threshold is the 75th percentile of the distribution of variation in the cash transfers, the difference-in-difference approach described above would not be appropriate for those variables. A significant difference is also observed for robbery rates when the thresholds are the mean and the median. Nevertheless, I will keep those variables in the ensuing analysis to contrast with the results for the other variables.

Table 1.3: Selected variables for different levels in the implementation of Bolsa Familia (Perc. Var. in CCT > 75 percentile)

	Whole Sample			2003-2004		
	Low Gap ($D_i = 1$)	**High Gap** ($D_i = 0$)	**t-test on diff.**	**Low Gap** ($D_i = 1$)	**High Gap** ($D_i = 0$)	**t-test on diff.**
Murder Rate	20.8405	28.4798	-4.52	19.9952	27.7501	-2.51
Homicide Rate (Health)	24.1419	34.1217	-5.72	23.7479	33.2542	-2.76
Robbery Rate	352.6788	379.1966	-0.54	341.8772	392.7397	-0.66
Theft Rate	1130.9132	1037.4245	0.74	1160.1571	1011.8172	0.67
Kidnapping Rate	0.3667	0.3482	0.17	0.3339	0.2698	0.47
Poverty Rate	0.3495	0.3573	-0.31	0.3894	0.3991	-0.18
Extreme Poverty Rate	0.1407	0.1356	0.36	0.1653	0.1634	0.06
Gini Index	0.5572	0.5447	1.85	0.5616	0.5488	1.23
Average Labour Income	742.1845	749.7728	-0.17	703.7415	716.9822	-0.16
Informality in the Labour Market	56.1774	55.7940	0.22	56.7759	56.7265	0.01
Single-Parent Households	0.2764	0.2877	-1.46	0.2763	0.28968	-1.08

Notes: High Gap and Low Gap as defined in equation 1.4. *Data Sources*: See appendix 1.A.2.

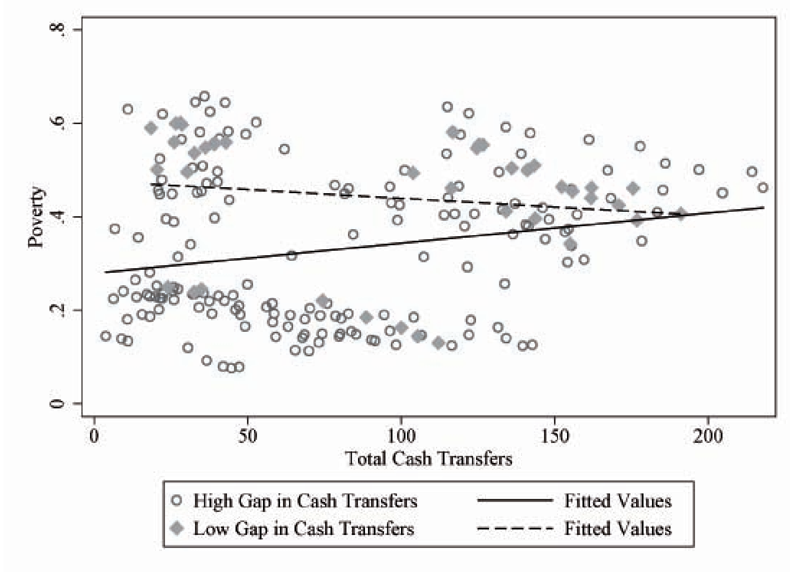


Figure 1.6: Poverty X Cash Transfers for different levels of variation in Cash Transfers *Data Source*: MDS/IPEA/PNAD.

1.5.4 Effect of CCT on Poverty

Figure 1.6 presents the effects of CCT on Poverty for different levels of variation. The reduction of poverty is slightly stronger for states that had a faster increase in total cash transfers after the program *Bolsa Familia*.

The estimation of differences-in-differences effect described by equation 1.6 is presented by table 1.4.[41]

The estimate of the relevant parameter δ is -0.0312. This 3.12 percentage points represents approximately an additional reduction of 7.5% in poverty rates in states that distributed the resources of CCT more quickly. It is a sizable magnitude, but it is even more significant if it is taken into consideration that the programme is also present in the "control" states. It is worth noting that poverty rates are not statistically different between the groups.[42]

Table 1.5 displays alternative procedures in the estimation of δ in order to check the robustness of the results.[43]

Similar effects are found when variations and percent variations are used, as well as

[41]One possible alternative would be to use $D_i \cdot 1(After\ Programme)_t$ as an instrumental variable for Pov_{it} in the crime equation. However, this approach would make the measure of the effect of CCT on crime to be lost.

[42]It should be noticed that because the number of clusters is 27, the underlying statistical test can over-reject the significance of the coefficients. As suggested by Cameron, Gelbach, and Miller (2008), I use bootstrapped standard errors to mitigate this problem.

[43]Reporting a sensitivity analysis in this context is essential to the legitimacy of the results as pointed out by Angrist and Pischke (2010).

Table 1.4: Difference-in-Difference Estimates: Poverty Rates for different levels in the implementation of *Bolsa Familia* Program, before and after 2004

	Low Gap ($D_i = 1$) (i)	**High Gap** ($D_i = 0$) (ii)	Difference: (i) - (ii) (iii)
Poverty Rate **Before** 2004	0.3960***	0.3809***	0.0150
	(0.0213)	(0.0360)	(0.0437)
Poverty Rate **After** 2004	0.3217***	0.3378***	-0.0161
	(0.0159)	(0.0275)	(0.0223)
Change in Poverty	-0.0742***	-0.0430***	-0.0312***
	(0.0263)	(0.0452)	(0.0135)

Notes: Standard errors robust to heteroscedasticity and clustering at the state level in parentheses. High Gap and Low Gap as defined in equation 1.4. Similar results for D_i defined using mean and median as threshold. Similar significance when standard errors are bootstrapped. Number of obs. before: 81, after: 135. Without considering groups, the total decline in poverty was -0.0672 (-6.72 perc. points).
*** p<0.01, ** p<0.05, * p<0.1
Data Sources: MDS/IPEA/PNAD.

other thresholds to define the levels of gap. The estimated δ does not vary significantly to the addition of different covariates, providing additional evidence of the exogeneity of the variation of the CCT expenditures between 2004 and 2008 relative to poverty.

It is also possible to argue that these results are based on a convergence effect, where poorer states would improve faster anyway. In order to control for that, three strategies are considered. The first controls for intensity of poverty - P(1)[44], which is a measure of distance from poverty line. The second one uses poverty rates in 2001 as a control. However this has the downside of preventing the use of fixed effects, since it is a constant over time. To get around this issue a slightly different approach is considered. In this third way, a new variable is defined so that it assumes the values of 2001's poverty rates if the year is before 2002 (inclusive) and 2003's values if the year is after 2003 (inclusive). In all cases, the estimates of the effort of cash transfers on poverty rates are quite similar to the previous approaches.

This relative robustness corroborates the assumption of random distribution of states into the different speeds of implementation of the program and consequently the causal interpretation of the estimates.

Because the values of extreme poverty are generally less than half of those of poverty and the estimated coefficient have similar sizes, the effect of CCT on extreme

[44] As defined in subsection 1.3.3.

Table 1.5: Difference-in-Difference Estimates - Poverty and Extreme Poverty: Estimates for δ under Alternative Specifications

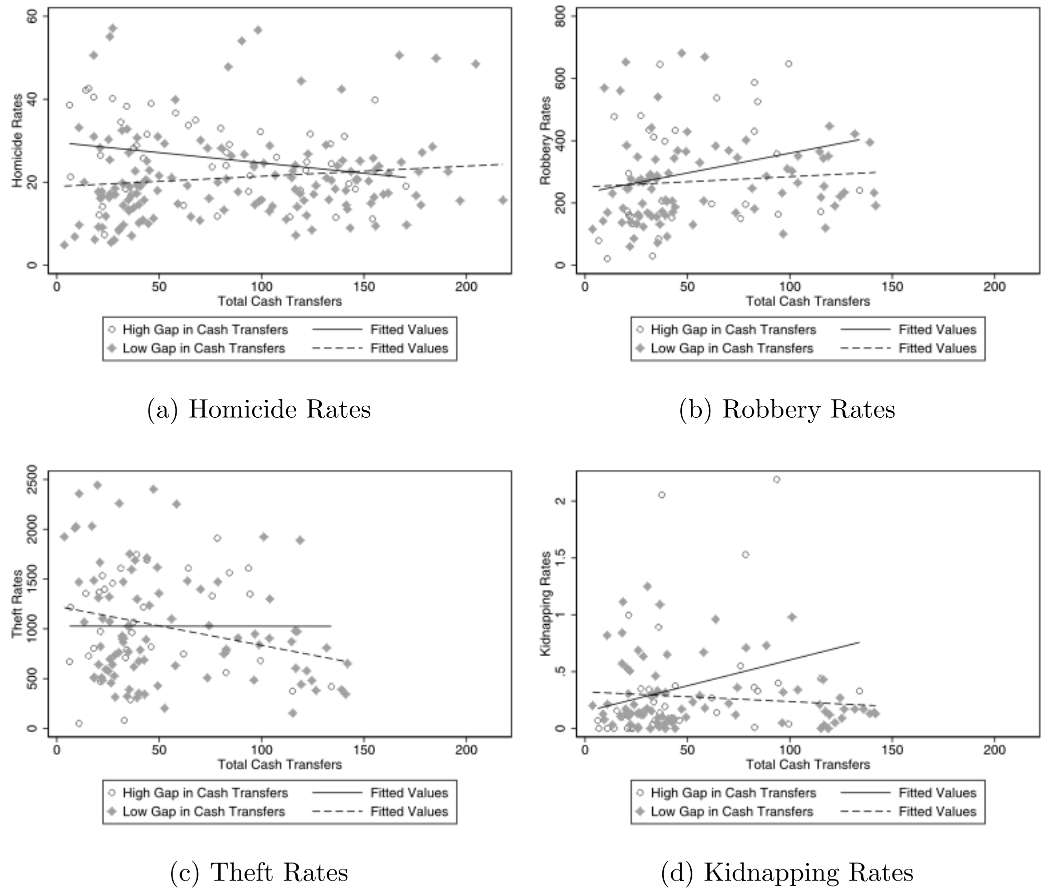
Left hand side variable:	(1) Poverty	(2) Extreme Poverty
State-specific trends as controls (Equation 1.8)	-0.0434** (0.0253)	-0.0255** (0.0132)
Median as threshold	-0.0306*** (0.0105)	-0.0214*** (0.0098)
2003/2004 comparison	-0.0354* (0.0184)	-0.0283** (0.0151)
Socio-economic variables as controls	-0.0216** (0.0109)	-0.0194** (0.0088)
Intensity of Poverty as Control (P(1) - Distance from Poverty Line)	-0.0312*** (0.0127)	-0.0256** (0.0092)
Poverty in 2001 as Control (No Fixed Effects)	-0.0321** (0.0141)	-0.0199** (0.0089)
Poverty in 2001/2003 as Control (Fixed Effects) (Poverty in 2001 if year $\leq$ 2002, 2003 if year $\geq$ 2003)	-0.0298** (0.0142)	-0.0201*** (0.0097)

Notes: All coefficients refer to the DD estimates ($\hat{\delta}$). Standard errors robust to heteroscedasticity and clustering at the state level in parentheses. Similar significance when standard errors are bootstrapped. "Poverty" reads "Extreme poverty" for control variables in column (2). All control variables used are significant at least at 0.05 level. Regression with socio-economic variables controlled for average labour income, average years of schooling and degree of informality in the labour market. For details about those variables, see table 1.7.
*** p<0.01, ** p<0.05, * p<0.1
Data Sources: MDS/IPEA/PNAD.

poverty is much more intense. This corroborates the hypothesis that many people are still poor even after receiving the benefit.

1.6 Estimating the Impact of CCT on Crime

The following figures show the association between crime rates and cash transfers for states with different speeds in the implementation of *Bolsa Familia* programme as defined by equation 1.4.



(a) Homicide Rates

(b) Robbery Rates

(c) Theft Rates

(d) Kidnapping Rates

Figure 1.7: Crime rates X Cash Transfers: Different levels of variation in Cash Transfers *Bolsa Familia*

Data Sources: SENASP/IPEA/PNAD.

Figure 1.7a shows a weak positive relationship between Homicide and Cash Transfers and no significant difference between high and low treatment.[45] There is also a positive association for Robbery, however with a significantly smaller slope for states with faster spending (Figure 1.7b).

In the case of theft and kidnapping rates, different relationships between crime and cash transfers for different levels of treatment emerge: increased cash transfers is associated with lower theft and kidnapping rates for states with lower gap in CCT (Figures 1.7c and 1.7d).

The estimation of the effect of CCT on crime rates using the differences-in-differences (DD) framework used in the previous section and other robustness checks are presented in table 1.6.

As discussed in section 1.5.3, the averages of murder and homicide are statistically different across the high and low gap groups defined in terms of the 75th percentile of the distribution of variation in the cash transfers, and therefore, the results for those variables are not valid. They are only kept in the analysis to contrast the result for those variables with the results obtained for the other variables. The only valid result for those categories of crime would be for the thresholds defined in terms of the mean and the median. In any case, it can be observed that they are not statistically significant in any specification.

Significant negative effects are found for robbery, theft and kidnapping for some definitions of low gap in the implementation of the programme. That is in line with Johnson, Kantor, and Fishback (2007) that find that relief spending after the great depression in the US has a negative effect on burglary, robbery and auto theft, but no significant impact on murder, aggravated assault and rape.

In order to test the robustness of the DD estimates, it is necessary to account for socio-economic variables and law enforcement expenditures. It is well established in the economics of crime literature that variables associated to the probability of punishment are endogenous.[46] I use tax revenue in the states in the previous year as an instrument to law enforcement, as those two variables are correlated and it can be argued that tax revenue would only affect crime rates through the extent of law enforcement expenditures.

[45] Very similar picture for murder rates (not reported here).

[46] See Levitt (1997) and McCrary (2002) for discussion.

Table 1.6: Difference-in-Difference Estimates: Crime Rates for different levels in the implementation of *Bolsa Familia* Program, before and after 2004 - Alternative Specifications

	(1) Murder	(2) Homicide	(3) Robbery	(4) Theft	(5) Kidnapping
75th percentile as threshold in D_i	-0.739	-0.489	-39.525**	-43.685*	-0.066*
	(3.183)	(0.307)	(21.561)	(26.426)	(0.043)
Median as threshold in D_i	2.779	4.623	-16.809*	-37.849	-0.109**
	(2.779)	(2.955)	(9.033)	(55.590)	(0.062)
2003/2004 comparison	-0.328	-0.399	-52.581	-81.932	-0.701*
	(0.424)	(0.831)	(135.563)	(91.477)	(0.413)
Socioeconomic variables and	-1.125	-1.692	-8.511	-56.145*	-0.064
Law Enforcement as controls	(3.220)	(2.838)	(13.981)	(30.436)	(0.228)
(IV = Tax Revenue(t-1))					
Socioeconomic variables and	-1.382	-1.338	-32.612	-51.196	0.138
Law Enforcement(t-1) as controls	(2.880)	(1.761)	(39.404)	(39.052)	(0.131)

Notes: All coefficients refer to the DD estimates ($\hat{\delta}$). Standard errors robust to heteroscedasticity and clustering at the state level in parentheses. Similar significance when standard errors are bootstrapped. Regression with socio-economic variables controlled for income inequality, average labour income, years of schooling, unemployment rate, % of male aged between 15 and 24, % of households with only one parent, and degree of informality in the labour market. For details about those variables, see table 1.7.
*** p<0.01, ** p<0.05, * p<0.1
Data Sources: SENASP/MDS/IPEA/PNAD.

The fourth line of table 1.6 present the DD estimates when socio-economic variables and law enforcement expenditures are used as controls and tax revenue as an instrumental variable. Only theft still has a significant from a more prompt implementation of the *Bolsa Familia* programme, when other relevant factors that affect crime are taken into account.

It can be argued that tax revenue is not a suitable instrument in this context as it could be that case that it is not excluded from the crime equation. One alternative approach is simply to use law enforcement expenditures lagged in one period.

The use of lagged law enforcement would minimise that problem. Furthermore as argued by Chalfin and McCrary (2013), this endogeneity is less likely to be the case as the response from the policy makers in the short run is very limited.

The model is re-estimated using this approach and it is presented in the last line of table 1.6. No significant effect is observed with this approach.[47]

This result is in line with other papers that find little or no evidence of the effect of social welfare spending on criminal behaviour, namely Chamlin, Cochran, and Lowenkamp (2002), Burek (2005) and Worrall (2005).

[47] First differences lead to similar results, but with even larger standard errors.

1.7 Conclusions

Heterogenous implementation of *Bolsa Familia* programme is used in order to identify the effect of this CCT programme on poverty and crime rates. *Bolsa Familia* has a significant effect on poverty reduction. States that reached the level of cash transfers expenditures proposed by the guidelines of the programme more promptly had a more significant reduction in poverty rates. However, many recipients of the programme are still below the poverty line.

Some results suggest that CCT expenditures contribute to reduction in robbery, theft and kidnapping rates, while no significant effect was found for homicide and murder. These findings would indicate that property crime would be more sensitive to CCT programmes. However, the results are not robust to different specifications and the inclusion of socio-economic variables and law enforcement expenditures. The positive or insignificant effect for the other types of crimes may suggest that the proposed approach is not efficient to correct the endogeneity problems in the crime equations.

Another possibility might be related to the fact that many people receiving the benefits are still poor or below a threshold of "acceptable" income, making the illicit activities still worth the risk. Moreover, two other parallel effects might dominate the relationships. More cash transfers can reduce willingness to work, increasing the probability of involvement with illicit activities. It can also positively affect informality to the detriment of formal labour, in the situations where the family would lose the benefit otherwise. This could potentially put individuals in contact with people that commit petty crimes affecting their probability of involvement in criminal activities.

It can also be argued that because individuals engaged in the criminal career have lower costs to commit a crime, they are not affected by variations in their legal income and a CCT programme would only influence the decision of individuals at the margin of the threshold between crime and no crime. That possibility will be analysed in the next chapter.

References

ANGRIST, J. D., AND J.-S. PISCHKE (2009): *Mostly Harmless Econometrics*. Princeton University Press.

——— (2010): "The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics," Discussion Paper 4800, IZA.

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

BENOIT, J.-P., AND J. M. OSBORNE (1995): "Crime, Punishment, and Social Expenditure," *Journal of Institutional and Theoretical Economics*, 151, 326–347.

BESLEY, T., AND A. CASE (2000): "Unnatural Experiments? Estimating the Incidence of Endogenous Policies," *The Economic Journal*, 110(467), F672–F694.

BOURGUIGNON, F. (1998): "Crime as a social cost of poverty and inequality: a review focusing on developing countries," Working Papers Mimeogr., DELTA.

BUONANNO, P., D. MONTOLIO, AND P. VANIN (2009): "Does Social Capital Reduce Crime?," *Journal of Law and Economics*, 52.

BURDETT, K., R. LAGOS, AND R. WRIGHT (2003): "Crime, Inequality, and Unemployment," *The American Economic Review*, 93(5), 1764–1777.

BUREK, M. W. (2005): "Now Serving Part Two Crimes: Testing the Relationship Between Welfare Spending and Property Crimes," *Criminal Justice Policy Review*, 16(3).

CAMERON, A. C., J. B. GELBACH, AND D. L. MILLER (2008): "Bootstrap-Based Improvements for Inference with Clustered Errors," *Journal of Economics and Statistics*, 90(3), 414–427.

CAMERON, A. C., AND P. K. TRIVEDI (2005): *Microeconometrics: Methods and Applications*. Cambridge University Press.

CHALFIN, A., AND J. MCCRARY (2013): "The Effect of Police on Crime: New Evidence from US Cities, 1960-2010," Discussion paper, National Bureau of Economic Research.

CHAMLIN, M., J. COCHRAN, AND C. LOWENKAMP (2002): "Longitudinal Analysis of the Welfare-Homicide Relationship," *Homicide Studies*, 6(1), 39–60.

CHIODA, L., J. DE MELLO, AND R. SOARES (2012): "Spillovers from conditional cash transfer programs: Bolsa Família and crime in urban Brazil," Discussion Paper 6371.

CORNWELL, C., AND W. N. TRUMBULL (1994): "Estimating the Economic Model of Crime with Panel Data," *The Review of Economics and Statistics*, 76(2), 360–66.

DILLS, A. K., J. A. MIRON, AND G. SUMMERS (2008): "What Do Economists Know About Crime?," Working Paper 13759, NBER.

FISZBEIN, A., AND N. SCHADY (2009): *Conditional Cash Transfers: Reducing Present and Future Poverty*. Washington, DC: World Bank.

FOLEY, C. F. (2011): "Welfare payments and crime," *The Review of Economics and Statistics*, 93(1), 97–112.

FOSTER, J., J. GREER, AND E. THORBECKE (1984): "A class of decomposable poverty measures," *Econometrica*, 52, 761765.

GERTLER, P. (2004): "Do Conditional Cash Transfers Improve Child Health? Evidence from PROGRESA's Control Randomized Experiment," *American Economic Review*, 94(2), 336–341.

GLEWWE, P., AND A. L. KASSOUF (2008): "The Impact of the Bolsa Escola/Familia Conditional Cash Transfer Program on Enrollment, Grade Promotion and Drop out Rates in Brazil," Proceedings of the 36th brazilian economics meeting, ANPEC - Brazilian Association of Graduate Programs in Economics.

HUANG, C.-C., D. LAING, AND P. WANG (2004): "Crime and Poverty: A Search-Theoretic Approach," *International Economic Review*, 45(3), 909–938.

IMBENS, G. M., AND J. M. WOOLDRIDGE (2008): "Recent Developments in the Econometrics of Program Evaluation," Working Paper 14251, NBER.

IMROHOROGLU, A., A. MERLO, AND P. RUPERT (2000): "On the Political Economy of Income Redistribution and Crime," *International Economic Review*, 41(1), 1–25.

JOHNSON, R. S., S. KANTOR, AND P. V. FISHBACK (2007): "Striking at the Roots of Crime: The Impact of Social Welfare Spending on Crime During the Great Depression," NBER Working Papers 12825, National Bureau of Economic Research, Inc.

LEVITT, S. (1997): "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime," *American Economic Review*, 87(3), 270–290.

LOCHNER, L., AND E. MORETTI (2004): "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports," *The American Economic Review*, 94(1).

MCCRARY, J. (2002): "Using electoral cycles in police hiring to estimate the effect of police on crime: Comment," *American Economic Review*, pp. 1236–1243.

MEDEIROS, M., T. BRITO, AND F. V. SOARES (2008): "Targeted Cash Transfer Programmes in Brazil: BPC and the Bolsa Familia," Working Paper 46, International Poverty Centre, UNDP.

MERLO, A. (2003): "Income Distribution, Police Expenditures, and Crime: A Political Economy Perspective," *Journal of the European Economic Association*, 1, 450–458.

PRATT, T., AND T. GODSEY (2002): "Social Support and Homicide: A Cross-National Test of an Emerging Criminological Theory," *Journal of Criminal Justice*, 30.

RAWLINGS, L. B., AND G. M. RUBIO (2003): "Evaluating the impact of conditional cash transfer programs: lessons from Latin America," Policy Research Working Paper Series 3119, The World Bank.

REIS, M. (2010): "Cash transfer programs and child health in Brazil," *Economics Letters*, 108, 22–25.

RESENDE, A. C. C., AND A. M. H. C. D. OLIVEIRA (2008): "Avaliando resultados de um programa de transferência de renda: o impacto do Bolsa-Escola sobre os gastos das famlias brasileiras," *Estudos Econômicos*, 38.

SAVAGE, J., R. BENNETT, AND M. DANNER (2008): "Economic Assistance and Crime," *European Journal of Criminology*, 5(3).

SICKLES, R. C., AND J. WILLIAMS (2008): "Turning from crime: A dynamic perspective," *Journal of Econometrics*, 145, 158173.

SKOUFIAS, E., AND V. D. MARO (2008): "Conditional Cash Transfers, Adult Work Incentives, and Poverty," *Journal of Development Studies*, 44(7), 935–960.

STEPHEN MACHIN, O. M., AND S. VUJIĆ (2010): "The Crime Reducing Effect of Education," Discussion Paper 5000, IZA.

WITTE, A. D., AND R. WITT (2001): "What we spend and what we get: Public and private provision of crime prevention and criminal justice," *Fiscal Studies*, 22(1), 1–40.

WOOLDRIDGE, J. M. (2002): *Econometric Analysis of Cross Section and Panel Data*, MIT Press Books. The MIT Press.

WORRALL, J. L. (2005): "Reconsidering the Relationship Between Welfare Spending and Serious Crime: A Panel Data Analysis with Implications for Social Support Theory," *Justice Quarterly*, 22(3), 364–391.

——— (2009): "Social Support and Homicide," *Homicide Studies*, 13(2), 124–143.

ZHANG, J. (1997): "The Effect of Welfare Programs on Criminal Behavior: A Theoretical and Empirical Analysis," *Economic Inquiry*, 35(1), 120–37.

1.A Appendix

1.A.1 Inconsistency in the estimation of the effect of CCT on Crime

An alternative to estimate the effect of cash transfers on crime would be to omit poverty rates from crime equation and take the endogeneity of the redistribution policy and its direct effect on crime into account. [48] If this is the case, we use equation 1.2 and rewrite equation 1.3 as:

$$Crime_{it} = \psi_i + \omega_t + \mathbf{z}'_{it}\beta + \gamma Law_{it} + \delta CCT_{it} + \epsilon_{it} \tag{1.9}$$

where $\epsilon_{it} = \eta Pov_{it} + \zeta_{it}$.

In order to analyse the bias that may stem from the inconsistency that emerges from the estimation of equation 1.9, consider the variables in equation 1.2 net of $\mathbf{x}_{it}$ and the specific unobserved effects ϕ_i and λ_t denoted by a tilde[49]:

$$\tilde{CCT}_{it} = \theta \tilde{Pov}_{it} + \nu_{it}. \tag{1.10}$$

A similar procedure for equation 1.9, but now net of ψ_i, ω_t, Law_{it} and $\mathbf{z}_{it}$ yields:

$$\tilde{Crime}_{it} = \delta \tilde{CCT}_{it} + \varepsilon_{it} \tag{1.11}$$

where $\varepsilon_{it} = \eta \tilde{Pov}_{it} + \xi_{it}$.

Note that θ and δ, the parameters of interest, remain unchanged. The OLS estimator estimator of the effect of CCT on Crime can be written as:

$$\hat{\delta} = (\sum_{i=1}^{n}\sum_{t=1}^{T} \tilde{CCT}'_{it}\tilde{CCT}_{it})^{-1}(\sum_{i=1}^{n}\sum_{t=1}^{T} \tilde{CCT}'_{it}\tilde{Crime}_{it}) \tag{1.12}$$

or in a more convenient matrix form:

$$\hat{\delta} = (\tilde{CCT}'\tilde{CCT})^{-1}(\tilde{CCT}'\tilde{Crime}). \tag{1.13}$$

Plugging the matrix correspondent of equation 1.11, with respective error term, on the previous expression results in:

---

[48] This analysis is based on Besley and Case (2000), that provides a more general analysis on endogenous policies.

[49] Frisch-Waugh-Lovell theorem assures that the estimated parameters are identical.

$$\hat{\delta} - \delta = \eta(C\tilde{C}T^{'}C\tilde{C}T)^{-1}(C\tilde{C}T^{'}P\tilde{o}v) + (C\tilde{C}T^{'}C\tilde{C}T)^{-1}(C\tilde{C}T^{'}\xi) \tag{1.14}$$

and using 1.10 and applying the probability limit, we get:

$$\begin{aligned} plim(\hat{\delta} - \delta) = plim & \left[\left(\frac{C\tilde{C}T^{'}C\tilde{C}T}{n} \right)^{-1} \eta\theta' \left(\frac{P\tilde{o}v^{'}P\tilde{o}v}{n} \right) \right] \\ & + plim \left[\left(\frac{C\tilde{C}T^{'}C\tilde{C}T}{n} \right)^{-1} \left(\frac{\nu'\xi}{n} \right) \right] \end{aligned} \tag{1.15}$$

where plims involving Pov and the error terms jointly are equal to zero, by construction.

From equation 1.15 we conclude that $\hat{\delta}$ is inconsistent due to two non-null quantities: (1) $plim \left[\eta\theta' \left(\frac{P\tilde{o}v^{'}P\tilde{o}v}{n} \right) \right]$ which depends on the observable poverty rate that is very frequently omitted in the crime equation and the parameters that describe the effect of Pov on crime and CCT; (2) $plim \left[\frac{\nu'\xi}{n} \right]$, which is a function of the unobserved omitted variables that determine both crime rates and the amount of resources spent on cash transfers.

1.A.2 Additional Tables

Table 1.7: Definition and Sources of Variables

Variable	Description	Source
Murder	Murder Rate (100,000 inhabitants)	SENASP
Homicide	Homicide Rate (100,000 inhabitants)	Health Ministry
Robbery	Robbery Rate (100,000 inhabitants)	SENASP
Theft	Larceny Rate (100,000 inhabitants)	SENASP
Kidnapping	Kidnapping Rate (100,000 inhabitants)	SENASP
Poverty	People below poverty line (%) (IPEA)	IPEA/PNAD
Extreme Poverty	People below extreme poverty line (%) (IPEA)	IPEA/PNAD
Gini	Income Gini Coefficient	IPEA/PNAD
Labour Income	Mean Per capita Household Income from Labour (INPC corrected - R$ 2007)	IPEA/PNAD
Education	Mean number of years of schooling (people aged 25 or more)	IPEA/PNAD
Young Male	Percentage of male aged between 15 and 24 in relation to all population	PNAD
Unemployment	Unemployment Rate	PNAD
Parent	Percentage of households with only one parent (Female Headed)	PNAD
Informality	Informality degree in the labour market (%)	PNAD
Law Enforcement	Per capita government expenditure on law enforcement (INPC corrected - R$ 2007)	STN
Cash Transfers	Per capita government expenditure on Bolsa Familia and Continuous Cash Benefit (INPC corrected - R$ 2007)	MDS, Portal da Transparência
Revenue	Per capita Tax revenue of the States (INPC corrected - R$ 2007)	STN

Notes: All variables constructed by the author with data from the mentioned sources.
SENASP - *Secretaria Nacional de Segurança Pública* (National Secretariat of Public Security)
IPEA - *Instituto de Pesquisa Econômica Aplicada* (Institute of Applied Economic Research)
PNAD - *Pesquisa Nacional por Amostra de Domicílios* (National Survey by Household Sampling)
STN - *Secretaria do Tesouro Nacional* (National Treasury Secretariat)
TSE - *Tribunal Superior Eleitoral* (Supreme Electoral Court)
MDS - *Ministério do Desenvolvimento Social e Combate à Fome*
(Ministry of Social Development and Hunger Eradication)

Table 1.8: Selected variables for different levels in the implementation of Bolsa Familia (Perc. Var. in CCT > Mean)

	Whole Sample			2003-2004		
	Low Gap ($D_i = 1$)	**High Gap** ($D_i = 0$)	**t-test on diff.**	**Low Gap** ($D_i = 1$)	**High Gap** ($D_i = 0$)	**t-test on diff.**
Murder Rate	21.3008	24.7215	-0.85	20.1200	24.3625	-1.08
Homicide Rate (Health)	24.2891	29.7795	-1.27	23.4567	29.6574	-1.38
Robbery Rate	276.5995	460.9220	-1.87	287.0144	440.1256	-1.96
Theft Rate	969.9526	1276.5770	-1.11	1026.4530	1240.7560	-0.79
Kidnapping Rate	0.3380	0.3928	-0.58	0.3683	0.2501	0.88
Poverty Rate	0.3630	0.3374	0.42	0.4028	0.3784	0.38
Extreme Poverty Rate	0.1474	0.1296	0.54	0.1713	0.1569	0.36
Gini Index	0.5568	0.5505	0.51	0.5594	0.5570	0.19
Average Labour Income	671.8287	834.5558	-1.60	635.7222	796.4894	-1.64
Informality in the Labour Market	57.4600	54.3505	0.78	58.2300	54.9296	0.83
Single-Parent Households	0.2719	0.2886	-1.36	0.2725	0.2889	-1.27

Notes: High Gap and Low Gap as defined in equation 1.4 with the **mean** of percentage variation in CCT as the threshold.
Data Sources: See Table 1.7.

Table 1.9: Selected variables for different levels in the implementation of Bolsa Familia (Perc. Var. in CCT > Median)

	Whole Sample			2003-2004		
	Low Gap ($D_i = 1$)	**High Gap** ($D_i = 0$)	**t-test on diff.**	**Low Gap** ($D_i = 1$)	**High Gap** ($D_i = 0$)	**t-test on diff.**
Murder Rate	21.0835	24.6925	-0.90	19.7821	24.4000	-1.19
Homicide Rate (Health)	24.5064	29.1232	-1.06	23.6735	28.9469	-1.17
Robbery Rate	278.6980	446.3674	-1.66	290.0293	425.1010	-1.71
Theft Rate	932.3053	1296.7950	-1.44	969.8768	1285.1990	-1.19
Kidnapping Rate	0.3469	0.3789	-0.36	0.3597	0.2693	0.67
Poverty Rate	0.3589	0.3437	0.25	0.3982	0.3852	0.20
Extreme Poverty Rate	0.1448	0.1337	0.34	0.1682	0.1614	0.17
Gini Index	0.5542	0.5538	0.03	0.5575	0.5592	-0.13
Average Labour Income	668.8614	825.2339	-1.54	631.0525	789.1516	-1.62
Informality in the Labour Market	57.6327	54.4037	0.81	58.2441	55.1684	0.78
Single-Parent Households	0.2692	0.2903	-1.78	0.2721	0.2881	-1.25

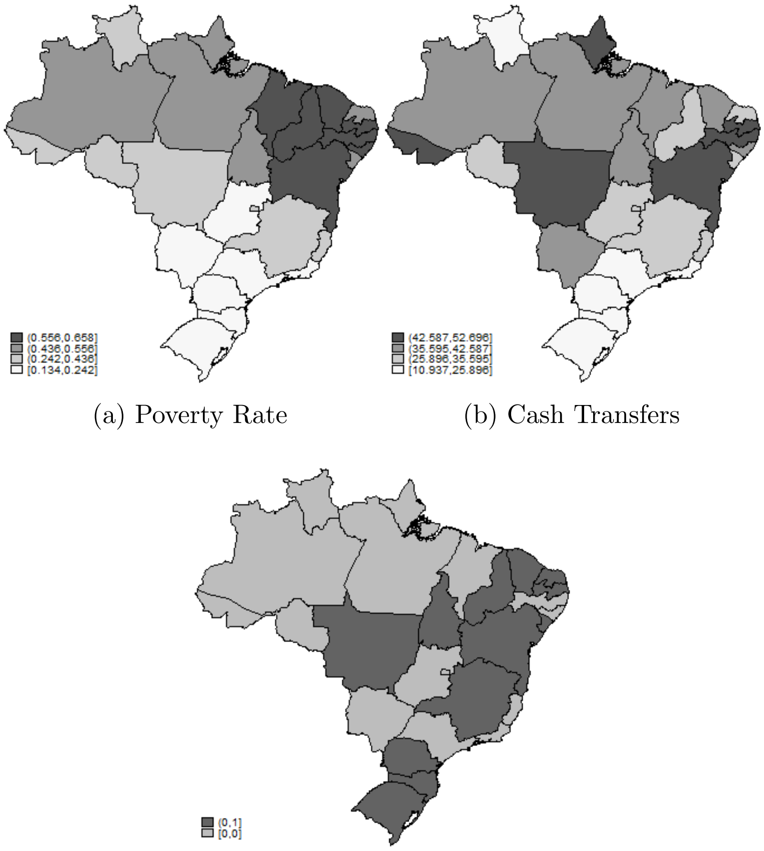
Notes: High Gap and Low Gap as defined in equation 1.4 with the **median** of percentage variation in CCT as the threshold.
Data Sources: See Table 1.7.

Table 1.10: Effect of Poverty on the lag in the implementation

	CCT perc. var. 04/08	CCT perc. var. 04/08	CCT perc. var. 03/04	CCT perc. var. 03/04
Poverty in 2001	-0.4114		0.4489***	
	(0.3145)		(0.0907)	
Poverty		-0.2737		0.4029***
		(0.3748)		(0.0869)
Constant	0.9251***	0.8376***	0.3849***	0.4065***
	(0.1326)	(0.1128)	(0.0382)	(0.0363)
R^2	0.0266	0.0209	0.4944	0.4623

*** p<0.01, ** p<0.05, * p<0.1 *Data Sources*: MDS/IPEA/PNAD.

1.A.3 Additional Figures



(a) Poverty Rate (b) Cash Transfers

(c) Gap in CCT Variation

Figure 1.8: Poverty, Cash Transfers and lag in its variation in Brazil - 2004
Note: Classes based on quartiles. *Data Sources*: SENASP/MDS/IPEA/PNAD.

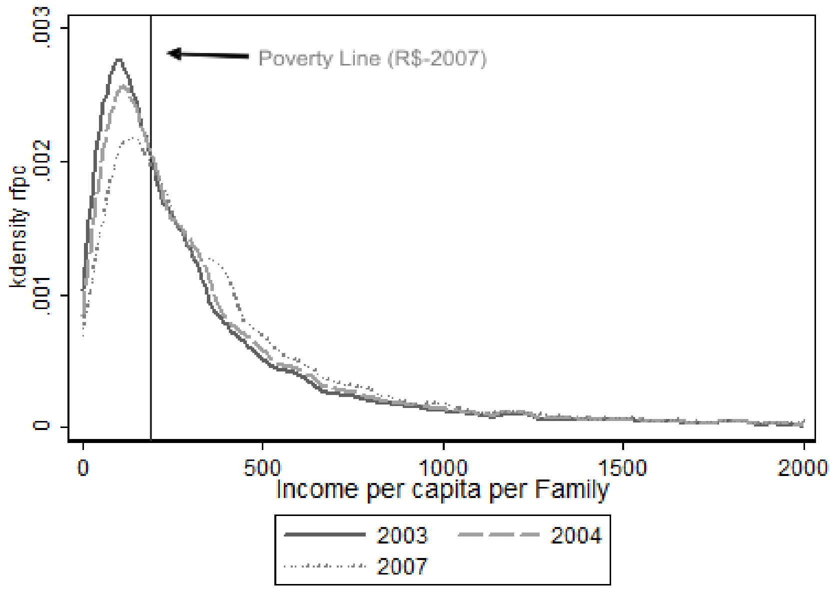


Figure 1.9: Income Distribution 2003, 2004 and 2007 (*R$* - 2007)
Note: Corrected for Inflation *Data Sources*: IPEA/PNAD.

Chapter 2

Hysteresis in Crime

2.1 Introduction

There is a significant gap between theory and empirical evidence about policies to reduce crime rates. The majority of the theoretical analyses predict a sharp decrease in crime rates when there are significant improvements in the economic conditions, such as a fall in unemployment and poverty rates or when legal market income has a significant increase. A similar prognosis is established in the case of a substantial increase in the probability of punishment. However, a predominant part of the empirical literature in the economics of crime observe a lower than expected effect on crime rates from exogenous variations in economic variables.[1]

This chapter argues that a possible reason for this is the fact that the current literature of economics of crime overlooks a likely hysteresis effect in the criminal behaviour, a situation in which, ceteris paribus, individuals that have a criminal past are more prone to take crime opportunities than someone that has never committed one, if the original conditions are subsequently restored, a subset of these agents will continue in their career in crime. As it will be discussed below, a relevant consequence of hysteresis in criminal behaviour is the fact that social policies to reduce crime will have a more important impact on potential criminals than on existing criminals, which is a more realistic representation of criminal decision, where the current crime decision of individuals are influenced by their previous choices.

Most of the theory treats potential and existing criminals as having the same propensity to commit a crime.[2] That is a very restrictive assumption. As it will be

[1] Freeman (1999), Dills, Miron, and Summers (2008) and Mustard (2010) show the great discrepancy between economic theory and empirical evidence.

[2] The few exceptions are given by Imrohoroglu, Merlo, and Rupert (2004), Mocan, Billups, and

discussed in this chapter, when potential and existing criminals are treated differently, contrasting conclusions are reached in terms of crime reduction policies.

If hysteresis is present in the criminal behavior, crime reduction policies will have a diminished impact when compared to the expected impact where individuals with a criminal past will behave similarly to individuals without a criminal past. That asymmetric effect will be very clear in a situation in which the crime reduction policy in a given period is simply a reversal of a deterioration of one determinant of crime. A concrete example would be a situation where part of the police officers in a city are dismissed in a given year, resulting in a escalation in crime. If all sacked police officers are readmitted in the following year in an attempt to restore the original crime levels and hysteresis is present in the criminal behavior, that policy will result in a lower crime rate, but higher than the original one.

I will show that depending on whether one assumes homogeneity or heterogeneity of agents with respect to their intrinsic costs of crime, two distinct types of hysteresis will arise.[3] In this chapter, I will refer to weak hysteresis as a phenomenon that occurs at the individual level only if specific threshold levels are surpassed. As it will be seen in the next sections, if the individuals are assumed to be homogeneous and their decisions are aggregated, that type of hysteresis also occur in the aggregate level. However, as it will be shown, with heterogeneous agents, the aggregate variation is reinforced generating a strong hysteresis effect. At every level of the input variable, positive variations result in different effects on the output variable when compared to effects from negative variations. Furthermore, as it will be seen in the next sections, unlike the situation of weak hysteresis, the amplitude of the remanence will depend on the magnitude of the shock.

I claim that there are two types of sources of hysteresis in crime: 1. **External** to the individuals (extrinsic). Weak hysteresis with this source has recently been considered in the literature by introducing social stigma into the labour market: Lower wage offers (Imrohoroglu, Merlo, and Rupert (2004)); higher duration in unemployment (Engelhardt (2010)); social capital depreciation (Sickles and Williams (2008)). Hysteresis would only be a result in this context in the situations where the criminal is caught *and* convicted and/or this fact is known by the others individuals in his social life. Nevertheless, data on countries like the US and the UK show that the punishment for the majority of committed crimes do not involve incarceration and the

Overland (2005), Sickles and Williams (2008), Santos (2009) and Engelhardt (2010). However, as it will be shown, they all assume specific types of non-stationarity associated to stigma in the labour market or human capital accumulation.

[3]This should not be confused with others uses of the term hysteresis in economics.

apprehension rate is very low.[4]

2. **Internal** to the individuals (intrinsic). The claim in this paper is that hysteresis could also occur as a result of sources internal to the individuals. As it will be shown in the following sections, sunk moral cost (fallacy) and learning in crime can lock individuals in the criminal career. Internal sources are crucial to an explicit characterisation of strong hysteresis in criminal choice.

This chapter argues that stigma in the labour market is only one of the sources of hysteresis in criminal behaviour[5] and a more relevant explanation of hysteresis stems from internal sources to the individual. To date and to my knowledge, there has been no formal model or empirical study on internal hysteresis in criminal behaviour. It is also the first work to place hysteresis at the heart of the analysis of criminal behaviour.[6] Moreover, this is the first effort to embed hysteresis (intrinsic or extrinsic) in an equilibrium search model.

In the following section, I discuss the related literature in hysteresis and crime. The relevance of hysteresis in criminal decisions is discussed in section 2.3. A benchmark framework is established in section 2.4, while I develop a non-stationary model to show the hysteresis effect in sections 2.5 and 2.6. In section 2.7, I analyse the consequences of hysteresis in the case of forward-looking agents. Section 2.8 concludes.

2.2 Related Literature

2.2.1 Hysteresis in Economics

Hysteresis has been a phenomenon identified in economic contexts like foreign investment (Dixit (1989, 1992)) and unemployment (Blanchard and Summers (1987) and Røed (2002)).[7] However, the use of the word hysteresis is not consistent in economics and it often used as a synonym to persistence.

Cross (1993) and Amable, Henry, Lordon, and Topol (1994) formally define hysteresis in the economic context, showing the difference between its weak and

[4] In some cities of the US and the UK, the clearance rate - crimes with a charge being laid divided by the total number of crimes recorded - is as low as 5%. Nationwide the fraction of cleared crimes varies between 1/5 and 1/4. See FBI's Uniform Crime Reporting (UCR) website for data on the US and the Home Office Statistical Bulletin for information about clearance rates in the UK.

[5] I provide a detailed review of that literature in the next section.

[6] Note that hysteresis is not necessarily the term used in the existing literature, nor is hysteresis the main focus in those papers.

[7] For a more recent exposition of the use of the concept of hysteresis in economics see Göcke (2002).

strong versions. It is pointed out the usual improper use of the word in economics to describe persistence stemming from unit root for discrete processes or zero eigenvalue for linear dynamic systems.

Amable, Henry, Lordon, and Topol (1994) let clear that a shock exerted on a state variable in a unit root process will lead to a new steady state, but will keep the number of equilibria and its respective locus unchanged. If crime rates are unit root processes, two opposite shocks with the same magnitude will leave the crime rate unaffected, whereas in a hysteric system they will lead the crime rate to a new equilibrium. In a system with hysteresis, the current behaviour depends on the dominant extremum of past shocks, whereas in a system with unit root, all past shocks matter. That implies that hysteresis is associated to local structural stability, whereas unit root processes are associated to global structural stability.

Unlike unit/zero root processes, both types of hysteresis display path dependence, asymmetric effects from the determinant variables and the remanence property. In the context of crime, remanence corresponds to crime rates not going back to its original value when the probability of punishment or the average legal wage is transitorily changed.

2.2.2 Crime Decision in Economics

Since the seminal paper of Becker (1968), economists consider formally the possible effects of socioeconomic variables on criminal behaviour. The standard fully rational crime decision model establishes that an individual periodically faces a decision whether to commit an illicit act or not, based on the expected return of the criminal market and associated probability and severity of punishment when compared to the expected stream of legal income.[8]

Just recently the economics of crime has formally considered the inherent intertemporal nature of criminal choice.[9] McCrary (2010) highlights the importance of considering a dynamic setting in order to properly describe the choice of a potential criminal. Burdett, Lagos, and Wright (2003, 2004) develop an on-the-job search model to analyse the interrelations between crime, unemployment and inequality. The authors find multiple equilibria, which suggest there is a hysteresis process in crime

[8]Stigler (1970), Ehrlich (1973) and Heineke and Block (1975) contributed to establish the economics of crime literature. Another relevant paper that applied Becker's paradigm in a purely economic law-breaking decision making was Allingham and Sandmo (1972) that proposed a model for tax evasion. Garoupa (1997) and Polinsky and Shavell (2007) provide thorough reviews of the models of optimal law enforcement.

[9]See McCrary (2010) for a extensive survey on dynamic models in the economics of crime literature.

decisions. The result helps to understand why higher levels of welfare benefits could lead to higher crime rates as higher taxes could discourage some individuals to work. Other relevant articles with dynamic settings are Imrohoroglu, Merlo, and Rupert (2004), Imai and Krishna (2004), Lochner (2004), Engelhardt, Rocheteau, and Rupert (2008), Sickles and Williams (2008) and Engelhardt (2010).

Nevertheless, it can be argued that even though the criminal decision takes place in a dynamic environment, especially regarding the future net stream of income in the legal market, that is not necessarily taken into account for individuals at the margin. Akerlof (1991) argues that any analysis that considers potential criminals evaluating all periods in the future is based on an unreasonable assumption.[10] The individual at the margin would be likely, in most cases, to be very myopic. Another ubiquitous questionable assumption in the dynamic models to describe the crime decision is that once the illicit option is chosen, the probabilities of an individual choosing the possible alternatives remain unaltered in the following period.

There are different reasons why hysteresis arises in the criminal choice. These sources can be classified into two groups: external to the individual (extrinsic) and internal to the individual (intrinsic). The main source of external hysteresis stems from the stigma in the labour market. Individuals with criminal records would have a higher duration in unemployment or lower wage offers. A more comprehensive source of external hysteresis is the depreciation of social capital. There are at least three sources of hysteresis internal to the individual that emerge in the criminal choice: sunk moral costs of the criminal activity, overly optimistic bias and criminal technology learning.

It is important to distinguish hysteresis in crime from recidivism, which has a specific literature, that is mostly empirical.[11] The studies on this topic focus on measuring the "treatment effect" of imprisonment and the duration before an individual commits a crime again once he/she is released from prison. While hysteresis emerges from the analysis of a situation in which the probability of committing a crime is higher if a crime was committed in the past, recidivism consists in a subset, since it focuses only on caught/arrested, and crucially, convicted individuals. More specifically, in a recidivism analysis, criminals that were never arrested and convicted will not be considered. It also deals with an intrinsic missing data problem, as the researcher does not necessarily observe all individuals that committed a crime again, since this observation depends on a new capture.[12] Another

[10] I thank Vasileios Vlaseros for making me aware of this paper.

[11] Recent studies are given by Bierens and Carvalho (2011), Carvalho (2009), Bierens and Carvalho (2007), Bowles and Florackis (2007), Freeman (2003) and Escarela, Francis, and Soothill (2000).

[12] Another relevant issue to be dealt with is right-censoring, as nothing prevents the ex-inmate

important difference lies in the fact that a significant part of crimes do not result in imprisonment, even if the offender is caught.[13] The punishments for these types of crimes are usually fines or community sentences. In most countries, sanctions that do not involve captivity are the most common punishment.[14] More importantly, the recidivism literature does not provide a full-fledged theoretical decision model.

One important source of hysteresis, is the social stigma that the criminal career entails and crucially results in additional frictions in the labour market. In this vein, Imrohoroglu, Merlo, and Rupert (2004) embedded stigma in the labour market, in the sense that a criminal record reduces their legal earnings. In a model with heterogeneous agents with respect to their income-earning ability, the existence of stigma results in a lower amount of juvenile delinquency and a higher amount of recidivism of older agents. It must be noted, however, that stigma in the model described above does not imply frictions in the labour market (or from a criminal career).

Freeman (1991) suggests that stigma affects the unemployment duration more significantly than it does on the decreases in wages. This is considered by Santos (2009), that embeds a Markov chain in a general equilibrium model in order to introduce a lower probability of getting a job for individuals with a criminal history. It concludes that harsher stigmatisation in the labour market implies lower crime rates (because agents are forward-looking) and a higher rate of recidivism. Similar conclusions are reached by Engelhardt (2010), as the author considers a search model of crime that incorporates heterogeneous agents with respect to their evaluations of leisure.[15] Using a time allocation framework, where time can be spent on three possible activities, namely, legitimate work, leisure and income producing crime and individuals accumulate social capital, Sickles and Williams (2008) concludes that the depreciation of this type of capital reduces the employability of individuals with criminal history. Mocan, Billups, and Overland (2005) also develop a formal model with similar aspects and find that individuals fall in a crime trap during a recession due to simultaneous depreciation of legal human capital and appreciation of criminal human capital. Empirical evidence is given in this direction by Imai and Krishna (2004), Buonanno, Montolio, and Vanin (2009), Mocan and Bali (2010) and

to commit a crime after the period of the collected data.

[13]The great majority of recidivism articles analyse individuals after they are released from prison, but there is also a recidivism literature that focus on petty crimes.

[14]The fraction of sentences involving confinement is approximately 5% in the UK and the US. See Sentencing Statistics - UK Ministry of Justice and US Sentencing Commission.

[15]Which implies heterogeneity in reservation wages in the legal sector. Worker heterogeneity is crucial to account for individual employment histories, especially the fact that hazard rates tend to decrease with the length of unemployment spell. (Rogerson, Shimer, and Wright (2005)) Additionally, it allows the model to explain why some individuals do not engage in crime.

Engelhardt (2010).

Criminal Interactions and Multiplicity of Equilibria

Sah (1991) was the first author to argue that criminal rates exhibit inertia. People estimate the actual probability of being punished by observing their peers, incorporating present and past information in their inferences generating inertia. Therefore, a reduction in the actual probability of being punished will have only a limited effect on peoples inferences and consequently on the current criminal rate in the short run. This inertia would lead to multiple steady states. However, there is no multiplicity of equilibria, as "there are no agents in the economy who can or wish to eliminate period-to-period changes." [Sah (1991)]

Calvó-Armengol and Zenou (2004) develop a model of networks where ex ante homogenous criminals compete with each other but benefit from being friends with other criminals by learning about the crime business. In a game in which individuals decide first to work or to become a criminal and then the crime effort provided if criminals, the authors show that multiple equilibria with different numbers of active criminals and levels of involvement in crime activities may coexist and are only driven by the geometry of the pattern of links connecting criminals. Similar conclusions are reached by Calvó-Armengol, Verdier, and Zenou (2007) in a model where non-criminal interact with criminals.

Bjerk (2010) develops a model analyzing how criminal behavior is associated with individual and neighborhood poverty. The model shows that even under relatively minimal assumptions, a connection between individual poverty and both property and violent crimes will arise, and moreover, neighborhood effects can develop, but will differ substantially in nature across crime types. A key implication is that greater economic segregation in a city should have no effect or a negative effect on property crime, but a positive effect on violent crime.

Multiple equilibria is also a result when the effect that level of crime c can have on the probability of punishment π is considered. Fender (1999) presents a general equilibrium model with this feature discussing the cases in which multiple equilibria may arise. In Bar-Gill and Harel (2001) multiple equilibria occurs when the authors consider not only the potential effect of c on π, but also the effect of c on s in an economy with risk-neutral agents. Using a different approach Kim (2012) get to similar conclusions. Ferrer (2010) analyses a model that explicitly takes into account the fact that the probability of punishment decreases with the total number of individuals in the criminal career in a society, implying multiple equilibria.

2.3 Stationarity in the Criminal Decision and Hysteresis

A crucial aspect of the dynamic economic models of criminal decision with a stochastic environment[16] is that, regardless of whether the illicit action in period t is taken or not, the probability of committing a crime in $t + 1$ remains unchanged. This is the assumption of stationarity, which assumes that the individual's perception of the future is independent of time or the duration of the criminal career.

Let θ_t be the propensity to commit a specific crime in a given period t. For a given probability of punishment, at any period t, an individual commit a crime if the utility associated to θ_t is larger than the one obtained from a threshold θ_t^*. The crime decision in two consecutive periods can then be depicted in a general fashion as in figure 2.1.

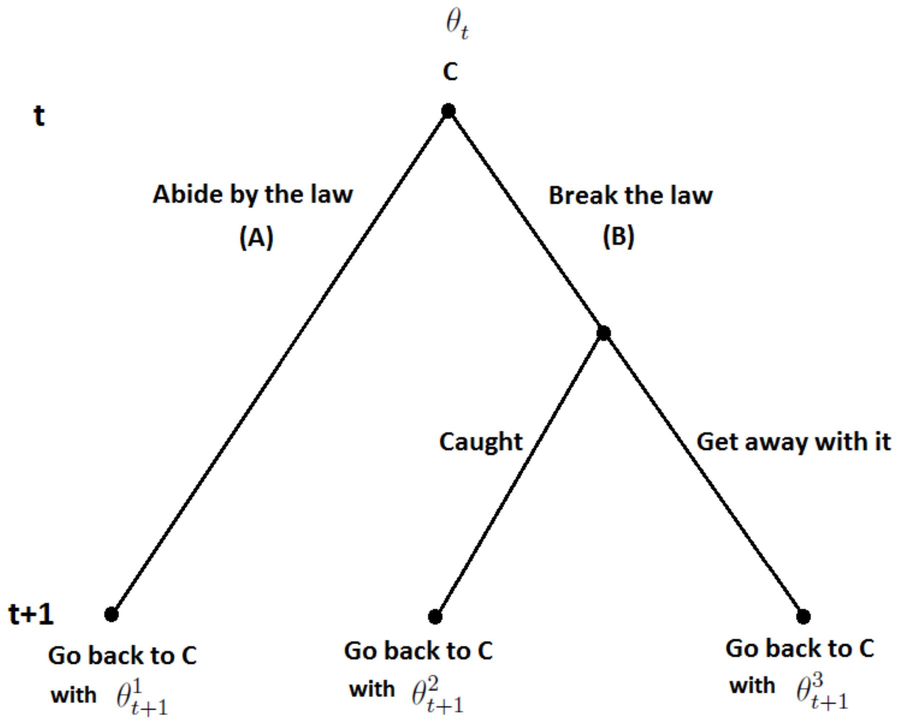


Figure 2.1: Crime Decision

For a given propensity to commit a crime and a probability of punishment, the stationary criminal decision of the individuals is the special case in figure 2.1 when

$$\theta^1_{t+1} = \theta^2_{t+1} = \theta^3_{t+1}. \tag{2.1}$$

[16]Stochastic transitions between states is a necessary assumption not only in order to account for hysteresis, but to allow the individual to determine a contingency plan. Models with deterministic transitions will lead to a situation where the decision makers pick a sequence of actions a priori, leaving them without any flexibility given the realised states.

However, it is reasonable to argue that depending on the choice and the response of the environment, the individuals reevaluate their choices. Being caught will reduce the propensity of individual below the previous propensity and getting away with the illicit act will increase the original propensity. That specific non-stationary case arise in figure 2.1 if

$$\theta^2_{t+1} < \theta^1_{t+1} < \theta^3_{t+1}. \tag{2.2}$$

In the case where condition 2.2 holds, figure 2.1 represents a non-stationary Markov process and must hold to the different types of crime. However, it can be argued that for types of crime where punishment entails incarceration, the net effect is ambiguous as the propensity to violate the norm will not necessarily decrease. Indeed, individuals can become more inclined to commit crimes due to social stigma and criminal learning in prison, increasing the probability of recidivism.[17] One way to circumvent this would be to focus on petty crimes, where the punishment usually involves a fine, rather than incarceration.

The propensity to break the law is inherently a latent variable and it can be captured in many different ways depending on the theoretical framework. A concrete situation would be an environment where individuals instead of observing the probability of punishment, they only observe a noisy signal correlated to the true probability of punishment and update their priors according to previous experience. If they are caught, they update their probability punishment estimates downwards, and upwards otherwise.

Another possibility where the crime decision is non-stationary occurs if all criminal acts entail an intrinsic cost, borne irrespective of whether the punishment occurs or not. If that cost is a decreasing function of criminal history, the relationship between the propensities to commit a crime would be

$$\theta^1_{t+1} < \theta^2_{t+1} < \theta^3_{t+1}. \tag{2.3}$$

In what follows, I will focus on a less abstract concept where the decision to commit a crime is determined by individuals comparing the probability of punishment and relevant economic variables to their corresponding thresholds.

[17] There is some evidence that time spent in the conventional jail increase the probability of recidivism. See Carvalho (2009) and Escarela, Francis, and Soothill (2000).

2.3.1 Hysteresis in the Individual and Aggregate Levels

If the crime decision in a given period is determined not only by the current expected costs and benefits entailed by the illicit act, but it is also affected by criminal decisions taken in the past, hysteresis at the individual level will arise. I will follow the hysteresis literature and label that type of phenomenon as weak hysteresis.

Figure 2.2 depicts the concept of weak hysteresis in the context of crime decision for variations in the probability of punishment. A sufficiently low probability of punishment ($\pi < \bar{\pi}^{Entry}$) will encourage an individual to commit a crime[18] ($\phi(\pi) = 1$), whereas a sufficiently high probability of punishment ($\pi \geq \bar{\pi}^{Exit}$) will unambiguously deter an illicit act. However, as it will be shown in this chapter, for one specific level of the probability of punishment between the two thresholds ($\bar{\pi}^{Entry} \leq \pi < \bar{\pi}^{Exit}$), both situations of crime and no crime are possible.

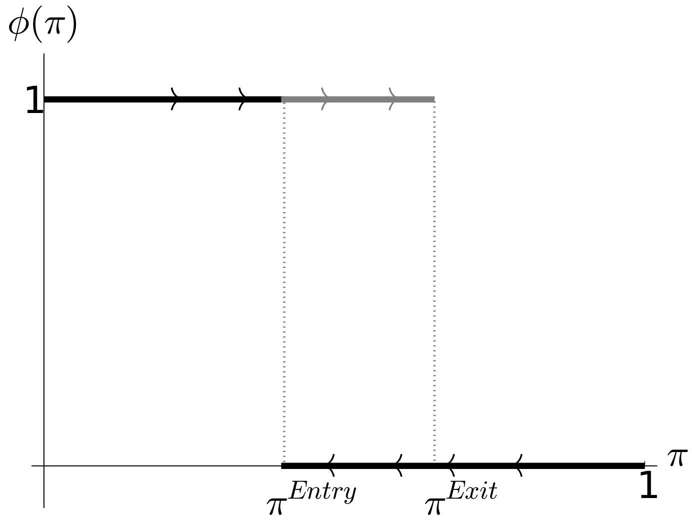


Figure 2.2: Weak Hysteresis - Individual Level

If the previous levels of π were all above $\bar{\pi}^{Entry}$, a situation in which $\bar{\pi}^{Entry} \leq \pi < \bar{\pi}^{Exit}$ will always result in no criminal acts. However, if at least one of the past values of π was below $\bar{\pi}^{Entry}$, a π in the region $\bar{\pi}^{Entry} \leq \pi < \bar{\pi}^{Exit}$ will always result in crime.

The aggregation of homogenous agents subject to hysteresis also displays the weak hysteresis effect and the pattern of hysteresis is similar to the individual level. However, with heterogeneous agents the aggregate variation is reinforced, since there will several different thresholds, generating a strong hysteresis effect. At every level of the input

[18] I assume that an individual faced with indifference will always abstain from crime.

variable, positive variations result in different effects on the output variable when compared to effects from negative variations.

Figure 2.3 illustrates the effect of a temporary reduction in the probability of punishment π on crime rates in a system with strong hysteresis. Initially there is an increase in the crime rate from c_0 to c_1 as a consequence of a fall in the probability of punishment from π_0 to π_1. However, as it will be shown in the next sections, if the probability of punishment is restored to its previous level from π_1 to π_0, the crime rate falls, but it does not return to its original level, but to a higher level $c_2 > c_0$.

The difference between c_2 and c_0 is a measure of the remanence in crime rates associated to all additional individuals that are not deterred from crime at $\pi = \pi_0$ *after* the temporary fall in π, even though the same individuals were not criminals at $\pi = \pi_0$ *before* the reduction in π.

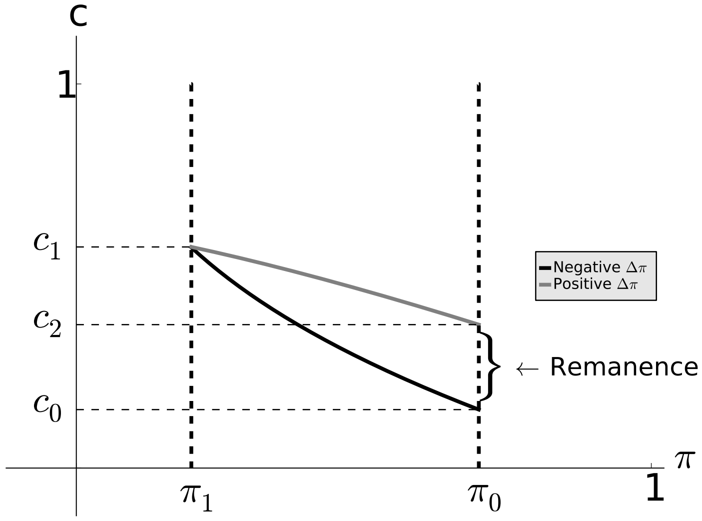


Figure 2.3: Strong Hysteresis - Aggregate level

Decrease in the probability of punishment π followed by an increase to its original

In the model that I develop in the next sections, I will focus on a less abstract concept where the decision to commit a crime is determined by individuals comparing the probability of punishment and relevant economic variables to their corresponding thresholds.

2.4 Baseline Framework

2.4.1 Homogenous Agents

There is a $[0, 1]$ continuum of ex ante homogenous[19] risk-neutral[20] myopic infinitely-lived individuals that decide whether to commit a crime or not. For sake of simplicity, I also assume that breaking the law is a victimless act. The wage of individuals abiding by the law is drawn from an exogenous distribution $F(w)$. Committing a crime can increase utility by the illicit (pecuniary and/or nonpecuniary) gain g, but it also entails a moral (or intrinsic) cost m and an exogenous[21] probability π [22] of facing a punishment s (sanction). The trade-off between the time spent in legal and criminal activities is captured by a reduction in w by $\gamma \in [0, 1]$.[23] The introduction of the parameter γ relaxes the assumptions present in all previous models in the economics of crime that either the law breakers have to entirely forfeit a legal income or not at all. As it will be seen, that parameter plays a key role in the crime decision.

Every criminal act entails an intrinsic cost m which is borne irrespective of whether the punishment occurs or not. This cost can have at least two non-mutually exclusive interpretations: 1. A moral cost associated with an illicit activity, either a disutility because the individuals feel guilt about the harm caused to the victim by the act (internal) or the disapproval of their peers (social). 2. The actual cost of a criminal act (entry fee or criminal technology/inputs) which is decreasing with the number of crimes committed in the past (individual learning) and/or with the number/strength of links in a criminal network (social learning).

The expected utility of crime (EUC) is then given by:

$$\begin{aligned} EUC &= \underbrace{(g + \gamma w - m - s)}_{\text{caught}}\pi + \underbrace{(g + \gamma w - m)}_{\text{get away with it}}(1 - \pi) \\ &= g + \gamma w - m - s\pi \end{aligned} \tag{2.4}$$

[19]The assumption of homogeneity will relaxed afterwards and is assumed to show that the results are not driven by ex ante heterogeneity.

[20]This assumption keeps the tractability of the model, not in detriment of the qualitative results.

[21]As it will be discussed in chapter 3, that is a reasonable assumption in the short run.

[22]The probability of punishment (fine, prison, etc) is conditional on conviction, which is itself conditional on being caught/arrested. Here it is assumed that being caught implies conviction and punishment.

[23]$(1 - \gamma)$ can be interpreted as the depreciation of the ability to earn w when and individual is engaged in crime. This reduction in w reflects stigma in the labour market and/or less time available to work in legal activities.

At every period t, agents choose to engage in criminal activity $\phi = 1$ or not $\phi = 0$. An agent will choose to engage in criminal activity if the expected gains of committing in crime outweigh the income without crime.

This decision is determined by:

$$Max_{\phi}\{w, EUC\} \tag{2.5}$$

that boils down to:

$$\phi = \begin{cases} 1 & if \quad g - (1-\gamma)w - m - s\pi > 0 \\ 0 & \text{otherwise} \end{cases} \tag{2.6}$$

Notice from equation 2.6 that if there is no conflict between legal and criminal activities ($\gamma = 1$) the crime decision will not depend on w. Conversely, if $\gamma = 0$, crime and no crime are two mutually exclusive activities and consequently different states. That implies that γ is closer to 0 for more serious crimes and closer to 1 for petty crimes.

It should also be noted that in the extreme case where $\pi = 0$ and there is no opportunity cost ($\gamma = 1$), crime only happens if the gain associated to the act is strictly higher than the intrinsic cost: $g > m$.

Definition 2.1. For a given combination (g, m, w, s), there is a unique π, the *deterrent threshold of punishment*, that will deter all individuals from crime:

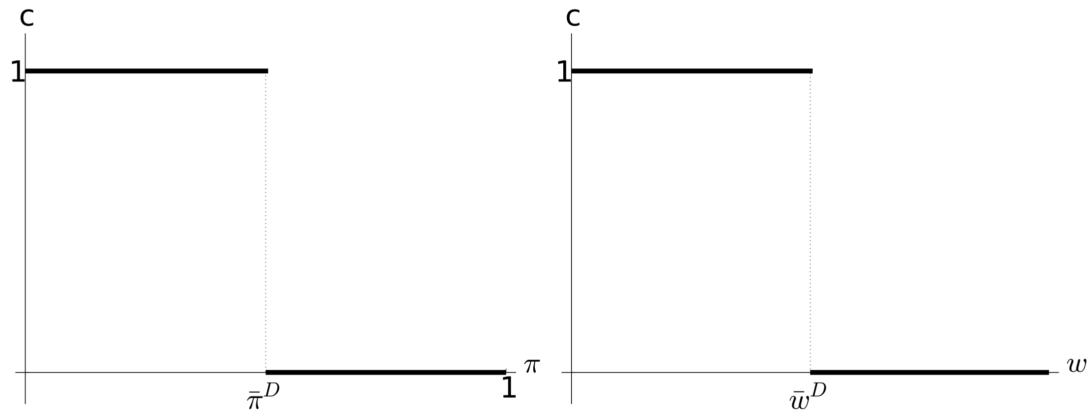
$$\bar{\pi}^D = \frac{g - (1-\gamma)w - m}{s} \tag{2.7}$$

Definition 2.2. For a given combination (g, m, π, s), there is a unique w, the *reservation wage of crime*, that will discourage all individuals from crime.

$$\bar{w}^D = \frac{g - m - s\pi}{1-\gamma} \tag{2.8}$$

Under homogeneity of the agents and assuming that each criminal commits only one type of crime per period, [24] the crime rates can be depicted for a given $\bar{\pi}^D$ and $\bar{w}^D$, respectively.

[24] This a reasonable assumption for a sufficiently short period. In the case of a longer period, the crime rate is obtained by also taking into account the distribution of the number of crimes each criminal commits per period. A simple formulation is to assume that the number of crimes of each criminal is given by a uniform probability mass function. If an additional assumption that all individuals commit the same number of crimes $q > 1$ is made, the crime rate can be obtained simply by multiplying the fraction of offenders by q. Note that in that case the crime rate can be greater than one.

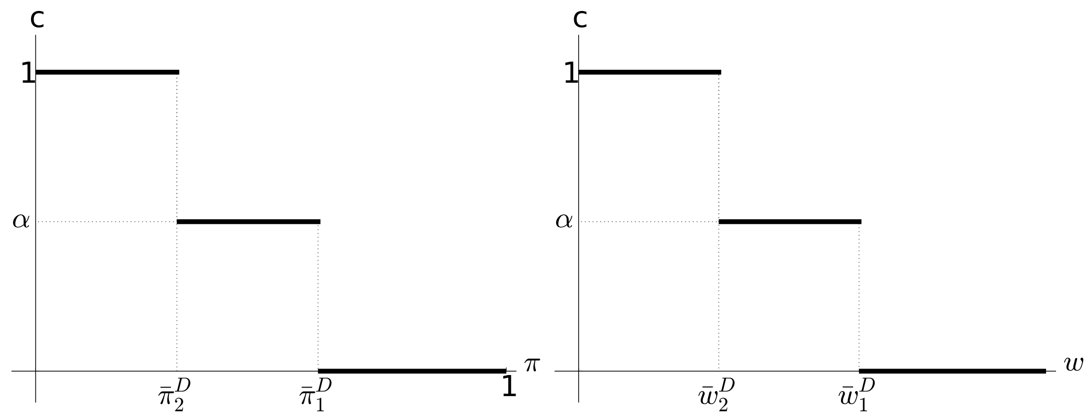


(a) Probability of Punishment (b) Legal Wage

Figure 2.4: Crime Rates for Homogenous Agents

2.4.2 Heterogeneous Agents

Heterogeneity allows for a more realistic description of the relationship between crime and its deterrents as it allows crime rates to be a value different from 0 and 1.[25]



(a) Probability of Punishment (b) Legal Wage

Figure 2.5: Crime Rates for Heterogeneous Agents ($K = 2$)

Let the individuals to vary with respect to their moral cost m_j, $j = \{1, ..., K\}$.

[25]What follows is based on exogenous ex ante heterogeneity. An analysis with ex ante homogeneity, but ex post heterogeneity will lead to similar results. It can be carried out if I assume that individuals instead of observing the probability of punishment, they only observe a noisy signal correlated to the true probability of punishment and update their priors according to previous experience. Sah (1991), Calvó-Armengol and Zenou (2004) and Calvó-Armengol, Verdier, and Zenou (2007) provided the first models were agents have an imperfect observation of the probability of punishment. Empirical evidence in this direction is given by Rincke and Traxler (2011).

(K types). For K=2, m is either m_L or m_H with probability $(\alpha, 1-\alpha)$. For a given combination (g, w, m_j, s), there are two distinct deterrent π's: $\pi_2^D = \frac{g-m_H-(1-\gamma)w}{s} < \pi_1^D = \frac{g-m_L-(1-\gamma)w}{s}$. Similarly, for a given combination (g, m_j, π, s), there are two distinct reservation wages of crime: $\bar{w}_2^D = \frac{g-m_H-s\pi}{1-\gamma} < \bar{w}_1^D = \frac{g-m_L-s\pi}{1-\gamma}$.

For a continuum of types, m_j can be represented by a distribution $f_m(m)$. The simplest non-trivial case is when the fraction of each type is equally likely and m follows a uniform distribution: $m \sim U(m_{min}, m_{max})$. For a given positive (negative) variation in π there is a negative (positive) variation in the crime rate correspondent to the fraction of the population with the level of m.[26]

The values of m_{min} and m_{max} can be normalised so that $\bar{\pi}^D \sim U(0,1)$: $\pi_{min}^D = 0 = \frac{g-(1-\gamma)w-m_{max}}{s} \Rightarrow m_{max} = g-(1-\gamma)w$ and $\pi_{max}^D = 1 = \frac{g-(1-\gamma)w-m_{min}}{s} \Rightarrow m_{min} = g-(1-\gamma)w-s$.

That implies that in terms of π, m has pdf given by:[27]

$$f_m(m) = \frac{1}{s} \quad , \quad m_{min} \le m \le m_{min} + s. \tag{2.9}$$

Similarly for w, for a given positive (negative) variation in w, there is a negative (positive) variation in the crime rate correspondent to the fraction of the population with the level of m. Therefore, in terms of w, m has pdf:[28]

$$f_m(m) = \frac{1}{1-\gamma} \quad , \quad m_{min} \le m \le m_{min} + 1 - \gamma. \tag{2.10}$$

Because $\bar{\pi}^D$ has pdf given by: $f_{\bar{\pi}^D}(\bar{\pi}^D) = 1, \quad 0 \le \bar{\pi}^D \le 1$, for a given π, crime rate can be computed by two equivalent ways:

$$c(\pi) = \int_{m_{min}}^{m(\pi)} \frac{1}{s}\,\mathrm{d}m = \int_{\pi}^{1} \mathrm{d}\bar{\pi}^D = 1 - \pi \tag{2.11}$$

The first integral in equation 2.11 corresponds to the aggregation of all individuals with moral cost below the moral cost associated to the prevailing probability of punishment π. That is equivalent to the deterrence probability of punishment $\bar{\pi}^D$ being greater than π, given by the second integral of the equation.

Similarly for w, $\bar{w}^D$ has pdf given by: $f_{\bar{w}^D}(\bar{w}^D) = \frac{1}{\bar{w}_{max}}, \quad 0 \le \bar{w}^D \le \bar{w}_{max}$, for a given w, crime rate is given by:

[26] Notice that if the highest level of m is sufficiently high, $c < 1$ even if $\pi = 0$.

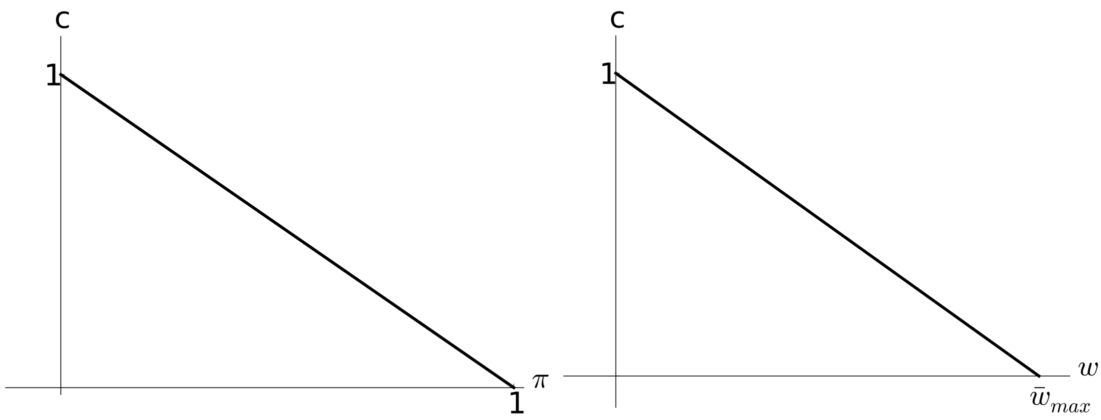
[27] The same result is obtained by applying the theorem of inverse transformation of random variables to the affine transformation $m = g - (1-\gamma)w - s\bar{\pi}^D$.

[28] Transformation: $m = g - s\pi - (1-\gamma)\bar{w}^D$

$$c(w) = \int_{m_{min}}^{m(w)} \frac{1}{1-\gamma} \, \mathrm{d}m = \int_{w}^{\bar{w}_{max}} \frac{1}{\bar{w}_{max}} \, \mathrm{d}\bar{w}^D = 1 - \frac{w}{\bar{w}_{max}} \tag{2.12}$$

Equation 2.12 corresponds to the aggregation of all individuals with moral cost below the moral cost associated to the prevailing legal wage w, or equivalently, the aggregation of all individuals with reservation wage of crime $\bar{w}^D$ greater than w.

The crime rates in terms of probability of punishment π and legal wage w are depicted in figure 2.6.



(a) Probability of Punishment (b) Legal Wage

Figure 2.6: Crime Rates for a Continuum of Heterogeneous Agents - Uniform

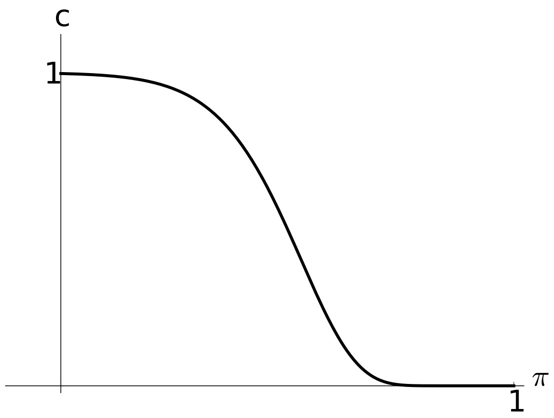
A more realistic representation of the distribution of the intrinsic cost would relax the assumption of equally likely values of m. For simplicity, assume $m_{min} = 0$, so that $f_m(m)$ has a support on the $[0, m_{max}]$ interval. Crime rate in terms of π can then be computed analogously to the uniform case using the distribution of m:

$$\begin{aligned} c(\pi) &= \int_0^{m(\pi)} f_m(m) \, \mathrm{d}m \\ &= F_m(m(\pi)) \\ &= F_m(g - (1-\gamma)w - s\pi) \end{aligned} \tag{2.13}$$

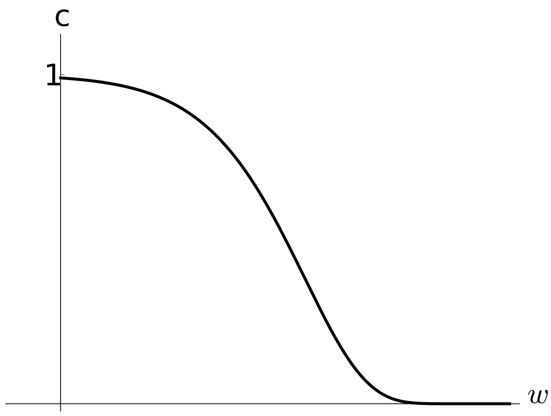
As the cdf F_m in non-decreasing, equation 2.13 in a non-linear decreasing function on π. This is also true to w, as it can be analogously shown. Figures 2.7a and 2.7b depict these relationships for a log-normal distribution.[29]

Notice that for extreme high values of π and w, crime rates are lower than the case that the intrinsic cost has a uniform distribution. Similarly, extreme low values of π and w are associated to higher crime rates than the uniform case.

[29] Which has support $[0, \infty)$. For a sufficiently small σ^2, the distribution is relatively symmetric.



(a) Probability of Punishment



(b) Legal Wage

Figure 2.7: Crime Rates for a Continuum of Heterogeneous Agents - Log-normal

2.5 Homogeneous Agents and Weak Hysteresis

Consider the previous framework and define criminal history at time t as the sum of criminal choices in the T past periods:

$$h^t = \sum_{\tau=0}^{T} \phi_{t-\tau} \tag{2.14}$$

All variables associated with the crime decision could potentially be affected by criminal history. The intrinsic cost m and the opportunity cost of crime given by $(1-\gamma)w$ are unambiguously decreasing in h^t, as was discussed in previous sections. The gain from crime g is increasing in h^t, as crime experience and networking allows a better targeting at loot with higher values. The severity of punishment s is also increasing in h^t, as most countries impose heftier sanctions for individuals with a criminal past. The effect of h^t on the probability of punishment π depends on whether the learning process dominates the higher number of traces left by the criminal acts.

To simplify the analysis, assume that the variables above only depend whether an individual has ever committed a crime or not. Define the binary variable $h = 1[h^{t-1} \geq 1]$, where $1[\cdot]$ is an indicator function. This assumption will subsequently be relaxed.

As all variables above, apart from s and π, increase the propensity of committing a crime given $h = 1$, the positive effect of crime history on crime choice (including any reduction in π through learning) can be encapsulated by m. The negative effect of h on crime choice (including any increase in π given the higher number of traces) is subsumed in s. A simple linear specification[30] is then given by:

$$m_t = \bar{m} - \tilde{m}h \tag{2.15}$$

$$s_t = \bar{s} + \tilde{s}\varpi h \tag{2.16}$$

where $\varpi \in [0, 1]$ captures the fact that only a fraction of the criminal past is known by the judicial system.

Using the equations above, equations 2.4 and 2.6 can be rewritten respectively as:

$$EUC = g + \gamma w - \bar{m} - \bar{s}\pi + (\tilde{m} - \varpi\tilde{s}\pi)h \tag{2.17}$$

and

$$\phi = \begin{cases} 1 & if \quad g - (1-\gamma)w - \bar{m} - \bar{s}\pi + (\tilde{m} - \varpi\tilde{s}\pi)h > 0 \\ 0 & \text{otherwise} \end{cases} \tag{2.18}$$

Solving equation 2.18 for π yields the deterrent threshold of punishment $\bar{\pi}^D$:

$$\bar{\pi}^D = \frac{g - (1-\gamma)w - \bar{m} + \tilde{m}h}{\bar{s} + \varpi\tilde{s}h} \tag{2.19}$$

And the reservation wage of crime $\bar{w}^D$:

$$\bar{w}^D = \frac{g - \bar{m} - \bar{s}\pi + (\tilde{m} - \varpi\tilde{s}\pi)h}{1-\gamma} \tag{2.20}$$

Then equation 2.18 can be rewritten as:

$$\phi_t(\pi) = \begin{cases} 1 & if \quad \pi < \bar{\pi}^D \\ 0 & \text{otherwise} \end{cases} \tag{2.21}$$

or

$$\phi_t(w) = \begin{cases} 1 & if \quad w < \bar{w}^D \\ 0 & \text{otherwise} \end{cases} \tag{2.22}$$

Therefore, the deterrent threshold of punishment and the crime reservation wages are functions of the criminal history.

It is clear from equation 2.20 that $\bar{w}^D$ is increasing in criminal history h if and only if

$$\tilde{m} > \varpi\tilde{s}\pi. \tag{2.23}$$

[30]A exponential specification ($\bar{m}e^{-\tilde{m}h}$ and $\bar{s}e^{\tilde{s}\varpi h}$) will lead to similar results.

And as shown in the appendix, $\bar{\pi}^D$ is increasing in h if

$$\tilde{m} > \varpi \tilde{s}, \tag{2.24}$$

in addition to the assumption that the part of sanction that does not depend on the known criminal past is proportional to the net gain of crime ($\bar{s} = \bar{g}$).[31]

Because $0 \leq \pi \leq 1$, the latter condition implies the former. Condition 2.24 is a reasonable assumption as only the convicted past crimes should increase the severity of punishment, whereas all past crimes should affect the intrinsic cost of crime ($\varpi < 1$). In what follows, it is assumed that condition 2.24 holds, and without loss of generality under this assumption, I consider the simplest case where this is true by setting $\tilde{s} = 0$.

Additionally, the focus of the analysis in the remaining of this section will be on the probability of punishment π, as the results for wages are analogous. The following proposition establishes weak hysteresis in the criminal decision:

Proposition 2.1. ***Weak Hysteresis*** *If individuals are homogeneous and maximize expected utility, their intrinsic cost of crime m evolves according to equation (2.15), with criminal history $h = \{0, 1\}$ and defining $\bar{g} \equiv g - \bar{m} - (1-\gamma)w$, for any $\tilde{m} > 0$ and any probability of punishment $\pi \in (0,1)$, the criminal decision is given by*

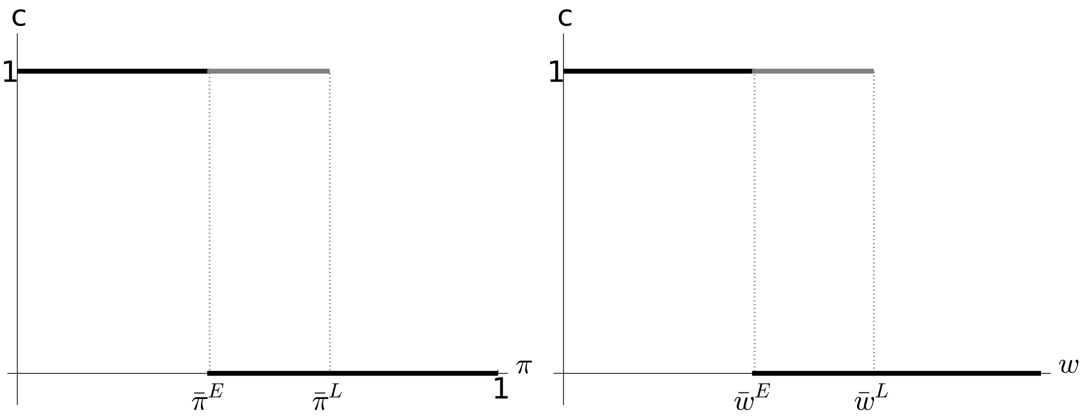
$$\phi_t(\pi) = \begin{cases} 1 & if \quad \pi < \frac{\bar{g}}{s} \\ 0 & if \quad \pi \geq \frac{\bar{g}}{s} + \frac{\tilde{m}}{s} \\ \phi_{t-1}(\pi) & if \quad \frac{\bar{g}}{s} \leq \pi < \frac{\bar{g}}{s} + \frac{\tilde{m}}{s} \end{cases} \tag{2.25}$$

with:

$$\phi_0(\pi) = \begin{cases} 1 & if \quad \pi < \frac{\bar{g}}{s} \\ 0 & if \quad \pi \geq \frac{\bar{g}}{s}. \end{cases} \tag{2.26}$$

Proof. Note that regardless of whether or not individuals have committed crime in the past, they will always commit a crime in period t if $\pi < \frac{\bar{g}}{s}$ (*entry threshold*) and they will never commit a crime if $\pi \geq \frac{\bar{g}}{s} + \frac{\tilde{m}}{s}$ (*exit threshold*). However, if π is between $\frac{\bar{g}}{s}$ and $\frac{\bar{g}}{s} + \frac{\tilde{m}}{s}$, the decision is conditional on the crime choice in the previous period. As $h = \{0, 1\}$, the range of decision described in 2.21 can be partitioned into 3 regions, and for any π, the crime decision is given by equations 2.25 and 2.26. ■

[31]This is a consequence of the proportionality principle, observed by criminal justice system in most countries, by which the harshness of the penalty for each offense is proportional to the damage the perpetrator causes and the gain from the illicit act.



(a) Probability of Punishment (b) Legal Wage

Figure 2.8: Crime Rates for Homogenous Agents - Weak Hysteresis

Equations 2.25 and 2.26 together represent a relay function, where the choice in period t is state-dependent.

At this point it is clear that there are two main aspects that determine the degree of hysteresis in crime decision. One is the degree of memory in the criminal history and the other is the size of $\tilde{m}$ relative to the severity of punishment. In the extreme cases where $\tilde{m}$ is very high or all criminal history matters, there is a ratchet effect where criminals are locked-in.

It is useful to have a new notation for the different levels of deterrent thresholds of punishment $\bar{\pi}^D$. Denote the entry threshold by $\bar{\pi}^E$ and the threshold that individuals leave the criminal career by $\bar{\pi}^L$. Therefore $\bar{\pi}^E = \frac{\bar{g}}{s}$ and $\bar{\pi}^L = \frac{\bar{g}}{s} + \frac{\tilde{m}}{s}h^t$.

If individuals are homogeneous the aggregation of individual crime decision into crime rates is trivial. The relationship between π and crime rates is depicted in figure 2.8.

It is clear that hysteresis is a phenomenon that occurs at the individual level only if specific threshold levels are reached. Hysteresis does not happen if individuals face probability of punishment and legal wages far from the critical thresholds. That is one of the main characteristics of weak hysteresis.

2.5.1 Weak Hysteresis versus Unit Root Process

If criminal history in the last T periods matter, the exit deterrent threshold of punishment will then be written as:

$$\bar{\pi}_t^L = \frac{\bar{g}}{s} + \frac{\tilde{m}}{s} \sum_{\tau=0}^{T} \phi_{t-\tau} \tag{2.27}$$

and the range of the relay function will have $T + 1$ partitions.

The deterrent threshold of punishment can return to the original level $\bar{\pi}^E = \frac{\bar{g}}{s}$ if individuals do not commit crimes for T periods.

It is important to contrast the history dependence that stems from this relay function with history dependence that emerges from a unit root in a discrete process.

If all criminal history is relevant, the deterrent threshold of punishment is non-decreasing in the criminal history:

$$\bar{\pi}_t^L = \frac{\bar{g}}{s} + \frac{\tilde{m}}{s} \sum_{\tau=0}^{\infty} \phi_{t-\tau} \tag{2.28}$$

Note that equation 2.28 is the steady state level of $\bar{\pi}_t^L$ following an $AR(1)$ process with unit root given by:

$$\bar{\pi}_t^L = \psi \bar{\pi}_{t-1}^L + \frac{\tilde{m}}{s} \phi_t \tag{2.29}$$

with $\psi = 1$ and $\bar{\pi}_0^L = \bar{\pi}^E = \frac{\bar{g}}{s}$.[32]

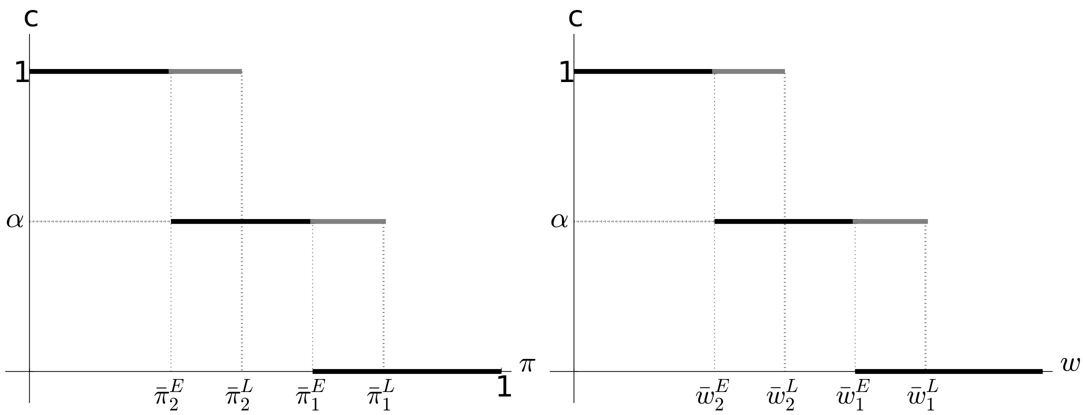
As a unit root process has a impulse-response function constant and equal to one, any temporary shock has a permanent effect. A shock in this setting is extremely simple and corresponds to $\frac{\tilde{m}}{s}\phi_t$ switching from 0 to $\frac{\tilde{m}}{s}$ or vice versa.

Unlike this unit root discrete process, where current behaviour depends weakly on all past shocks, in a hysteric process, current behaviour strongly depends only on non-dominated extrema of past shocks.

[32] Another way to see this is to notice that equation 2.28 is a $MA(\infty)$, which is equivalent to an $AR(1)$.

2.6 Heterogeneous Agents and Strong Hysteresis

For 2 types of agents, the different exit and entry thresholds occur in both levels of crime rates. The relationship between π and c for $k = 2$ is shown in figure 2.9.



(a) Probability of Punishment

(b) Legal Wage

Figure 2.9: Crime Rates for Heterogeneous Agents ($K = 2$) - Weak Hysteresis

For a continuum of heterogeneous agents, I proceed as in the previous section, where the intrinsic cost of crime follow a uniform distribution, but with the bounds taking into account the reduced intrinsic costs for individuals with criminal history: $\bar{m} \sim U(\bar{m}_{min}, \bar{m}_{max})$, with $\bar{m}_{max} = g - (1-\gamma)w + \tilde{m}$ and $\bar{m}_{min} = g - (1-\gamma)w - s$.

That implies that $\bar{\pi}^D \sim U(0, 1 + \frac{\tilde{m}}{s})$, with density:

$$f_{\bar{\pi}^D}(\bar{\pi}^D) = \frac{1}{1 + \frac{\tilde{m}h}{s}} = \frac{s}{s + \tilde{m}h} \quad , \quad 0 < \bar{\pi}^D < 1 + \frac{\tilde{m}h}{s} \tag{2.30}$$

Aggregating in terms of $\bar{\pi}^D$ yields the crime rate as a function of π:

$$c(\pi) = \int_{\pi}^{\frac{s+\tilde{m}h}{s}} \frac{s}{s + \tilde{m}h} \,\mathrm{d}\bar{\pi}^D = 1 - \frac{s}{s + \tilde{m}h}\pi \tag{2.31}$$

Aggregating in terms of $\bar{w}^D$ yields crime rate:

$$c(w) = \int_{w}^{\frac{s\bar{w}_{max}+\tilde{m}h}{s}} \frac{s}{s\bar{w}_{max} + \tilde{m}h} \,\mathrm{d}\bar{w}^D = 1 - \frac{s}{s\bar{w}_{max} + \tilde{m}h} w \tag{2.32}$$

If there was crime in the past, present crime rate is less sensitive to the impact of policies affecting π and w when compared to the situation where there is no crime in the past.

As seen above, by relaxing the assumption of a equally likely distribution of intrinsic cost of crime m, there is a non-linear relationship between crime, π and w.

I formally define some concepts related to hysteresis in crime rates in terms of the probability of punishment π and summarize the results in two propositions.

Definition 2.3. Let Ω be the *criminal remanence*, the increase in the crime rate when π returns to its original level π_0.

Definition 2.4. Let π^C be the *coercive probability of punishment*, the level of π necessary to return the crime rate to its original level.

Definition 2.5. Let $\xi = \pi^C - \pi_0$ be the *coercive force* necessary to return the crime rate to its original level.

Definition 2.6. Let $H(\pi^t)$ be the *maximum historical crime rate* in the last T periods:

$$H(\pi^t) = \max\{c(\pi_{t-T}), ..., c(\pi_{t-1})\}.$$

Consider the situation set choice of the policy maker deciding the level of the probability of punishment is simply: $\{\pi^H, \pi^L\}$, with $\pi^H > \pi^L$.

Proposition 2.2. ***Strong Hysteresis*** *If individuals are heterogeneous with respect to their initial intrinsic cost of crime $\bar{m}$ and the intrinsic cost of crime evolves according to equation 2.15, for any two levels of probability of punishment $\pi^L < \pi^H \in [0,1]$, an exogenous reduction in π from π^H to π^L ($\Delta^-\pi$) will increase crime by Δ^+c. If it is followed by an exogenous increase in π from π^L to π^H ($\Delta^+\pi = -\Delta^-\pi$), it will decrease crime by $\Delta^-c = -\Delta^+c + \Omega$, where*

$$\Omega = F_{\bar{m}}(g - (1-\gamma)w - \tilde{s}\pi + \tilde{m}\ H(\pi^L)) - F_{\bar{m}}(g - (1-\gamma)w - \tilde{s}\pi + \tilde{m}\ H(\pi^H)).$$ [33]

Proof. Proof in the appendix. ■

Corollary 2.1. *For any given combination of the parameters of the model, as the the probability of the punishment $\pi \to 0$ or $\pi \to 1$, both the criminal remanence $\Omega \to 0$ and the coercive force $\xi \to 0$.*

Proof. Direct result from the fact that $F_{\bar{m}}$ is a CDF. ■

[33] Proof of this theorem arose from joint work with Vasileios Vlaseros.

Corollary 2.1 states that the remanence effect Ω is close to 0 when π are close to their lower and upper bounds.

Proposition 2.3. *If the level of severity s is sufficiently high, both the criminal remanence $\Omega \to 0$ and the coercive force $\xi \to 0$.*

Proof. This is a direct result from equations 2.60 and 2.61, and the fact that $F_{\bar{m}}$ is a CDF. ■

The result established in proposition 2.3 is very intuitive since there is a trade-off between severity and probability of punishment. However, increasing the latter involves the allocation of resources to costly monitoring and undesirable risk of punishing an innocent person (type II error). That would imply setting the probability of punishment as low as possible and the severity of punishment as high as possible. However, the harshness of the penalty for each offense cannot be increased beyond the level that the society considers a fair punishment (proportionality principle) or in case of a pecuniary punishment beyond the ability of the lawbreaker to pay the fine. Besides, the process of punishment (apprehension, prosecution and actual punishment) entails significant costs as well. For that reason, a crime minimising policy must balance these two variables.

Example

As an example and without loss of generality,[34] assume that m follows a log-normal distribution and plug its CDF into equation 2.13 yielding:

$$c(\pi) = \frac{1}{2} + \frac{1}{2}\,\mathbf{erf}\,[\frac{ln(g-(1-\gamma)w-s\pi)-\mu}{\sqrt{2\sigma^2}}] \tag{2.33}$$

where $\mathbf{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}\,\mathrm{d}t$.

From the equation 2.15 describing the effect of criminal history on intrinsic cost m, it is clear that the mean and the variance of m can be written respectively as:

$$E(m) = E(\bar{m}) - \tilde{m}E(h) \tag{2.34}$$

$$var(m) = var(\bar{m}) + \tilde{m}^2 var(h) \tag{2.35}$$

It is clear that a higher proportion of individuals with criminal record decreases the mean and increases the variance of m. That is translated in a shift of $f_m(m)$ to the left.

[34]Similar results with several other distributions.

That implies that the curve given by equation 2.33 is displaced to the right, as show figures 2.10a and 2.10b.

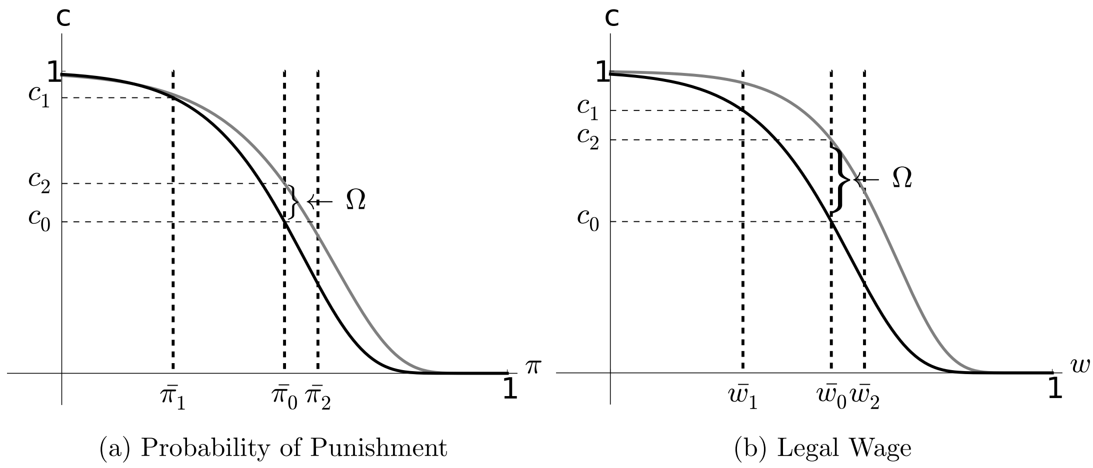


Figure 2.10: Crime Rates for different levels of π and w

The thick lines illustrate the full variation of π or w between 0 and 1 in both directions. If π is exogenously reduced from π_0 to π_1, crime rate goes up from c_0 to c_1. If π is restored to its original value π_0, crime rate falls to $c_2 = c_0 + \Omega$. Only if π is increased to π_2 will make the crime rate to return to its original level c_0. A similar explanation applies to variations in w. For this smaller range of variation in π or w, there is a smaller loop inside the one plotted in both figures.

It is also clear from the figure 2.10 that ex ante heterogeneity leads to hysteresis at every point of the input variable, instead of trigger thresholds in the case of weak hysteresis when agents are homogeneous.

2.7 Forward-Looking Agents with Search

In the previous sections, it was implicitly assumed that at every period, individuals deciding whether to commit a crime or not, would have a crime opportunity and a legal wage with probability one. It was also assumed that individuals are myopic with respect with the costs and benefits of committing a crime.

Using the concepts developed in the benchmark above, I present in this section a more general view, where individuals are forward-looking and search for economic activities, taking into account the net expected present value of income from crime and job opportunities. This setting allows a better characterization of the dynamic aspect of hysteresis in criminal decision with respect to the opportunity cost of the legal market for the types of crimes where the individuals are more likely to take into account the loss of the stream of legal income. The other crucial difference is the use of prison as the associated punishment. This section builds on the models presented in Burdett, Lagos, and Wright (2003, 2004) and Engelhardt (2010) by exploring the consequences of criminal learning on crime choice and crime rates.

Consider now a $[0, N]$ continuum of infinitely lived homogeneous firms and as in the previous sections, a $[0, 1]$ continuum of infinitely lived risk-neutral individuals. Homogeneity is assumed for simplicity and to show that results do not stem from ex ante heterogeneity. An individual can be in three states: employed, unemployed or in jail, respectively e, u, and j. Firms post an i.i.d. wage w from the distribution of legal wages $F(w)$. In each state, there is an associated value function for an individual employed, unemployed or in jail given respectively by $V_e(w)$, V_u and V_j.

Unemployed individuals consume the benefit b and receive offers from $F(w)$ at rate λ_u. Employed workers consume w and also receive offers from $F(w)$ at rate λ and are dismissed at rate δ.

Crime is defined here as an illicit activity that entails in a probability $\pi > 0$ of an individual being incarcerated. Individuals encounter state-dependent opportunities to commit a crime at a rate μ_u if unemployed and rate μ_e if employed, choosing to undertake it with conditional probability ϕ_u and ϕ_e respectively. Individuals that undertake the crime opportunity get a loot g and face a probability π of being caught and going to jail.

Jailed individuals consume z and are released at an exogenous rate ρ. These two parameters capture the severity of the punishment. A crucial assumption here is that, if the criminal is not caught in the period the crime is perpetrated, the probability of punishment in the next periods will be null.[35] Individuals fall victim to crime at rate ξ. In order to close the model, ξ must be related to the endogenous crime decisions. I assume $\xi = 0$, which can be interpreted as the crimes considered are victimless.

As in the previous section, criminal history at time t is given by the sum of criminal choices in the past described in equation 2.14 and the binary variable $h = 1[h^t \geq 1]$ indicates if an individual has committed a crime in the last T periods. As before, the intrinsic cost of crime is a function of criminal history described by equation 2.15

2.7.1 Crime and Legal Job Decisions

The expected gain from crime for an unemployed individual is given by:

$$H_u = g - \bar{m} + \tilde{m}h + \pi V_j + (1 - \pi)V_u \tag{2.36}$$

And for an employed individual is given by:

$$H_e(w) = g - \bar{m} + \tilde{m}h + \pi V_j + (1 - \pi)V_e(w). \tag{2.37}$$

An unemployed and employed individual will commit a crime given an opportunity if $H_u \geqslant V_u$ and $H_e(w) \geqslant V_e(w)$, respectively. That implies the criminal decision for unemployed and employed individuals to be described respectively by:

$$\phi_u = \begin{cases} 1 & \text{if } V_u - V_j \leq (g - \bar{m} + \tilde{m}h)/\pi \\ 0 & \text{if } V_u - V_j > (g - \bar{m} + \tilde{m}h)/\pi \end{cases} \tag{2.38}$$

and

$$\phi_e(w) = \begin{cases} 1 & \text{if } V_e(w) - V_j \leq (g - \bar{m} + \tilde{m}h)/\pi \\ 0 & \text{if } V_e(w) - V_j > (g - \bar{m} + \tilde{m}h)/\pi. \end{cases} \tag{2.39}$$

Defining $\beta = \frac{1}{1+r}$ as the discount factor, so that r is the time preference rate, and assuming that only one event occur at a time, the Bellman's equation for an unemployed individual is given by:

[35] This is a ubiquitous assumption in the literature. However, as discussed in section 3, it is very unlikely that this assumption holds for serious crimes which reinforces the hysteresis effect. As the criminal knows that even though he was not caught so far, he can go to jail at any moment. For this reason, he may fall in the sunk cost fallacy and consider that a second (or subsequent) criminal act will not affect the probability of being caught or the punishment significantly.

$$rV_u = b + \mu_u \phi_u [H_u - V_u] + \lambda_u \int max\{V_e(x) - V_u, 0\} dF(x). \tag{2.40}$$

Equation 2.40 can be interpreted as the instantaneous return to being unemployed, which equals b plus the expected value of receiving either a crime or job opportunity. The Bellman's equation for an employed individual is given analogously by:

$$\begin{aligned} rV_e(w) = w + \delta[V_u - V_e(w)] + \mu_e \phi_e(w)[H_e(w) - V_e(w)] \\ + \lambda_e \int max\{V_e(x) - V_e(w), 0\} dF(x). \end{aligned} \tag{2.41}$$

Individuals accept a job offer if the expected present value of the returns from the new job exceeds the value of being unemployed or the expected present value of the returns from the current job and the net value of a crime opportunity.

Similarly, for a criminal in jail:

$$rV_j = z + \rho[V_u - V_j]. \tag{2.42}$$

Several results are unaltered in comparison with Burdett, Lagos, and Wright (2003, 2004). The first one is that there is a unique reservation wage $\bar{w}^R$, since $V_e(w)$ in equation 2.41 is strictly increasing in w.[36] Moreover, note from equation 2.37 that $H_e(w) - V_e(w)$ is decreasing in w. That implies workers are less likely to commit a crime when their wages are higher.

It can also be shown that $H_u - V_u = H_e(\bar{w}^R) - V_e(\bar{w}^R)$. For that reason, unemployed individuals engage in crime if and only if workers employed at $\bar{w}^R$ do. Therefore, as in the aforementioned articles, $\forall w$, $\phi_u = 0$ implies $\phi_e = 0$ and if $\phi_u = 1$ then $\phi_e = 1$ for $w < \bar{w}_D$ and $\phi_e = 0$ for $w \geq \bar{w}^D$, where $\bar{w}^D$ is the reservation wage for committing a crime, defined by $H_e(\bar{w}^D) = V_e(\bar{w}^D)$.

The following proposition establishes the reservation wage of crime as a function of the parameters of the model. This is a similar result from Burdett, Lagos, and Wright (2003, 2004) when the criminal history is taken into account.

[36]Derivation in the appendix.

Proposition 2.4. *If the ex ante criminal choice of individuals is governed by equations (2.38) and (2.39), and they maximise utility overtime according Bellman's equations (2.40) through (2.42), the reservation wage of crime is given by*

$$\bar{w}^D = z + (r+\delta)\frac{g - \bar{m} + \tilde{m}h}{\pi} + (\rho - \delta)\Gamma(\bar{w}^R) - \lambda_e \Psi(\bar{w}^D) \tag{2.43}$$

where

$$\Gamma(\bar{w}^R) = \frac{b - z + (g - \bar{m} + \tilde{m}h)\mu_u + \lambda_u \Psi(\bar{w}^R)}{r + \rho + \mu_u \pi} \tag{2.44}$$

and

$$\Psi(\bar{w}^R) = \int_{\bar{w}^R}^{\bar{w}^D} \frac{[1 - F(x)]dx}{r + \delta + \mu_e \pi + \lambda_e[1 - F(x)]} + \int_{\bar{w}^D}^{\infty} \frac{[1 - F(x)]dx}{r + \delta + \lambda_e[1 - F(x)]} \tag{2.45}$$

Proof. Proof in the appendix. ∎

Note that equation 2.43 is an implicit function of the reservation wage of crime and several parameters of the model. Therefore, the effect of h on $\bar{w}^D$ is obtained by implicity differentiating equation 2.43. Proposition 2.5 establishes the relationship between the reservation wage of crime and the criminal history. To simplify the analysis below, I assume $\lambda_e = 0$ and $\lambda_u = \lambda$, i.e., there is no on-the-job search. The results are qualitatively unaltered if this assumption is relaxed.

Proposition 2.5. *If the ex ante criminal choice of individuals is governed by equations (2.38) and (2.39), and they maximise utility overtime according Bellman's equations (2.40) through (2.42), $\forall\, \pi > 0,\ 0 \leq \bar{w}^D < w_{max},\ \tilde{m} > 0$, the reservation wage of crime $\bar{w}^D$ is a increasing function of the criminal history h.*

Proof. Proof in the appendix. ∎

As the criminal choice ultimately depends on legal wage being below or above the reservation wage of crime, individuals with criminal history otherwise identical to individuals without criminal past, require a strictly higher legal wage to opt out of crime than his/her counterparts that have never broken the law.

It also clear from proposition 2.5 that the impact of criminal history on the reservation wage of crime is a function of several parameters of the model. Crucially, it is decreasing in the probability of punishment and the reservation wage of crime, as well as it is increasing in the time preference rate, firing rate, job offer rate and the sensitivity of history with respect the intrinsic criminal cost.

2.7.2 Steady State Equilibria and Transition Rates

This section analyses the effect on crime rates as a consequence of the state-dependence at the individual level discussed in the previous section. It also departs from the model described in Burdett, Lagos, and Wright (2003, 2004).

In order to solve for the steady state, let U denote the number of workers unemployed, assume $\phi_u = 0$ and consider the exit rate from unemployment $\psi = 1 - F(\bar{w}^R)$ and equate the flow of workers into employment $\lambda_u[1 - F(\bar{w}^R)]U$ to the flow from unemployment to employment $\delta(1 - U)$, and therefore

$$U = \frac{\delta}{\delta + \lambda_u[1 - F(\bar{w}^R)]}, \tag{2.46}$$

as in the standard model search model without crime.

Assuming $w \geq \bar{w}^R$ with probability 1, we have $U = \frac{\delta}{\delta+\lambda_u} = \frac{1}{1+\kappa_u}$, where $\kappa_u = \lambda_u/\delta$, which characterise the degree of search friction from unemployment to employment.

If $\phi_u = 1$, E_L can be defined as the number of workers employed at $w < \bar{w}^D$, E_H can be defined as the number of workers employed at $w \geq \bar{w}^D$ and $\sigma = 1 - F(\bar{w}^D)$ is the fraction of firms offering at least $\bar{w}^D$. In that case, the flows in this model can be depicted as in figure 2.11, where J denote the number of individuals in jail and dotted arrows depict involuntary transitions.[37]

[37] As in Burdett, Lagos, and Wright (2003, 2004).

$$
\begin{tikzcd}
E_H \arrow[d, dotted, bend right, "\delta"'] & E_L \arrow[l, "\lambda_e\sigma"'] \arrow[d, "\mu_e\pi"] \arrow[dl, dotted, bend right, "\delta"'] \\
U \arrow[u, "\lambda_u\sigma"] \arrow[ur, "\lambda_u(1-\sigma)"] \arrow[r, "\mu_u\pi"] & J \arrow[l, dotted, bend left, "\rho"]
\end{tikzcd}
$$

Figure 2.11: Flows between states

The unemployment rate is then $U^* = \frac{U}{1-J}$ and, in any equilibrium with $\phi_u = 1$, the crime rate is:

$$C^* = \frac{U\mu_u + E_L\mu_e}{1-J}, \tag{2.47}$$

since the number of individuals in the economy is normalised to 1 and individuals in jail should not be considered in the rates.

It is also clear that because $\bar{w}^D$ is increasing in h, as shown in proposition 2.5, $\sigma = 1 - F(\bar{w}^D)$ is decreasing in h. Therefore, individuals with a criminal history will have lower transition intensities from U to E_H and E_L to E_H, as well as higher transition intensities from U to E_H. That implies that a society with a larger fraction of individuals with a criminal past has a higher crime rate than a society with a lower fraction of individuals with a criminal past.

Note that in this model, hysteresis occurs even if ex-convicts are not stigmatised in the labour market and have the same probability of employment as individuals who have never been incarcerated. It is clear that if such assumption is relaxed, it will reinforce the hysteresis effect.

Another implicit assumption, closely related to the previous one, is the fact that ex-inmates have the same intrinsic/moral cost of crime (m) as the other offenders that have not been in prison. Because individuals that have been in prison increase their interactions with other criminals, it is likely that they have a lower m when they are released from prison, making than more prone to re-offend. However, if the prison system is sufficiently effective and the reduction in m would be negligible and the punishment would make individuals less likely to commit a crime again. The impact of that feature in the hysteresis effect would depend on the quality of the legal system.

2.8 Conclusions

This chapter explores a theoretical model to explicitly account for hysteresis in both criminal behaviour and crime rates in order to fill the gap between the theoretical predictions and the empirical evidence about the efficiency of policies in reducing crime rates.

When the probability of crime deterrence decreases or there is a fall in the real income obtained in the labour market, some previously law-abiding agents will start committing crimes. Because, ceteris paribus, individuals that have a criminal past are more prone to take crime opportunities than someone that has never committed one, if the original conditions are subsequently restored, a subset of these agents will continue in their career in crime. At the individual level, if criminal activity is associated with intrinsic sunk costs and learning, then the cost of leaving a criminal career is higher than entering it. At the aggregate level with homogeneous agents, this is translated into a hysteresis effect that will only occur if a specific threshold is surpassed. With heterogeneous agents in terms of their intrinsic cost of crime, this phenomenon is reinforced generating a hysteresis effect that exists for all possible values of the variable affecting the crime decision. Only when punishment is extremely severe will the effect of hysteresis cease to exist.

The main consequence of hysteresis is that variations in the determinants of crime will have asymmetric effects, depending on the sign of the changes. That result provides a more suited theoretical explanation to the reason the empirical literature have found lower than expected effects from variations in the factor that affect crime. Nevertheless, it must be emphasised that the results hinge on some simplifying assumptions.

The existence of hysteresis in crime has at least two direct implications. The first one regards the policies to reduce crime rates. If hysteresis has a relevant effect in criminal decision, policies to reduce crime should be focused on crime prevention rather than mitigation. The second one is relevant to any future empirical analyses of the impact of policies to reduce crimes. The asymmetric nature of positive and negative shocks on crime rates must be taken into account in order to obtain a more precise estimation of the policy effect.

References

AKERLOF, G. A. (1991): "Procrastination and Obedience," *The American Economic Review*, 81(2), pp. 1–19.

ALLINGHAM, M. G., AND A. SANDMO (1972): "Income tax evasion: a theoretical analysis," *Journal of Public Economics*, 1(3-4), 323–338.

AMABLE, B., J. HENRY, F. LORDON, AND R. TOPOL (1994): "Strong hysteresis versus zero-root dynamics," *Economics Letters*, 44(1), 43–47.

BAR-GILL, O., AND A. HAREL (2001): "Crime Rates and Expected Sanctions: The Economics of Deterrence Revisited," *The Journal of Legal Studies*, 30(2), 485–501.

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

BIERENS, H. J., AND J. R. CARVALHO (2007): "Semi-nonparametric competing risks analysis of recidivism," *Journal of Applied Econometrics*, 22(5), 971–993.

——— (2011): "Job Search, Conditional Treatment and Recidivism: The Employment Services for Ex-Offenders Program Reconsidered," *The B.E. Journal of Economic Analysis & Policy*, 11(1).

BJERK, D. (2010): "Thieves, thugs, and neighborhood poverty," *Journal of Urban Economics*, 68(3), 231–246.

BLANCHARD, O., AND L. SUMMERS (1987): "Hysteresis in unemployment," *European Economic Review*, 31(1-2), 288–295.

BOWLES, R. A., AND C. FLORACKIS (2007): "Duration of the time to reconviction: Evidence from UK prisoner discharge data," *Journal of Criminal Justice*, 35(4), 365–378.

BUONANNO, P., D. MONTOLIO, AND P. VANIN (2009): "Does Social Capital Reduce Crime?," *Journal of Law and Economics*, 52.

BURDETT, K., R. LAGOS, AND R. WRIGHT (2003): "Crime, Inequality, and Unemployment," *The American Economic Review*, 93(5), 1764–1777.

——— (2004): "An On-The-Job Search Model of Crime, Inequality, and Unemployment," *International Economic Review*, 45(3), 681–706.

CALVÓ-ARMENGOL, A., T. VERDIER, AND Y. ZENOU (2007): "Strong and weak ties in employment and crime," *Journal of Public Economics*, 91(1), 203–233.

CALVÓ-ARMENGOL, A., AND Y. ZENOU (2004): "Social Networks and Crime Decisions: The Role of Social Structure in Facilitating Delinquent Behavior," *International Economic Review*, 45(3), 939–958.

CARVALHO, J. R. (2009): *Microeconometrics of Labor Markets and Criminal Behavior: Recidivism, Duration Analysis, and Program Evaluation*. VDM Publishing.

CROSS, R. (1993): "On the foundations of hysteresis in economic systems," *Economics and Philosophy*, 9(01), 53–74.

DILLS, A. K., J. A. MIRON, AND G. SUMMERS (2008): "What Do Economists Know About Crime?," Working Paper 13759, NBER.

DIXIT, A. (1989): "Entry and exit decisions under uncertainty," *Journal of political Economy*, pp. 620–638.

——— (1992): "Investment and hysteresis," *The Journal of Economic Perspectives*, pp. 107–132.

EHRLICH, I. (1973): "Participation in Illegitimate Activities: A Theoretical and Empirical Investigation," *Journal of Political Economy*, 81(3), 521–65.

ENGELHARDT, B. (2010): "The Effect of Employment Frictions on Crime," *Journal of Labor Economics*, 28(3), 677–718.

ENGELHARDT, B., G. ROCHETEAU, AND P. RUPERT (2008): "Crime and the labor market: A search model with optimal contracts," *Journal of Public Economics*, 92(10-11), 1876 – 1891.

ESCARELA, G., B. FRANCIS, AND K. SOOTHILL (2000): "Competing Risks, Persistence, and Desistance in Analyzing Recidivism," *Journal of Quantitative Criminology*, 16, 385–414.

FENDER, J. (1999): "A general equilibrium model of crime and punishment," *Journal of Economic Behavior & Organization*, 39(4), 437–453.

FERRER, R. (2010): "Breaking the law when others do: A model of law enforcement with neighborhood externalities," *European Economic Review*, 54(2), 163–180.

FREEMAN, R. B. (1991): "Crime and the Employment of Disadvantaged Youths," NBER Working Papers 3875, National Bureau of Economic Research, Inc.

——— (1999): "The economics of crime," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 3 of *Handbook of Labor Economics*, chap. 52, pp. 3529–3571. Elsevier.

——— (2003): "Can We Close the Revolving Door?: Recidivism vs. Employment of Ex-Offenders in the U.S.," Urban institute reentry round table, New York University Law School.

GAROUPA, N. (1997): "The theory of optimal law enforcement," *Journal of Economic Surveys*, 11(3), 267–295.

GÖCKE, M. (2002): "Various concepts of hysteresis applied in economics," *Journal of Economic Surveys*, 16(2), 167–188.

HEINEKE, J., AND M. BLOCK (1975): "A labor theoretic analysis of criminal choice," *The American Economic Review*, 65(3), 314–325.

IMAI, S., AND K. KRISHNA (2004): "Employment, Deterrence, And Crime in a Dynamic Model," *International Economic Review*, 45(3), 845–872.

IMROHOROGLU, A., A. MERLO, AND P. RUPERT (2004): "What Accounts For The Decline In Crime?," *International Economic Review*, 45(3), 707–729.

KIM, J.-Y. (2012): "A note on the non-maximality of the optimal fines when the apprehension probability depends on the offense rate," *European Journal of Law and Economics*, Published online:June.

LOCHNER, L. (2004): "Education, Work, and Crime: A Human Capital Approach," *International Economic Review*, 45(3), 811–843.

MCCRARY, J. (2010): "Dynamic Perspectives on Crime," in *Handbook on the Economics of Crime*, chap. 4, pp. 82–108. MPG Books Group, UK.

MOCAN, H. N., AND T. G. BALI (2010): "Asymmetric Crime Cycles," *The Review of Economics and Statistics*, 92(4), 899–911.

MOCAN, H. N., S. C. BILLUPS, AND J. OVERLAND (2005): "A Dynamic Model of Differential Human Capital and Criminal Activity," *Economica*, 72(288), 655–681.

MUSTARD, D. B. (2010): "How Do Labor Markets Affect Crime? New Evidence on an Old Puzzle," IZA Discussion Papers 4856, Institute for the Study of Labor (IZA).

POLINSKY, A. M., AND S. SHAVELL (2007): "The Theory of Public Enforcement of Law," in *Handbook of Law and Economics, Volume 1*, chap. 6, pp. 403–454. Elsevier B.V.

RINCKE, J., AND C. TRAXLER (2011): "Enforcement spillovers," *Review of Economics and Statistics*, 93(4), 1224–1234.

RØED, K. (2002): "Hysteresis in unemployment," *Journal of Economic Surveys*, 11(4), 389–418.

ROGERSON, R., R. SHIMER, AND R. WRIGHT (2005): "Search-Theoretic Models of the Labor Market: A Survey," *Journal of Economic Literature*, 43(4), 959–988.

SAH, R. K. (1991): "Social Osmosis and Patterns of Crime," *Journal of Political Economy*, 99(6), 1272–95.

SANTOS, M. R. D. (2009): "Labor Supply, Criminal Behavior and Income Redistribution," *Brazilian Review of Econometrics*, 29(2).

SICKLES, R. C., AND J. WILLIAMS (2008): "Turning from crime: A dynamic perspective," *Journal of Econometrics*, 145, 158173.

STIGLER, G. J. (1970): "The optimum enforcement of laws," *Journal of Political Economy*, 78, 526–536.

2.A Appendix

Derivation of condition 2.24

Define $\bar{g} \equiv g - \bar{m} - (1-\gamma)w$ so that equation 2.19 can be rewritten as:

$$\bar{\pi}^D = \frac{\bar{g} + \tilde{m}h}{\bar{s} + \varpi\tilde{s}h} \tag{2.48}$$

Then differentiate equation 2.48 with respect to h:

$$\begin{aligned}\frac{\partial\bar{\pi}^D}{\partial h} &= \frac{\tilde{m}(\bar{s} + \varpi\tilde{s}h) - \varpi\tilde{s}(\bar{g} + \tilde{m}h)}{(\bar{s} + \varpi\tilde{s}h)^2} \\ &= \frac{\tilde{m}\bar{s} + \tilde{m}\varpi\tilde{s}h - \varpi\tilde{s}\bar{g} - \varpi\tilde{s}\tilde{m}h}{(\bar{s} + \varpi\tilde{s}h)^2} \\ &= \frac{\tilde{m}\bar{s} - \varpi\tilde{s}\bar{g}}{(\bar{s} + \varpi\tilde{s}h)^2}\end{aligned} \tag{2.49}$$

Note that $\frac{\partial\bar{\pi}^D}{\partial h} > 0$ if and only if :

$$\tilde{m}\bar{s} > \varpi\tilde{s}\bar{g}, \tag{2.50}$$

as all parameters above are greater or equal to zero, and $\bar{g}$, the net gain of crime independent of punishment is also strictly positive (otherwise the criminal act to would not be rational). As the part of sanction that does not depend on the known criminal past ($\bar{s}$) is proportional to the net gain of crime, the assumption that $\bar{s} = \bar{g}$ implies that the derivative in 2.49 can be rewritten as:

$$\frac{\partial\bar{\pi}^D}{\partial h} = \frac{(\tilde{m} - \varpi\tilde{s})\bar{g}}{(\bar{s} + \varpi\tilde{s}h)^2} \tag{2.51}$$

And therefore $\frac{\partial\bar{\pi}^D}{\partial h} > 0$ if

$$\tilde{m} > \varpi\tilde{s}. \tag{2.52}$$

Derivation of equation 2.63

This is a variation of the derivation in Burdett, Lagos, and Wright (2003, 2004), where the criminal history is taken into account. The reservation wage, as very well established in the search literature, is obtained by solving $V(w) = V_u$ for w.[38] If equation 2.41 is set equal to equation 2.40 and solved for w:

$$\bar{w}^R = b + (\lambda_u - \lambda_e)\Psi(\bar{w}^R) + (\mu_u - \mu_e)\phi_u[g - \bar{m} + \tilde{m}h - \pi(V_u - V_J)] \tag{2.53}$$

where

$$\Psi(\bar{w}^R) = \int_{\bar{w}^R}^{\infty}[V_e(x) - V_e(\bar{w}^R)]dF(x) \tag{2.54}$$

If the integral in the last equation is computed by parts:

$$\Psi(\bar{w}^R) = \int_{\bar{w}^R}^{\infty} V_e'(x)[1 - F(x)]dx \tag{2.55}$$

After differentiating the Bellman's equation for employed workers (eq. 2.41) and plugging into the previous expression:

$$\Psi(\bar{w}^R) = \int_{\bar{w}^R}^{\bar{w}^D} \frac{[1 - F(x)]dx}{r + \delta + \mu_e\pi + \lambda_e[1 - F(x)]} + \int_{\bar{w}^D}^{\infty} \frac{[1 - F(x)]dx}{r + \delta + \lambda_e[1 - F(x)]} \tag{2.56}$$

Subtract 2.42 and 2.40 to obtain:

$$V_u - V_j = \frac{b - z + (g - \bar{m} + \tilde{m}h)\mu_u + \lambda_u\Psi(\bar{w}^R)}{r + \rho + \mu_u\pi} \tag{2.57}$$

Plugging the last two equations into equation 2.53 yields the reservation wage.

[38]See Rogerson, Shimer, and Wright (2005).

2.A.1 Proofs

Proof of proposition 2.2

For a given distribution of initial intrinsic cost of crime, $\bar{m} \sim F(\bar{m})$, crime rates are computed by:

$$c(\pi) = \int_{\bar{m}_{min}}^{m(\pi)} \mathrm{d}F_{\bar{m}} = F_{\bar{m}}(\bar{m}(\pi)) - F_{\bar{m}}(\bar{m}_{min}) \tag{2.58}$$

Using equation 2.19 and definition 2.6 in the previous equation yields:

$$c(\pi) = F_{\bar{m}}(g - (1-\gamma)w - s\pi + \tilde{m}H(\pi)) \tag{2.59}$$

Since $F_{\bar{m}}$ is an non-decreasing function and $H(\pi^L) > H(\pi^H)$, we have that, for any x,

$$\Omega = F_{\bar{m}}(x + \tilde{m}H(\pi^L)) - F_{\bar{m}}(x + \tilde{m}H(\pi^H)) \geq 0 \tag{2.60}$$

and

$$\Omega = F_{\bar{m}}(g-(1-\gamma)w-\tilde{s}\pi+\tilde{m}H(\pi^L))-F_{\bar{m}}(g-(1-\gamma)w-\tilde{s}\pi+\tilde{m}H(\pi^H)) \quad \blacksquare \tag{2.61}$$

Proof of proposition 2.4

When individuals are indifferent between committing crime, from equation 2.39, this is equivalent to:

$$V_e(\bar{w}^D) = \frac{g - \bar{m} + \tilde{m}h}{\pi} + V_j. \tag{2.62}$$

Multiply the previous equation by r and insert Bellman's equations 2.42 2.41 at $w = \bar{w}^D$ to get:

$$\begin{aligned} \bar{w}^D = z + (r+\delta)\frac{g - \bar{m} + \tilde{m}h}{\pi} + (\rho - \delta)(V_u - V_j) \\ -\lambda_e \int_{\bar{w}^D}^{\infty} \frac{[1-F(x)]dx}{r + \delta + \lambda_e[1-F(x)]}. \end{aligned} \tag{2.63}$$

From the derivation of the reservation wage in the appendix we have that:

$$V_u - V_j = \frac{b - z + (g - \bar{m} + \tilde{m}h)\mu_u + \lambda_u \Psi(\bar{w}^R)}{r + \rho + \mu_u \pi} \tag{2.64}$$

where

$$\Psi(\bar{w}^R) = \int_{\bar{w}^R}^{\bar{w}^D} \frac{[1 - F(x)]dx}{r + \delta + \mu_e \pi + \lambda_e [1 - F(x)]} + \int_{\bar{w}^D}^{\infty} \frac{[1 - F(x)]dx}{r + \delta + \lambda_e [1 - F(x)]} \tag{2.65}$$

Plugging the previous two equations into equation 2.63, in addition to the Bellman's equation for employed workers (eq. 2.41) and using the fact that when individuals are indifferent between committing a crime, $g - \bar{m} + \tilde{m}h = \pi[V_e(\bar{w}^D) - V_j]$ yields:

$$\bar{w}^D = z + (r + \delta)\frac{g - \bar{m} + \tilde{m}h}{\pi} + (\rho - \delta)\Gamma(\bar{w}^R) - \lambda_e \Psi(\bar{w}^D) \tag{2.66}$$

where equation 2.64 and equation 2.65 are substituted in the last equation. ■

Proof of proposition 2.5

Rewrite equation 2.43 in order to define:

$$P(\bar{w}^D, h, \cdot) = -\bar{w}^D + z + (r + \delta)\frac{g - \bar{m} + \tilde{m}h}{\pi} + (\rho - \delta)\Gamma(\bar{w}^R) - \lambda \Psi(\bar{w}^D) \tag{2.67}$$

where

$$\Gamma(\bar{w}^R) = \frac{b - z + (g - \bar{m} + \tilde{m}h)\mu_u + \lambda \Psi(\bar{w}^R)}{r + \rho + \mu_u \pi} \tag{2.68}$$

and

$$\Psi(\bar{w}^R) = \int_{\bar{w}^R}^{\bar{w}^D} \frac{[1 - F(x)]dx}{r + \delta + \mu_e \pi + \lambda [1 - F(x)]} + \int_{\bar{w}^D}^{\infty} \frac{[1 - F(x)]dx}{r + \delta + \lambda [1 - F(x)]} \tag{2.69}$$

Therefore $\frac{\partial \bar{w}^D}{\partial h}$ is given by:

$$\frac{\partial \bar{w}^D}{\partial h} = -\frac{\partial P(\bar{w}^D, h, \cdot)/\partial h}{\partial P(\bar{w}^D, h, \cdot)/\partial \bar{w}^D} \tag{2.70}$$

Differentiate $P(\bar{w}^D, h, \cdot)$ with respect to both arguments to obtain:

$$\frac{\partial P(\bar{w}^D, h, \cdot)}{\partial h} = \frac{(r+\delta)\tilde{m}}{\pi} + \frac{(\rho-\delta)\tilde{m}\mu_u}{r+\rho+\mu_u\pi} \tag{2.71}$$

$$\frac{\partial P(\bar{w}^D, h, \cdot)}{\partial \bar{w}^D} = -1 + \frac{\frac{[1-F(\bar{w}^D)][r+\delta+\mu_u\pi]\lambda\pi}{\lambda[1-F(\bar{w}^D)]+r+\delta} - \frac{[1-F(\bar{w}^D)][\delta-\rho]\lambda\pi}{\lambda[1-F(\bar{w}^D)]+r+\delta+\mu_u\pi}}{[r+\rho+\mu_u\pi]\pi}, \tag{2.72}$$

with the Leibniz integral rule and the fact that $F(\infty) = 1$ being used to compute the last partial derivative.

The effect of h on $\bar{w}^D$ is obtained by plugging 2.71 on 2.72 into equation 2.70 and simplifying:

$$\begin{aligned}
\frac{\partial \bar{w}^D}{\partial h} &= -\frac{(r+\delta)\tilde{m}(r+\rho+\mu_u\pi) + (\rho-\delta)\tilde{m}\mu_u\pi}{\frac{[1-F(\bar{w}^D)][r+\delta+\mu_u\pi]\lambda\pi}{\lambda[1-F(\bar{w}^D)]+r+\delta} - \frac{[1-F(\bar{w}^D)][\delta-\rho]\lambda\pi}{\lambda[1-F(\bar{w}^D)]+r+\delta+\mu_u\pi} - (r+\rho+\mu_u\pi)\pi} \\
&= \frac{(r+\delta)\tilde{m}(r+\rho+\mu_u\pi) + (\rho-\delta)\tilde{m}\mu_u\pi}{\frac{[1-F(\bar{w}^D)][\delta-\rho]\lambda\pi}{\lambda[1-F(\bar{w}^D)]+r+\delta+\mu_u\pi} - \frac{[1-F(\bar{w}^D)][r+\delta+\mu_u\pi]\lambda\pi}{\lambda[1-F(\bar{w}^D)]+r+\delta} + (r+\rho+\mu_u\pi)\pi} \qquad (2.73) \\
&= \frac{(r+\delta)(r+\rho+\mu_u\pi)[1-F(\bar{w}^D)+r+\delta][1-F(\bar{w}^D)+r+\delta+\mu_u\pi]}{(r+\delta+\mu_u\pi)\left\{\lambda(r+\rho)[1-F(\bar{w}^D)+r+\delta]+(r+\delta)\mu_u\pi\right\}}\frac{\tilde{m}}{\pi}
\end{aligned}$$

Thus, $\frac{\partial \bar{w}^D}{\partial h} > 0, \ \forall\, \pi > 0, \ 0 \leq \bar{w}^D < w_{max}, \ \tilde{m} > 0.$ ■

Chapter 3

Asymmetric Effects and Hysteresis in Crime Rates: Evidence from the United States

3.1 Introduction

This chapter empirically examines the predictions in the previous chapter that the processes governing criminal behaviour are inherently permeated by hysteresis and, consequently, the absolute value of the magnitude of the impact of variables such as the probability of punishment and income on crime rates will depend on whether the variations of those determinant variables are positive or negative.

It is shown that if criminal activity is associated with intrinsic sunk costs and learning, then the cost of leaving a criminal career is higher than entering it and there is hysteresis at the individual level, where the crime decision in a given period is determined not only by the current expected costs and benefits entailed by the illicit act, but it is also affected by criminal decisions taken in the past.

If there is hysteresis at the individual level, the aggregation of all crime decisions of the agents will render hysteresis in crime rates. For any sufficiently large society, the smallest reduction in the probability of punishment will make some individuals at the margin between committing a crime or not to choose the illegal option. A subsequent increase in the probability of punishment, will make some of the new criminals to stop committing crimes, but some individuals will prefer to continue in the criminal career, even if the probability of punishment is exactly the same as the original one.

Therefore, hysteresis in crime rates emerges from the aggregation of individual crime decisions that display the hysteresis effect and can be understood as a path-dependent process, where the current level of crime depends not only on the current levels of variables like the number of police officers and income, but also whether their levels in the previous periods were below or above the current levels. If there is hysteresis in crime, policies aiming to reduce crime rates will have a diminished impact when compared to the impact where individuals with a criminal past will behave similarly to individuals without a criminal past.

Chapter 2 concluded that there is hysteresis in crime rates for at least two of the factors that affect them: probability of punishment and legal income, where the latter captures the opportunity cost of crime.

One important consequence of hysteresis is that the effect on an outcome variable from positive exogenous variations in the determining variables has a different magnitude from negative variations. That asymmetric effect is clearer in a situation in which the crime reduction policy in a given period is simply a reversal of a deterioration of one determinant of crime. A concrete example would be a situation where part of the police officers in a city are dismissed in a given year, resulting in a escalation in crime. If all sacked police officers are readmitted in the following year in an attempt to restore the original crime levels and hysteresis is present in the criminal behavior, that policy will result in a lower crime rate, but higher than the original one. [1]

That prediction is empirically investigated by analysing US crime data at the state level between 1977 and 2010, using police force size as a proxy for the probability of punishment[2] and real average income of unskilled workers as the opportunity cost of crime.[3] It is, to my knowledge, the first paper to consider the existence of asymmetric effects of variations in the probability of punishment and in the opportunity cost of crime.

I find that two of the main determinants of crime - police force size and real average income of unskilled workers - have significant asymmetric effects on most types of crimes. The average impact of positive variations in the income of unskilled workers

[1]Note that this a simplifying example to convey the concept of hysteresis in crime rates. It is very unlikely that a negative variation will be followed by a positive value with the exact same size in absolute terms. However, if crime decision is indeed permeated by hysteresis, crime rates will also display the hysteresis property, even in the case of the absolute value of increases/decreases are different, provided that increase are not much larger than the decreases in absolute terms.

[2]From the Uniform Crime Reports - UCR Crime data compiled by the Federal Bureau of Investigation - FBI and Annual Survey of State and Local Government Employment and Census of Governments.

[3]From the Current Population Survey - CPS

on property crime rates are statistically smaller than the absolute value of the average effect of negative variations. That asymmetry is also observed for the law enforcement variable in both property and violent crime.

The results are robust under several models and specifications. As will be discussed, no theoretical or empirical analysis would be complete without careful consideration of that important feature of the relationships between crime, police and legal income.

In the following section, I present the empirical literature of the economics of crime focused on the impact of the probability of punishment and legal wages on crime. The data used in the present analysis is discussed in section 3.3, while the empirical framework is presented in section 3.4. This is followed by the results in section 3.5 and a nonparametric analysis in section 3.6. Section 3.7 concludes.

3.2 Related Literature

3.2.1 Hysteresis and Asymmetric Effects in Crime

The economics of crime literature is fairly extensive as economists play a prominent role in the pursuit to establish causality of socioeconomic variables on aggregate crime measures. Polinsky and Shavell (2007) and Dills, Miron, and Summers (2008) provide a relatively recent overview of contributions provided by economists since Becker's seminal paper. However, only very recently have the study of persistence and asymmetric cycles been the primary focus of a study in the empirical crime literature. The first authors (and to date, the only ones) to explicitly test for the existence of asymmetric cycles in crimes rates were Mocan and Bali (2010). The authors compare the effects on crime rates of positive and negative variations on unemployment rates and obtain statistical evidence for the existence of asymmetries in that relationship.

Hysteresis and asymmetric effects have also been the objects of empirical studies in other areas in economics, especially in unemployment.[4] Because variables like unemployment and inflation are closely connected with GDP cycles, those studies frequently refer to this analysis as asymmetric cycles.

3.2.2 Effect of the Probability of Punishment on Crime

The relationship between police and crime is one of the classical examples of the pernicious effects of simultaneity on the interpretation of correlations. Even though a

[4] For a recent survey on hysteresis in unemployment, see O'Shaughnessy (2011).

larger police force is expected to increase the probability of punishment and consequently reduce crime rates, the empirically observed relationship could be largely dominated by the positive response of the policy maker to the higher levels of crime in the previous periods.[5]

Ehrlich (1973) provided the first empirical analysis in the economics of crime in which the author was also the first to empirically estimate the relationship between the probability of punishment (proxied by the per capita expenditure on police) and crime. However, only almost a quarter of a century later, was the relevant issue of simultaneity between police and crime rates the main focus of a research paper.

The first paper to use a panel data set in order to mitigate the crucial problem of unobserved heterogeneity to estimate the impact of deterrence variables on crime was Cornwell and Trumbull (1994), that made use of an offense ratio of face-to-face crimes to non-face-to-face crimes and tax revenues per capita. Levitt (1997) uses electoral cycles in the police expenditures to identify the impact of police on crimes. However, McCrary (2002) points out some coding errors in Levitt (1997) and showed that Levitt's results were not statistically significant, and police would not have any effect on crime with the instrumental variables based on electoral cycles. Levitt (2002) provides a reply by using the number of firefighters as an instrument for police force and finds significant effects of police on crime rates. Lin (2009) uses state tax rates as an instrumental variable for local police numbers. Evans and Owens (2007) and Worrall and Kovandzic (2010) instrument police levels with two types of federal law enforcement grants.

Corman and Mocan (2000) get around the use of instrumental variables by using monthly data and exploring the fact that the training of a police officer takes at least six months to identify the impact of police on crime rates in New York City. Unfortunately, that approach is not possible for a nationwide estimation in the US, since, although monthly crime data is available at many levels of aggregation since the first years of the Uniform Crime Reports - UCR, compiled by the Federal Bureau of Investigation - FBI, the number of police officers are aggregated by year.

[5]One alternative perspective is to examine the impact on crime of exogenous variations on the level of harshness of the penalties, rather than probability of punishment, since, as seen in chapter 2, there is a trade-off between those variables. However, due to the fact that it would involve the use of very restricted data and that there is little variation in the severity of punishment across areas and over time, it is empirically challenging to examine impacts of variations in the severity of punishment on crime. A rare empirical study on this issue is Lee and McCrary (2009), that explores the discontinuous increase in the severity of punishment for individuals at the age of 18. They find that a negative, but very small impact of harsher penalties on criminal behaviour. That result corroborates the hypothesis of hysteresis in crime from a different perspective.

Buonanno and Mastrobuoni (2012) circumvent the endogeneity issues by exploring a centralised and lengthy process to hire police officers in Italy.

Di Tella and Schargrodsky (2004) and Draca, Machin, and Witt (2011) are two prominent studies that exploit natural experiments (terrorist attacks, followed by sharp increase in the police force in the affect city) to detect the impact of police on crime rates.

Chalfin and McCrary (2013) show that all previous analyses suffer from severe and unexpected measurement error in the data on the size of the police force. This measurement error, rather than the simultaneity between police and crime, is the main source of upwards bias in the police-crime nexus. Supported by thorough anecdotal evidence, public administration and political economy studies, the authors make a very compelling case for the relative exogeneity of the law enforcement variable as the response from policy makers in the short term is extremely limited and is, to a great extent, idiosyncratic. The authors find a consistent estimator without measurement error by using data on the number of police from the Annual Survey of Government (ASG) as an instrument for the standard UCR data on police. Their estimates are five times larger in absolute value when compared to the estimates that do not correct for the measurement error.

3.2.3 Effect of Income on Crime

Economic theory predicts that labour income has a negative effect on crime as it encapsulates the opportunity cost of the criminal career.[6] Nevertheless, many empirical studies that analyse the impact of income on crime rates find ambiguous signs for the associated coefficients.[7] One of the reasons is related to the fact that the usual measure of income is the overall income in a location, such as GDP per capita, which captures the labour market expected benefits, but also the expected gain from criminal activity.

Another problem associated with the use of income measures of an area, rather then the individual in the empirical crime equation is the possibility that the associated coefficient is also capturing the higher propensity to report crime of richer areas, as pointed out by Soares (2004) and Soares and Naritomi (2010).

It is a stylised fact in the crime literature that most crimes are committed by young men. Gould, Weinberg, and Mustard (2002) explore that fact and analyse the impact

[6] One alternative perspective to capture the opportunity cost of crime, particularly for developing countries, is to examine poverty rates. Loureiro (2013) provides a recent example of that approach.

[7] For a survey on the studies that analyse labour market outcomes on crime, see Mustard (2010).

of wages of unskilled men on crime rates using a panel data set of US counties. The authors show that the long-run trend in crime rates can be better explained by the long-run trend in wages of unskilled men rather than by the trend in unemployment.[8]

It can be argued that level of income is an endogenous variable in the crime equation if areas with high crime rates receive less investment and have less job opportunities when compared to areas with lower crimes rates. That is less likely to happen at the state level. However, if that was still the case, it was more likely that this bias would be stronger in violent crimes than property crimes. However, the fact that none of the coefficients associated to violent crime are statistically significant provides little support for the hypothesis of simultaneity between income levels and crime in the present analysis.

Another possibility in this direction is that individuals with better job prospects in terms of wages, emigrate from areas with high levels of crime. Cullen and Levitt (1999) show that every additional reported crime is associated with one-person person decline in a average city with at least 100,000 inhabitants in the US. However, most migration occurred within the same Metropolitan Statistical Area (MSA). As the great majority of the MSAs in the US are within one state, it is less likely that such effect occur at the state level.

Doyle, Ahmed, and Horn (1999) and Gould, Weinberg, and Mustard (2002) show that OLS and GMM estimates do not vary significantly, suggesting that endogeneity of income is not a concern in the crime equation.

3.3 Data

I analyse data from the 50 US states between 1977 and 2010. Six crime categories are used. Three types of property crime: burglary, larceny and motor vehicle theft; and three types of violent crime: murder, aggravated assault and forcible rape. All crime variables are calculated in terms of rates per 100,000 inhabitants. These data are obtained from the Uniform Crime Reports - UCR Crime data compiled by the Federal Bureau of Investigation - FBI, and obtained from the Inter-university Consortium for Political and Social Research - ICPSR, which also provides data on police officers per 100,000 inhabitants.

A usual concern with respect to crime data is the existence of measurement error in crime rates, especially underreporting. This problem can be greatly mitigated by

[8]Freeman (1996) provides an inquiry about the reasons that make young unskilled men more prone to commit crimes.

the use of panel data techniques that control for unobserved heterogeneity. Any measurement error in the dependent variable not captured by this approach will only affect the standard errors of the coefficients, reducing the probability of rejecting the null of no relationship.

A more pernicious problem is related to measurement error of one of the explanatory variables.[9] The studies in the empirical literature of the economics of crime displayed no concern about the existence of measurement error in the explanatory variables in the crime equation. Chalfin and McCrary (2013) show the great deal of measurement error in the police force rates obtained from the UCR data. There are two sources of measurement error in the law enforcement variable generally used in the literature: 1. Misclassification in what constitutes a sworn police officer; 2. Underestimation of the population size used in the calculation in the police officers rate, especially in the years prior to the census.

I tackle one of the sources of measurement error by correcting the population used in both UCR and ASG data sets by nonparametric smoothing. This issue is especially crucial in the present study, as it explore the positive and negative variations in the police officers rate.

Because the theoretical sign of the coefficient associated to police is negative, the measurement error in the police variable will cause the estimated coefficient to be overestimated, increasing the probability of the type II error.

I also used the number of police officers from the Annual Survey of State and Local Government Employment and Census of Governments - hereafter ASG data - run by the US Census Bureau. The data for federal and state employees are the actual numbers, however, as the local government statistics are based on a sample of local governments, and the same reasons that can lead to measurement error in the UCR data are also present, the variable from this source is also noisy.

Because the police data from the UCR and ASG data in each year are a snapshot of October,[10] I follow the usual procedure in the literature to use the measures for the previous year in all regressions. In addition, because the years included in the analysis are restricted only by the CPS data (before 1977, some states are not in the sample), I can keep the 34 years by using police data in 1976.

I generate the average real income of young males without college and the other control variables by using microdata from the Current Population Survey - CPS,

[9] If it is the classical measurement error problem, the estimate of the coefficient will suffer from attenuation bias. See Fuller (1987), Hausman (2001) and Wooldridge (2002).

[10] After 1997 the month was March in the ASG data. I include a dummy variable to account for that fact in the regressions involving the two measures.

provided by the Bureau of Labor Statistics and which I retrieved from Integrated Public Use Microdata Series - IPUMS.[11] Details on the extraction and organisation of the data is outlined in the appendix 3.A.1.

3.3.1 Descriptive Statistics and Trends in Crime Rates, Police Force and Wages

Figure 3.1 shows the trends for the proportion of police officers by 100,000 inhabitants, real total income for unskilled workers and for the overall population in the United States between 1977 and 2010. There is a sustained increase in the three variables over the period, with exceptions given by a pronounced decrease in law enforcement in the beginning of the 80s and a significant decline in the total real income after the economic crisis in 2008. A closer look at the police variable reveals the short term cyclicality of this variable, first pointed out by Levitt (1997).[12]

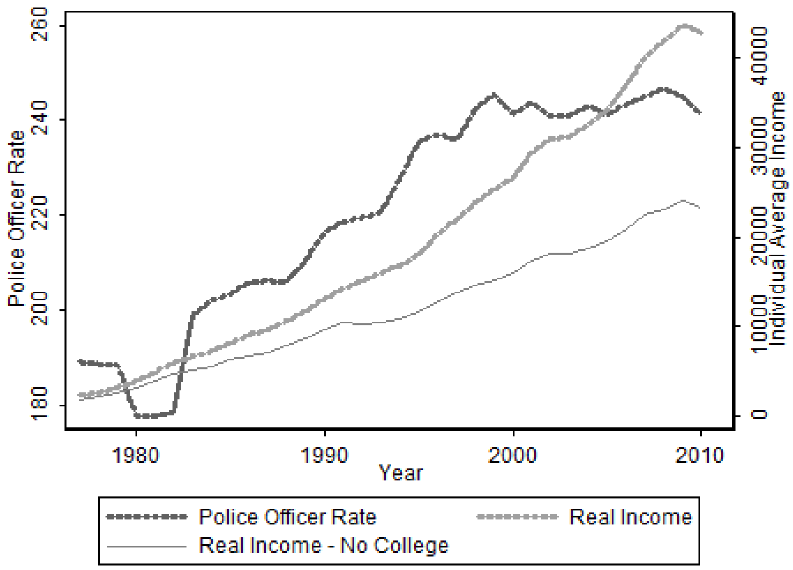


Figure 3.1: Police Rate and Overall Average Income and Unskilled Workers Income - USA - 1977-2010

Data Source: UCR/FBI and CPS.

[11] King, Ruggles, Alexander, Flood, Genadek, Schroeder, Trampe, and Vick (2012)

[12] Figure 3.1 suggests that the variables are non-stationary. Panel data tests that account for heterogeneity in the autoregressive coefficient of the Dickey-Fuller regression, cross-section dependence and level shifts reject the hypothesis of unit roots for police, income and crime rates. Additionally, most series for the individual states are trend-stationary (stationary after the trend is accounted for) and this result is even stronger when at least one structural break is allowed.

Figures 3.2a and 3.2b display crime rates and clearance rates[13] for respectively property and violent crimes in the US between 1970 and 2010. There is an unequivocal turn in the trend of both types of crimes that does not coadunate with clearance rates over the period.

Table 3.1 provides the descriptive statistics for the variables under analysis for the 50 US states between 1977 and 2010.[14] The first part of the table presents the statistics related to crime rates. It is clear that larcenies correspond to the bulk of property crimes in the US, whereas aggravated assaults represents the great majority of crimes registered as violent.

It is also possible to observe that the two alternative measures of police force, data from the UCR/FBI and the ASG have similar means and standard deviations. That is also captured by the scatter plot between the two measures in figure 3.15 in the appendix. It is clear that police and population data from both the UCR and the ASG have very similar distributions, but there is still significant noise in the measurement of those variables in both sources.

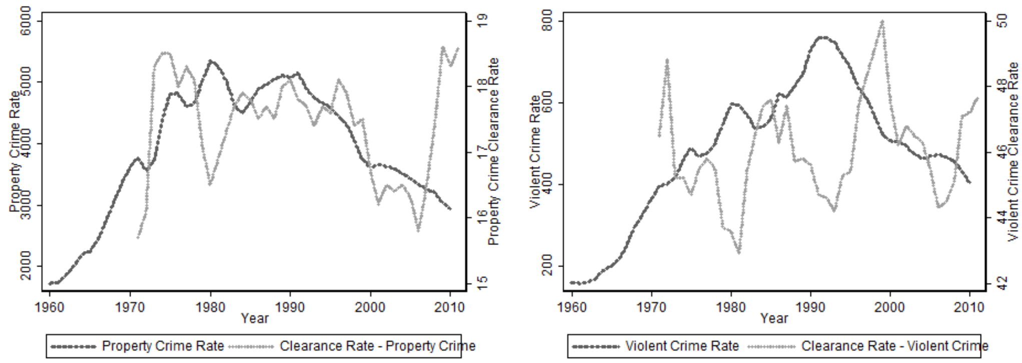
Another important aspect regarding the UCR police variable, the reference measure of law enforcement in this paper, regards the fraction of observations across the states and over the years that had positive variations. Table 3.1 shows that approximately 61% of the states/years had positive variations in police in the period under consideration. That is confirmed by figure 3.3a which shows the distribution of variations in police force for the uncorrected data. Figure 3.3b displays the same information for the corrected data, which has a very similar pattern. As expected, the variations are concentrated around the smallest values. The variations in the police force can also be observed in figure 3.5 which displays the law enforcement sizes for all states between 1977 and 2010.

Also from table 3.1, it is possible to observe that the level of income of unskilled workers was larger, when compared to the previous year in 89% of the cases. This fact can also be observed from figure 3.6. Even tough the number of negative variations is relatively small, figures 3.4a and 3.4b shows that the variations for both the income of unskilled workers and all workers were larger in absolute terms than the ones observed for police.

The spatial distribution across the continental states of the variables discussed in this section are displayed in appendix 3.A.2.

[13]Clearance rates refer to the proportion of crimes that are solved by the arrest of the perpetrator.

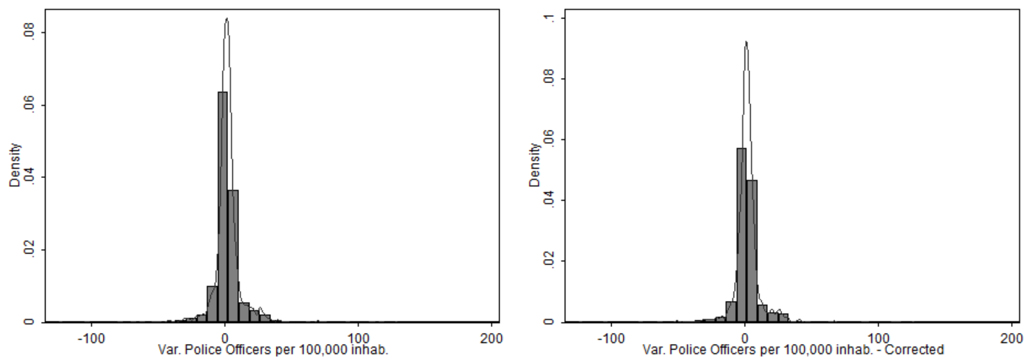
[14]Washington DC is excluded from the sample for constituting a classical outlier in the analysis, since its law enforcement size is considerably larger than all other states.



(a) Property Crime (b) Violent Crime

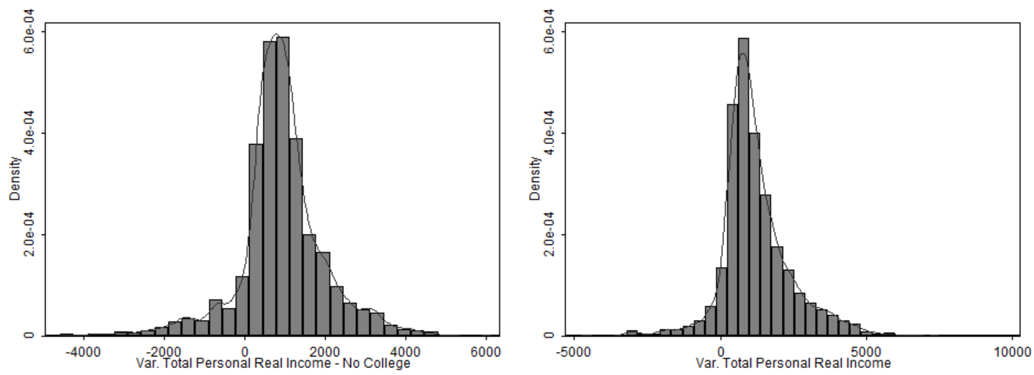
Figure 3.2: Crime rates and Clearance Rates - USA - 1970-2010

Data Source: UCR/FBI.



(a) ΔPolice (b) ΔPolice Corrected

Figure 3.3: Histograms and Kernel Densities (Epanechnikov): Police



(a) ΔIncome - No College (b) Δ Income

Figure 3.4: Histograms and Kernel Densities (Epanechnikov): Income

Data Source: CPS.

Table 3.1: Descriptive Statistics - States

	N	mean	sd	min	max
Property Crime Rate	1700	4089.4	1177.44	1705.3	7996
Burglary Rate	1700	985.22	414.26	292.3	2906.7
Larceny Rate	1700	2723.07	725.48	1235.6	5106.1
Motor Vehicle Theft Rate	1700	381.11	204.62	70.5	1157.7
Violent Crime Rate	1700	450.61	228.27	47	1244.3
Murder Rate	1700	6.13	3.53	0.2	20.3
Aggravated Assault Rate	1700	278.79	146.62	31.3	785.7
Forcible Rape Rate	1700	34.45	13.54	7.3	102.2
Police Officers rate - UCR	1700	205.68	54.9	68.84	434.24
Police Officers per - ASG	1700	203.85	42.6	122.40	403.08
Variation in UCR Police > 0	1700	0.61	0.49	0	1
Total Real Income	1700	19041.11	12954.5	1686.25	56685.22
Total Real Income - No College	1700	15823.39	9540.4	1686.25	41398.51
Variation in Income - No College > 0	1700	0.89	0.31	0	1
Unemployment Rate	1700	0.06	0.02	0.02	0.19
Unemployment Rate - No College	1700	0.07	0.02	0.02	0.19
Perc. Urban Area	1700	0.61	0.26	0	1
Fraction of Female Headed HH	1700	0.27	0.03	0.11	0.36
Perc. of Young Males	1700	0.09	0.01	0.05	0.15
Total Real Income of HH	1700	43366.54	29683.26	3994.7	142940.75
Population - CPS	1700	5.08E+06	5.55E+06	365149	3.57E+07
Population - UCR	1700	5.29E+06	5.82E+06	410986	3.73E+07
Population - ASG	1700	5.22E+06	5.70E+06	402000	3.73E+07

Notes: All monetary variables adjusted by CPI in 2010. *Data Source*: UCR/FBI/ASG/CPS.

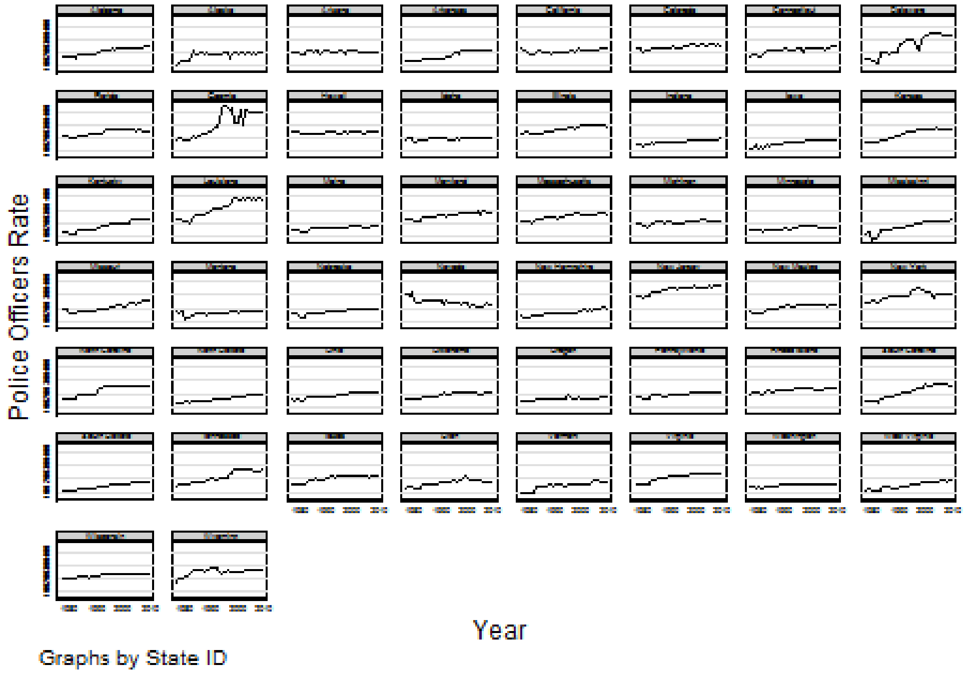


Figure 3.5: Police Rate by State

Data Source: UCR/FBI.

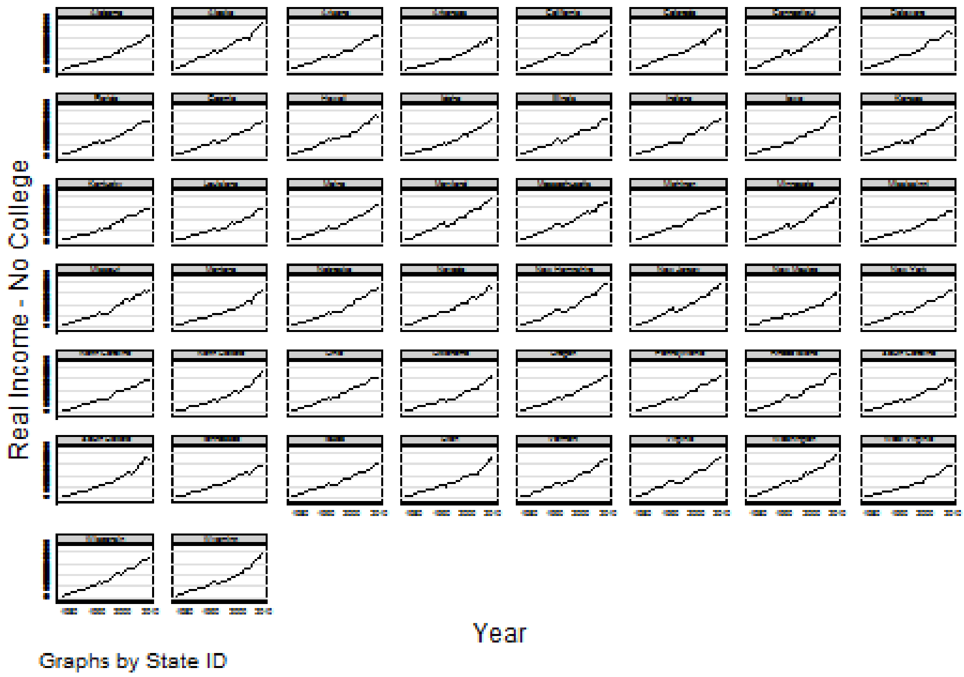


Figure 3.6: Real Income - Unskilled Workers by State

Data Source: CPS.

3.4 Empirical Framework

I focus attention on two variables susceptible to public policy and that potentially affect crime rates: income levels and number of police officers, with the last variable being potentially endogenous. However, as discussed in section 3.2, the uncoordinated efforts of overlapping police agencies make this variable relatively exogenous, especially at the state level. As in Chalfin and McCrary (2013), I focus attention at another source of endogeneity: measurement error of the police variable.[15] The other variables present in the theoretic model of crime and relevant to the decision of the individual at the margin are either controlled by observable variables or assumed to vary across states (but constant over time) and captured by the fixed effects, such as the severity of the punishment. Other variables that are not explicitly considered in the theoretical economic model, but can also affect crime, like the urbanization rate and fraction of young males are also included in the regressions.

I start by investigating the existence of asymmetric effects in variations of wages. As it will be seen in the next subsection, a similar approach is not feasible for the police variable as it would require the availability of two different instruments for each direction of variation. To examine the hypothesis of asymmetric cycles in the number of police officers, I will split the sample in two sub-samples, according to the predominance of positive or negative variations in the police force.

3.4.1 Random Coefficients Model

I initially estimate the following equation:[16]

$$Crime_{it} = \psi_i + \lambda_t + \mathbf{x}'_{it}\beta + \gamma Police_{it-1} + \delta_0 + \delta_1 Income^{+}_{it} + \delta_2 Income^{-}_{it} + \epsilon_{it} \quad (3.1)$$

where

$$\delta_0 = \begin{cases} 1 & if\ \Delta Income_{it} \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (3.2)$$

[15]The third possible source of endogeneity is omitted variable problem. That is a less likely possibility in the present analysis after controlling for several relevant variables and state and fixed effects.

[16]This is similar to the procedure adopted by Mocan and Bali (2010), that analyses the asymmetric effects of unemployment. However, the authors do not relax the assumption of a common intercept and as it will be seen, that is crucial to the analysis of asymmetric effects of wages and police force.

$$Income^{+}_{it} = \begin{cases} Income_{it} & if\ \Delta Income_{it} \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{3.3}$$

$$Income^{-}_{it} = \begin{cases} Income_{it} & if\ \Delta Income_{it} < 0 \\ 0 & \text{otherwise} \end{cases} \tag{3.4}$$

$Crime_{it}$ denotes one of the crime rates described in the previous section, $Police_{it-1}$ is the proportion of police officers relative to the state population in the previous period, $Income_{it}$ is the average real[17] income level for unskilled workers, $\mathbf{x}'_{it}$ is a vector of socioeconomic variables and ψ_i and λ_t are respectively state and year fixed effects. [18]

If individuals get locked in crime and there is hysteresis in the relationship between crime and income, a test of the hypothesis that $|\delta_1| = |\delta_2|$ must be rejected. However, note that the converse is not necessarily true. It must be emphasized that the detection of asymmetric effects is only indicative of the existence hysteresis in crime rates.

Additionally, note that the inclusion of the term δ_0 as defined by equation 3.2 relaxes the assumption of common intercepts for positive and negative changes. If this coefficient is not included, a linear specification might not reject the null hypothesis of homogeneous slopes for positive and negative variations, even if this is the case in the data.[19] As it will be seen in the following section, the assumption of heterogenous intercepts for positive and negative changes in the wage rate is crucial to correctly specify equation 3.1.

One would be worried about multicolinearity problem arising from the fact of estimating $Income^{+}$ and $Income^{-}$ in the same equation. That would be the case in a cross-sectional analysis, but the use of panel data dilute this problem.

3.4.2 Sample Split

An important implicit assumption under the specification given by equations 3.1-3.4 is that all other coefficients and state and time fixed effects are common to both states with predominantly positive variations and states with predominantly

[17] Deflated using the CPI with base in 2000.

[18] Baltagi, Matyas, and Sevestre (2008) show that an over-specification of the error components model (estimate a two-way model when the true model is one-way) provides consistent estimates, whereas under-specification of the error components model (estimate a one-way model when the true model is two-way) generates inconsistent estimates.

[19] To some extent, this is equivalent to the situation of a simple linear regression where the intercept is not specified when the data display an association that does not depart from the origin. In the case of where the true relationship is positive and has a positive intercept, the omission of the intercept in the regression will lead to an overestimation of the coefficient and potentially to an incorrect non-rejection of the null of no relationship.

negative variations. To check the robustness of the results, I split the sample according to the predominance of positive variations (number of positive variations above the average). [20] Because the threshold here is known, the inference is based on the standard distributions.[21] The empirical model is then given by the following equations:

$$Crime_{it} = \psi_{1i} + \lambda_{1t} + \mathbf{x}'_{it}\beta_1 + \gamma_1 Police_{it-1} + \delta_1 Income_{it} + \epsilon_{1it}, \text{ if } i \in \Gamma \quad (3.5)$$

$$Crime_{it} = \psi_{2i} + \lambda_{2t} + \mathbf{x}'_{it}\beta_2 + \gamma_2 Police_{it-1} + \delta_2 Income_{it} + \epsilon_{2it}, \text{ if } i \notin \Gamma \quad (3.6)$$

where Γ is the set of states with positive variations in income above the average.

That model allows to use the instrumental variable approach[22] to estimate the heterogeneity of negative and positive variations in police force using the same instrument:

$$Crime_{it} = \phi_{1i} + \lambda_{1t} + \mathbf{x}'_{it}\beta_1 + \theta_1 Police_{it-1} + \varrho_1 Income_{it} + \mu_{1it}, \text{ if } i \in \Omega \quad (3.7)$$

$$Crime_{it} = \phi_{2i} + \lambda_{2t} + \mathbf{x}'_{it}\beta_2 + \theta_2 Police_{it-1} + \varrho_2 Income_{it} + \mu_{2it}, \text{ if } i \notin \Omega \quad (3.8)$$

where Ω is the set of states with positive variations in police above the average.

To simplify the analysis, I exclude the police variable from equations 3.5 and 3.6 in the estimations presented in section 3.5.

3.4.3 Measurement Error in Police and IV Estimation

As recently emphasised by Chalfin and McCrary (2013), there is considerable measurement error in the law enforcement variable caused by misclassification and underestimation of the population size, especially in years prior to the census. That

[20] Chan (1993) shows that this type of estimator is strongly consistent.

[21] To see the properties of sample splitting and threshold estimators, see Hansen (1999) and Hansen (2000).

[22] Caner and Hansen (2004) show that if the threshold is exogenous, the IV/GMM estimator is consistent, but not necessarily efficient.

measurement error would lead to the attenuation bias, which in the present case would be an upward bias.

The authors claim that measurement error, rather than the simultaneity between police and crime would the main source of upwards bias in the relationship between police and crime. They claim that the police variable would be relatively exogenous, since the response from policy makers in the short term is extremely limited. That argument is even more compelling for state level data as the police department efforts are combinations of overlapping and uncoordinated local police agencies.

In order to tackle the measurement error in the police officers rates, I will follow Chalfin and McCrary (2013) and instrument the UCR/FBI law enforcement measure with the ASG measure. This procedure in which a noisy measure of police is used as an instrument for another measure of police will provide consistent estimates, conditional on the fact that the assumptions in the classical measurement error model hold.

Classical Measurement Error Model

Consider a crime equation with all variables and time and fixed effects netted out, apart from the correctly measured police size $Police^*_{it}$:

$$Crime_{it} = \eta_0 + \eta_1 Police^*_{it} + \epsilon_{it} \tag{3.9}$$

and $E(Police^*_{it}\epsilon_{it}) = 0$. However, if $Police^*_{it}$ is not observed, and all that can be observed is a noisy measure of police, $Police_{it}$, and the corresponding measurement error is given by:

$$e_{it} = Police_{it} - Police^*_{it} \tag{3.10}$$

It is assumed that $E(e_{it}) = 0$ and since $E(Police^*_{it}\epsilon_{it}) = 0$ as well as $E(Police_{it}\epsilon_{it}) = 0$, we have that $E(e_{it}\epsilon_{it}) = 0$.

A crucial assumption is that:

$$E(e_{it}Police^*_{it}) = 0 \tag{3.11}$$

Equation 3.11 is the classical errors-in-variables (CEV)[23]

The OLS estimation of the crime equation with the noisy police variable will yield:

$$\begin{aligned} Crime_{it} &= \gamma_0 + \gamma_1(Police_{it} - e_{it}) + \epsilon_{it} \\ &= \gamma_0 + \gamma_1 Police_{it} + \kappa_{it} \end{aligned} \tag{3.12}$$

[23] For details, see Wooldridge (2002).

with $\kappa_{it} = \epsilon_{it} - \gamma_1 e_{it}$ and consequently $E(Police_{it} \cdot \kappa_{it}) \neq 0$

If condition 3.11 holds, then $E(e_{it}Police_{it}) = E(e_{it}Police^*_{it}) + E(e^2_{it}) = \sigma^2_e$ and

$$plim(\hat{\gamma_1}) = \gamma_1(\frac{\sigma^2_{Police^*}}{\sigma^2_{Police^*} + \sigma^2_e}) = \gamma_1 \frac{\sigma^2_{Police^*}}{\sigma^2_{Police}} \neq \gamma_1 \tag{3.13}$$

for $\sigma^2_{Police^*} \neq \sigma^2_{Police}$.

IV Solution

With the use of another independent and noisy measure of police $Police^b_{it}$ and respective measurement error given by:

$$u_{it} = Police^b_{it} - Police^*_{it} \tag{3.14}$$

with $E(u_{it}) = 0$ and similarly to the previous imperfect measure of police the following conditions hold:

$$\begin{aligned} &E(Police_{it}\epsilon_{it}) = 0 \\ &E(u_{it}Police^*_{it}) = 0 \\ &E(u_{it}\epsilon_{it}) = 0 \\ &E(u_{it}e_{it}) = 0 \end{aligned} \tag{3.15}$$

$Police^b_{it}$ is an appropriate instrument to $Police_{it}$ since the conditions above imply that:

$$\begin{array}{lclcl} E(Police^b_{it}\epsilon_{it}) & = & E(Police^*_{it}\epsilon_{it}) + E(u_{it}\epsilon_{it}) & = & 0 \\ E(Police^b_{it}Police_{it}) & = & E(Police^*_{it}Police_{it}) + E(u_{it}Police_{it}) & & \\ & = & E(Police^*_{it}Police_{it}) + 0 & > & 0 \end{array} \tag{3.16}$$

which are respectively the conditions of exclusion and relevance for the instrument.

3.5 Results

This section presents the estimates of the models specified in the previous section. All regressions control for state and year effects and for the variables discussed in the data section.[24] The standard errors are clustered at the state level and all coefficients are standardised. All regressions with instruments reject the null of weak instruments and underidentification.[25]

Table 3.2 provides the estimates for the baseline model with respect to the level of income of unskilled workers.[26] Tables with the same approach where the overall level of income is used are presented in the appendix.

A higher level of income for unskilled workers has a negative impact on property crime, burglary and larceny, with an increase in one standard deviation in income contributing to reduce those crimes by respectively 0.84, 0.76 and 0.87 deviations. Those are very large magnitudes if one considers the size of the standard deviations of those variables. There is no statistically significant effect on motor vehicle theft or any violent crime.

The fact that impact of income is only observed on property crimes is in line with the empirical literature that find some evidence in this direction. This is also an expected result from the theoretical perspective, as the effect of the level of wages would play a role as an opportunity cost of crime in the less serious crimes.

[24]The year effects capture the trends and common shocks such as changes in federal legislation and efficiency of the police to deter crime. The models were estimated with First Differences as well, yielding similar coefficients to the Fixed effects estimates.

[25]The tests are respectively the Cragg-Donald Wald F statistic and the Kleibergen-Paap rank LM statistic.

[26]In order to simplify the analysis, I exclude the police variable from the estimations focused on the impact of income on crime.

Table 3.2: Crime Equations - States - Unskilled Workers Total Income - Baseline

	Property	Burglary	Larceny	MVT	Violent	Murder	Assault	Rape
Unskilled Workers								
Total Income	-0.85***	-0.77***	-0.90***	-0.18	-0.00	0.30**	0.07	0.05
	(0.22)	(0.22)	(0.23)	(0.25)	(0.13)	(0.10)	(0.17)	(0.28)
Constant	-1.21***	-0.54	-1.57***	-0.34	-0.33	0.47**	-0.36	-0.33
	(0.34)	(0.34)	(0.34)	(0.34)	(0.20)	(0.16)	(0.27)	(0.40)
State Effects?	yes	yes	yes	yes	yes	yes	yes	yes
Time Effects?	yes	yes	yes	yes	yes	yes	yes	yes
Observations	1700	1700	1700	1700	1700	1700	1700	1700
R^2	0.63	0.71	0.58	0.34	0.35	0.48	0.34	0.26

Notes: Standardised coefficients. Weighted by population in 2010. Standard errors robust to heteroscedasticity and clustering on state in parentheses. p-values for the test for unobserved state, time and two-way effects, respectively: 0.0000 ; 0.0000; 0.0000 in all regressions. Regressions controlled for the Proportion of black people,
Proportion of population between 15-19, 20-24, 25-29, 30-34 and 35-39 years old
Urbanization, Proportion of one-parent households and Unemployment rate.
*** p<0.01, ** p<0.05, * p<0.1 *Data Source*: UCR/FBI/CPS.

3.5.1 Random Coefficients Model

Tables 3.3 and 3.4 present the estimates of the empirical models presented at section 3.4.1 with respect to the level of income of unskilled workers. Tables with the same approach where the overall level of income is used are presented in the appendix.

Table 3.3 shows the results of the estimation of equation 3.1 with $\delta_0 = 0$.[27] There is no evidence of asymmetries between positive and negative variation with a common intercept.[28] However, as discussed in chapter 2, if there is a hysteresis loop in the relationship, it would be nonlinear. For a sample of observed values within a relatively short range of income values, the assumption of linearity is fairly reasonable, but the assumption of common intercept is very restrictive.[29]

That assumption is relaxed in the estimation presented in Table 3.4. There are asymmetric effects on positive and negative variations in wages for all property crime rates, apart from motor vehicle theft. Reductions of one standard deviation in the level of income of unskilled workers increase those crime rates by approximately one standard deviation. Increases in those variables produce a decrease in crime rates, however, that reduction is significantly smaller in absolute value than the ones observed when wages increase.

As in the previous approach, the impact of income is only observed on property crimes. It is also important to note that the estimated coefficients are now all larger then the ones observed when the assumption of common slopes and intercepts for positive and negative variations on income is implicitly assumed.

[27] I also exclude the police variable from the estimations focused on the impact of income on crime, in order to simplify the analysis and the exposition of the results.

[28] When testing the hypothesis of asymmetric effects, I use the more conservative two-sided test rather then the one-sided test, since the latter is more powerful to reject the null than the former.

[29] That assumption was present in Mocan and Bali (2010) in their analysis using unemployment.

Table 3.3: Crime Equations - States - Unskilled Workers Total Income - Asymmetric Effects - Common Intercept

	Property	Burglary	Larceny	MVT	Violent	Murder	Assault	Rape
Unskilled Workers								
Total Income (Negative Δ)	-0.50	-0.31	-0.69*	0.11	0.14	0.80***	0.12	0.21
	(0.34)	(0.36)	(0.33)	(0.42)	(0.27)	(0.23)	(0.29)	(0.35)
Unskilled Workers								
Total Income (Positive Δ)	-0.50	-0.33	-0.66*	0.10	0.12	0.75***	0.11	0.21
	(0.32)	(0.34)	(0.31)	(0.39)	(0.25)	(0.21)	(0.27)	(0.33)
State Effects?	yes	yes	yes	yes	yes	yes	yes	yes
Time Effects?	yes	yes	yes	yes	yes	yes	yes	yes
Observations	1700	1700	1700	1700	1700	1700	1700	1700
R^2	0.67	0.76	0.60	0.48	0.55	0.66	0.56	0.35
Positive Δ=Negative Δ test	1.01	0.79	0.84	2.67	2.87	1.53	4.03	0.37
p-value	0.31	0.77	0.36	0.10	0.09	0.22	0.05	0.54

Notes: Standardised coefficients. Weighted by population in 2010. Standard errors robust to heteroscedasticity and clustering on state in parentheses. p-values for the test for unobserved state, time and two-way effects, respectively: 0.0000 ; 0.0000; 0.0000 in all regressions.
Regressions controlled for the proportion of population between 15-19, 20-24, 25-29, 30-34 and 35-39 years old, Urbanization, Proportion of one-parent households and Unemployment rate.
*** p<0.01, ** p<0.05, * p<0.1 *Data Source*: UCR/FBI/CPS.

Table 3.4: Crime Equations - States - Unskilled Workers Total Income - Asymmetric Effects

	Property	Burglary	Larceny	MVT	Violent	Murder	Assault	Rape
Unskilled Workers								
Total Income (Negative Δ)	-1.03***	-0.94***	-1.05***	-0.38	0.01	0.33*	0.14	-0.07
	(0.26)	(0.24)	(0.28)	(0.30)	(0.16)	(0.12)	(0.21)	(0.32)
Unskilled Workers								
Total Income (Positive Δ)	-0.88***	-0.79**	-0.92***	-0.29	0.00	0.34**	0.08	0.03
	(0.24)	(0.23)	(0.25)	(0.29)	(0.15)	(0.11)	(0.20)	(0.29)
Positive Δ Unskilled Workers								
Total Income =1	-0.27*	-0.32**	-0.21*	-0.18	0.02	-0.08	0.15	-0.22
	(0.11)	(0.11)	(0.10)	(0.15)	(0.08)	(0.06)	(0.10)	(0.20)
Constant	-1.01**	-0.25	-1.43***	-0.28	-0.33	0.59**	-0.46	-0.18
	(0.37)	(0.38)	(0.38)	(0.42)	(0.24)	(0.18)	(0.31)	(0.47)
State Effects?	yes	yes	yes	yes	yes	yes	yes	yes
Time Effects?	yes	yes	yes	yes	yes	yes	yes	yes
Observations	1700	1700	1700	1700	1700	1700	1700	1700
R^2	0.62	0.71	0.57	0.26	0.34	0.47	0.34	0.25
Positive Δ=Negative Δ test	8.88	10.96	7.95	1.84	0.26	0.34	2.78	1.10
p-value	0.00	0.00	0.00	0.18	0.61	0.55	0.10	0.29

Notes: Standardised coefficients. Weighted by population in 2010. Standard errors robust to heteroscedasticity and clustering on state in parentheses. p-values for the test for unobserved state, time and two-way effects, respectively: 0.0000 ; 0.0000; 0.0000 in all regressions.
Regressions controlled for the proportion of population between 15-19, 20-24, 25-29, 30-34 and 35-39 years old, Urbanization, Proportion of one-parent households and Unemployment rate.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$ *Data Source*: UCR/FBI/CPS.

3.5.2 Sample Split

In this section I estimate the empirical models proposed in section 3.4.2, correcting for measurement error in the police force variable. Table 3.5 present the estimates for asymmetric variations in wages.[30] The results are qualitative similar to the previous approach, indicating strong asymmetric effects for all property crime rates, apart from motor vehicle theft. Nevertheless, the asymmetries are sharper than the ones observed in the random coefficients approach, under the assumption of homogeneous effects.

I now turn the focus to test the existence of asymmetric effects of police on crime. As a benchmark, the first part of table 3.6 presents the estimates without the correction for the measurement error in the population and police force sizes, whereas the second part display the results when both issues are tackled according to the strategy specified in section 3.4.3.

Table 3.7 present the estimates for asymmetric variations in police size. The asymmetric effects are even sharper than the ones observed for wages. There are asymmetric effects on positive and negative variations in the number of police officers for all crime rates, except larceny and rape. Reductions of one standard deviation in the police force size increase those crime rates produce increases in crime rates with varying magnitudes in terms of standard deviations. The largest effect is observed for motor vehicle theft, where reducing the police presence can increase crime rates by 2.59 standard deviations. More strikingly, increases in the number of police officers produce a decrease only in murder rates, with reduction being significantly smaller in absolute value than the ones observed in a police increase. All the other crime rates are either not statistically sensible to positive variations in police or positively affected by increase in police in the short run. That suggests that there still simultaneity between police and crime. Nevertheless, there is no reason to believe that this asymmetric effects would not persist if that issue is taken into account.

Table 3.8 present a robustness check. The results are robust to alternative estimation strategies, especially in terms of the sign of the coefficients. The conclusions of the test for the differences of the coefficients in this table are very similar to the ones in the previous analysis and therefore omitted. The results are robust across the alternative estimation procedures.

The first part of the table shows the estimates when the District of Columbia (DC) is included in the sample (initially excluded for having a level of police much higher

[30] As in previous section, I exclude the police variable from the estimations focused on the impact of income on crime.

than all US states). Asymmetric effects are also found with this approach. However, unlike the estimates where DC is not included, the effects from police are intensified and the asymmetries are more evident. Additionally, unlike the previous approach, Larceny has a negative impact from negative variations in police.

It can also be seen from table 3.8 that the results are not greatly altered when other approaches are used: excluding all the other covariates; excluding very small variations in police; excluding all states with population size below the 25th percentile and when the regressions are not weighted by the population size of each state.

However, as suggested in chapter 2, the relationship between crime rates and its main determinants is more likely to be non-linear rather than linear as assumed in the preceding analysis.

Table 3.5: Crime Equations - States - Unskilled Workers Total Income - Split Sample

	Property	Burglary	Larceny	MVT	Violent	Murder	Assault	Rape
Unskilled Workers								
Total Income (Negative Δ)	-1.23***	-1.14***	-1.25***	-0.39***	-0.12	0.16**	0.05	-0.44**
	(0.12)	(0.12)	(0.13)	(0.11)	(0.07)	(0.06)	(0.09)	(0.17)
Constant	-1.93***	-1.26***	-2.19***	-0.86***	-0.49***	0.14	-0.27	-0.88**
	(0.19)	(0.20)	(0.21)	(0.18)	(0.11)	(0.09)	(0.15)	(0.27)
State Effects?	yes	yes	yes	yes	yes	yes	yes	yes
Time Effects?	yes	yes	yes	yes	yes	yes	yes	yes
Observations	816	816	816	816	816	816	816	816
R^2	0.636	0.702	0.582	0.211	0.307	0.488	0.262	0.280
Unskilled Workers								
Total Income (Positive Δ)	-0.44***	-0.33**	-0.48***	-0.19	0.10	0.53***	0.05	0.65***
	(0.12)	(0.12)	(0.12)	(0.16)	(0.09)	(0.06)	(0.10)	(0.13)
Constant	-0.48*	0.28	-0.90***	-0.13	-0.18	0.88***	-0.49**	0.38
	(0.20)	(0.19)	(0.20)	(0.26)	(0.15)	(0.10)	(0.17)	(0.22)
State Effects?	yes	yes	yes	yes	yes	yes	yes	yes
Time Effects?	yes	yes	yes	yes	yes	yes	yes	yes
Observations	884	884	884	884	884	884	884	884
R^2	0.621	0.742	0.546	0.310	0.361	0.493	0.385	0.237
Positive Δ=Negative Δ test	4.63	14.29	4.48	40.92	9.98	2.29	0.19	0.78
p-value	0.0316	0.0002	0.0244	0.0000	0.0016	0.1303	0.6628	0.3760

Notes: Standardised coefficients. Weighted by population in 2010. Standard errors robust to heteroscedasticity and clustering on state in parentheses. p-values for the test for unobserved state, time and two-way effects, respectively: 0.0000 ; 0.0000; 0.0000 in all regressions.
Regressions controlled for the proportion of population between 15-19, 20-24, 25-29, 30-34 and 35-39 years old, Urbanization, Proportion of one-parent households and Unemployment rate.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$ *Data Source*: UCR/FBI/ASG/CPS.

Table 3.6: Crime Equations - States - Police & Total Income - OLS & IV

OLS	Property	Burglary	Larceny	MVT	Violent	Murder	Assault	Rape
Police	0.47*	0.37	0.63**	-0.21	0.00	-0.16*	0.11	0.01
	(0.26)	(0.26)	(0.26)	(0.21)	(0.16)	(0.09)	(0.16)	(0.20)
Total Income	-1.01***	-0.85***	-0.93***	-0.79*	-0.48***	0.23***	-0.41**	-0.53**
	(0.22)	(0.18)	(0.21)	(0.40)	(0.18)	(0.07)	(0.20)	(0.26)
State Effects?	yes	yes	yes	yes	yes	yes	yes	yes
Time Effects?	yes	yes	yes	yes	yes	yes	yes	yes
Observations	1700	1700	1700	1700	1700	1700	1700	1700
R^2	0.791	0.856	0.742	0.510	0.606	0.726	0.580	0.621

IV & Population Correction	Property	Burglary	Larceny	MVT	Violent	Murder	Assault	Rape
Police	0.51**	0.15	1.10***	-1.12***	-0.43**	-0.67***	-0.21	0.57***
	(0.21)	(0.22)	(0.20)	(0.27)	(0.17)	(0.12)	(0.13)	(0.14)
Total Income	-0.86***	-0.75***	-0.72***	-0.88***	-0.54***	0.12***	-0.46***	-0.37***
	(0.09)	(0.09)	(0.09)	(0.14)	(0.07)	(0.04)	(0.08)	(0.09)
State Effects?	yes	yes	yes	yes	yes	yes	yes	yes
Time Effects?	yes	yes	yes	yes	yes	yes	yes	yes
Observations	1700	1700	1700	1700	1700	1700	1700	1700
R^2	0.796	0.851	0.735	0.462	0.601	0.654	0.566	0.583

Notes: Standardised coefficients. Weighted by population in 2010. Standard errors robust to heteroscedasticity and clustering on state in parentheses. p-values for the test for unobserved state, time and two-way effects, respectively: 0.0000 ; 0.0000; 0.0000 in all regressions.
Regressions controlled for the proportion of population between 15-19, 20-24, 25-29, 30-34 and 35-39 years old, Urbanization, Proportion of one-parent households and Unemployment rate.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$ *Data Source*: UCR/FBI/ASG/CPS.

Table 3.7: Crime Equations - States - Police - Split Sample - IV & Population Correction

	Property	Burglary	Larceny	MVT	Violent	Murder	Assault	Rape
Police (Negative Δ)	-0.35	-0.55*	-0.54**	-2.59***	-1.74***	-0.71***	-1.37***	0.62**
	(0.28)	(0.30)	(0.25)	(0.47)	(0.31)	(0.16)	(0.25)	(0.23)
Total Income	-0.42***	-0.45***	-0.34**	-0.30	-0.01	0.19***	-0.07	0.03
	(0.14)	(0.13)	(0.14)	(0.22)	(0.13)	(0.06)	(0.12)	(0.13)
State Effects?	yes	yes	yes	yes	yes	yes	yes	yes
Time Effects?	yes	yes	yes	yes	yes	yes	yes	yes
Observations	714	714	714	714	714	714	714	714
R^2	0.865	0.902	0.829	0.514	0.616	0.725	0.623	0.659
Police (Positive Δ)	1.44***	0.86***	1.67***	0.64*	0.20	-0.54***	0.02	0.58**
	(0.29)	(0.23)	(0.30)	(0.28)	(0.14)	(0.11)	(0.19)	(0.21)
Total Income	-0.88***	-0.85***	-0.76***	-0.67***	-0.40***	-0.04	-0.19*	-0.81***
	(0.15)	(0.13)	(0.14)	(0.18)	(0.08)	(0.05)	(0.09)	(0.13)
State Effects?	yes	yes	yes	yes	yes	yes	yes	yes
Time Effects?	yes	yes	yes	yes	yes	yes	yes	yes
Observations	986	986	986	986	986	986	809866	986
R^2	0.646	0.788	0.544	0.552	0.615	0.772	0.500	0.459
Positive Δ=Negative Δ test	3.13	6.42	0.22	9.38	3.73	9.61	12.09	2.24
p-value	0.0767	0.0113	0.6376	0.0022	0.0534	0.0019	0.0005	0.1341

Notes: Standardised coefficients. Weighted by population in 2010. Standard errors robust to heteroscedasticity and clustering on state in parentheses. p-values for the test for unobserved state, time and two-way effects, respectively: 0.0000 ; 0.0000; 0.0000 in all regressions.
Regressions controlled for the proportion of population between 15-19, 20-24, 25-29, 30-34 and 35-39 years old, Urbanization, Proportion of one-parent households and Unemployment rate.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$ *Data Source*: UCR/FBI/ASG/CPS.

Table 3.8: Crime Equations - States - Police - Split Sample - Robustness Check

	Property	Burglary	Larceny	MVT	Violent	Murder	Assault	Rape
Include DC								
Police (Negative Δ)	-1.29***	-1.89***	0.25	-4.18***	-2.24***	-2.06***	-1.06***	0.51*
Police (Positive Δ)	1.19***	0.65**	1.37***	0.67**	0.50*	0.53	0.61*	-0.07
No covariates								
Police (Negative Δ)	-0.94**	-1.26***	0.44	-4.13***	-2.05***	-2.04***	-0.99***	0.40
Police (Positive Δ)	1.31***	0.77***	1.48***	0.78***	0.24*	-0.31***	0.17	0.22
Exclude $\lvert\Delta police\rvert < 1$								
Police (Negative Δ)	-1.00**	-1.63***	0.40	-3.61***	-1.94***	-1.81***	-0.87***	0.52
Police (Positive Δ)	0.99**	0.45*	1.24**	0.40*	0.14	-0.33**	0.17	-0.05
Exclude pop. < 25 perc.								
Police (Negative Δ)	-1.29***	-1.91***	0.22	-4.13***	-2.16***	-2.09***	-0.95***	0.59*
Police (Positive Δ)	0.98***	0.55**	1.22**	0.35	0.13	-0.29**	0.08	0.11
No population weights								
Police (Negative Δ)	-0.43	-1.04***	1.12**	-3.99***	-1.73***	-0.80***	-1.27***	1.22*
Police (Positive Δ)	1.08***	0.51***	1.30***	0.58***	0.06	-0.36***	0.09	-0.21

Notes: Standardised coefficients. Standard errors robust to heteroscedasticity
and clustering on state in parentheses. p-values for the test for unobserved state, time and two-way effects, respectively: 0.0000 ; 0.0000; 0.0000 in all regressions.
*** p<0.01, ** p<0.05, * p<0.1 *Data Source*: UCR/FBI/ASG/CPS.

3.6 Semiparametric Estimation

All the analyses in the previous sections, and indeed all longitudinal studies in the crime literature, assumed linearity of the investigated relationships and all coefficient comparisons were in terms of the means. In order to test if those assumptions were very restrictive I estimate the asymmetric effects of positive and negative variations in wages and police on crime nonparametrically.

That also allows to reconcile the observed asymmetric effects with the concept of hysteresis, by taking into account the shape of the relationships over the observed support of both wages and police and the theoretical possible values of those variables. If there is hysteresis the confidence intervals for the negative and positive variations relationships should not overlap in at least a significant part of the independent variable.

In the nonparametric approach, I keep the same structure of the previous sections controlling linearly for the state/time effects, covariates and instrument. For that reason the procedure in this section is effectively a semiparametric estimation.

The models discussed in section 3.4.2 can be now restated as semiparametric single index models:

$$Crime_{it} = \varphi_1(\psi_{1i} + \lambda_{1t} + \mathbf{x}_{it}'\beta_1 + \gamma_1 Police_{it-1} + \delta_1 Income_{it}) + \epsilon_{1it} \text{ if } i \in \Gamma \quad (3.17)$$

$$Crime_{it} = \varphi_2(\psi_{2i} + \lambda_{2t} + \mathbf{x}_{it}'\beta_2 + \gamma_2 Police_{it-1} + \delta_2 Income_{it}) + \epsilon_{2it} \text{ if } i \notin \Gamma \quad (3.18)$$

$$Crime_{it} = \varphi_3(\phi_{1i} + \lambda_{1t} + \mathbf{x}_{it}'\beta_1 + \theta_1 Police_{it-1} + \varrho_1 Income_{it}) + \mu_{1it} \text{ if } i \in \Omega \quad (3.19)$$

$$Crime_{it} = \varphi_4(\phi_{2i} + \lambda_{2t} + \mathbf{x}_{it}'\beta_2 + \theta_2 Police_{it-1} + \varrho_2 Income_{it}) + \mu_{2it} \text{ if } i \notin \Omega \quad (3.20)$$

where Γ is the set of states with positive variations in income above the average, Ω is the set of states with positive variations in police above the average and $\varphi_k,\ k = \{1,2,3,4\}$ are estimated nonparametrically.[31]

[31] For the properties of the semiparametric single index model, see Ichimura (1993), that established the estimator.

As in the parametric estimation, I simplify the analysis by excluding the police variable from the estimations of equations 3.17 and 3.18.

I use the the Lowess estimator - Locally Weighted Scatterplot Smoothing that uses the tricubic kernel:

$$K(z) = \frac{70}{81}(1 - |z|^3)^3 \times \mathbf{1}(|z| < 1). \tag{3.21}$$

One of its advantages lies on the fact that, unlike most nonparametric estimators based on other kernel functions, it is a robust estimator against outliers.[32] Because the results of nonparametrically estimated models are much more sensible to the bandwidth choice, rather than the kernel choice, even for the Lowess estimator, I test the robustness of the estimates for different values of bandwidth.

In order to implement the semiparametric single index models for panel data, I partial out the control variables and the fixed and time effects where the residuals, rather than the variables in levels are used in order to account for state and year effects, in addition to all control variables and the instrument used in the previous section.

Equations 3.17-3.20 are then rewritten as:

$$Resid(Crime_{it}) = \varphi_1(\delta_1 Resid(Income_{it})) + \varepsilon_{1it} \text{ if } i \in \Gamma \tag{3.22}$$

$$Resid(Crime_{it}) = \varphi_2(\delta_2 Resid(Income_{it})) + \varepsilon_{2it} \text{ if } i \notin \Gamma \tag{3.23}$$

$$Resid(Crime_{it}) = \varphi_3(\theta_1 Resid(Police_{it-1})) + \zeta_{1it} \text{ if } i \in \Omega \tag{3.24}$$

$$Resid(Crime_{it}) = \varphi_4(\theta_2 Resid(Police_{it-1})) + \zeta_{2it} \text{ if } i \notin \Omega \tag{3.25}$$

The Frisch-Waugh-Lovell theorem ensures that the estimated coefficients δ_1, δ_2, θ_1 and θ_1 in equations 3.17-3.20 are identical to the ones in equations 3.22-3.25.[33]

Figures 3.7 and 3.8 show the nonparametric estimates using the Lowess estimator with 0.5 bandwidth. The results are qualitatively unaltered for a different variations of levels of bandwidth and presented in the appendix.

Figure 3.7 provides a clear evidence of asymmetric effects for property crime rates. The degree of asymmetry is particularly strong for higher values of wages and nonexistent for lower levels of the distribution. This result is consistent with the

[32] For details, see Fan and Gijbels (1996) and Cameron and Trivedi (2005).

[33] See Frisch and Waugh (1933) and Lovell (2008).

hypothesis of hysteresis that predicts a larger gap between directions of variations as we move away from lower values of the independent variable.

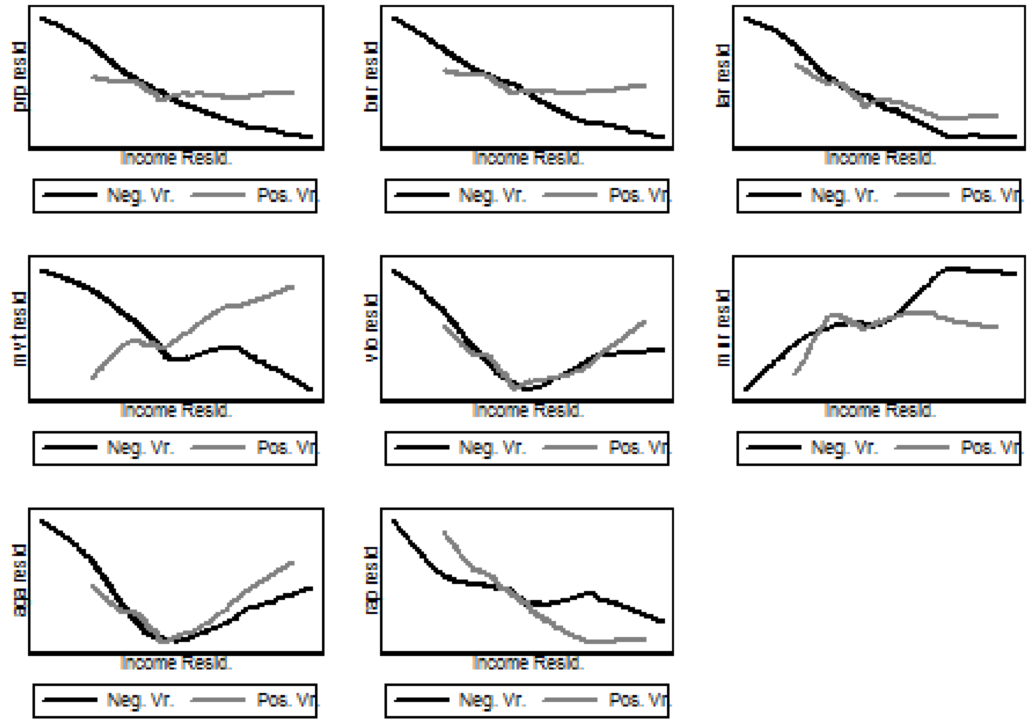


Figure 3.7: Positive vs Negative Variation in Wages - Lowess estimator - Bandwidth=0.5
Data Source: UCR/FBI/ASG/CPS.

Figure 3.8 displays the lowess estimates for the residuals of the IV regressions for the different crime rates and the police size. The results are similar to the ones observed for wages, but with the evidence of asymmetric effects present for all types of crimes. One important conclusion that can be taken form figure 3.8 is that all crime rates are negatively associated to police when this variable is fairly monotonically decreasing, but the relationship is highly nonlinear for positive variations.

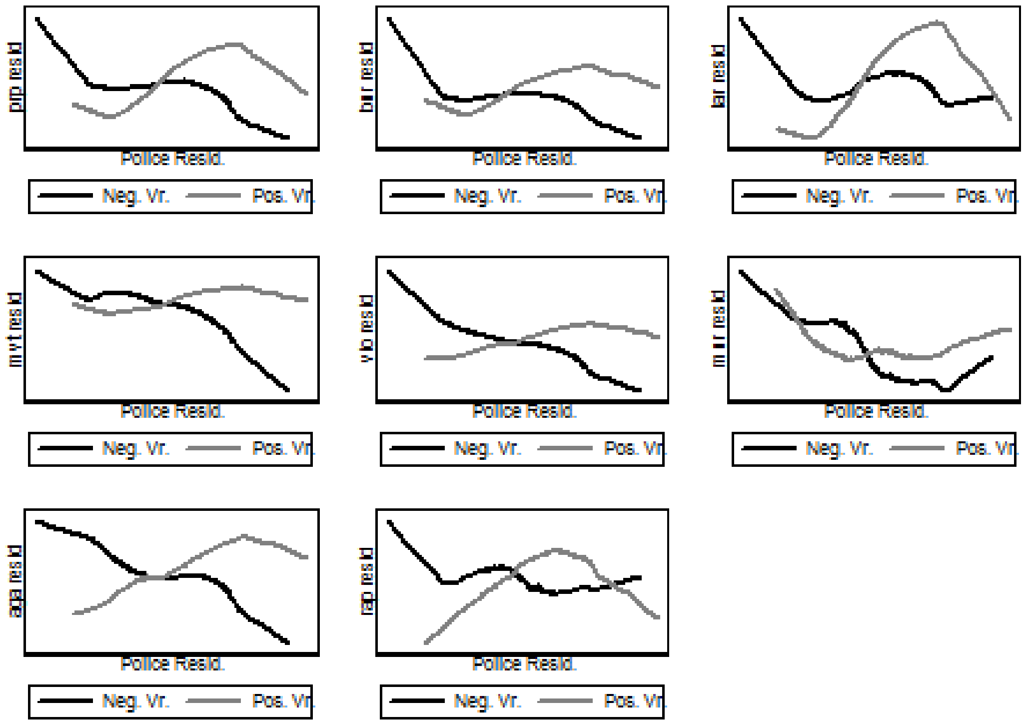


Figure 3.8: Positive vs Negative Variation in Police - Lowess estimator - Bandwidth=0.5
Data Source: UCR/FBI/ASG/CPS.

It is also important to note from 3.8 that the positive coefficients associated to positive variations in police officers obtained in the linear specification of the previous section were crucially stemming from increases when the size of the police force were smaller than the average in the sample. For higher numbers of police officers, those relationships are all negative, except to murder, but the slopes are still significantly larger in absolute value when compared to negative variations in police, respectively to each type of crime.

3.7 Conclusions

This chapter estimates the asymmetric impact of police and income on property and violent crime rates in the US between 1997 and 2010. This is the first effort to capture these asymmetries and it is also the first to correct for the measurement error at a higher level of aggregation. Furthermore, this is also the first analysis to take into account the existence of nonlinearities in the main determinants of crime with longitudinal aggregate crime data.

There is evidence of asymmetric effects of positive and negative variations in the level of income for both unskilled workers and the general population for property crime rates. In general, reductions of one standard deviation in the level of real income of unskilled workers increase those crime rates by approximately one standard deviation. Increases in the levels of real income produce a decrease in crime rates, however, that reduction is significantly smaller in absolute value than the ones observed when real income increase.

Those asymmetric effects are also observed for positive and negative variations in the number of police officers for all crime rates, except larceny and rape. Reductions of one standard deviation in the size of the police force increase those crime rates by one to four standard deviations, depending on the type of crime. For some types crimes and under some specifications, increases in the police size produce a decrease is significantly smaller in absolute value than the ones observed when law enforcement increase. For most specifications and types of crime, police and crime are positively associated when only positive variations are considered. A closer inspection with a semiparametric estimation reveals that the slopes are positive only for lower levels of police officers and they are negative for higher levels, except to murder, but the slopes are still significantly larger in absolute value when compared to negative variations in police, respectively to each type of crime.

The magnitude of the asymmetries found in this chapter supports the hypothesis of hysteresis in crime and suggests that no theoretical or empirical analysis would be complete without careful consideration of that important feature of the relationships between crime, police and legal income. These results are relevant for any empirical analysis of policies at crime reduction, but they are particularly important for evaluations of policies based on increases of police force size.

References

BALTAGI, B., L. MATYAS, AND P. SEVESTRE (2008): "Error Components Models," in *The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice*, ed. by L. Matyas, and P. Sevestre. Springer.

BUONANNO, P., AND G. MASTROBUONI (2012): "Police and crime: Evidence from Dictated Delays in Centralized Police Hiring," Discussion Paper 6477, IZA Discussion Paper.

CAMERON, A. C., AND P. K. TRIVEDI (2005): *Microeconometrics: Methods and Applications*. Cambridge University Press.

CANER, M., AND B. E. HANSEN (2004): "Instrumental variable estimation of a threshold model," *Econometric Theory*, 20(5), 813–843.

CHALFIN, A., AND J. MCCRARY (2013): "The Effect of Police on Crime: New Evidence from US Cities, 1960-2010," Discussion paper, National Bureau of Economic Research.

CHAN, K.-S. (1993): "Consistency and limiting distribution of the least squares estimator of a threshold autoregressive model," *The Annals of Statistics*, 21(1), 520–533.

CORMAN, H., AND H. N. MOCAN (2000): "A Time-Series Analysis of Crime, Deterrence, and Drug Abuse in New York City," *American Economic Review*, 90(3), 584–604.

CORNWELL, C., AND W. N. TRUMBULL (1994): "Estimating the Economic Model of Crime with Panel Data," *The Review of Economics and Statistics*, 76(2), 360–66.

CULLEN, J. B., AND S. D. LEVITT (1999): "Crime, urban flight, and the consequences for cities," *Review of economics and statistics*, 81(2), 159–169.

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*, pp. 115–133.

DILLS, A. K., J. A. MIRON, AND G. SUMMERS (2008): "What Do Economists Know About Crime?," Working Paper 13759, NBER.

DOYLE, J. M., E. AHMED, AND R. N. HORN (1999): "The effects of labor markets and income inequality on crime: evidence from panel data," *Southern Economic Journal*, pp. 717–738.

DRACA, M., S. MACHIN, AND R. WITT (2011): "Panic on the streets of London: crime, police and the July 2005 terror attacks," *American Economic Review*, 101(5), 2157–81.

EHRLICH, I. (1973): "Participation in Illegitimate Activities: A Theoretical and Empirical Investigation," *Journal of Political Economy*, 81(3), 521–65.

EVANS, W. N., AND E. G. OWENS (2007): "COPS and Crime," *Journal of Public Economics*, 91(1), 181–201.

FAN, J., AND I. GIJBELS (1996): *Local polynomial modelling and its applications*, vol. 66. Chapman & Hall/CRC.

FREEMAN, R. B. (1996): "Why Do So Many Young American Men Commit Crimes and What Might We Do About It?," *The Journal of Economic Perspectives*, 10(1), 25–42.

FRISCH, R., AND F. V. WAUGH (1933): "Partial time regressions as compared with individual trends," *Econometrica*, pp. 387–401.

FULLER, W. (1987): *Measurement Error Models*. Wiley-Blackwell.

GOULD, E. D., B. A. WEINBERG, AND D. B. MUSTARD (2002): "Crime rates and local labor market opportunities in the United States: 1979-1997," *The Review of Economics and Statistics*, 84(1), 45–61.

HANSEN, B. E. (1999): "Threshold effects in non-dynamic panels: Estimation, testing, and inference," *Journal of econometrics*, 93(2), 345–368.

——— (2000): "Sample Splitting and Threshold Estimation," *Econometrica*, 68(3), 575–603.

HAUSMAN, J. (2001): "Mismeasured variables in econometric analysis: Problems from the right and problems from the left," *Journal of Economic Perspectives*, pp. 57–67.

ICHIMURA, H. (1993): "Semiparametric least squares (SLS) and weighted SLS estimation of single-index models," *Journal of Econometrics*, 58(1), 71–120.

KING, M., S. RUGGLES, J. T. ALEXANDER, S. FLOOD, K. GENADEK, M. B. SCHROEDER, B. TRAMPE, AND R. VICK (2012): "Integrated Public Use Microdata Series, Current Population Survey [Machine-readable database]," *University of minnesota*.

LEE, D. S., AND J. MCCRARY (2009): *The deterrence effect of prison: Dynamic theory and evidence*. Industrial Relations Section, Princeton University.

LEVITT, S. (1997): "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime," *American Economic Review*, 87(3), 270–290.

LEVITT, S. D. (2002): "Using electoral cycles in police hiring to estimate the effects of police on crime: Reply," *American Economic Review*, pp. 1244–1250.

LIN, M.-J. (2009): "More police, less crime: Evidence from US state data," *International Review of Law and Economics*, 29(2), 73–80.

LOUREIRO, A. O. F. (2013): "Can Conditional Cash Transfers Reduce Poverty and Crime? Evidence from Brazil," Proceedings, Royal Economic Society Annual Conference.

LOVELL, M. C. (2008): "A simple proof of the FWL theorem," *The Journal of Economic Education*, 39(1), 88–91.

MCCRARY, J. (2002): "Using electoral cycles in police hiring to estimate the effect of police on crime: Comment," *American Economic Review*, pp. 1236–1243.

MOCAN, H. N., AND T. G. BALI (2010): "Asymmetric Crime Cycles," *The Review of Economics and Statistics*, 92(4), 899–911.

MUSTARD, D. B. (2010): "How Do Labor Markets Affect Crime? New Evidence on an Old Puzzle," IZA Discussion Papers 4856, Institute for the Study of Labor (IZA).

O'SHAUGHNESSY, T. (2011): "Hysteresis in unemployment," *Oxford Review of Economic Policy*, 27(2), 312–337.

POLINSKY, A. M., AND S. SHAVELL (2007): "The Theory of Public Enforcement of Law," in *Handbook of Law and Economics, Volume 1*, chap. 6, pp. 403–454. Elsevier B.V.

SOARES, R. R. (2004): "Development, crime and punishment: accounting for the international differences in crime rates," *Journal of Development Economics*, 73(1), 155–184.

SOARES, R. R., AND J. NARITOMI (2010): "Understanding High Crime Rates in Latin America: The Role of Social and Policy Factors," in *The Economics of Crime: Lessons for and from Latin America*, pp. 19–55. National Bureau of Economic Research, Inc.

WOOLDRIDGE, J. M. (2002): *Econometric Analysis of Cross Section and Panel Data*, MIT Press Books. The MIT Press.

WORRALL, J. L., AND T. V. KOVANDZIC (2010): "Police levels and crime rates: An instrumental variables approach," *Social Science Research*, 39(3), 506–516.

3.A Appendix

3.A.1 Extraction and Organisation of the Data

In this section I explain the details of extraction of the data from three used sources: Uniform Crime Reports - UCR, Current Population Survey - CPS and Annual Survey of State and Local Government Employment and Census of Governments - ASG.

UCR Data

The crime and police data are provided by the Federal Bureau of Investigation - FBI that runs the Uniform Crime Reports - UCR. Because the FBI webpage provide only the most recent years of data, I used the historical series archived by the Inter-university Consortium for Political and Social Research - ICPSR that dates back to 1966. However, years prior 1975 follow a very different methodology of organisation. And because the CPS data was not representative to all states before 1977, I restrict the sample between 1977 and 2010. This compilation of local government data provides a snapshot in October for the number of police personnel, disaggregated by police officers and administrative workers. To calculate the crime variables, I use the number of offenses that are available at monthly basis and are disaggregated by several categories of crime, where each category and each month is recorded in an individual variable. I group the data onto the crime types used in this paper by year. I use all main categories of crime: burglary, larceny, motor vehicle theft, murder, aggravated assault and forcible rape. I calculate the crime rates by dividing each crime variable by the respective population in each state and multiply by 100,000. A similar procedure is carried out to obtain the number of police officers by 100,000 inhabitants.

CPS Data

I use the data from the Current Population Survey - CPS, provided by the Bureau of Labor Statistics to construct most of the covariates in the estimated model, including the personal level of income. The microdata for the interviews carried out in march is available at the Integrated Public Use Microdata Series (IPUMS) website since 1962. However, the data is representative to all states only after 1977. The micro data is aggregated into sample-weighted variables by state for each year between 1977 and 2010.

ASG Data

In order to obtain the second measurement of police size, I also used the number of police officers from the Annual Survey of State and Local Government Employment and Census of Governments, U.S. Census Bureau. The data for federal and state employees are the actual numbers, however, as the local government statistics are based on a sample of local governments. As in the UCR data, the numbers refer to October of each year, but after 1997 the reference month was changed to March. As carried out in the UCR data, I calculate the number of police officers by 100,000 inhabitants by using the corrected population sizes provided by the dataset.

3.A.2 Maps

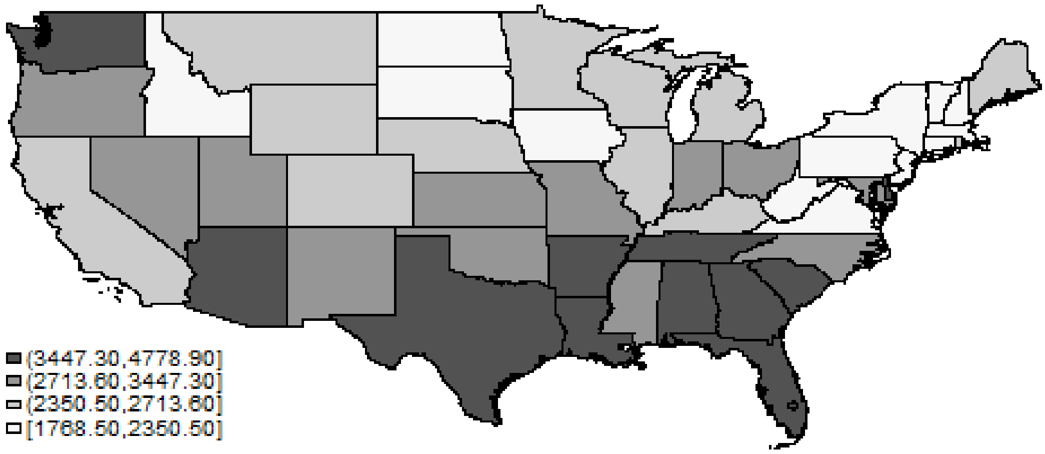


Figure 3.9: Property Crime Rates - States - 2010

Data Source: UCR/FBI.

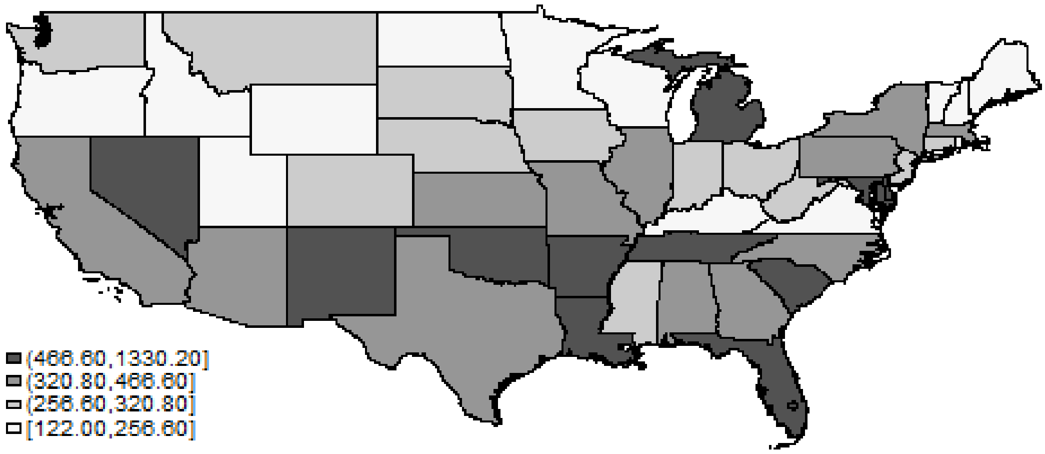


Figure 3.10: Violent Crime Rates - States - 2010

Data Source: UCR/FBI.

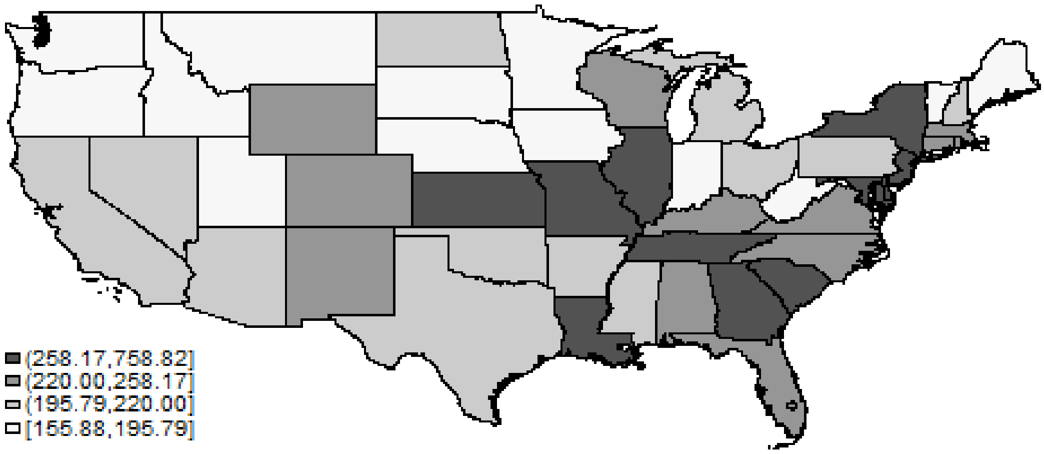


Figure 3.11: Police Officers per 100,000 inhab. - States - 2010
Data Source: UCR/FBI.

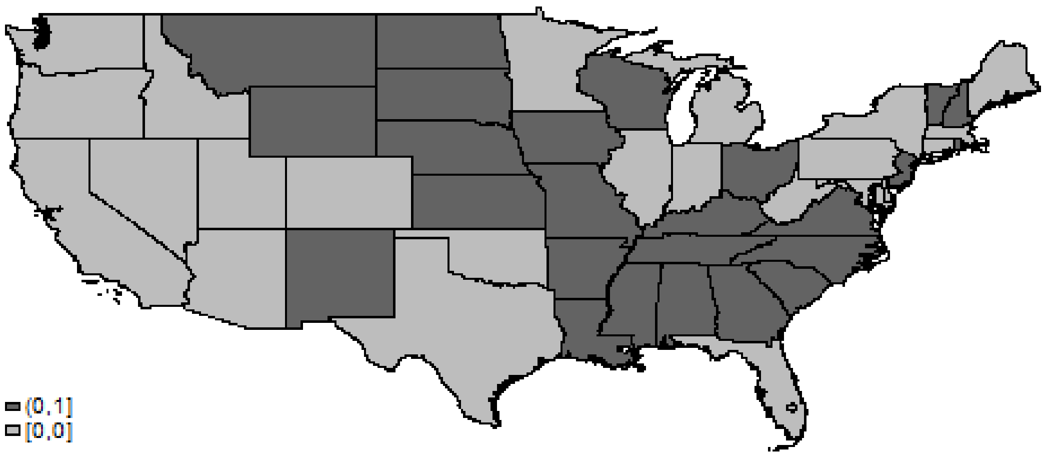


Figure 3.12: Positive Variations in Police Officers (pc) above Average in the 1977 - 2010 period - States
Data Source: UCR/FBI.

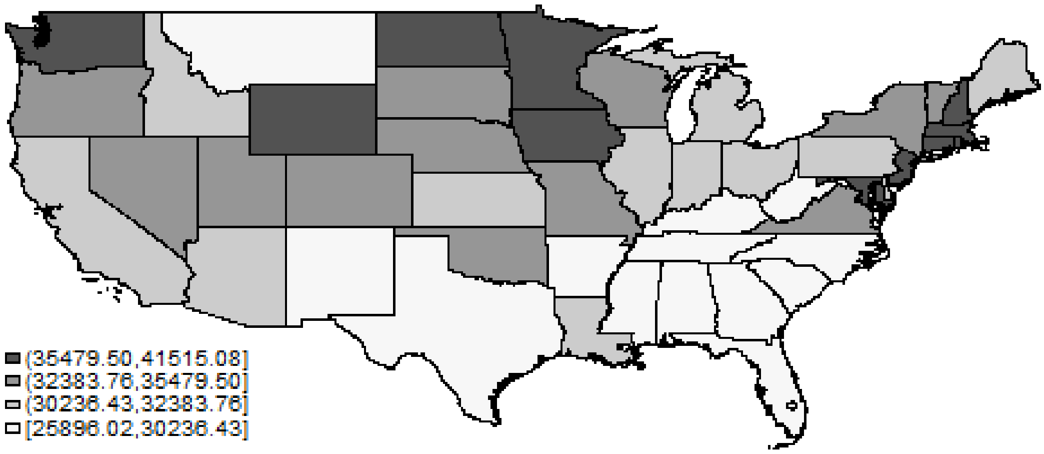


Figure 3.13: Real Income - Unskilled Workers - States - 2010
Data Source: CPS.

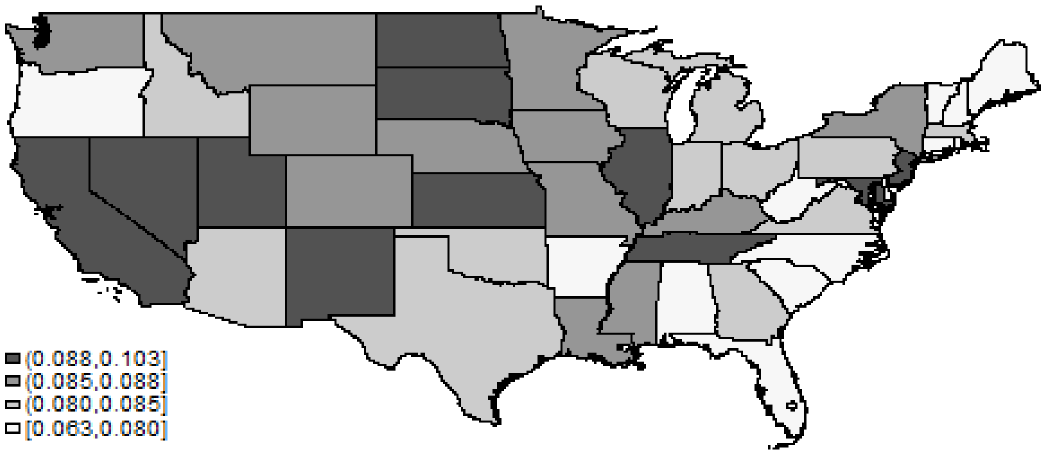


Figure 3.14: Young Male (%) - States - 2010
Data Source: CPS.

3.A.3 Additional Figures

Figure 3.15: Police UCR vs Police ASG - US States - 1977 - 2010

Data Source: UCR/FBI/ASG.

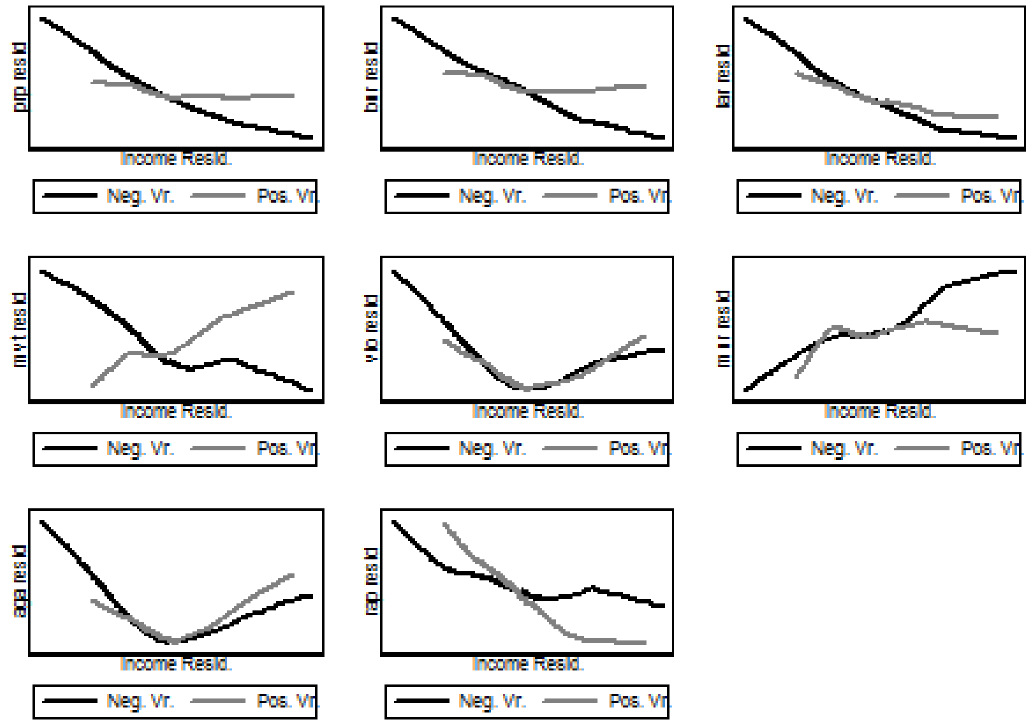


Figure 3.16: Positive vs Negative Variation in Wages - Lowess estimator - Bandwidth=0.8

Data Source: UCR/FBI/ASG/CPS.

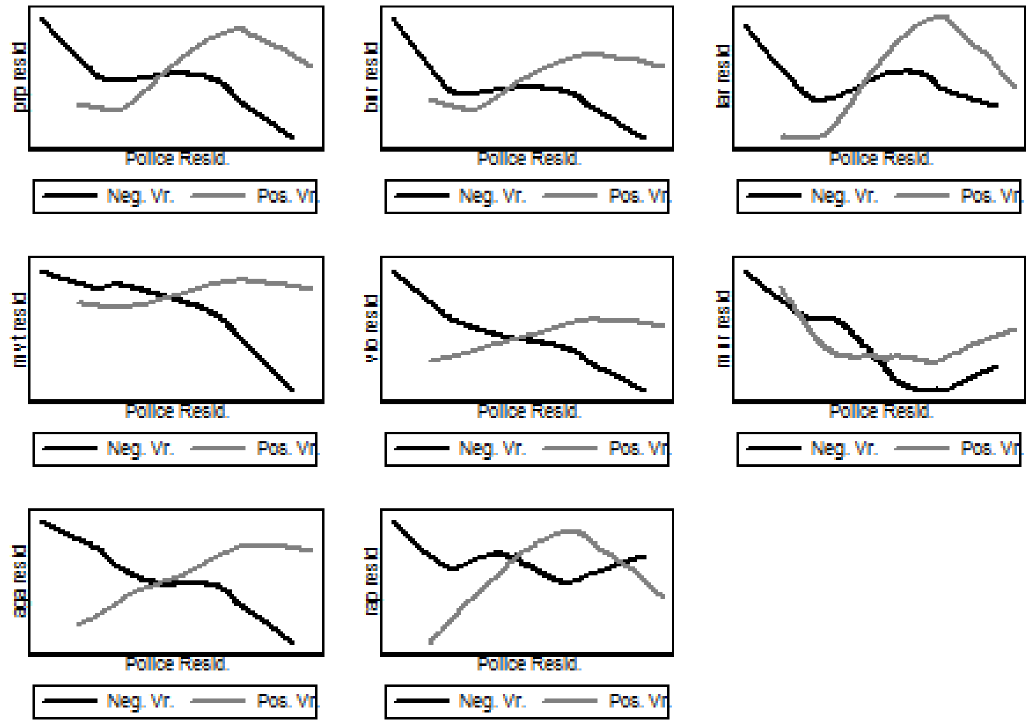


Figure 3.17: Positive vs Negative Variation in Police - Lowess estimator - Bandwidth=0.8

Data Source: UCR/FBI/ASG/CPS.

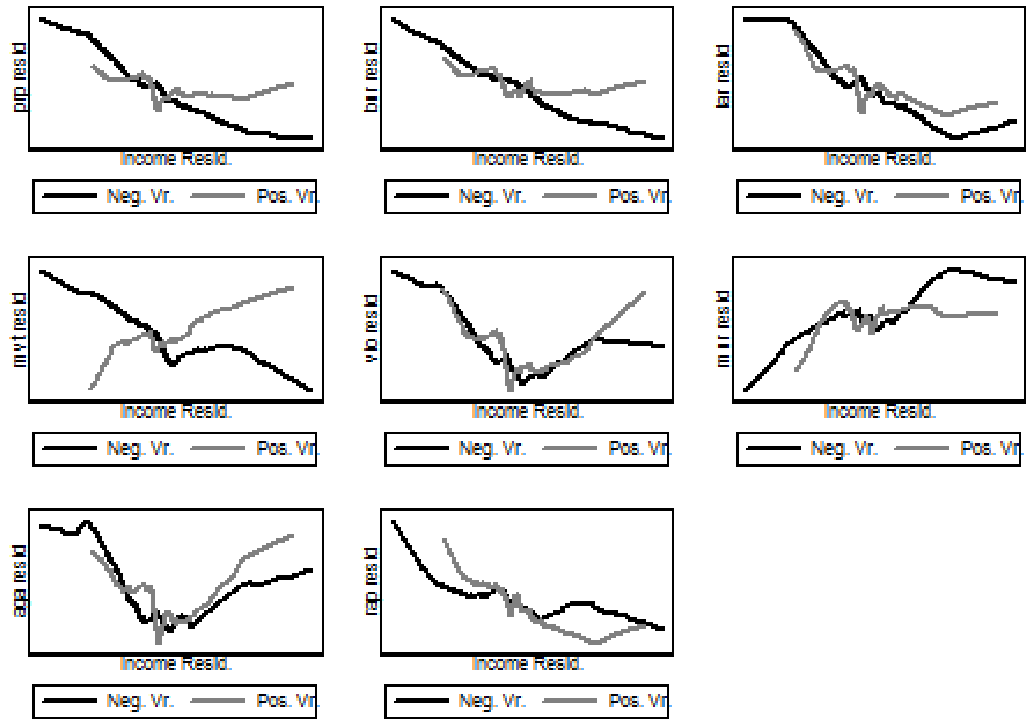


Figure 3.18: Positive vs Negative Variation in Wages - Lowess estimator - Bandwidth=0.2

Data Source: UCR/FBI/ASG/CPS.

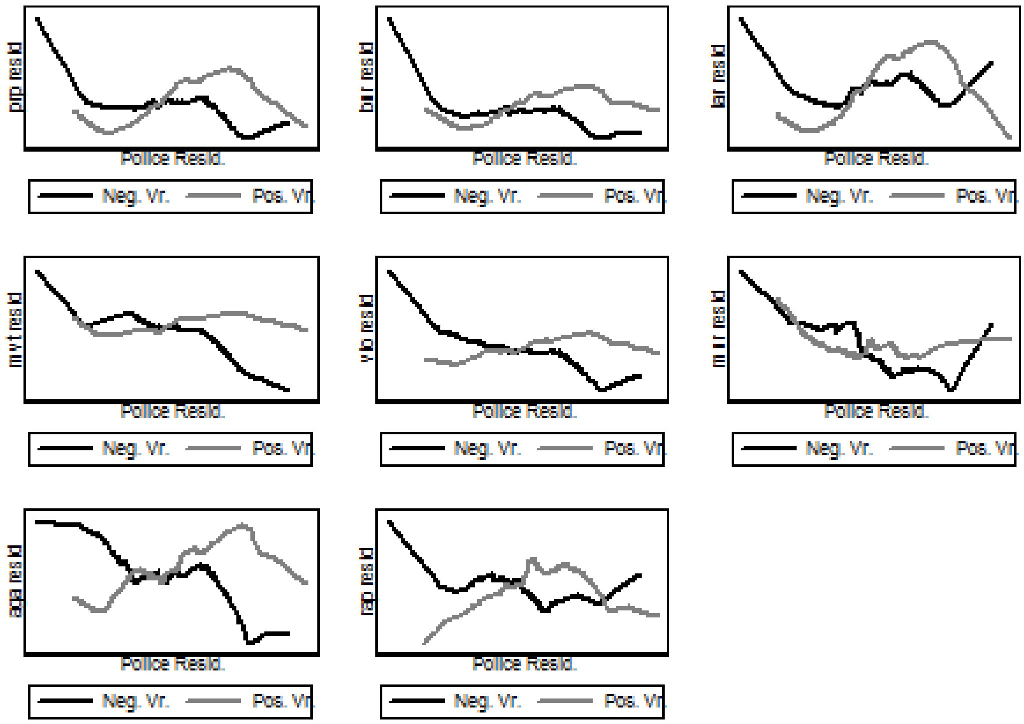


Figure 3.19: Positive vs Negative Variation in Police - Lowess estimator - Bandwidth=0.2

Data Source: UCR/FBI/ASG/CPS.

3.A.4 Additional Tables

Table 3.9: Crime Equations - States - Workers Total Income

	Property	Burglary	Larceny	MVT	Violent	Murder	Assault	Rape
Total Income	-1.21***	-1.00**	-1.16***	-0.86*	-0.48*	0.34	-0.37	-0.50
	(0.29)	(0.31)	(0.28)	(0.41)	(0.18)	(0.17)	(0.20)	(0.29)
Constant	-1.04**	-0.05	-1.33**	-1.13*	-0.63*	0.74***	-0.70*	-0.82*
	(0.36)	(0.38)	(0.39)	(0.53)	(0.24)	(0.20)	(0.30)	(0.39)
State Effects?	yes	yes	yes	yes	yes	yes	yes	yes
Time Effects?	yes	yes	yes	yes	yes	yes	yes	yes
Observations	1700	1700	1700	1700	1700	1700	1700	1700
R^2	0.70	0.78	0.63	0.50	0.57	0.62	0.57	0.36

Notes: Standardised coefficients. Weighted by population in 2010. Standard errors robust to heteroscedasticity and clustering on state in parentheses. p-values for the test for unobserved state, time and two-way effects, respectively: 0.0000 ; 0.0000; 0.0000 in all regressions.
Regressions controlled for the proportion of population between 15-19, 20-24, 25-29, 30-34 and 35-39 years old, Urbanization, Proportion of one-parent households and Unemployment rate.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$ *Data Source*: UCR/FBI/CPS.

Table 3.10: Crime Equations - States - Total Income - Asymmetric Effects

	Property	Burglary	Larceny	MVT	Violent	Murder	Assault	Rape
Total Income (Negative Δ)	-1.34***	-1.12***	-1.25***	-1.04*	-0.52**	0.32	-0.38	-0.54
	(0.29)	(0.31)	(0.29)	(0.40)	(0.19)	(0.17)	(0.21)	(0.30)
Total Income (Positive Δ)	-1.19***	-0.97**	-1.15***	-0.84	-0.48*	0.36*	-0.39	-0.48
	(0.29)	(0.32)	(0.28)	(0.42)	(0.18)	(0.18)	(0.20)	(0.30)
Positive Δ Total Income =1	-0.33	-0.40	-0.19	-0.42*	0.00	-0.16	0.13	-0.19
	(0.21)	(0.20)	(0.21)	(0.20)	(0.12)	(0.10)	(0.13)	(0.19)
Constant	-0.71	0.36	-1.15*	-0.71	-0.65*	0.92**	-0.86**	-0.61
	(0.47)	(0.48)	(0.49)	(0.62)	(0.28)	(0.27)	(0.32)	(0.49)
State Effects?	yes	yes	yes	yes	yes	yes	yes	yes
Time Effects?	yes	yes	yes	yes	yes	yes	yes	yes
Observations	1700	1700	1700	1700	1700	1700	1700	1700
R^2	0.70	0.78	0.63	0.50	0.57	0.63	0.58	0.36
Positive Δ=Negative Δ test	3.99	4.69	1.76	7.92	0.48	2.05	0.60	0.65
p-value	0.05	0.03	0.18	0.00	0.48	0.15	0.80	0.42

Notes: Standardised coefficients. Weighted by population in 2010. Standard errors robust to heteroscedasticity and clustering on state in parentheses. p-values for the test for unobserved state, time and two-way effects, respectively: 0.0000 ; 0.0000; 0.0000 in all regressions.
Regressions controlled for the proportion of population between 15-19, 20-24, 25-29, 30-34 and 35-39 years old, Urbanization, Proportion of one-parent households and Unemployment rate.
*** p<0.01, ** p<0.05, * p<0.1   *Data Source*: UCR/FBI/CPS.

Table 3.11: Crime Equations - States - Total Income - Split Sample

	Property	Burglary	Larceny	MVT	Violent	Murder	Assault	Rape
Total Income (Negative Δ)	-2.10***	-1.90***	-2.06***	-1.02***	-0.59***	0.18**	-0.66***	-0.23
	(0.13)	(0.12)	(0.15)	(0.16)	(0.09)	(0.06)	(0.13)	(0.19)
Constant	-2.93***	-2.06***	-3.25***	-1.31***	-0.99***	0.32***	-1.12***	-0.65*
	(0.19)	(0.18)	(0.22)	(0.24)	(0.13)	(0.10)	(0.19)	(0.29)
State Effects?	yes	yes	yes	yes	yes	yes	yes	yes
Time Effects?	yes	yes	yes	yes	yes	yes	yes	yes
Observations	612	612	612	612	612	612	612	612
R^2	0.698	0.781	0.608	0.206	0.323	0.462	0.262	0.322
Total Income (Positive Δ)	-0.90***	-0.84***	-0.77***	-0.76***	-0.25***	0.21***	-0.11	-0.47***
	(0.09)	(0.09)	(0.09)	(0.11)	(0.06)	(0.05)	(0.07)	(0.11)
Constant	-1.04***	-0.40**	-1.16***	-1.07***	-0.67***	0.31***	-0.67***	-1.09***
	(0.14)	(0.14)	(0.14)	(0.18)	(0.10)	(0.08)	(0.12)	(0.18)
State Effects?	yes	yes	yes	yes	yes	yes	yes	yes
Time Effects?	yes	yes	yes	yes	yes	yes	yes	yes
Observations	1088	1088	1088	1088	1088	1088	1088	1088
R^2	0.673	0.750	0.621	0.328	0.370	0.447	0.377	0.229
Positive Δ=Negative Δ test	10.84	6.05	4.72	5.06	5.42	15.62	0.77	15.22
p-value	0.0035	0.0102	0.0300	0.0246	0.0200	0.0001	0.3811	0.0001

Notes: Standardised coefficients. Weighted by population in 2010. Standard errors robust to heteroscedasticity and clustering on state in parentheses. p-values for the test for unobserved state, time and two-way effects, respectively: 0.0000 ; 0.0000; 0.0000 in all regressions.
Regressions controlled for the proportion of population between 15-19, 20-24, 25-29, 30-34 and 35-39 years old, Urbanization, Proportion of one-parent households and Unemployment rate.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$ *Data Source*: UCR/FBI/CPS.

Index