# Crime and Punishment Revisited

Kelaher, Richard and Sarafidis, Vasilis

The University of Sydney

11 January 2011

# Crime and Punishment Revisited*

Richard Kelaher†  Vasilis Sarafidis‡

University of Sydney  University of Sydney

This version: January 2011

**Abstract**

Despite an abundance of empirical evidence on crime spanning over forty years, there exists no consensus on the impact of the criminal justice system on crime activity. We argue that this may be due to the combined effect of simultaneity, omitted variable bias and aggregation bias that may confound many of these studies. We construct a new panel data set of local government areas in Australia and develop a testing framework for the implications of economic theory on crime behaviour. The empirical results suggest that the criminal justice system can potentially exert a much greater influence on crime activity than is the common view in the literature. In addition, we find that increasing the risk of apprehension and conviction is more influential in reducing crime than raising the expected severity of punishment. Violent crime is more persistent and relatively less responsive to law enforcement policies compared to non-violent crime.

Key words: Crime, deterrence, simultaneity, omitted variable bias, aggregation bias, panel data, GMM.

*We are grateful to Eddie Anderson for useful comments and suggestions. The usual disclaimer applies.

†Faculty of Economics and Business, University of Sydney, NSW 2006, Australia. e-mail: richard.kelaher@sydney.edu.au.

‡Corresponding author. Faculty of Economics and Business, University of Sydney, NSW 2006, Australia. e-mail: vasilis.sarafidis@sydney.edu.au.

# 1 Introduction

Crime, originating from the root of Latin cernō ("I decide, I give judgment"), is the behavior judged by the State to be in violation of the prevailing norms that underpin the moral code of society. Where informal social controls are not sufficient to deter such behavior, the State may intervene to punish or reform those responsible through the criminal justice system. The precise sanctions imposed depend on the type of crime and the prevailing cultural norms of the society. For offences deemed to be serious, criminal justice systems have historically imprisoned those responsible, in the hope that a combination of deterrence and incapacitation may lower the crime rate. Currently, more than 9.8 million people in the world are institutionalised for punishment, almost half of which are held in America, China and the U.K. (Walmsley, 2009). Over the past thirty years, the American prison population has more than quadrupled. Such massive increases in the U.S. prison population may be explained almost entirely by an increase in punitiveness rather than an increase in crime rates (see e.g. Raphael and Stoll, 2009), leading some to label this extraordinary measure one of the largest scale policy experiments of the century.

Other countries such as the U.K. and Australia have also experienced rising prison populations. For instance, the incarceration rate in NSW, which is the most populous state in Australia, has increased over 23 percent in the last 10 years and is currently higher than that of Germany. The NSW prison system now costs taxpayers more than $1 billion per year.[1] At the same time crime rates have remained relatively stable, leading some to declare such high rates of incarceration a policy failure.

How effective is the criminal justice system in deterring crime? To what extent do changes in the expected punishment influence the motivation of individuals to engage in illegal pursuits? How much wrong-doing does each additional prisoner avert?

[1]Steering Committee for the Review of Government Service Provision (2010), Report on Covernment Services 2010, Chapter 8, Productivity Commission, Canberra.

In order to address these questions in a constructive way it is important to recognise that changes in the aggregate crime rate stem from individual behavior. Policies such as increased sentence lengths may lower the crime rate through two possible channels; deterrence and incapacitation. It is well accepted in the literature that for a particular policy to be effective it cannot operate on incapacitation effects alone (Durlauf and Nagin, 2010). In turn, for a policy to deter criminal behavior it must be designed with an understanding of what causes individuals to engage in criminal activity.

During the early part of the twentieth century criminal behavior was viewed as a type of social illness. For example, the strain theory of Merton (1938) suggests that crime is a behavioral response to social inequality. The seminal work of Becker (1968) changed this view, postulating that individuals engage in such activity simply because the subjective expected benefit exceeds the expected cost of doing so. Criminals, therefore, do not differ from the rest of society in their basic motivation but in their appraisal of benefits and costs. On this view a rational criminal behaves in a calculated manner, considering the benefit of the illegal act together with the risk of apprehension and conviction as well as the likelihood and severity of potential punishment, which are a function of three separate stages of processing through the criminal justice system pertaining to the roles of police, courts and prison system respectively. The idea of a rational criminal represents a major step forward in criminology and forges an important link with the deterrence hypothesis that underpins the criminal justice system − the notion that the crime rate can be reduced by raising the expected cost of criminal activity.

Since the seminal work of Becker, a large empirical literature has developed, seeking to inform public policy by collecting data on various populations and building econometric models that describe criminal behavior of individuals. The public concern about crime is well justified given the pernicious effects that it has on economic activity, as well as on the quality of one's life in terms of a reduced sense of personal and proprietary security. However, despite the rich history of econometric modeling spanning over forty years, it appears there is largely a disconnect between theory and evidence as we have not been able to identify a single study that examines statistically the implications of economic theory on crime behavior of individual agents. Furthermore, there is arguably no consensus on whether there is a strong deterrent effect of law enforcement policies on crime activity. For example, Hirsch (1988) argues:

> "Estimates of the magnitude of the deterrent effect vary... Further empirical investigation is necessary in order to gain a more accurate estimate of the magnitude of this deterrent effect coefficient, though the true value of the coefficient is probably closer to 1 than to 0.3." (Hirsch (1988) p. 271).

In contrast, Cornwell and Turmbull (1994) conclude:

> "The ability of the criminal justice system to deter crime is much weaker than previous results indicate... A fundamental flaw in each of the [previous] studies is an inability to control for unobserved heterogeneity in the unit of observation." (Cornwell and Turmbull, 1994, p. 361).

Recent studies also provide mixed evidence that are insufficient to draw clear conclusions (see Section 3). The present paper revisits the economics of crime, deterrence and punishment and provides new findings. In particular, we specify a full econometric model of crime and develop a testing framework for the implications of economic theory on crime activity. The resulting restrictions appear to be supported by the data. Our results show a much stronger effect of the criminal justice system compared to the common view in the literature.

In addition, we find that increasing the risk of apprehension and conviction exhibits a much larger effect in reducing crime compared to raising the expected severity of punishment. This may have significant policy implications. For example, if it were estimated that the cost of keeping a prisoner incarcerated for a year was roughly equivalent to the cost of making a single additional arrest, then one could justify a redirection of resources from prisons to policing.

The difference between our results and those reported in a substantial body of literature may be attributed to several reasons. First, omitted variable bias; it is rarely the case that empirical studies of crime specify a complete econometric model with all deterrence variables included. Using the available data we show that the parameter estimates of the economic model of crime can be very sensitive to model mis-specification and the exclusion of relevant deterrence variables, which can lead practically to under-estimating the true effect of the criminal justice system.

Second, aggregation bias; while the economic model of crime purports to represent individual behavior, most data involve some form of aggregation − often, measurement takes place at the country or state level. This is likely to yield results that are inconsistent with economic theory. For example, Levitt (2001) argues that relying on national time series data can be particularly problematic since averaging across all of the locales removes useful variation, which may potentially result in misleading inferences. In the present study we are able to achieve a relatively low level of aggregation since the unit of observation is the Local Government Area (LGA) level in NSW. In addition, as it is shown, economic theory bears direct implications on the econometric specification of the model of crime behavior, which are testable using data constructed at the aggregate level. We find that these restrictions are not rejected in our case, which indicates that the level of aggregation used in this study is not harmful to modeling individual behavior.

Third, often the identification strategy employed in the literature is rather problematic since the deterrence variables are treated as exogenous, or if otherwise, the instruments used may not be orthogonal to the error term as they are likely to be correlated with the deterrence variables omitted from the regression. This study makes use of panel data analysis, which provides natural instruments with respect to sufficiently lagged values of the endogenous regressors. Furthermore, panel data analysis allows capturing different sources of unobserved heterogeneity, which makes the orthogonality conditions more likely to be satisfied. The validity of the instruments used is testable based on standard methods.

Finally, we deviate from the majority of the literature by specifying a dynamic model, which captures the essential feature of the behavior of individuals towards crime in that in practice it takes time to adjust fully to changes in law enforcement policies due to habit formation and costs of adjustment. This is important because it permits distinguishing between the effect of law enforcement policies in the short- and the long-run, and deriving equilibrium conditions as well as other meaningful dynamic quantities such as mean and median lag length of the effects.

We note that our estimated results are specific to the type of crime considered in the analysis. In particular, violent crime appears to be much more persistent and relatively less responsive to changes in law enforcement policies compared to non-violent crime. Furthermore, while the restrictions implied by economic theory are found to be supported by the data for non-violent crime, this is not the case for violent crime. These results can be considered natural given that Becker's idea of a criminal behaving in a rational manner is less likely to apply to violent crime, which is often influenced by feelings of anger and jealousy rather than rational behavior.

The remainder of this paper is as follows. Section 2 reviews the economic model of crime and motivates the key deterrence variables to be used in our empirical analysis. Section 3 discusses the empirical evidence pertaining to crime deterrence and analyses the problems inherent in using aggregate data to estimate the economic model of crime. Section 4 presents the econometric specification employed in the paper and its relationship with the underlying theory. Section 5 discusses the empirical results. A final section concludes.

## 2 A review of the economic model of crime

This section reviews the economic model of crime in order to motivate the theoretical relationship between crime and deterrence and analyse some of its implications. The framework extends Becker's (1968) representation of criminal behavior as a choice based on maximisation of expected utility. Consider an individual $i$ who engages in an illegal act based on a comparison of the expected benefits and the expected costs of doing so. The expected cost of criminal activity comprises both the direct inputs of the criminal justice system (the likelihood and severity of punishment), as well as the opportunity cost of activity in the legal sector foregone. Individual $i$ will therefore engage in criminal activities if

$$p_i U_i^{NL}(Y_i - (C_i + S_i)) + (1 - p_i) U_i^{NL}(Y_i) > U_i^L(I_i), \tag{1}$$

where $p_i$ is the unconditional probability of conviction[2], $Y_i$ is the "income" flowing from the criminal act, material or otherwise, $C_i$ is the collateral costs of criminal charges, $S_i$ is the cost to the individual of the sanction imposed as punishment and $I_i$ is the income from legal activity. The $C_i$ are costs that are incurred upon being charged with a crime but not necessarily punished; for example, social stigmatisation and diminished employment prospects. The $U_i$ are utility functions representing the way in which individual $i$ subjectively values benefits and costs associated with legal ($U_i^L$) and illegal ($U_i^{NL}$) activities respectively. Rearranging, we obtain the equivalent condition

$$p_i \left( U_i^{NL}(Y_i - (C_i + S_i)) - U_i^L(I_i) \right) + (1 - p_i) \left( U_i^{NL}(Y_i) - U_i^L(I_i) \right) > 0. \tag{2}$$

The term on the left-hand side of the inequality is the expected net utility flowing from criminal activity. With probability $(1 - p_i)$ the individual realises the full benefit of criminal activity over and above the opportunity cost of legal activity. With probability $p_i$ the benefit of criminal activity is deflated by $C_i + S_i$. If the expected utility of criminal activity net of the opportunity cost imposed by the legal sector is positive, a rational individual $i$ will engage in criminal behavior. Therefore, under the assumption of rational, utility maximising agents, individual $i$ will engage in criminal behavior if and only if (2) holds. Formally the criminal

[2]It is assumed that an individual is punished if they are found guilty.

decision is

$$\psi_i = \begin{cases} 1 \text{ if } p_i \left(U_i^{NL}(Y_i - (C_i + S_i)) - U_i^L(I_i)\right) + (1 - p_i)\left(U_i^{NL}(Y_i) - U_i^L(I_i)\right) > 0 \\ 0 \text{ otherwise,} \end{cases}$$

where $\psi_i$ is an indicator function taking the value of 1 if individual $i$ chooses to commit a crime and 0 otherwise.

Following Ehrlich (1975) the probability of punishment is decomposed into its three component parts: the probability of arrest, the probability of conviction given arrest, and the probability of imprisonment given conviction. Such an extension of the Becker model more realistically represents the risk posed by the criminal justice system. The transition from committing a criminal act to the realisation of punishment involves multiple stages of processing through the criminal justice system, none of which is certain. In order for punishment to occur, an individual must first be caught and arrested, then be found guilty by a judiciary. A further source of uncertainty follows since a judge must decide both the specific sanction imposed as punishment (eg. imprisonment, fine, home detention) and its severity. Expected utility from criminal activity is therefore represented as a function of the probability of arrest ($P_A$), the probability of conviction given arrest ($P_{C|A}$), the probability of imprisonment conditional on conviction ($P_{P|C}$) and the expected prison sentence length ($S$). These variables are the standard deterrence variables that appear in the literature and are the focus of our analysis. For ease of exposition, the opportunity cost flowing from the legal sector is set to zero. Hence, the expected utility from criminal activity can be written as

$$\begin{aligned} E(U_i^{NL}) &= (1 - P_{A_i})U_i^{NL}(Y_i) + P_{A_i}(1 - P_{C|A_i})U_i^{NL}(Y_i - C_i) \\ &\quad + P_{A_i}P_{C|A_i}P_{P|C_i}U_i^{NL}(Y_i - C_i - S_i) \\ &\quad + P_{A_i}P_{C|A_i}(1 - P_{P|C_i})U_i^{NL}(Y_i - C_i - S_i'), \end{aligned} \quad (3)$$

where the first term on the right hand side represents the full benefit of criminal activity in the case that one is not caught, which occurs with probability (1 −

$P_{A_i}$), the second term represents the benefit from criminal activity in the event that one is arrested but not convicted of the crime (deflated by $C_i$), occurring with probability $P_{A_i}(1 - P_{C|A_i})$,[3] while the third term represents the benefit from criminal activity in the event that one gets caught, convicted and therefore is punished (deflated by $C_i + S_i$), occurring with probability $P_{A_i}P_{C|A_i}P_{P|C_i}$. The fourth term captures all cases where the criminal is caught and found guilty (as with the previous term), but where an alternative to imprisonment is used. This occurs with probability $P_{A_i}P_{C|A_i}(1-P_{P|C_i})$, and the benefit from criminal activity is deflated by $C_i + S'_i$, where $S'_i$ is the cost to the individual of this alternative punishment. It is assumed that imprisonment is the most severe punishment possible for any given crime − that is, $S_i > S'_i \ \forall \ i$.

The theoretical model has a number of implications for individual behavior towards crime. These can be summarised by the following propositions.

**Proposition 1** *Increases in $P_{A_i}$, $P_{C|A_i}$ or $P_{P|C_i}$ decrease the expected utility derived from criminal activity.*

**Proof.** See Appendix A.1. ■

**Proposition 2** *The marginal deterrence effects of the criminal justice system are ordered such that the effect of $P_{A_i}$ is larger than that of $P_{C|A_i}$, which in turn is larger than the effect of $P_{P|C_i}$.*

**Proof.** See Appendix A.2. ■

Proposition 1 outlines that a potential criminal behaves in a calculated manner, taking into account the risk of apprehension and conviction as well as the likelihood and severity of punishment for a given level of benefit of the criminal act. Hence, the crime rate can be reduced by increasing the expected cost of criminal activity. The intuition of Proposition 2 lies in that the price of being arrested and convicted includes the cost incurred upon being charged but not necessarily punished, such as social stigmatisation and diminished employment opportunities.

[3] Although it is likely that the collateral costs of criminal charges are greater if the individual is actually convicted of crime, assuming that the full extent of these costs are incurred immediately upon arrest greatly simplifies the exposition of the analysis.

Thus, economic theory suggests that, ceteris paribus, policies targeting the probability of arrest and the probability of conviction can be more effective in deterring criminal activity than those targeting the probability of imprisonment, assuming all policies are equally costly. Despite these important implications, there appears to be a large disconnect between economic theory and empirical evidence in the literature for reasons that may be attributed to factors analysed in Section 3.

# 3 Evidence on Crime Deterrence

The nature of crime data available render the analysis of the effect of law enforcement policies on criminal activity inherently problematic. Criminological research logically began with the analysis of data collected from individuals. However such data are self-reported and are doubtlessly affected by significant measurement error. Moreover, the time and cost involved in surveying a representative population can be prohibitively large. An alternative is to use some form of aggregate data, which describe crime in locales (for example local areas, states or countries) and are based on official records rather than self-reported information.

However, empirical studies based on aggregate data are also not without problems, leading some to suggest that the use of individual and aggregate data may be regarded as two complementary approaches (Trumbull, 1989). In particular, aggregate data may inherently introduce a form of bias by invoking the "representative agent" assumption, which implies that all individuals are homogeneous and thus they behave in a similar manner. Furthermore, the use of aggregate data introduces a problem of simultaneity that makes the causal effect of law enforcement policies on crime more difficult to identify. For example, an exogenous upward shift in crime rate may overwhelm police resources, given that police resources are fixed in the short term, causing the probability of arrest to decrease. Increases in crime may also cause overcrowding in courts, leading individuals to enter guilty pleas with the understanding that alternatives to imprisonment are more likely to be used (Nagin, 1978). To the extent that courts behave in this way in response to overcrowding, both the conditional probabilities of conviction and imprisonment are endogenous. Still another argument holds that an exogenous increase in the crime rate may cause courts to increase sentence lengths in an attempt to combat high rates of crime.

Despite the strong potential for simultaneity when the relation between crime, policing and justice is modeled using aggregate data, many studies fail to control for endogeneity, which casts serious doubt on their results (see e.g. Blumstein, Cohen and Nagin, 1978). It is well known that in the presence of endogeneity, least-squares based estimates of the economic model of crime are contaminated by the reverse effect that crime may exhibit on law enforcement policies, and hence are biased and inconsistent. Dills, Miron and Summers (2008) use aggregate data to demonstrate that raw correlations between crime rates and deterrence variables are frequently weak or even perverse due to the problem of simultaneity, and note that any identification strategy would need to be powerful enough to partial out the effect of deterrence on the crime rate and provide a result consistent with economic theory.

A further problem that may arise in empirical studies that use aggregate data is the potential for omitted variable bias in the estimated parameters. In particular, it is hardly ever the case that a complete model is specified that includes all deterrence variables prescribed by economic theory. This is likely to be due to lack of data or the fact that certain experimental designs intended to combat endogeneity preclude the possibility of examining all deterrence variables of interest. Whatever the appropriate explanation is, the evidence on crime deterrence has come to conform broadly to several distinct sub-literatures, in which the effect of the probability of arrest, the probability of conviction, the probability of imprisonment and the length of average sentence are rarely examined together.

Table 1 summarises the empirical results for some of the most widely cited contributions to the crime deterrence literature using aggregate data. For each of the studies noted, the table reports the sampling population, the unit of observation, the structure of the data followed by the sample size[4], the method used to estimate the model, the type of crime analysed and finally the actual results. Clearly, there is a paucity of studies that estimate a fully specified economic model of crime, with notable exceptions being the papers by Pyle (1984), Trumbull (1989) and Cornwell and Trumbull (1994). However, in both Pyle (1984) and Trumbull (1989) all deterrence variables are treated as exogenous and therefore least-squares based methods are used to obtain estimates of the parameters. Trumbull justifies this choice claiming that simultaneity is not a salient feature of the existing dataset, based on the results of a Wu-Hausman specification test. Cornwell and Trumbull (1994) treat the probability of arrest as endogenous but all remaining variables as exogenous. The authors fail to find a statistically significant relationship between the deterrence variables and crime using a 2SLS procedure. Nevertheless, they produce inferences based on least-squares, arriving at a conclusion similar to Trumbull in that, as it is argued, the probability of arrest is exogenous.

However, even if reverse causality were not present in these data, the probability of arrest (when defined as number of arrests divided by the number of crime incidents) is endogenous in the crime equation by construction, since the numerator of the dependent variable (number of crime incidents) is the denominator in the probability of arrest, which artificially induces a negative correlation between the two variables (Nagin, 1978) − a phenomenon that is known as "ratio bias" in the literature (see e.g. Dills, Miron and Summers, 2008). Figure 1 illustrates this phenomenon using real time series data for NSW, aggregated across Local Government Areas (LGAs).

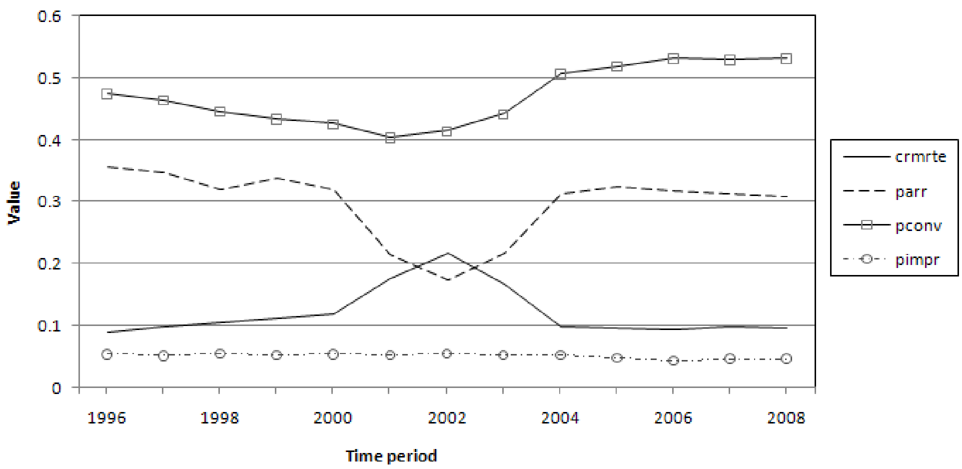
Finally, although Cornwell and Trumbull (1994) observe that the ordering of deterrence coefficients matches the ordering predicted by economic theory, they do not explicitly test the significance of this ordering.[5]

The remaining studies restrict their attention to a particular variable of in-

[4]For panel data models the cross-sectional dimension, $N$, is given first, followed by the time dimension, $T$.

[5]The reported coefficient estimates and standard errors in Cornwell and Trumbull (1994) suggest that such a test would fail to reject the null hypothesis of no difference between successive deterrent elasticities.

Figure 1: Crime and deterrence, New South Wales (1996 - 2008)




terest. Failing to include all deterrence variables fosters a disconnect between economic theory and empirical analysis. In order for a criminal to be punished, the person must be arrested and found guilty first; omitting the probability of arrest and conviction clearly ignores a fundamental aspect of the criminal decision and is likely to lead to biased inferences. Furthermore, omitted variables may invalidate estimation based on instrumental variables, as instruments may not be orthogonal to the deterrence variables omitted from the regression. For example, Mustard (2003) shows that arrest rates are likely to be negatively correlated with the probability of conviction and sentence length. As a result the author concludes that previous estimates of the marginal effect of the probability of arrest may understate the true effect of the arrest rate by as much as fifty percent. The following section analyses the econometric specification employed in this paper and discusses some of the implications of the economic model of crime.

Table 1: Empirical estimates of the elasticity of the crime rate with respect to policing and justice

| Author | Year | Population | Unit of observation | Data (sample size) | Method | Crime type | Arrest | Conviction | Imprisonment | Sentence length |
|---|---|---|---|---|---|---|---|---|---|---|
| Panel A: Complete economic models of crime | | | | | | | | | | |
| Cornwell and Trumbull | 1994 | North Carolina | County | Panel (90, 7) | OLS, 2SLS | Total | -.455 | -.336 | -.196 | -.03 |
| Trumbull | 1989 | North Carolina | County | Cross section (98) | OLS | Total | -.217 | -.451 | -.325 | -.149 |
| Pyle | 1984 | England and Wales | Police authorities | Cross section (41) | OLS | Robbery | -.5 | .73 | -.48 | -.57 |
| | | | | | | Property | -.32 | .4 | -.55 | -.85 |
| Panel B: Arrest | | | | | | | | | | |
| Klick and Tabarrok | 2005 | Washington D.C. | Police district | Panel (7,506) | OLS | Violent | -.3 | | | |
| | | | | | | Burglary | -.3 | | | |
| Fajnzylber et al. | 2002 | United Nations | Country | Panel (45, 5) | GMM | Robbery | .08 | | | .035 |
| | | | | | | Homicide | -.09 | | | -.346 |
| Corman and Mocan | 2000 | New York | City | Time series (108) | OLS | Murder | -.336 | | | |
| | | | | | | Burglary | -.355 | | | |
| Bodman and Maultby | 1997 | Australia | State | Cross section (60) | 2SLS | Robbery | -.258 | | | -.621 |
| | | | | | | Burglary | -.367 | | | |
| Levitt | 1997 | United States | City | Panel (59, 23) | 2SLS | Violent | -0.9* | | | |
| | | | | | | Property | -0.24* | | | |
| Marvell and Moody | 1996 | United States | City | Panel (56, 22) | Granger | Total | -.133 | | | |
| | | | | | | Homicide | -.241 | | | |
| | | | | | | Burglary | -.151 | | | |
| Sampson and Cohen | 1988 | United States | City | Cross section (171) | 2SLS | Robbery | -.28 | | | |
| | | | | | | Burglary | -.12 | | | |
| Car-Hill and Stern | 1973 | England and Wales | Police districts | Cross section (64) | FIML | Total | -.59 | | | -.17 |
| Panel C: Imprisonment | | | | | | | | | | |
| Johnson and Raphael | 2006 | United States | State | Panel (51, 27) | 2SLS | Violent | | | -.11 | |
| | | | | | | Property | | | -.21 | |
| Liedka et al. | 2006 | United States | State | Panel (51, 29) | Granger | Total | | | -.245 | |
| | | | | | | Murder | | | -.13 | |
| | | | | | | Burglary | | | -.136 | |
| Levitt | 2002 | United States | City | Panel (100, 21) | 2SLS | Violent | -.435 | | -.171 | |
| | | | | | | Property | -.501 | | -.305 | |
| Witt and Witte | 2000 | United States | Country | Time series (38) | VAR | Total | | | -.55 | |
| Levitt | 1996 | United States | State | Panel (51, 23) | 2SLS | Violent | | | -.261 | |
| | | | | | | Property | | | -.379 | |
| Marvell and Moody | 1994 | United States | State | Panel (49, 19) | Granger | Total | | | -.159 | |
| | | | | | | Homicide | | | -.065 | |
| | | | | | | Burglary | | | -.253 | |
| Ehrlich | 1973 | United States | State | Cross section (47) | 2SLS | Total | | | -.991 | -1.123 |
| Panel D: Conviction and other studies | | | | | | | | | | |
| Haas | 1980 | New Jersey | Municipality | Cross section (181) | 2SLS | Total | | -.02 | | |
| Withers | 1984 | Australia | State | Cross section (104) | OLS | Violent | | .29 | .09 | |
| | | | | | | Total | | -.62 | -.6 | |
| | | | | | | Property | | -.59 | -.56 | |
| Sjoquist | 1973 | United States | Municipality | Cross section (53) | OLS | Theft | -.342 | -.678 | | -.212 |

* indicates author provided multiple estimates, in which case the median is reported.

# 4 Econometric Specification

The starting point of our analysis is to relate the number of offences committed in a given LGA to the number of arrests, the number of convictions, the number of imprisonments and the average sentence length. Since the level of crime in each LGA depends on the size of the population, the latter is also included in the aggregate regression. It will be demonstrated that many of the models estimated in the literature are in fact a restricted version of this aggregate model. We remark that we deviate from the literature in a significant way in that we also allow for persistence in the level of crime due to habit formation and costs of adjustment, thus specifying a dynamic model of crime. In contrast, common practice in the literature presumes a static relation, where the entire effect of law enforcement policies is assumed to be realised immediately within the same time period.

Our aggregate model is given by

$$\begin{aligned}
\ln crm_{it} \quad = \quad & \delta \ln crm_{it-1} + \delta_1 \ln arr_{it} + \delta_2 \ln conv_{it} + \delta_3 \ln impr_{it} + \\
& \delta_4 \ln avsen_{it} + \delta_5 \ln income + \delta_6 \ln unemp + \delta_7 \ln pop_{it} + \\
& \delta_8 \ln pop_{it-1} + u_{it}, \qquad (4)
\end{aligned}$$

where $crm_{it}$ denotes the number of crime offences in LGA $i$ at time $t$, and the remaining variables are self-explanatory. Due to incapacitation effects alone, one naturally expects that $\delta_j < 0$ for $j \leq 4$, while $\delta \in (0, 1)$ to ensure stationarity.

The economic model of crime postulates that criminals are rational individuals who assess the risk of apprehension and conviction as well as the likelihood of punishment prior to committing an offence, and ultimately evaluate the expected benefit and cost associated with an illegal activity. Therefore, our hypothesis is that the model can be expressed in the following form:

$$\begin{aligned}
\ln\left(\frac{crm_{it}}{pop_{it}}\right) &= \alpha\ln\left(\frac{crm_{it-1}}{pop_{it-1}}\right)+\beta_1\ln\left(\frac{arr_{it}}{crm_{it}}\right)+\beta_2\ln\left(\frac{conv_{it}}{arr_{it}}\right) \\
&\quad +\beta_3\ln\left(\frac{impr_{it}}{conv_{it}}\right)+\beta_4\ln avsen_{it}+\beta_5\ln income_{it} \\
&\quad +\beta_6\ln unemp_{it}+\upsilon_{it}, \quad 0<\alpha<1, \qquad (5)
\end{aligned}$$

which can be rewritten as

$$\begin{aligned}
\ln crmr_{it} &= \alpha\ln crmr_{it-1}+\beta_1\ln prbarr_{it}+\beta_2\ln prbconv_{it}+\beta_3\ln prbimpr_{it} \\
&\quad +\beta_4\ln avsen_{it}+\beta_5\ln income_{it}+\beta_6\ln unemp_{it}+\upsilon_{it}. \qquad (6)
\end{aligned}$$

Precise definitions of all variables used in our regression modelling are provided in Table 2. Therefore, crime rate is a function of the empirical probability of arrest, $prbarr$, the probability of conviction given arrest, $prbconv$, and the probability of imprisonment given conviction, $prbimpr$, albeit with a dynamic effect of the deterrence variables on the crime rate, for which the speed of adjustment is determined by the coefficient of the lagged value of the dependent variable. The inclusion of sentence length, income and unemployment in the equation captures the expected gains from the illegal and legal sectors.

In this case, we have

$$\delta=\frac{\alpha}{1+\beta_1},\ \delta_1=\frac{\beta_1-\beta_2}{1+\beta_1},\ \delta_2=\frac{\beta_2-\beta_3}{1+\beta_1},\ \delta_7=\frac{1}{1+\beta_1},\ \delta_8=-\frac{\alpha}{1+\beta_1}, \qquad (7)$$

while the remaining coefficients can be reparameterised conveniently such that $\delta_j=\beta_j/(1+\beta_1)$ for $j=3,...,6$. Therefore, the null hypothesis can be formulated as follows:

$$H_0:\ \delta+\delta_1+\delta_2+\delta_3+\delta_7+\delta_8=1. \qquad (8)$$

This set of restrictions is testable in the general model in a standard way. Acknowledging these restrictions leads to an interesting point of symmetry between economic theory and the econometric specification, which is outlined below.

**Proposition 3** *The marginal deterrence effects are ordered, such that the coefficient of prbarr in* (6) *is more negative than the coefficient of prbconv, which in turn is more negative than the coefficient of prbimpr – that is,* $|\beta_1| > |\beta_2| > |\beta_3|$.

**Proof.** The restrictions in (7) imply that $\beta_j = (1 + \beta_1)\, \delta_j + \beta_{j+1}$ for $j = 1, 2$, and also $-1 < \beta_1 < 0$ because $\alpha,\ \delta \in (0, 1)$. Since $\delta_j < 0$, $\beta_{j+1} < 0$ for $j = 1, 2$, the result follows directly. ■

Thus, Proposition 3 implies the same ordering predicted by economic theory and therefore provides a further restriction that can be used to test the econometric specification.

Many of the models used in the literature (see e.g. Table 1) are restricted versions of (4). For example, static models impose $\alpha = 0$. Omitting the probability of arrest and using the unconditional empirical probability of conviction instead, defined as the ratio between the number of convictions over the number of crime offences, imposes $\delta_1 = 0$. This restriction seems unrealistic and is likely to lead to omitted variable bias if arrests are correlated with convictions and imprisonments. In general, models that use unconditional probabilities are subject to omitted variable bias since they impose invalid restrictions in the aggregate model. Obviously, models that omit conditional probabilities are also subject to the same problem. For instance, conviction rates are rarely studied in the literature and the focus is often on the risk of apprehension and the severity of punishment. Table 5 shows that the parameter estimates of the model of crime can be very sensitive to omitted variables and typically the effect is to underestimate the impact of the criminal justice system as a whole.

Notice that the restricted model induces by construction some form of endogeneity and therefore warrants estimation based on instrumental variables. This is outlined in the following remark.

**Remark 4** *Even in the absence of simultaneity, least-squares based estimation methods of the restricted model are subject to "ratio bias" that does not appear in the original model. This is because the denominator in prbarr is endogenous. Therefore,* $E\left(\varepsilon_{it} | \ln prbarr_{it}\right) \neq 0$. *This ratio bias does not necessarily arise in*

Table 2: Definitions of variables included in the econometric model of crime

| Variable | Definition |
|---|---|
| *crm* | Number of criminal incidents divided by total population |
| *prbarr* | Number of arrests divided by criminal incidents |
| *prbconv* | Number of convictions divided by arrests |
| *prbimpr* | Number of imprisonments divided by convictions |
| *avsen* | Average non-parole period (months) imposed for prison sentences |
| *income* | Average wage and salary earner income |
| *unemp* | Unemployment rate (%) |

*other econometric models that involve ratios. For example, consider a cost function,* $\ln cost_{it} = \beta_1 \ln price_{k,it} + \beta_2 \ln price_{l,it} + \beta_3 \ln price_{e,it} + \gamma \ln output_{it} + u_{it}$, *where* $price_k$, $price_l$ *and* $price_e$ *denote the price of capital, labour and energy, respectively, and the remaining variables are self-explanatory. Under linear homogeneity in input prices,* $\beta_1 + \beta_2 + \beta_3 = 1$, *which implies that the model can be written as* $\ln\left(\frac{cost_{it}}{price_{e,it}}\right) = \beta_1 \ln \frac{price_{k,it}}{price_{e,it}} + \beta_2 \ln\left(\frac{price_l}{price_{e,it}}\right) + \gamma \ln output_{it} + u_{it}$. *Assuming that firms are price takers (competitive markets),* $price_e$ *is strongly exogenous and therefore the transformed regressors remain strongly exogenous; for example,* $E\left(u_{it} | \ln \frac{price_{k,it}}{price_{e,it}}\right) = 0$.

# 5 Data Analysis, Estimation and Results

## 5.1 Data

We construct a new dataset containing information on criminal activity and deterrence for all 153 local government areas in New South Wales, each one observed over a period of thirteen years from 1995/96 to 2007/08. The Australian Standard Geographic Classification (ASGC) defines the LGA as the lowest level of aggregation following the census Collection District (CD) and Statistical Local Area (SLA).[6] Thus, the LGA represents a low level of aggregation compared to standard practice in the literature, where regressions using city-, state- and country-level data are common. To the best of our knowledge, this is the first panel data

[6] Each CD contains on average about 225 households (2001 Census). There are about 37,000 CDs throughout Australia. The boundaries of an SLA are designed to be typically coterminous with Local Government Areas unless the LGA does not fit entirely into a Statistical Subdivision, or is not of a comparative nature to other LGA's. There are 193 SLAs in NSW.

model of crime that has been constructed for Australia. The raw data for crime offences and deterrence variables have been purchased from the NSW Bureau of Crime Statistics and Research. Income and population data have been obtained from the Australian Bureau of Statistics (ABS) website, while the unemployment data have been purchased from the Small Area Labour Markets division of the Department of Education, Employment and Workplace Relations (DEEWR).

The NSW Bureau of Crime Statistics and Research provides two alternative definitions for average prison sentence; average non-parole period and average head sentence. We use the non-parole period in the analysis because this represents more closely the actual amount of time spent in confinement. The raw data for income and population are not readily comparable with the crime data because they are based on different ASGC standards, i.e. LGA boundaries are defined slightly differently by the NSW Bureau and the ABS. To achieve consistency, we mapped the data to a common ASGC standard (2006) using a series of concordance tables. Similarly, the unemployment data were first mapped to the same ASGC standard (2006) to account for name and boundary changes that occured in the LGAs over the sample period. The resulting SLA data were then aggregated to the LGA level to be directly comparable to the other data.

Table 3 reports descriptive statistics for the different categories of crime considered in our analysis. As expected, the mean value of the rate of violent crime is smaller than that of non-violent crime and it exhibits a much smaller dispersion as well, which indicates that violent crime occurs less frequently and is more localised. The empirical probability of arrest and the probability of imprisonment are both higher on average for violent crime, although the opposite occurs for the probability of conviction, which is perhaps reflective of the fact that for violent crime, police are more likely to bring a prosecution when a case is weak and more jury trials. The mean value of average sentence length is much larger than the value in the 90th percentile, which shows that there is a relatively small number of very big sentences in the sample.

Table 3: Descriptive statistics

| Variable | Crime type | Mean | Standard deviation | 10th Percentile | 90th Percentile |
|---|---|---|---|---|---|
| Crime rate | | | | | |
| | Total | .133 | .088 | .064 | .218 |
| | Non-violent | .100 | .070 | .043 | .176 |
| | Violent | .034 | .024 | .016 | .049 |
| Probability of arrest | | | | | |
| | Total | .313 | .117 | .169 | .466 |
| | Non-violent | .308 | .124 | .156 | .471 |
| | Violent | .344 | .128 | .198 | .505 |
| Probability of conviction | | | | | |
| | Total | .489 | .144 | .325 | .673 |
| | Non-violent | .506 | .177 | .301 | .739 |
| | Violent | .340 | .140 | .200 | .500 |
| Probability of imprisonment | | | | | |
| | Total | .071 | .040 | .031 | .118 |
| | Non-violent | .071 | .046 | .031 | .119 |
| | Violent | .159 | .129 | .060 | .290 |
| Average sentence (days) | | | | | |
| | Total | 280.1 | 4767.9 | 5.7 | 15 |
| | Non-violent | 37.9 | 1013.6 | 4.5 | 11.6 |
| | Violent | 608.1 | 9672.3 | 2 | 25.6 |
| Income ($ '000) | — | 34.01 | 9.4 | 25.18 | 44.03 |
| Unemployment (%) | — | 7.07 | 5.1 | 3.05 | 12.36 |

Descriptive statistics computed for the variables used in regression analysis. $N = 153$ and $T = 13$, yielding a total of 1,989 observations.

## 5.2 Estimation method and results

We analyse total crime, non-violent and violent crime sequentially, based on the econometric model studied in the previous section. The composite error term is specified as follows:

$$\upsilon_{it} = \eta_i + \tau_t + \mu_{it} + \varepsilon_{it}.$$

Therefore, $\upsilon_{it}$ allows for unobserved regional-level effects that may be correlated with the regressors, $\eta_i$, such as geographical location and crime reporting conventions, as well as time effects that capture common variations in crime across regions, $\tau_t$. $\mu_{it}$ reflects a serially uncorrelated measurement error and $\varepsilon_{it}$ is the usual random error component.[7] The results are obtained using the Generalised Method of Moments (GMM) estimator developed originally by Hansen (1982) and extended for dynamic panel data models by Arellano and Bond (1991), Ahn and

[7]Some measurement error is likely to be present – especially in measuring average sentence length, as this does not control for the criminal history of the offender, or the type of the offence.

Schmidt (1995), Arellano and Bover (1995) and Blundell and Bond (1998), among others. GMM is a natural choice when multiple explanatory variables are endogenous. Furthermore, the GMM approach has the advantage that it avoids full specification of the serial correlation and heteroskedasticity properties of the error, or indeed any other distributional assumptions. Our model specifies all explanatory variables as endogenous. The underlying reason for such treatment of the deterrence variables has already been motivated in Section 3. Errors in measurement may also contribute to endogeneity of the regressors. In addition, we treat average income as endogenous since crime has a direct effect on economic activity and thereby on employment. Similar considerations apply to the unemployment rate.

Table 4 shows the results for the model of total crime. For comparison purposes we also report results based on the within-group (WG), or fixed effects, estimator, which although frequently used is inconsistent under endogeneity. The long-run estimates are computed by dividing the short-run slope coefficients by one minus the estimated autoregressive parameter. Robust standard errors are reported in parentheses, which are valid under arbitrary forms of heteroskedasticity and serial correlation. Furthermore, for GMM specifically we perform the correction proposed by Windmeijer (2005) for the finite-sample bias of the standard errors of the two-step GMM estimator.[8] The standard errors of the long run estimated parameters are subsequently obtained using the formula for the approximation of the variance of a ratio of coefficients.[9] For GMM we also report the p-value of Hansen's test of overidentifying restrictions and the p-value of Arellano and Bond's (1991) test of serial correlation of the disturbances up to third order. The former is used to determine empirically the validity of the overidentifying restrictions in the GMM model. The null hypothesis is that the model is correctly specified. The latter is useful because it provides an indication of the appropriate lag length of the instruments to be used, since instruments are required to be orthogonal to

[8]All results have been obtained using David Roodman's *xtabond2* algorithm in Stata 11. The interested reader may refer to Roodman (2009).

[9]This is given by $Var(a/b) = (1/b^2)Var(a) + (a^2/b^4)Var(b) - 2(a/b^3)Cov(a,b)$, where $a = \beta_j, j = 1-6$ and $b = (1-\alpha)$.

the error term. On the bottom of the table, we also report the p-value of the test statistic of the joint null hypothesis formulated in (8), as well as the p-value of the hypothesis that the marginal deterrence effects are ordered sequentially. Together, these hypotheses provide a testing framework for the implications of economic theory on crime behavior.

Table 4: Estimated marginal elasticities for total crime

| | WG | | GMM | |
|---|---|---|---|---|
| | Short-run | Long-run | Short-run | Long-run |
| Lagged crime rate | .334*** | — | .350*** | — |
| | (.023) | | (.082) | |
| Probability of arrest | -.551*** | -.827*** | -.865*** | -1.33*** |
| | (.035) | (.074) | (.095) | (.249) |
| Probability of conviction | -.540*** | -.811*** | -.575*** | -.885*** |
| | (.036) | (.068) | (.093) | (.178) |
| Probability of imprisonment | -.039*** | -.059*** | -.218** | -.335** |
| | (.011) | (.017) | (.123) | (.196) |
| Average sentence | .004 | .005 | -.251*** | -.386** |
| | (.003) | (.004) | (.107) | (.176) |
| Income | -.079 | -.119 | -1.03** | -1.584** |
| | (.100) | (.150) | (.452) | (.833) |
| Unemployment | .028** | .042** | .626*** | .962*** |
| | (.013) | (.020) | (.162) | (.283) |
| p-value overidentifying restrictions | — | | .674 | |
| p-value serial correlation | | | | |
| - Lag 1 | — | | .017 | |
| - Lag 2 | — | | .199 | |
| - Lag 3 | — | | .131 | |
| p-value of $H_0$ in (8) | .033 | | .814 | |
| p-value of ordered effects | | | | |
| - $H_0 : \beta_A - \beta_{C\|A} = 0$ | .732 | | .008 | |
| - $H_0 : \beta_{C\|A} - \beta_{P\|C} = 0$ | .000 | | .034 | |

Robust and bias-corrected standard errors reported in parentheses. Each regression includes LGA-specific effects and time-specific effects. $NT = 153 \times 12 = 1,836$ observations. * indicates significance at the 10 percent level; ** and *** indicate significance at the 5 and 1 percent levels respectively, using one-tail tests.

Clearly there is substantial difference between the WG and GMM estimates; in particular, the former appears to significantly underestimate the effect of all explanatory variables, which demonstrates the importance of accounting for endogeneity in crime activity. The GMM estimates of the parameters are statistically significant and of the expected sign in the short- and the long-run. Thus, one percent increase in the probability of arrest appears to decrease the expected value of the crime rate by .865 percent in the short-run and 1.33 percent in the long-run, ceteris paribus. Likewise, the elasticity of the probability of conviction is about

-.575 and -.885 in the short- and long-run respectively. The fact that the estimated elasticities are larger in the long-run is well anticipated, since typically one needs time to adjust fully to changes in law enforcement policies, due to habitual behavior, imperfect knowledge and uncertainty. In particular, the value of the autoregressive parameter indicates that it takes about 2.5 time periods for ninety percent of the total impact of either one of the explanatory variables on crime to be realised, all else being constant.

The estimated coefficients of the probability of imprisonment and the average sentence length are not statistically different. This is consistent with the result we would expect if criminals responded to the expected length of sentence as a single factor. In particular, define the expected length of sentence as $e = prbimpr \times avsen$. Taking logs yields $\ln e = \ln prbimpr + \ln avsen$. This implies that the likelihood of imprisonment and the severity of punishment bear equal importance in reducing crime. If this is true then policies targeting prison sentence length as opposed to imprisonment probability do not differ in their effectiveness. Furthermore, they both appear to exhibit a much smaller effect on crime compared to the probability of arrest and the probability of conviction. This shows that imprisoning more criminals, or imprisoning them for longer, is not as effective as increasing the risk of apprehension or conviction once captured.

The above provides support to the idea that the consequences of being arrested and found guilty of a criminal offence include the indirect sanctions imposed by society and not just the punishment meted out by the criminal justice system. A convicted individual may no longer enjoy the same opportunities in the labour market or the same treatment by their peers, and so the opportunity cost of lost income and the cost to the individual of social stigmatisation is implied in the event of conviction. For example, Zimring and Hawkins (1973, pg. 174) argue:

> "Official actions can set off societal reactions that may provide potential offenders with more reason to avoid conviction than the officially imposed unpleasantness of punishment" (Zimring and Hawkins, 1973).

The results suggest that the lost social standing resulting from a conviction may well outweigh the effects of prison sentence, let alone a fine or community service order.

The effect of income and unemployment appears to be large and statistically significant. However, these coefficients are substantially different, suggesting that in formulating the crime-no crime decision individuals may consider more the level than the certainty of income flowing from the legal sector, which is consistent with Becker's conception of criminals as risk-seeking. Furthermore, crime appears to be unit-elastic in the short-run with respect to income but elastic in the long-run. On the other hand, crime is inelastic with respect to changes in unemployment.

The joint hypothesis formulated in (8) is not rejected when the test statistic is based on the GMM estimated parameters, while the hypothesis that the deterrence coefficients are statistically the same is rejected at the 5 percent level. This is in stark contrast to the WG-based tests, which yield the opposite results in both cases. Again, this demonstrates the importance of accounting for endogeneity in order to obtain findings consistent with economic theory.

Table 5 reports results with respect to a number of models that are subject to different sources of mis-specification error. For comparison purposes we also include the full model, estimated before and containing all variables prescribed by economic theory. Model (2) is similar to many of the models estimated in the literature, in that it specifies a static relation between crime and deterrence. Clearly the exclusion of the lagged dependent variable results in underestimating all coefficients without exception, which is not surprising given that crime and therefore its lagged value are negatively correlated with most of the regressors. Hansen's test statistic rejects the null hypothesis that the model is correctly specified at the five percent level, while the null of serially uncorrelated disturbances is rejected even for up to third order serial correlation. Model (3) omits the probability of arrest and makes use of the unconditional empirical probability of conviction instead, defined as number of convictions divided by crime offences. The results are similar in the sense that all estimated coefficients are smaller in absolute value compared to the full model. Furthermore, Hansen's test statistic rejects the null hypothesis that the instruments used are valid at the five percent level. This is expected given that omitted variables may invalidate the restriction that the instruments are orthogonal to the error term. Model (4) commits a common mis-specification error in that it includes only the probability of arrest and the average sentence length from the set of deterrence variables. Finally, Model (5) restricts attention to the likelihood and severity of punishment.[10] It is worth noting that the impact of omitted variables appears to be absorbed by the lagged dependent variable in all of the relevant mis-specified models, thus producing invalid inferences for the dynamic process of the crime equation. In summary, we can see that the estimated parameters of the total crime equation can be very sensitive to the specification of the model, and typically mis-specification errors result in obtaining estimates that show a smaller effect of the deterrence variables compared to the full model, which corroborates the results of Mustard (2003). This is not surprising because as shown in Table 6, most of the explanatory variables are mutually negatively correlated in the sample.

The following table reports estimates of the parameters by type of crime. As we can see, there are some stark differences between non-violent and violent crime. The former seems to resemble relatively more closely the model for total crime, which is expected given that about three quarters of all crime is non-violent. The null hypothesis of correct specification is not rejected in either of the two models. Likewise, the null hypothesis formulated in (8) is supported empirically in both cases. However, the hypothesis of sequential ordering of the coefficients of deterrence variables is supported only for non-violent crime. This indicates that the idea of rational behavior towards illegal activity may apply only to non-violent crime. This is also manifestated through the actual estimates of the model parameters. For example, the effect of punishment, both in terms of likelihood and severity, is statistically significant only for non-violent crime, and even so it remains small compared to the effect of the likelihood of arrest and conviction. Moreover, income and unemployment appear to have an appreciably smaller effect

[10]Similarly as before, the likelihood of imprisonment refers to the unconditional probability, defined as number of imprisonments divided by crime offences.

Table 5: Sensitivity of parameter estimates to omitted variables[†]

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Lagged crime rate | .350*** | | .517*** | .617*** | .577*** |
| | (.080) | | (.108) | (.101) | (.108) |
| Probability of arrest | -.865*** | -.652*** | | -.485*** | |
| | (.090) | (.223) | | (.150) | |
| Probability of conviction | -.575*** | -.380*** | -.357*** | | |
| | (.090) | (.149) | (.107) | | |
| Probability of imprisonment | -.218** | -.194* | .122** | | -.077 |
| | (.120) | (.122) | (.055) | | (.080) |
| Average sentence | -.251*** | .129 | .077* | -.071 | .048 |
| | (.110) | (.126) | (.056) | (.093) | (.078) |
| Income | -1.03** | -.991*** | -.389 | | |
| | (.450) | (.395) | (.308) | | |
| Unemployment | .626*** | .246* | .266** | .245*** | .325*** |
| | (.160) | (.150) | (.124) | (.098) | (.133) |
| p-value overidentifying restrictions | .002 | .032 | .011 | .004 | .009 |
| p-value serial correlation | | | | | |
| - Lag 1 | .017 | .151 | .000 | .000 | .000 |
| - Lag 2 | .199 | .072 | .687 | .792 | .458 |
| - Lag 3 | .131 | .035 | .007 | .150 | .003 |

Robust and bias-corrected standard errors reported in parentheses. Each regression includes LGA-specific effects and time effects, and is estimated by GMM using $NT = 153 \times 12 = 1,836$ observations. * indicates significance at the 10 percent level; ** and *** indicate significance at the 5 and 1 percent levels respectively.

Table 6: Correlation matrix for explanatory variables in total crime equation

| | ln prbarr | ln prbconv | ln prbimpr | ln avsen | ln income | ln unemp |
|---|---|---|---|---|---|---|
| ln prbarr | 1 | | | | | |
| ln prbconv | -.319 | 1 | | | | |
| ln prbimpr | .274 | -.212 | 1 | | | |
| ln avsen | -.174 | .111 | -.039 | 1 | | |
| ln income | -.542 | .340 | -.436 | .163 | 1 | |
| ln unemp | .282 | -.251 | .299 | -.021 | -.519 | 1 |

on violent crime.

Table 7: Elasticity estimates by type of crime

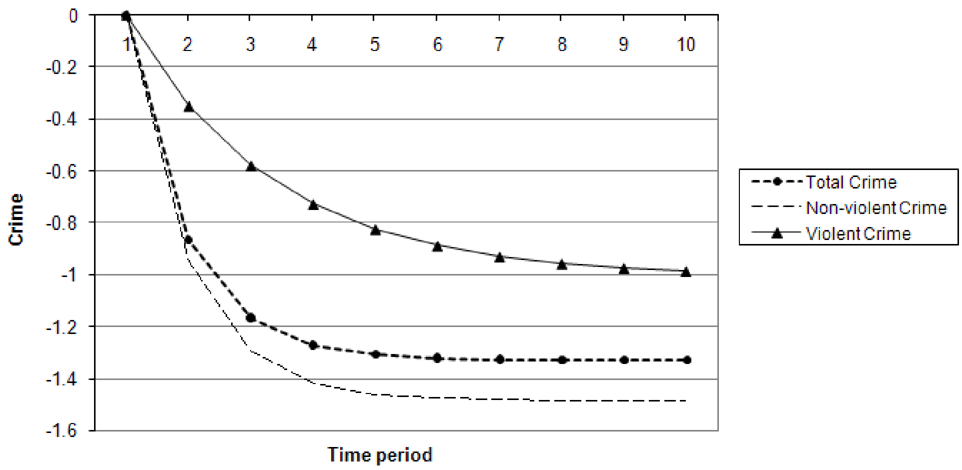
| | Violent crime | | Non-violent crime | |
|---|---|---|---|---|
| | Short-run | Long-run | Short-run | Long-run |
| Lagged crime rate | .642*** | — | .366*** | — |
| | (.125) | | (.087) | |
| Probability of arrest | -.258*** | -.720** | -.920*** | -1.45*** |
| | (.093) | (.367) | (.161) | (.383) |
| Probability of conviction | -.273*** | -.763*** | -.581** | -.916*** |
| | (.044) | (.258) | (.103) | (.219) |
| Probability of imprisonment | -.002 | -.005 | -.179** | -.282** |
| | (.054) | (.150) | (.102) | (.170) |
| Average sentence | .008 | .023 | -.210*** | -.331** |
| | (.032) | (.087) | (.107) | (.161) |
| Income | -.268 | -.748 | -1.115*** | -1.759** |
| | (.294) | (1.029) | (.436) | (.832) |
| Unemployment | .198*** | .554** | .305*** | .481*** |
| | (.065) | (.306) | (.111) | (.196) |
| p-value overidentifying restrictions | .514 | | .726 | |
| p-value serial correlation | | | | |
| - Lag 1 | .000 | | .005 | |
| - Lag 2 | .631 | | .316 | |
| - Lag 3 | .099 | | .303 | |
| p-value of $H_0$ in (8) | .881 | | .794 | |
| p-value of ordered effects | | | | |
| - $H_0 : \beta_A - \beta_{C\|A} = 0$ | .856 | | .012 | |
| - $H_0 : \beta_{C\|A} - \beta_{P\|C} = 0$ | .000 | | .000 | |

Robust and bias-corrected standard errors reported in parentheses. Each regression includes LGA-specific effects and time effects, and is estimated by GMM using $NT = 153 \times 12 = 1,836$ observations. * indicates significance at the 10 percent level; ** and *** indicate significance at the 5 and 1 percent levels respectively.

Finally, it is worth emphasising that violent crime is characterised by higher persistence, since the value of the autoregressive coefficient is almost double that of non-violent crime. Thus, while for non-violent crime it takes about 2.5 periods for ninety percent of the total (i.e. long-run) effect of either one of the deterrence variables to be realised, all other things being constant, violent crime requires about 6 periods for the same effect to occur. This is anticipated since violent crime may be attributed to factors which are fundamentally different and less prone to be calculated in the type of manner implied by economic theory. The following figure illustrates the dynamic path of the various types of crime following a one percent increase in the probability of arrest. One can see that the long-run estimated elasticity of the probability of arrest is smaller for violent crime and it also takes much longer to adjust to equilibrium compared to non-violent crime.

Similar results hold for the remaining explanatory variables, albeit they converge to a different equilibrium level.

Figure 2: Dynamic path of total crime: probability of arrest




## 6 Concluding Remarks

We estimate an econometric model for crime using a new panel of data set containing information on illegal activity and deterrence variables for local government areas in NSW. Our findings suggest that the criminal justice system can potentially exert a much larger impact on crime compared to previous estimates in the literature, particularly for non-violent crime. We show that the estimated parameters can be very sensitive to model mis-specification, the exclusion of relevant deterrence variables and the lack of a proper identification strategy under endogeneity. These factors typically tend to understate the effect of law enforcement policies, which may explain the difference between the results presented in this paper and those reported in a large body of empirical work.

From a policy design perspective, it appears that targeting the risk of apprehension and conviction are more effective strategies than increasing the severity of punishment, which indicates that the increasingly higher rates of incarceration observed across the world are not justified.

Moreover, violent crime appears to be more persistent and relatively less responsive to changes in law enforcement policies compared to non-violent crime. This result is natural given that Becker's idea of a criminal behaving in a rational manner is less likely to apply to violent crime, which is often influenced by feelings of anger and jealousy.

There are several interesting issues that remain to be explored. In particular, given our analysis it would be useful to measure the effectiveness of different police activities in influencing the risk of apprehension and determining the empirical probability of arrest following an offence. Furthermore, from an economic point of view it is inviting to examine the costs and benefits associated with crime prevention. We intend to pursue both of these issues in future research.

# References


Ahn, S. and Schmidt, P. (1995). Efficient estimation of models for dynamic panel data. *Journal of Econometrics*, 68(1):5–27.

Arellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58(2):277–297.

Arellano, M. and Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, 68(1):29–51.

Banks, G. (2010). Report on government services 2010. Technical report, Steering Committee for the Review of Government Service Provision, Productivity Commission, Canberra.

Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, 76(2):169–217.

Blumstein, A., Cohen, J., and Nagin, D. S., editors (1978). *Deterrence and Incapacitation - Estimating the Effect of Criminal Sanctions on Crime Rates*. National Academies Press, Washington, DC.

Blundell, R. and Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1):115–143.

Bodman, P. M. and Maultby, C. (1997). Crime, punishment and deterrence in australia: A further empirical investigation. *International Journal of Social Economics*, 24:884–901.

Car-Hill, R. and Stern, N. (1973). An econometric model of the supply and control of recorded offences in england and wales. *Journal of Public Economics*, 2(4):289–318.

Corman, H. and Mocan, H. N. (2000). A time-series analysis of crime, deterrence, and drug abuse in new york city. *The American Economic Review*, 90(3):584–604.

Cornwell, C. and Trumbull, W. N. (1994). Estimating the economic model of crime with panel data. *The Review of Economics and Statistics*, 76(2):360–366.

Dills, A. K., Miron, J. A., and Summers, G. (2008). What do economists know about crime? *National Bureau of Economic Research Working Paper Series*, No. 13759.

Durlauf, S. N. and Nagin, D. S. (2010). The deterrent effect of imprisonment. In *Controlling Crime: Strategies and Tradeoffs*, NBER Chapters. National Bureau of Economic Research, Inc.

Ehrlich, I. (1973). Participation in illegitimate activities: A theoretical and empirical investigation. *Journal of Political Economy*, 81(3):521–65.

Ehrlich, I. (1975). The deterrent effect of capital punishment: A question of life and death. *American Economic Review*, 65(3):397–417.

Fajnzylber, P., Lederman, D., and Loayza, N. (2002). What causes violent crime? *European Economic Review*, 46(7):1323–1357.

Haas, E. (1980). The supply and demand for municipal crime prevention. *American Economist*, 24(1).

Hansen, L. P. (1982). Large sample properties of generalized methods of moments estimators. *Econometrica*, 50(4):1029–1054.

Hirsch, W. Z. (1988). *Law and Economics.* Academic Press, San Diego, 2nd edition.

Johnson, R. and Raphael, S. (2006). How much crime reduction does the marginal prisoner buy? *Mimeo.*

Klick, J. and Tabarrok, A. (2005). Using terror alert levels to estimate the effect of police on crime. *Journal of Law and Economics*, 48(1):267–79.

Levitt, S. D. (1996). The effect of prison population size on crime rates: Evidence from prison overcrowding litigation. *The Quarterly Journal of Economics*, 111(2):319–351.

Levitt, S. D. (1997). Using electoral cycles in police hiring to estimate the effect of police on crime. *The American Economic Review*, 87(3):270–290.

Levitt, S. D. (2001). Alternative strategies for identifying the link between unemployment and crime. *Journal of Quantitative Criminology*, 17(4):377–390.

Levitt, S. D. (2002). Using electoral cycles in police hiring to estimate the effect of police on crime: Reply. *The American Economic Review*, 92(4):1244–1250.

Liedka, R. V., Piehl, A. M., and Useem, B. (2006). The crime-control effect of incarceration: Does scale matter? *Criminology and Public Policy*, 5(2):245–276.

Marvell, T. and Moody, C. (1994). Prison population growth and crime reduction. *Journal of Quantitative Criminology*, 10(2):109–140.

Marvell, T. B. and Moody, C. E. (1996). Specification problems, police levels, and crime rates. *Criminology*, 34(4):609–646.

Merton, R. K. (1938). Social structure and anomie. *American Sociological Review*, 3(5):672–682.

Mustard, D. B. (2003). Reexamining criminal behavior: The importance of omitted variable bias. *The Review of Economics and Statistics*, 85(1):205–211.

Nagin, D. S. (1978). General deterrence: A review of the empirical evidence. In Blumstein, A., Cohen, J., and Nagin, D. S., editors, *On the Feasibility of Identifying the Crime Function in a Simultaneous Model of Crime Rates and Sanction Levels*. National Academies Press, Washington, DC.

Pyle, D. (1984). Combating crime: Do the police give value for money? *Public Money*, 4(1):27 – 30.

Raphael, S. and Stoll, M. A. (2009). Why are so many americans in prison? In Raphael, S Stoll, M. A., editor, *Do Prisons Make Us Safer? The Benefits and Costs of the Prison Boom*. Russell Sage Foundation, New York.

Roodman, D. (2009). How to do xtabond2: An introduction to difference and system gmm in stata. *Stata Journal*, 9(1):86–136.

Sampson, R. J. and Cohen, J. (1988). Deterrent effects of the police on crime: A replication and theoretical extension. *Law and Society Review*, 22(1):163–189.

Sjoquist, D. L. (1973). Property crime and economic behavior: Some empirical results. *The American Economic Review*, 63(3):439–446.

Trumbull, W. N. (1989). Estimations of the economic model of crime using aggregate and individual level data. *Southern Economic Journal*, 56(2):423–439.

Walmsley, R. (2009). World prison population list. Technical report, King's College London.

Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step gmm estimators. *Journal of Econometrics*, 126(1):25–51.

Withers, G. (1984). Crime, punishment and deterrence in australia: An empirical investigation. *Economic Record*, 60(2):176–185.

Witt, R. and Witte, A. (2000). Crime, prison, and female labor supply. *Journal of Quantitative Criminology*, 16(1):69–85.

Zimring, F. E. and Hawkins, G. J. (1973). *Deterrence - The Legal Threat In Crime Control*. University of Chicago Press, Chicago.

# A Proof of propositions

## A.1 Proof of Proposition 1

Differentiating (3) with respect to each deterrence variable we obtain:

$$\begin{aligned}
\frac{\partial E(U_i^{NL})}{\partial P_{Ai}} &= -[U_i^{NL}(Y_i) - P_{C|Ai}[P_{P|Ci}U_i^{NL}(Y_i - C_i - S_i) + \\
&\quad (1 - P_{P|Ci})U_i^{NL}(Y_i - C_i - S_i')] - (1 - P_{C|Ai})U_i^{NL}(Y_i - C_i)] \quad (9) \\
\frac{\partial E(U_i^{NL})}{\partial P_{C|Ai}} &= -P_{Ai}[U_i^{NL}(Y_i - C_i) - P_{P|Ci}U_i^{NL}(Y_i - C_i - S_i) - \\
&\quad (1 - P_{P|Ci})U_i^{NL}(Y_i - C_i - S_i')] \quad (10) \\
\frac{\partial E(U_i^{NL})}{\partial P_{P|Ci}} &= -P_{Ai}P_{C|Ai}[U_i^{NL}(Y_i - C_i - S_i') - U_i^{NL}(Y_i - C_i - S_i)]. \quad (11)
\end{aligned}$$

Proposition 1 implies that each of (9), (10) and (11) must be negative, and so in each case the terms inside brackets must be positive. For the term in brackets in (9), $U_i^{NL}(Y_i)$ is deflated by a weighted average of two numbers smaller than itself, since the first number in the average is itself a weighted average of two numbers smaller than $U_i^{NL}(Y_i)$ and the second number, $U_i^{NL}(Y_i - C_i)$, is also smaller than $U_i^{NL}(Y_i)$. For the bracketed term in (10), $U_i^{NL}(Y_i - C_i)$ is deflated by a weighted average of two numbers smaller than itself (assuming $S_i, S_i' > 0$). Therefore, (9) and (10) are negative assuming that $U_i$ is an increasing function of the income of criminal activity. The bracket term in (11) is negative if the earlier assumption that prison is the most severe sanction for the particular crime ($S_i > S_i'$) is met.

## A.2 Proof of Proposition 2

The negative point elasticities of expected utility with respect to unit increases in each of the deterrence variables are obtained from (9), (10) and (11) to allow a comparison of the marginal disutility associated with a unit increase in each probability.

$$
\begin{aligned}
&-\frac{\partial E(U_i^{NL})}{\partial P_{Ai}}\frac{P_{Ai}}{E(U_i^{NL})} = \frac{1}{E(U_i^{NL})}P_{Ai}U_i^{NL}(Y_i) \\
&-\frac{1}{E(U_i^{NL})}\left\{P_{Ai}P_{C|Ai}\left[P_{P|Ci}U_i^{NL}(Y_i - C_i - S_i) + (1 - P_{P|Ci})U_i^{NL}(Y_i - C_i - S_i')\right]\right\} \\
&-\frac{1}{E(U_i^{NL})}\left[P_{Ai}(1 - P_{C|Ai})U_i^{NL}(Y_i - C_i)\right] \\
= \ &\frac{1}{E(U_i^{NL})}\left[P_{Ai}U_i^{NL}(Y_i) - P_{Ai}P_{C|Ai}P_{P|Ci}U_i^{NL}(Y_i - C_i - S_i)\right] \\
&-\frac{1}{E(U_i^{NL})}\left[P_{Ai}P_{C|Ai}(1 - P_{P|Ci})U_i^{NL}(Y_i - C_i - S_i') + P_{Ai}(1 - P_{C|Ai})U_i^{NL}(Y_i - C_i)\right].
\end{aligned}
\tag{12}
$$

$$
\begin{aligned}
&-\frac{\partial E(U_i^{NL})}{\partial P_{C|Ai}}\frac{P_{C|Ai}}{E(U_i^{NL})} = \frac{1}{E(U_i^{NL})}P_{C|Ai}P_{Ai}\left[U_i^{NL}(Y_i - C_i) - P_{P|Ci}U_i^{NL}(Y_i - C_i - S_i)\right] \\
&-\frac{1}{E(U_i^{NL})}P_{C|Ai}P_{Ai}(1 - P_{P|Ci})U_i^{NL}(Y_i - C_i - S_i') \\
= \ &\frac{1}{E(U_i^{NL})}\left[P_{Ai}P_{C|Ai}U_i^{NL}(Y_i - C_i) - P_{Ai}P_{C|Ai}P_{P|Ci}U_i^{NL}(Y_i - C_i - S_i)\right] \\
&-\frac{1}{E(U_i^{NL})}\left[P_{Ai}P_{C|Ai}(1 - P_{P|Ci})U_i^{NL}(Y_i - C_i - S_i')\right].
\end{aligned}
\tag{13}
$$

$$
\begin{aligned}
&-\frac{\partial E(U_i^{NL})}{\partial P_{P|Ci}}\frac{P_{P|Ci}}{E(U_i^{NL})} = \\
= \ &\frac{1}{E(U_i^{NL})}P_{Ai}P_{C|Ai}P_{P|Ci}\left[U_i^{NL}(Y_i - C_i - S_i') - U_i^{NL}(Y_i - C_i - S_i)\right] \\
= \ &\frac{1}{E(U_i^{NL})}\left[P_{Ai}P_{C|Ai}P_{P|Ci}U_i^{NL}(Y_i - C_i - S_i') - P_{Ai}P_{C|Ai}P_{P|Ci}U_i^{NL}(Y_i - C_i - S_i)\right].
\end{aligned}
\tag{14}
$$

Subtracting (13) from (12) yields

$$\begin{aligned}
& \frac{1}{E(U_i^{NL})}\{P_{Ai}U_i^{NL}(Y_i) - P_{Ai}(1-P_{C|Ai})U_i^{NL}(Y_i - C_i) - P_{Ai}P_{C|Ai}U_i^{NL}(Y_i - C_i)\} \\
= & \frac{P_{Ai}}{E(U_i^{NL})}\{U_i^{NL}(Y_i) - P_{C|Ai}U_i^{NL}(Y_i - C_i) - (1-P_{C|Ai})U_i^{NL}(Y_i - C_i)\} \\
= & \frac{P_{Ai}}{E(U_i^{NL})}\{U_i^{NL}(Y_i) - U_i^{NL}(Y_i - C_i)\} \\
> & \ 0,
\end{aligned}$$

if $C_i > 0$ and $U_i^{NL}$ is an increasing function of the benefit of criminal activity. Under these conditions, the marginal disutility associated with an increase in $P_{Ai}$ is greater than that associated with a corresponding increase in $P_{C|Ai}$. Subtracting (14) from (13) yields

$$\begin{aligned}
& \frac{1}{E(U_i^{NL})}\{EP_{Ai}P_{C|Ai}U_i^{NL}(Y_i - C_i) - P_{Ai}P_{C|Ai}(1-P_{P|Ci})U_i^{NL}(Y_i - C_i - S_i') - \\
& P_{Ai}P_{C|Ai}P_{P|Ci}U_i^{NL}(Y_i - C_i - S_i')\} \\
= & \frac{P_{Ai}P_{C|Ai}}{E(U_i^{NL})}\{U_i^{NL}(Y_i - C_i) - P_{P|Ci}U_i^{NL}(Y_i - C_i - S_i') - \\
& (1-P_{P|Ci})U_i^{NL}(Y_i - C_i - S_i')\} \\
= & \frac{P_{Ai}P_{C|Ai}}{E(U_i^{NL})}\{U_i^{NL}(Y_i - C_i) - U_i^{NL}(Y_i - C_i - S_i')\} \\
> & \ 0,
\end{aligned}$$

provided that $S_i' > 0$ and $U_i^{NL}$ is an increasing function of the benefit from criminal activity. It follows that the deterrent effects of the criminal justice system are ordered according to Proposition 2 if $C_i, S_i' > 0$ and utility increases with the benefit of criminal activity.