

Crime, deterrence and punishment revisited

Maurice J. G. Bun¹ · Richard Kelaher² · Vasilis Sarafidis³ · Don Weatherburn⁴

Received: 17 March 2016 / Accepted: 22 June 2019 / Published online: 10 September 2019 © The Author(s) 2019

Abstract

Despite an abundance of empirical evidence on crime spanning over 40 years, there exists no consensus on the impact of the criminal justice system on crime activity. We construct a new panel data set that contains all relevant variables prescribed by economic theory. Our identification strategy allows for a feedback relationship between crime and deterrence variables, and it controls for omitted variables and measurement error. We deviate from the majority of the literature in that we specify a dynamic model, which captures the essential feature of habit formation and persistence in aggregate behaviour. Our results show that the criminal justice system exerts a large influence on crime activity. Increasing the risk of apprehension and conviction is more influential in reducing crime than raising the expected severity of punishment.

Keywords Crime \cdot Deterrence \cdot Feedback \cdot Omitted variable bias \cdot Measurement error \cdot Panel data \cdot GMM

JEL Classification $C23 \cdot C36 \cdot K42$

Maurice J. G. Bun m.j.g.bun@uva.nl

☑ Vasilis Sarafidis vasilis.sarafidis@monash.edu

Extended author information available on the last page of the article

This article greatly benefited from comments made by the editor and two anonymous referees. We are grateful to Eddie Anderson and Frank Windmeijer for useful comments and suggestions. We would also like to thank seminar participants at the Statistical Society of Australia (NSW branch) and the NSW Bureau of Crime statistics and Research. The usual disclaimer applies.

Maurice J. G. Bun: The research of the author has been funded by the Dutch Research Council (NWO) Vernieuwingsimpuls research grant 'Causal Inference with Panel Data'.

Vasilis Sarafidis: The research of the author has been by the Australian Research Council (ARC) under research grant number DP-170103135.

1 Introduction

Crime, originating from the root of Latin cern \bar{o} ('I decide, I give judgment'), is the behaviour 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 behaviour, 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. According to an estimate, about 10 million people in the world are institutionalized for punishment, almost half of which are held in America, China and the UK (Walmsley 2009). Over the past 30 years, the American prison population has more than quadrupled. Raphael and Stoll (2009) have shown that the increase in the US prison population between 1984 and 2002 may be explained almost entirely by an increase in punitiveness, rather than an increase in crime. This has led some to label the extraordinary growth in the US prison population as one of the largest scale policy experiments of the century (Spelman 2000). Others have recently argued that significant reductions in the size of prison populations are possible nowadays without endangering public safety (Sundt et al. 2016).

The USA is not the only country to have experienced upward trends in its imprisonment rate over the last 30 years of course; nor is it the only country where the increase in imprisonment rates was driven by more punitive law and order policies rather than by an increase in crime. The Australian imprisonment rate, for example, has risen by 88% over the last 30 years. In the period covered by this article, it rose by more than 30%. Since 2000, this rise in imprisonment rates in Australia has occurred against a backdrop of falling crime (Weatherburn 2016).

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 wrongdoing does each additional prisoner avert? In order to address these questions in a constructive way, it is important to recognize that changes in the aggregate crime rate stem from individual behaviour. 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 2011). In turn, for a policy to deter criminal behaviour it must be designed with an understanding of what causes individuals to engage in criminal activity.

During the early part of the twentieth century, most theories of crime tended to attribute criminal behaviour to defects in the individual or in society (Vold et al. 2002). The seminal papers by Becker (1968) and Ehrlich (1975) denied the existence of any qualitative difference between offenders and non-offenders and asserted that individuals engage in criminal activity whenever the expected benefit of doing so exceeds the expected cost. Therefore, criminals 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 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 (1968) and Ehrlich (1975), a large empirical literature has developed, seeking to inform public policy by collecting data on various populations and building econometric models that describe criminal behaviour 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 modelling spanning over 40 years, there is arguably no consensus on whether there is a strong deterrent effect of law enforcement policies on crime activity. Empirical studies provide mixed evidence that are insufficient to draw clear conclusions.¹

The present paper revisits the economics of crime and punishment and provides a case study for New South Wales (NSW), Australia. We focus on four individual crime categories—namely theft, robbery, assault and homicide incidents. These offences broadly span the classification of criminal activity often employed in the literature. In addition, we consider two broader crime categories, property and violent crime. In order to alleviate heterogeneity bias, which can potentially arise due to differences across individual crime types in terms of occurrence and level of seriousness (see e.g. Cherry and List 2002), we consider weighted sums of the aforementioned four individual crime categories.

Our empirical strategy relies on GMM estimation of dynamic panel data models, and it takes into account various important methodological issues arising in the empirical analysis of criminal behaviour. In particular, our identification strategy allows for endogeneity or weak exogeneity between crime and the deterrence variables. Due to the panel structure of our modelling framework, instruments naturally arise with respect to sufficiently lagged values of the regressors. The validity of the instruments is examined empirically using tests for weak identification and overidentifying restrictions. In addition, the dynamic specification of our model captures the essential feature of habit formation and costs of adjustment in aggregate behaviour. 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 lag length of the effects.²

The results of our analysis show that criminal activity is highly responsive to the prospect of arrest and conviction, but much less responsive to the prospect or severity of imprisonment, if at all. This provides support to the idea that the consequences of being arrested and found guilty of a criminal offence include indirect sanctions imposed by society and not just the punishment meted out by the criminal justice system. In

¹ See Table 1 for a highly selective overview of crime studies.

 $^{^2}$ For a recent overview on the dynamic panel data literature, see Bun (2015).

particular, a convicted individual may no longer enjoy the same opportunities in the labour market and so the cost of social stigmatization can already be substantial in the event of conviction.

The sensitivity of our results is analysed extensively. First, we examine different moment conditions, depending on whether the probability of arrest is treated as endogenous or (weakly) exogenous, and we test for the validity of each specification. Second, we apply the methodology of Griliches and Hausman (1986) in order to test for measurement error in the data. Third, we estimate the crime model using a range of estimators other than GMM. Finally, we examine the effect of omitted variables in our model. The conclusions of our analysis appear to be fairly robust.

The remainder of this paper is as follows. Section 2 reviews various methodological issues arising in the empirical analysis of criminal behaviour. Section 3 presents the econometric specification and the identification strategy employed. Section 4 discusses the data and reports the results. A final section concludes.

2 Methodological issues

Unfortunately, empirical analysis of the effect of law enforcement policies on criminal activity is inherently problematic due to the nature of crime data available. In particular, data collected from individuals are self-reported and are doubtlessly affected by significant measurement error (Freeman 1999). Moreover, the time and cost involved in surveying a representative population can be prohibitively large. As a result, empirical studies of crime typically use some form of aggregate data, which describe crime in locales (e.g. local areas, states or countries) and are based on official records rather than self-reported information.

However, aggregate data are also not without problems. This has led some to suggest that the use of individual and aggregate data may be regarded as two complementary approaches (Trumbull 1989).

To begin with, since the economic model of crime purports to describe illegal behaviour of individual agents rather than an empirical aggregate, summing up over crime offences over individuals might inherently introduce some form of so-called *aggregation bias*.

Furthermore, the use of aggregate data introduces a problem of *lack of exogeneity* for some regressors, making the causal effect of law enforcement policies on crime more difficult to identify. For example, an exogenous upward shift in crime rate may eventually overwhelm police resources, given that police resources are fixed in the short term, causing the probability of arrest to decrease. This property is known in the econometrics jargon as *reverse causality* or simultaneity.

Even if some feedback relationship is not present in the data, the empirical probability of arrest (when defined as number of arrests divided by the number of crime offences) suffers from the fact that the numerator of the dependent variable (number of crime offences) is the denominator in the probability of arrest. This artificially induces a negative correlation between the two variables (Nagin 1978)—a phenomenon that is known as *ratio bias* (see e.g. Dills et al. 2008). An additional issue arising with aggregate data might be *measurement error*. Measurement error can manifest itself in at least two ways. Firstly, crime data typically record reported crime offences, rather than actual ones. For instance, it has sometimes been argued that it does not always serve the victim's best interest to report an offence (Myers 1980). This type of measurement error may particularly affect the crime rate and probability of arrest variables (Levitt 1998), as they are both constructed based on the number of crime offences. Secondly, because of timing issues, judiciary variables (courts and prisons) might follow a different timing than other variables affecting the (individual) propensity to commit a crime. As a result, (say) the empirical probability of imprisonment, which is the ratio of the number of imprisonments over convictions, could reflect crime offences that occurred at previous points in time.³ To put this differently, if offenders have high discount rates (as it has been argued in, for example, Wilson and Hernstein 1985), the speed with which offenders are apprehended and punished is important. This issue can be particularly pronounced with relatively high-frequency data (e.g. monthly observations), but less so for yearly data.⁴

Finally, there is the potential for *omitted variable bias* in the estimated parameters. Omitted variables imply a bias if they are correlated with included regressors. 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-researches, in which the effect of the probability of arrest, the probability of conviction, the probability of imprisonment and the length of average sentence is rarely examined together.

The aforementioned issues—namely aggregation bias, reverse causality, ratio bias, measurement error and omitted variable bias, all render the deterrence explanatory variables *endogenous*, that is, correlated with the error term of the model. It is well known that in the presence of endogeneity, least-squares-based estimates of the economic model of crime are biased and inconsistent. Despite that, a majority of crime studies do not control for endogeneity, which casts doubt on their results (Blumstein et al. 1978). Dills et al. (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 reverse causality 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.

Table 1 summarizes the empirical results for some widely cited contributions to the crime deterrence literature using aggregate data.⁵ For each of the studies noted, the table reports the sampling population, the unit of observation, the structure of the

 $^{^{3}}$ We are grateful to an anonymous referee for pointing this out.

⁴ This last observation is confirmed in our empirical analysis of Sect. 4, where we examine the case of measurement error in the arrest/conviction probability. We find little evidence for it.

⁵ The empirical crime literature is vast and extends over to several fields in social sciences, including economics, psychology, criminology, to mention a few. As such, it is impossible to pay justice to it. However, Table 1 does provide a good picture of the main methodological issues typically arising in crime analyses.

	usti	
	-	2
	and	
	olicing	
	5	
	respect	
	e with	
	ne rate	
•	e crin	
	단	
	0	,
1	E	
	asti	
	e e	
	Ē	
	es c	
	Imat	
	est	
	ica.	
	DIL	
ļ	H	
,	-	
	able	
	_	

Table 1 Empirical estima	ates of tl	he elasticity of the crime rat	te with respect to pol	licing and justice					
Authors	Year	Population	Unit of observation	Data (sample size)	Method	Crime type	Arrest Convictio	n Imprisonme	nt Sentence length
Panel A: Complete econo	mic mo	dels of crime							
Entorf and Spengler	2015	Germany	West German states	Panel (10, 25)	FGLS	Violent	-0.225 - 0.251	-0.006	0.019
						Property	-0.155 - 0.228	-0.055	0.029
Cornwell and Trumbull	1994	North Carolina	County	Panel (90, 7)	OLS, 2SLS	Total	-0.455 - 0.336	-0.196	-0.03
Trumbull	1989	North Carolina	County	Cross section (98)	OLS	Total	-0.217 - 0.451	-0.325	-0.149
Pyle	1984	England and Wales	Police authorities	Cross section (41)	OLS	Robbery	-0.5 0.73	-0.48	-0.57
						Property	-0.32 0.4	-0.55	-0.85
Wolpin	1980	USA, England and Wales	Country	Panel (3, 17)	OLS	Robbery	-0.537 - 0.460	- 1.13	-1.27
Panel B: Arrest									
Klick and Tabarrok	2005	Washington D.C.	Police district	Panel (7,506)	OLS	Violent	-0.3		
						Burglary	-0.3		
Fajnzylber et al.	2002	United Nations	Country	Panel (45, 5)	GMM	Robbery	0.08		.035
						Homicide	-0.09		-0.346
Corman and Mocan	2000	New York	City	Time series (108)	OLS	Murder	-0.336		
						Burglary	-0.355		
Bodman and Maultby	1997	Australia	State	Cross section (60)	2SLS	Robbery	-0.258		-0.621
						Burglary	-0.367		
Levitt	1997	USA (City	Panel (59, 23)	2SLS	Violent	-0.9^{*}		
						Property	-0.24^{*}		
Marvell and Moody	1996	USA VSA	City	Panel (56, 22)	Granger	Total	-0.133		
						Homicide	-0.241		
						Burglary	-0.151		

Table 1 continued										
Authors	Year	Population	Unit of observation	Data (sample size)	Method	Crime type	Arrest	Conviction	Imprisonment	Sentence length
Sampson and Cohen	1988	USA	City	Cross section (171)	2SLS	Robbery	-0.28			
						Burglary	-0.12			
Car-Hill and Stern	1973	England and Wales	Police districts	Cross section (64)	FIML	Total	-0.59			-0.17
Panel C: Imprisonment										
Johnson and Raphael	2012	USA	State	Panel (51, 27)	2SLS	Violent			- 0.11	
						Property			-0.21	
Lofstrom and Raphael	2016	USA	County	Panel (19, 36)	2SLS	Violent			0.01	
						Property			- 0.09	
Liedka et al.	2006	USA	State	Panel (51, 29)	Granger	Total			-0.245	
						Murder			-0.13	
						Burglary			-0.136	
Levitt	2002	USA	City	Panel (100, 21)	2SLS	Violent	-0.435		-0.171	
						Property	-0.501		-0.305	
Witt and Witte	2000	USA	Country	Time series (38)	VAR	Total			- 0.55	
Levitt	1996	NSA	State	Panel (51, 23)	2SLS	Violent			-0.261	
						Property			-0.379	
Marvell and Moody	1994	USA	State	Panel (49, 19)	Granger	Total			-0.159	
						Homicide			-0.065	
						Burglary			-0.253	
Ehrlich	1973	NSA	State	Cross section (47)	2SLS	Total			- 0.991	- 1.123
Panel D: Conviction and	other st	udies								
Haas	1980	New Jersey	Municipality	Cross section (181)	2SLS	Total	I	- 0.02		

2309

 $\stackrel{{}_{\scriptstyle{\frown}}}{\underline{\frown}}$ Springer

Table 1 con	tinued									
Authors	Year	Population	Unit of observation	Data (sample size)	Method	Crime type	Arrest	Conviction	Imprisonment	Sentence length
Withers	1984	Australia	State	Cross section (104)	SIO	Violent		0.29	0.09	
						Total		-0.62	-0.6	
						Property		-0.59	-0.56	
Sjoquist	1973	USA	Municipality	Cross section (53)	OLS	Theft	-0.342	-0.678		-0.212
* Author pro	vided mul	ltiple estimates,	in which case the media	n is reported						

 $\underline{\textcircled{O}}$ Springer

data followed by the sample size⁶, 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 Wolpin (1978, 1980), Pyle (1984), Trumbull (1989), Cornwell and Trumbull (1994) and Entorf and Spengler (2008, 2015). In most of these studies, least-squares-based methods are used to obtain estimates of the parameters; hence, all deterrence variables are treated as exogenous. Trumbull (1989) justifies this choice claiming that endogeneity is not a salient feature of the existing data set, based on the results of a Wu–Hausman specification test. A few studies also apply IV regression and treat the probability of arrest as endogenous, but all remaining variables as exogenous (Cornwell and Trumbull 1994; Entorf and Spengler 2008, 2015). The authors fail to find a statistically significant relationship between the deterrence variables and crime using a 2SLS procedure.⁷

The remaining studies restrict their attention to a particular variable of interest. 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. For example, Mustard (2003) shows that arrest rates are likely to be negatively correlated with the probability of conviction and sentence length since arrest rates are often substitutes for conviction rates and sentences. As a result, Mustard (2003) 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 50%. Furthermore, omitted variables may invalidate estimation based on instrumental variables. Candidate instrumental variables may not be orthogonal to the deterrence variables omitted from the regression; hence, actually they do not constitute instruments.

3 Econometric specification

3.1 Model

The dependent variable is the rate of crime, which is defined as the ratio of the number of crime offences committed in a given local government area (LGA) *i* at time *t* (labelled crm_{it}) over population (pop_{it}). The rate of crime is not the same as the binary 'crime–no crime' decision an individual faces, but it is arguably the closest substitute one can observe at the aggregate level.

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, the crime rate is modelled as a function

⁶ For panel data models, the cross-sectional dimension, N, is given first, followed by the time dimension, T.

 $^{^7}$ Bun and Sarafidis (2015) shows that this is due to weak instruments for the Cornwell and Trumbull (1994) study.

VariableDefinitioncrmrNumber of criminal incidents divided by total populationprbarrNumber of arrests divided by criminal incidentsprbconvNumber of convictions divided by arrestsprbimprNumber of imprisonments divided by convictionsavsenAverage non-parole period (months) imposed for prison sentenceincomeAverage wage and salary earner incomeunempUnemployment rate (%)		
crmrNumber of criminal incidents divided by total populationprbarrNumber of arrests divided by criminal incidentsprbconvNumber of convictions divided by arrestsprbimprNumber of imprisonments divided by convictionsavsenAverage non-parole period (months) imposed for prison sentenceincomeAverage wage and salary earner incomeunempUnemployment rate (%)	Variable	Definition
prbarrNumber of arrests divided by criminal incidentsprbconvNumber of convictions divided by arrestsprbimprNumber of imprisonments divided by convictionsavsenAverage non-parole period (months) imposed for prison sentenceincomeAverage wage and salary earner incomeunempUnemployment rate (%)	crmr	Number of criminal incidents divided by total population
prbconvNumber of convictions divided by arrestsprbimprNumber of imprisonments divided by convictionsavsenAverage non-parole period (months) imposed for prison sentenceincomeAverage wage and salary earner incomeunempUnemployment rate (%)	prbarr	Number of arrests divided by criminal incidents
prbimprNumber of imprisonments divided by convictionsavsenAverage non-parole period (months) imposed for prison sentenceincomeAverage wage and salary earner incomeunempUnemployment rate (%)	prbconv	Number of convictions divided by arrests
avsenAverage non-parole period (months) imposed for prison sentenceincomeAverage wage and salary earner incomeunempUnemployment rate (%)	prbimpr	Number of imprisonments divided by convictions
income Average wage and salary earner income unemp Unemployment rate (%)	avsen	Average non-parole period (months) imposed for prison sentences
unemp Unemployment rate (%)	income	Average wage and salary earner income
	unemp	Unemployment rate (%)

Table 2 Definition of variables in the crime model

of the empirical probability of arrest, the empirical probability of conviction given arrest and the empirical probability of imprisonment given conviction.

This leads to the following specification:

$$\ln\left(\frac{crm_{it}}{pop_{it}}\right) = \alpha \ln\left(\frac{crm_{it-1}}{pop_{it-1}}\right) + \beta_1 \ln\left(\frac{arr_{it}}{crm_{it}}\right) + \beta_2 \ln\left(\frac{conv_{it}}{arr_{it}}\right) + \beta_3 \ln\left(\frac{impr_{it}}{conv_{it}}\right) + \beta_4 \ln avsen_{it} + \beta_5 \ln income_{it} + \beta_6 \ln unemp_{it} + \eta_i + \lambda_t + \varepsilon_{it},$$
(1)

for t = 1, ..., T time periods and i = 1, ..., N regions, where $(crm_{it}, arr_{it}, conv_{it}, impr_{it})$ denote the number of crime offences, arrests, convictions and imprisonments, respectively. The inclusion of sentence length $(avsen_{it})$, income $(income_{it})$ and unemployment $(unemp_{it})$ in the above equation captures the expected cost/gains from the illegal and legal sectors. Precise definitions of all variables used in our regression analysis are provided in Table 2.

Using short-hand notation, the model can be rewritten as:

$$\ln crmr_{it} = \alpha \ln crmr_{it-1} + \beta_1 \ln prbarr_{it} + \beta_2 \ln prbconv_{it} + \beta_3 \ln prbimpr_{it} + \beta_4 \ln avsen_{it} + \beta_5 \ln income_{it} + \beta_6 \ln unemp_{it} + \eta_i + \lambda_t + \varepsilon_{it}.$$
(2)

The error term in (2) allows for regional-specific effects (η_i), which may be correlated with the regressors, as well as time effects (λ_i) that capture common variations in crime across regions. The coefficient of the lagged value of the dependent variable, α , measures the combined effect of short-run dynamics and time-varying omitted regressors hidden in lagged crime rates.

We also considered the possibility that criminals may form expectations about conviction rates in an adaptive manner, implying that lags of these variables should also be included on the right-hand side. We tested specifications including such lagged effects but they were largely insignificant.

Finally, it is worth mentioning that many of the models used in the literature (see e.g. Table 1) are restricted versions of (2). For example, many studies do include the

probability of arrest, but exclude the probabilities of conviction and imprisonment and sentence lengths.

3.2 Identification strategy

We estimate model (2) by the generalized method of moments (GMM) developed originally by Hansen (1982) and adapted for estimation of dynamic panel data models by Arellano and Bond (1991), Arellano and Bover (1995) and Blundell and Bond (1998). The GMM approach has the advantage that, compared to maximum likelihood, it requires much weaker assumptions about the initial conditions of the data generating process and avoids full specification of the serial correlation and heteroskedasticity properties of the error, or indeed any other distributional assumptions. Moreover, GMM is a natural choice when multiple explanatory variables are endogenous. For the reasons discussed in Sect. 2, we treat the probability of arrest as an endogenous regressor. The lagged crime rate, which models the short-run dynamics, is an additional endogenous regressor.

To remove the region-specific effects, first differences are taken from the original model in levels (2) resulting in:

$$\Delta \ln crmr_{it} = \alpha \Delta \ln crmr_{it-1} + \beta_1 \Delta \ln prbarr_{it} + \beta_2 \Delta \ln prbconv_{it} + \beta_3 \Delta \ln prbimpr_{it} + \beta_4 \Delta \ln avsen_{it} + \beta_5 \Delta \ln income_{it} + \beta_6 \Delta \ln unemp_{it} + \Delta \lambda_t + \Delta \varepsilon_{it}.$$
(3)

GMM estimation of the first-differenced model (3) has been developed by Arellano and Bond (1991). Since dynamic panels are often largely overidentified, an important practical issue is how many moment conditions to use. It is well documented that numerous instruments can overfit endogenous variables in finite samples, resulting in a trade-off between bias and efficiency. There is substantial theoretical work on the overfitting bias of GMM coefficient estimators in panel data models (Ziliak 1997; Alvarez and Arellano 2003; Bun and Kiviet 2006). Furthermore, with many moment conditions the power of (mis)specification tests deteriorates rapidly (Bowsher 2002). Roodman (2009) compares two popular approaches for limiting the number of instruments: (i) the use of (up to) certain lags instead of all available lags and (ii) combining instruments into smaller sets. Using asymptotic expansion techniques, Bun and Kiviet (2006) show that the order of magnitude of bias is reduced when going from all moment conditions to using only nearest lags as instruments. Furthermore, their simulation results show that when the number of time periods is a double digit, like T = 13in the current study, the GMM estimator using nearest lags as instruments has much less finite-sample bias than the GMM estimator using all moment conditions. Therefore, we follow the recommendation of using a limited number of moment conditions and only employ the three nearest lagged instruments. Furthermore, we collapse them resulting in the following six moment conditions for the model in first differences:

$$E\left[\sum_{t=3+j}^{T}\ln crmr_{it-2-j}\Delta\varepsilon_{it}\right] = 0; \qquad E\left[\sum_{t=3+j}^{T}\ln prbarr_{it-2-j}\Delta\varepsilon_{it}\right] = 0, (4)$$

Springer

for j = 0, 1, 2. We note that another advantage of collapsed instruments is that the underlying time-specific moment conditions do not need to hold exactly for each time period, but only in sum. Regarding the probability of arrest, we will also estimate a specification under weak exogeneity, in which case we make use of

$$E\left[\sum_{t=3}^{T}\ln prbarr_{it-1}\Delta\varepsilon_{it}\right] = 0.$$
(5)

Conviction and imprisonment rates, as well as the severity of punishment, are determined by the judiciary system. In practice, it is hard to believe that the judiciary system is strictly exogenous to crime rates. Similarly, it is natural to think that recent changes in economic conditions may exert some impact on current crime rates. As a result, we allow in estimation for a feedback relationship between past values of crime and current values of the remaining deterrence regressors (*prbconv*, *prbimpr*, *avsen*), as well as of economic conditions (*income*, *unemp*).

As it is briefly mentioned in Sect. 2, since the number of crime offences enters into both the numerator of $\ln crmr$ and the denominator of $\ln prbarr$ in Eq. (3), the model is subject to ratio bias. However, so long as the idiosyncratic error term of the model, ε_{it} , is serially uncorrelated, then values of $\ln prbarr_{it-1}$ lagged two or more periods may serve as valid instruments to identify β_2 .⁸

It is well known that identification can be weak when the panel data are persistent or, more generally, when the correlation between endogenous regressors and instruments is close to zero. We therefore check the identification strength of the exploited moment conditions in various ways. First, we estimate pure autoregressive models for the endogenous regressors and check whether autoregressive dynamics are reasonably far away from the unit root. Second, we use the Kleibergen and Paap (2006) rank statistic to test for underidentification in the first-differenced and levels IV models.

To check the validity of the estimated specification, we report the p value of Hansen's (1982) J test of overidentifying restrictions and the p value of Arellano and Bond's (1991) test of serial correlation of the disturbances up to second order. The former is used to determine empirically the validity of the overidentifying restrictions in the GMM model. The latter is useful because the use of lagged values of the endogenous variables as instruments (in levels) requires that serial correlation in the idiosyncratic error term is only up to a certain order.⁹

Long-run estimates are computed by dividing the short-run slope coefficients by 1 minus the estimated autoregressive parameter. Robust standard errors are reported in parentheses, which are valid under arbitrary forms of heteroskedasticity and serial correlation. Furthermore, we perform the correction proposed by Windmeijer (2005) for the finite-sample bias of the standard errors of the two-step GMM estimator.¹⁰ The

⁸ That is, $\ln prbarr_{it-2}$ (and further lags) satisfies in this case the two conditions required for validity of an instrument, namely, it remains correlated with $\Delta \ln prbarr_{it}$ in (3) but it is uncorrelated with $\Delta \varepsilon_{it}$.

⁹ We note that instruments are more likely to be valid in panel data settings compared to time series or cross-sectional regressions because the multi-dimensionality of panel data allows one to capture richer sources of unobserved heterogeneity relative to time series and cross-sectional data alone.

¹⁰ All GMM results have been obtained using David Roodman's *xtabond2* algorithm in Stata 15.

4 Data and results

4.1 Data

We construct a new data set containing information on criminal activity and deterrence for all N = 153 local government areas in New South Wales, each one observed over a period of T = 13 years from 1995/1996 to 2007/2008. 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).¹¹ 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. LGAs in NSW range in size from over 350,000 people (3261.9 persons per km² to a little over 2,000 people (i.e. less than one person per km²). The average LGA has a population of 58,438 persons (sd. = 77,585 persons) (Australian Bureau of Statistics 2017). Although LGAs include both urban and rural areas, most of the population of NSW can be found in urban rather than rural areas. The three cities of Sydney, Newcastle and Wollongong account for three quarters of the NSW population (NSW Department of Industry 2017).

Our data on crime are drawn from COPS—the NSW Police Operational Policing System, to which the NSW Bureau of Crime Statistics and Research has online access. This system records each crime incident reported to or discovered by police. A crime incident is defined as an activity detected by or reported to police which:

- Involved the same offender(s);
- Involved the same victims(s);
- Occurred at one location;
- Occurred during one uninterrupted period of time;
- Falls into one offence category;
- Falls into one incident type (e.g. 'actual', 'attempted', 'conspiracy').

The data are categorized according to the date of reporting to or detection by police, not by the date of occurrence of the offence. The deterrence variables (probabilities of arrest, conviction and imprisonment, as well as average non-parole period length) are drawn from a separate database on court appearances and outcomes maintained by the NSW Bureau of Crime Statistics and Research. This database contains information, inter alia, on the charge(s) laid against each offender, which charge(s) resulted in a conviction, whether a prison sentence was imposed, the length of the aggregate (total) sentence and the length of the non-parole period. The aggregate sentence is the maximum time the offender can be held in custody for the offence(s) he/she

¹¹ 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 LGAs. There are 193 SLAs in NSW.

has committed. The non-parole period defines the minimum period the offender must spend in custody before being released on parole. Where the aggregate sentence is less than 6 months, no non-parole period can be specified. In this instance, the aggregate sentence defines the minimum period the offender must spend in custody. For the purposes of this analysis, we define the sentence length as the length of the non-parole period where one is specified and the length of the aggregate sentence where the aggregate sentence is less than 6 months.¹²

It is worth noting that in NSW there are three levels to the NSW court system, namely the Supreme, District and Local Courts. However, more than 90% of criminal cases are finalized in a Local Court, where guilt or innocence is determined by a magistrate rather than by a judge or jury. In 2016, more than 90% of Local Court cases resulted in a conviction on at least one charge, in most cases because the defendant pleaded guilty.¹³ The average time taken to finalize a Local Court matter in 2016 was 64 days. The only offences examined here that are exclusively dealt with by a higher court (i.e. the District or Supreme Court) are those involving robbery, homicide and some serious sexual offences. The majority of these cases also involve a guilty plea. The average time to finalize a guilty plea in the higher courts is around 14 months but the distribution is highly skewed, with the majority of guilty pleas being finalized in less than 12 months. In the vast majority of cases, then, only a short period elapses between arrest and sentence and most of those arrested will end up convicted.

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 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 this end, we mapped the data to a common ASGC standard (2006) using a series of concordance tables, in order to achieve consistency. Similarly, the unemployment data were first mapped to the same ASGC standard (2006) to account for name and boundary changes that occurred 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.

We distinguish between two broad crime categories, i.e. property and violent crime. Property crime is defined as any incident of robbery without a weapon, robbery with a firearm, robbery with a weapon not a firearm, stealing property in a dwelling house, motor vehicle theft, stealing from motor vehicle, stealing from retail store, stealing from dwelling, stealing from person, stock theft, other theft and fraud. Violent crime is defined as any incident of homicide, non-domestic violence-related assault, domestic violence-related assault, robbery without a weapon, robbery with a firearm, robbery

¹² Courts in NSW enjoy considerable discretion as to the type of penalty and—where prison is concerned the length of the prison term. Indeed, in most cases the only constraint is the statutory maximum penalty, which is usually well above the average sentence (see Weatherburn 1994). This implies some considerable variation in sentence length being exogenously driven, depending on individual characteristics of the judge.

¹³ The only offences examined here that are exclusively dealt with by a judge and jury are murder, sexual assault and robbery.

with a weapon not a firearm, sexual assault, indecent assault or act of indecency, or other sexual offence.¹⁴

Since both property and violent crime comprise individual crime categories that have quite different occurrence and level of seriousness, adding them all up together using simple (unweighted) summation can potentially invalidate inferences.¹⁵ For instance, a homicide incident occurs less often and can be far more severe compared to a robbery incident. Therefore, in what follows we analyse weighted sums of property and violent crime. The weights are computed using two different ways. Firstly, we compute variance weights based on the sample within standard deviation of the individual crime types. This set of weights reflects the fact that more severe types of crime occur less often. Secondly, we compute weights based on the 'average seriousness score per incident' reported in Heller and McEwen (1973) and Blumstein (1974).¹⁶ The seriousness scores for each of the aforementioned crime categories are as follows: theft, 2.29; robbery, 6.43; assault, 9.74; homicide, 33.29. These scores have been adjusted such that in each of the two broad crime categories, they add up to unity.¹⁷

In order to deal effectively with heterogeneity across different individual crime types, we also analyse in length each of the aforementioned four crime types separately, namely theft, robbery, assault and homicide. These four categories broadly span the classification of crime offences often used in the literature, see e.g. Table 1 in the present paper, as well as Table 1 in Cherry and List (2002).

The various deterrence variables (probabilities of arrest, conviction and imprisonment, as well as average prison length) are computed specifically for each type of crime analysed. This accommodates the expectation that, apart from having different values across crime types, these variables may potentially have a different deterrence effect across crime types.

Tables 3 and 4 report descriptive statistics for the various categories of crime considered in our analysis.¹⁸ As expected, the mean value of the rate of violent crime, as well as that of its individual crime components, is smaller than that of property crime and it exhibits a much smaller dispersion as well. This indicates that violent crime occurs less frequently and is more localized. The empirical probability of arrest

¹⁴ It is apparent that robbery is both a property and violent crime because it involves violence (or the threat of it) to unlawfully obtain property.

¹⁵ We are grateful to two anonymous referees for alerting us about this issue at first place.

¹⁶ On the computation of these scores, the interested reader may refer to Tables 6 and 2 of the aforementioned papers, respectively, as well as the associated discussion.

¹⁷ In terms of property crime, weights based on the within standard deviation of individual crime types are theft = 0.950; robbery = 0.050, whereas weights based on the seriousness score are theft = 0.528; robbery = 0.472.

In terms of violent crime, weights based on the within standard deviation of individual crime types are robbery = 0.272; assault = 0.718; homicide = 0.010, whereas weights based on the seriousness score are robbery = 0.099; assault = 0.387; homicide = 0.514.

¹⁸ We note that the results of the GMM estimated aggregate crime models are very similar irrespective of the method used to compute weights; see Tables 5 and 6. Therefore, unless otherwise stated, in what follows aggregate crime results correspond to variance weights computed based on the within standard deviation of individual crime types.

X 7 ' 1 1	0: /		0.1	10:1 D	00/1 D
variable	Crime type	Mean	Stdev	10th Perc	90th Perc
Crime rate					
	Property	0.047	0.028	0.023	0.074
	Violent	0.009	0.010	0.003	0.013
	Theft	0.049	0.029	0.024	0.092
	Robbery	0.001	0.002	0.000	0.002
	Assault	0.012	0.013	0.004	0.018
	Homicide	0.000	0.000	0.000	0.000
Probability of arrest					
	Property	0.087	0.049	0.040	0.149
	Violent	0.611	0.177	0.386	0.182
	Theft	0.087	0.049	0.040	0.149
	Robbery	0.396	0.247	0.000	0.975
	Assault	0.617	0.174	0.403	0.840
	Homicide	0.665	0.453	0.000	0.983
Probability of conviction	1				
	Property	0.366	0.309	0.138	0.594
	Violent	0.391	0.140	0.251	0.540
	Theft	0.360	0.278	0.135	0.588
	Robbery	0.604	0.659	0.250	0.967
	Assault	0.378	0.132	0.244	0.361
	Homicide	0.646	0.670	0.000	0.966
Probability of imprisonn	nent				
	Property	0.107	0.119	0.000	0.219
	Violent	0.098	0.083	0.000	0.239
	Theft	0.101	0.112	0.000	0.212
	Robbery	0.373	0.275	0.000	0.700
	Assault	0.083	0.083	0.000	0.177
	Homicide	0.522	0.222	0.000	1.000
Average sentence (days)					
	Property	3.742	3.089	0.000	6.950
	Violent	22.62	28.41	1.358	58.58
	Theft	5.065	2.437	2.500	7.496
	Robbery	11.62	7.497	4.500	19.00
	Assault	6.902	5.621	3.000	11.50
	Homicide	2043	12612	12.000	180.0
Income (\$ '000)		34.0	9.4	25.2	44.0
Unemployment (%)		7.1	5.1	3.0	12.4

Table 3 Descriptive statistics

Descriptive statistics computed for the variables used in regression analysis. N = 153 and T = 13, yielding a total of 1989 observations

	crmr	prbarr	prbconv	prbimpr	avsen	income	unemp
Property crime	;						
crmr	1						
prbarr	-0.003	1					
prbconv	-0.121	-0.390	1				
prbimpr	0.096	0.100	0.014	1			
avsen	0.182	-0.065	0.099	0.376	1		
income	0.073	-0.406	0.052	-0.030	0.195	1	
unemp	0.165	0.277	-0.067	0.082	0.089	-0.317	1
Violent crime							
crmr	1						
prbarr	0.161	1					
prbconv	-0.254	-0.194	1				
prbimpr	0.206	0.112	0.005	1			
avsen	-0.005	-0.048	0.008	0.025	1		
income	-0.147	-0.466	-0.085	-0.174	0.001	1	
unemp	0.131	0.170	-0.004	0.169	-0.002	-0.317	1

 Table 4
 Sample correlation coefficients

Descriptive statistics computed for the variables used in regression analysis. N = 153 and T = 13, yielding a total of 1989 observations

is higher on average for violent crime than property crime, which reflects the fact that in many cases violent crime involves face-to-face contact increasing the probability of apprehension. In regard to the category of homicides, the mean value of average sentence length is much larger than the value in the 90th percentile, which indicates that there are a relatively small number of very big sentences in the sample.

Figures 1 and 2 show the development over time of property and violent crime as well as their corresponding arrest rates. Cross-sectional averages have been plotted; hence, the line graphs depict average development across LGAs. Figure 1 shows that property crime increases gradually in the 1990s, peaks around 2000 and then falls sharply afterwards. The property crime arrest rate gradually decreases until 2003 and then stabilizes. By contrast as shown in Fig. 2, average violent crime increases steadily until 2002, and then, it exhibits a small downward trend. At the same time, the arrest rate declines until 2002 and it follows an upward trend after that date. Figure 3 shows that economic conditions, i.e. per capita income and unemployment, improved steadily over the whole sample period.

To get an idea of the time-series persistence in our data, we estimate pure autoregressive models for the crime rate and the probability of arrest, i.e. the endogenous regressors in our empirical analysis. System GMM estimates of autoregressive coefficients are in the range 0.2–0.6, which shows moderate persistence only. Not surprisingly panel unit root tests (Harris and Tzavalis 1999; Im et al. 2003) reject the null hypothesis of a unit root for both crime rates and probability of arrest and for both property and violent crime.



Fig. 1 Average property crime rate and arrest rate. *Note*: line graphs are cross-sectional averages over 153 LGAs in NSW



Fig. 2 Average violent crime rate and arrest rate. Note: see Fig. 1



Fig. 3 Average per capita income and unemployment rate. *Note*: per capita income in 1000 s of Australian dollars. Unemployment rate is percentage. Line graphs are cross-sectional averages over 153 LGAs in NSW

4.2 Baseline results

We analyse the weighted sums of property crime and violent crime, based on the econometric model presented in the previous section. First-differenced GMM estimates allowing for endogeneity of lagged crime and the probability of arrest are reported in Tables 5 and 6. We treat all remaining explanatory variables as weakly exogenous. In Table 5, aggregate crime series have been constructed using the within standard deviation of individual crime types, whereas in Table 6 aggregate crime series have been constructed using weights computed based on the average seriousness score per incident.

The results reported in Tables 5 and 6 are remarkably similar even if the corresponding weights take rather different values. Therefore, we only discuss the results in Table 5 based on variance weights. To begin with, the p values from the reported overidentifying restrictions test show no evidence of lack of validity of the estimated specification. Furthermore, the Kleibergen and Paap (2006) rank test indicates no underidentification in either equation as the null hypothesis is soundly rejected at the 5% level of significance.

The GMM estimates of the deterrence effects are of the expected sign with the exception of the coefficient of average sentence, which is largely insignificant. For property crime, a 1% increase in the probability of arrest appears to decrease the expected value of the crime rate by 0.283% in the short run and 0.365% in the long run, ceteris paribus. Likewise, the elasticity of the probability of conviction is about

	Property		Violent	
	Short run	Long run	Short run	Long run
Lagged crime rate	0.225		0.425	
	(0.125)		(0.143)	
Probability of arrest	-0.283	-0.365	-0.398	-0.692
	(0.113)	(0.181)	(0.169)	(0.173)
Probability of conviction	-0.148	-0.191	-0.325	-0.564
	(0.043)	(0.069)	(0.050)	(0.150)
Probability of imprisonment	-0.019	-0.025	-0.016	-0.027
	(0.013)	(0.018)	(0.013)	(0.023)
Average sentence	0.019	0.025	0.010	0.017
	(0.017)	(0.024)	(0.007)	(0.012)
Income	-0.452	-0.583	0.174	0.302
	(0.306)	(0.402)	(0.309)	(0.527)
Unemployment	-0.043	-0.055	-0.322	-0.500
	(0.095)	(0.122)	(0.145)	(0.333)
<i>p</i> value oir	0.465		0.152	
p value rank test	0.000		0.000	
p value serial correlation				
lag 1	0.000		0.000	
lag 2	0.047		0.960	

Table 5 GMM estimates of the (weighted) crime model

Weights are computed based on the within standard deviation of individual crime types. Each regression includes LGA- and time-specific effects

-0.148 and -0.191 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 behaviour, imperfect knowledge and uncertainty. In particular, the value of the autoregressive parameter indicates that it takes about 2 years for 90% of the total impact of either one of the explanatory variables on crime to be realized, all else being constant.

The estimated coefficients of the probability of imprisonment are much smaller compared to the coefficients of the probability of arrest and the probability of conviction, whereas the effect of average sentence is insignificant as mentioned above. This indicates that imprisoning more criminals, or imprisoning them for longer, is not as effective as increasing the risk of apprehension or conviction once arrested. In other words, criminal activity seems to be highly responsive to the prospect of arrest and conviction, but less responsive to the prospect or severity of imprisonment. This 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

	Property		Violent	
	Short run	Long run	Short run	Long run
Lagged crime rate	0.218		0.405	
	(0.119)		(0.145)	
Probability of arrest	-0.297	-0.380	-0.400	-0.672
	(0.105)	(0.172)	(0.170)	(0.184)
Probability of conviction	-0.215	-0.275	-0.318	-0.535
	(0.053)	(0.089)	(0.055)	(0.148)
Probability of imprisonment	-0.035	-0.044	-0.031	-0.052
	(0.016)	(0.024)	(0.014)	(0.027)
Average sentence	0.009	0.012	0.005	0.009
	(0.011)	(0.015)	(0.004)	(0.006)
Income	-0.578	-0.739	0.176	0.296
	(0.278)	(0.366)	(0.310)	(0.511)
Unemployment	-0.058	-0.075	0.322	0.542
	(0.093)	(0.118)	(0.145)	(0.329)
p value oir	0.456		0.471	
p value rank test	0.000		0.000	
p value serial correlation				
lag 1	0.000		0.000	
lag 2	0.052		0.789	

Table 6	GMM	estimates	of the	(weighted)	crime	model
---------	-----	-----------	--------	------------	-------	-------

Weights are computed based on the average seriousness score per incident. Each regression includes LGAand time-specific effects

of social stigmatization are implied in the event of conviction. Zimring and Hawkins (1973, p 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.

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.

Table 7 presents results for the four disaggregated crime categories. The conclusions are qualitatively similar to those of the weighted aggregates. In particular, the estimated deterrence effects are of the expected sign, with the exception of average sentence, the effect of which is largely statistically insignificant. Moreover, the coefficients of the risk of apprehension and conviction remain much larger than those of the probability of imprisonment and average sentence. Hence, increasing the risk of apprehension or conviction once arrested appears to be much more effective compared to the practice of imprisoning more criminals, or imprisoning them for longer.

	Theft	Robbery	Assault	Homicide
Lagged crime rate	0.366	0.132	0.526	0.024
	(0.201)	(0.107)	(0.151)	(0.091)
Probability of arrest	-0.365	-0.628	-0.348	-0.706
	(0.262)	(0.259)	(0.203)	(0.374)
Probability of conviction	-0.238	-0.582	-0.301	-0.480
	(0.118)	(0.120)	(0.063)	(0.157)
Probability of imprisonment	-0.026	-0.083	-0.047	-0.168
	(0.022)	(0.048)	(0.019)	(0.161)
Average sentence	0.018	0.063	0.022	0.005
	(0.020)	(0.043)	(0.013)	(0.042)
Income	-2.155	-0.447	-0.121	-1.871
	(1.824)	(2.076)	(0.907)	(2.499)
Unemployment	-0.047	0.154	-0.318	0.114
	(0.062)	(0.315)	(0.187)	(2.006)
p value oir	0.437	0.318	0.093	0.118
p value rank test	0.149	0.924	0.000	0.816
p value serial correlation				
lag 1	0.005	0.000	0.000	0.000
lag 2	0.002	0.383	0.856	0.413

Table 7 GMM estimates for disaggregated crime types

See Table 5

There are two notable differences in the results obtained between the disaggregated and aggregated crime offences. Firstly, the standard error of the estimated coefficients obtained from the disaggregated models is relatively larger in general. This reflects the fact that the sample size for individual crime categories is smaller because there are quite more zero-crime occurrences in this case.¹⁹ Secondly, arguably for the same reason, the p value of the rank test statistics indicates a potential weak instruments problem for the robbery and homicide equations, and less so for theft. This shows that analysing weighted aggregates of individual crime categories is quite appealing in this respect.

4.3 Sensitivity analysis

As discussed previously, the arrest probability is often seen as an endogenous regressor in the empirical crime literature. Hence, we have estimated our baseline crime model allowing for endogeneity of the probability of arrest. However, efficiency gains in the coefficient estimates may arise by imposing weak exogeneity on this main deterrence regressor. As such, we have re-estimated the model allowing for weak exogeneity in

¹⁹ This is especially so for LGAs covering rural areas.

Lagged crime rate

eity				
Violent	Theft	Robbery	Assault	Homicide
0.379	0.440	0.158	0.487	0.042
(0.127)	(0.159)	(0.094)	(0.172)	(0.100)
-0.233	-0.218	-0.528	-0.356	-0.688

Table 8 GMM estimates under weak exogeneity

Property

0.285

Probability of arrest -0.097(0.037)(0.069)(0.087)(0.111)(0.088)(0.181)Probability of conviction -0.085-0.279-0.189-0.529-0.286-0.468(0.027)(0.043)(0.056)(0.095)(0.052)(0.109)Probability of imprisonment -0.006-0.016-0.027-0.080-0.044-0.290(0.224)(0.009)(0.013)(0.019)(0.047)(0.019)Average sentence 0.001 0.008 0.009 0.061 0.016 0.032 (0.010)(0.008)(0.016)(0.043)(0.013)(0.038)Income -0.7080.329 -1.3400.625 0.375 4.894 (0.358)(0.334)(1.888)(2.563)(0.982)(26.54)-0.0100.028 Unemployment -0.038-0.311-0.4520.588 (0.224)(1.504)(0.096)(0.158)(0.065)(0.279)0.155 0.124 0.244 0.137 0.033 0.034 p value oir 0.000 0.000 0.002 0.032 0.000 0.463 p value rank test p value serial correlation lag 1 0.000 0.000 0.002 0.000 0.000 0.088 lag 2 0.072 0.785 0.005 0.351 0.711 0.848

See Table 5

the probability of arrest. In this case, we add (5) into the set of moment conditions employed by the GMM estimator. The results are reported in Table 8.

There are two main differences between Table 5 and Table 8. Firstly, in the latter case the standard error of the estimated coefficients is much smaller, often by a magnitude of less than a half. That is, imposing weak exogeneity substantially improves the efficiency of the estimated coefficients. Secondly, the p value of the overidentification test statistic is now smaller in general, although it remains larger than 0.05 in most cases. Furthermore, compared to Table 7 the p value of the rank test falls dramatically, such that this time a weak instruments problem is potentially an issue only for the homicide equation.

To further test the exogeneity of the probability of arrest and conviction, we apply the empirical test of Griliches and Hausman (1986) to detect the presence of measurement error. The idea is that long differences²⁰, as opposed to first differences, are less vulnerable to measurement error. Therefore, in the absence of measurement error the OLS estimator of the arrest/conviction rate elasticity in the differenced crime model should not show any systematic pattern across the different lengths. Levitt (1998) applied this test to investigate the extent of measurement error. Table 9 reports findings for the Australian crime data. To save space, we only report results

²⁰ For a variable y_{it} , the *j*th long difference is defined as $y_{it} - y_{i,t-j}$.

	Arrest				Conviction			
	Property	Violent	Theft	Assault	Property	Violent	Theft	Assault
First difference	-0.060	-0.186	-0.101	-0.207	-0.079	-0.228	-0.125	-0.208
	(0.012)	(0.032)	(0.025)	(0.036)	(0.015)	(0.025)	(0.017)	(0.027)
Second long difference	-0.072	-0.169	-0.094	-0.205	-0.100	-0.196	-0.135	-0.185
	(0.028)	(0.039)	(0.027)	(0.043)	(0.019)	(0.026)	(0.020)	(0.029)
Third long difference	-0.092	-0.173	-0.105	-0.216	-0.123	-0.194	-0.143	-0.193
	(0.033)	(0.045)	(0.031)	(0.048)	(0.021)	(0.030)	(0.022)	(0.033)
Fourth long difference	-0.120	-0.167	-0.111	-0.222	-0.160	-0.188	-0.153	-0.183
	(0.036)	(0.050)	(0.035)	(0.053)	(0.021)	(0.036)	(0.022)	(0.040)

Table 9 Different OLS estimators of the arrest and conviction rate elasticities

Cluster robust standard errors reported in parentheses. Control variables and time-specific effects included. The jth long difference of prbarr_{it} is defined as $prbarr_{it} - prbarr_{i,t-i}$

with respect to the two broad crime categories, as well as, for each one of them, their most dominant sub-category, namely theft and assault.

The results corroborate the findings of Levitt (1998) in that there is little evidence of measurement error and ratio bias in the probability of arrest and the probability of conviction. We took second, third and fourth long differences, and the arrest/conviction rate elasticity changes little across specifications.

Next, we analyse the sensitivity of our results to omitted deterrence variables. Mustard (2003) shows how excluding conviction rates and sentence length from the model leads to omitted variables bias. In particular, due to the negative correlation between these regressors and the probability of arrest, the true effect of arrest rates on crime may be underestimated. Table 10 reports results from specifications including only the probability of arrest as a deterrence variable, which is treated as endogenous.

The pattern of the estimates corroborates the findings of Mustard (2003), i.e. omitting other relevant deterrence variables lowers the arrest rate elasticity considerably in all cases. As expected, the overidentification test statistic suggests that the moment conditions used in GMM estimation are invalidated.²¹ The omitted deterrence variables are serially correlated and also correlated with the regressors (see Table 4).

Imposing (weak) exogeneity of the arrest probability implies that one can also use alternative inference methods. In particular, the standard least squares dummy variables (LSDV) estimator becomes a meaningful choice when T is large enough, and so does the mean group (MG) estimator proposed by Pesaran and Smith (1995), which allows for slope parameter heterogeneity across different LGAs. Here, T = 13is double digit; hence, we might apply such methods with some confidence. In addition to the aforementioned estimators, imposing weak exogeneity of the arrest probability implies that a seemingly unrelated regressions (SUR) estimator could also be a viable alternative, especially in situations where T is thought to be large enough. In the present case, SUR presents an alternative approach for estimating aggregated crime

 $^{^{21}}$ Again, the only exceptions appear to be the robbery and homicide equations, which, however, are weakly identified as discussed previously.

	Property	Violent	Theft	Robbery	Assault	Homicide
Lagged crime rate	0.364	0.390	0.324	0.053	0.409	-0.154
	(0.108)	(0.085)	(0.107)	(0.084)	(0.089)	0.089
Probability of arrest	-0.054	-0.154	-0.036	-0.291	-0.048	-0.252
	(0.065)	(0.101)	(0.066)	(0.092)	(0.114)	(0.125)
Income	-0.427	-0.319	-1.077	-2.186	-0.491	3.381
	(0.234)	(0.199)	(0.403)	(1.249)	(0.362)	(3.921)
Unemployment	-0.152	0.010	-0.161	0.269	0.001	0.163
	(0.079)	(0.050)	(0.820)	(0.221)	(0.057)	(0.361)
p value oir	0.009	0.016	0.007	0.122	0.025	0.683
p value rank test	0.035	0.000	0.007	0.109	0.001	0.752
p value serial correlation						
lag 1	0.000	0.000	0.000	0.000	0.000	0.000
lag 2	0.453	0.753	0.765	0.600	0.987	0.826

Table 10 Excluding conviction and sentencing data

Probability of conviction, imprisonment and average sentence length omitted Probability of arrest is treated as endogenous

Table 11	POLS	estimates	of the	crime	mode
Table 11	POLS	estimates	of the	crime	mode

	Property	Violent	Theft	Robbery	Assault	Homicide
Lagged crime rate	0.922	0.883	0.917	0.557	0.892	0.338
	(0.010)	(0.008)	(0.010)	(0.016)	(0.008)	(0.038)
Probability of arrest	-0.042	-0.058	-0.078	-0.413	-0.059	-0.476
	(0.012)	(0.019)	(0.014)	(0.023)	(0.019)	(0.056)
Probability of conviction	-0.072	-0.201	-0.100	-0.385	-0.183	-0.518
	(0.008)	(0.017)	(0.010)	(0.016)	(0.017)	(0.039)
Probability of imprisonment	-0.010	-0.007	-0.015	0.044	-0.005	-0.091
	(0.005)	(0.006)	(0.006)	(0.024)	(0.007)	(0.067)
Average sentence	0.012	0.004	-0.003	0.020	-0.007	0.014
	(0.005)	(0.005)	(0.007)	(0.021)	(0.013)	(0.013)
Income	-0.090	-0.164	-0.153	-0.772	-0.217	- 1.398
	(0.023)	(0.027)	(0.025)	(0.076)	(0.029)	(0.224)
Unemployment	0.038	0.046	0.026	0.081	0.021	-0.077
	(0.009)	(0.009)	(0.009)	(0.026)	(0.009)	(0.067)

See Table 5

models since it effectively weighs observations according to their (co-)variance (see e.g. Cherry and List 2002).

Tables 11, 12, 13 and 14 report results for the pooled OLS (POLS), LSDV, SUR and MG estimators, respectively. It is obvious that not accounting for region-specific effects (Table 11) leads to severe underestimation of the effect of the judicial system on crime, whereas the autoregressive coefficient is biased upwards, as expected. The estimated

	Property	Violent	Theft	Robbery	Assault	Homicide
Lagged crime rate	0.686	0.402	0.580	0.152	0.417	0.028
	(0.019)	(0.019)	(0.020)	(0.021)	(0.020)	(0.044)
Probability of arrest	-0.043	-0.266	-0.141	-0.451	-0.248	-0.659
	(0.014)	(0.024)	(0.017)	(0.024)	(0.025)	(0.058)
Probability of conviction	-0.088	-0.289	-0.168	-0.449	-0.271	-0.551
	(0.010)	(0.017)	(0.013)	(0.018)	(0.017)	(0.041)
Probability of imprisonment	-0.004	-0.012	-0.021	-0.079	-0.029	-0.286
	(0.005)	(0.006)	(0.007)	(0.020)	(0.007)	(0.066)
Average sentence	-0.004	0.012	-0.018	0.007	-0.001	0.009
	(0.006)	(0.004)	(0.007)	(0.017)	(0.007)	(0.012)
Income	-0.319	0.060	-0.699	-1.079	0.065	-1.564
	(0.033)	(0.036)	(0.039)	(0.087)	(0.043)	(0.321)
Unemployment	0.103	0.006	0.022	0.062	-0.001	0.020
	(0.013)	(0.013)	(0.013)	(0.030)	(0.014)	(0.099)

Table 12 LSDV estimates of the crime model

See Table 5

Table 13 SUR estimates of the crime model

	Property	Violent	Theft	Robbery	Assault	Homicide
Lagged crime rate	0.209	0.197	0.525	0.314	0.503	-0.046
	(0.019)	(0.029)	(0.042)	(0.043)	(0.047)	(0.042)
Probability of arrest	-0.284	-0.378	-0.186	-0.299	-0.212	-0.543
	(0.016)	(0.029)	(0.035)	(0.036)	(0.045)	(0.055)
Probability of conviction	-0.296	-0.394	-0.208	-0.365	-0.276	-0.477
	(0.012)	(0.023)	(0.032)	(0.028)	(0.043)	(0.039)
Probability of imprisonment	-0.014	-0.023	-0.004	-0.020	0.015	-0.214
	(0.007)	(0.016)	(0.012)	(0.031)	(0.014)	(0.062)
Average sentence	-0.000	0.011	-0.007	0002	0.012	0.023
	(0.008)	(0.012)	(0.015)	(0.033)	(0.013)	(0.012)
Income	-0.846	-0.719	-0.466	-0.621	-0.041	0.601
	(0.129)	(0.508)	(0.215)	(0.487)	(0.262)	(1.16)
Unemployment	0.016	0.027	-0.027	-0.073	0.060	0.040
	(0.013)	(0.024)	(0.017)	(0.039)	(0.021)	(0.092)

See Table 5

coefficients obtained from LSDV and MG (Tables 12 and 14) are quite similar to each other and are, overall, plausible in sign and magnitude. Since the pattern of the MG-based estimates is mostly in line with the LSDV estimates, this suggests that the assumption of common parameters across regions is not so restrictive in our sample. The estimated coefficients obtained from SUR for the aggregated crime categories are statistically similar to those obtained from GMM (see Tables 5 and 6), with the

	Property	Violent	Theft	Robbery	Assault	Homicide
Lagged crime rate	0.619	0.269	0.511	0.249	0.198	0.007
	(0.052)	(0.042)	(0.050)	(0.047)	(0.043)	(0.043)
Probability of arrest	-0.108	-0.195	-0.149	-0.379	-0.266	-0.403
	(0.044)	(0.059)	(0.053)	(0.069)	(0.063)	(0.221)
Probability of conviction	-0.245	-0.294	-0.212	-0.441	-0.279	-0.463
	(0.039)	(0.033)	(0.038)	(0.060)	(0.032)	(0.112)
Probability of imprisonment	-0.011	-0.022	-0.015	-0.006	-0.019	-0.175
	(0.014)	(0.016)	(0.018)	(0.039)	(0.015)	(0.278)
Average sentence	0.006	0.012	-0.036	-0.020	-0.015	0.171
	(0.029)	(0.015)	(0.020)	(0.044)	(0.016)	(0.079)
Income	-0.159	0.258	-0.588	-0.797	0.344	-1.972
	(0.074)	(0.108)	(0.103)	(0.199)	(0.107)	(1.838)
Unemployment	0.045	0.070	-0.003	0.044	0.103	0.369
	(0.048)	(0.049)	(0.050)	(0.092)	(0.042)	(1.394)

 Table 14 MG estimates of the crime model

See Table 5

main exception perhaps being the coefficient of the lagged dependent variable for violent crime, which appears to be somewhat smaller. The standard errors for all three estimators suggest much higher precision relative to GMM.

5 Concluding remarks

We estimate an econometric model for crime using a new panel data set containing information on illegal activity and deterrence variables for local government areas in New South Wales, Australia. We take into account various endogeneity concerns expressed previously in the literature. Our findings suggest that the criminal justice system can potentially exert a large impact on crime.

Our results show 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. This implies that imprisoning more criminals, or imprisoning them for longer, is not optimal from a policy perspective, assuming that the cost involved behind these activities is of similar magnitude.

In our analysis, we address the impact of feedback between crime and deterrence, and it controls for measurement error and omitted variables. The resulting dynamic panel data model of crime is estimated by GMM. We show that the detrimental effects of measurement error and ratio bias are largely absent in our data. In general, we do not find overwhelming evidence for endogeneity of arrest rates. Furthermore, we show the necessity of including all relevant deterrence variables from the judiciary system to avoid underestimation of the effect of law enforcement policies. Our conclusion that the deterrent effect of prison is rather limited will be regarded by some as controversial but it is entirely consistent with recent research on prison downsizing strategies that have been implemented over the last few years in the USA. For example, in the first year following the passage of California's Public Safety Realignment Act in 2011 (an Act designed to reduce the California prison population), the State's prison population fell by approximately 27,400, a roughly 17% decline on the population in 2010. With the exception of a slight and transient increase in autotheft, studies have found little evidence that the reduction in prisoner numbers produced an increase in crime (Sundt et al. 2016; Lofstrom and Raphael 2016). The extensive literature on the specific deterrent effect of prison on re-offending also reveals largely negative findings (Nagin et al. 2009).

There are, nonetheless, several issues that remain to be explored. As Nagin (2013) points out, the conclusion that risk of apprehension exerts a more significant effect than punishment severity is largely derived from research examining the risk of apprehension by police. Little has been done to see whether the risk of conviction, given arrest, also has a deterrent effect or on punishment celerity—that is, the effect which sanction speed has on risk of re-offending (Chalfin and McCrary 2017). We know comparatively little about the kinds of police activities (and other forms of intervention) that exert the strongest influence on the perceived risk of apprehension. Finally, much work remains to be done in assessing the costs and benefits of various criminal justice options for reducing crime.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Alvarez J, Arellano M (2003) The time series and cross-sectional asymptotics of dynamic panel data estimators. Econometrica 71:1121–1159
- Arellano M, Bond S (1991) Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. Rev Econ Stud 58:277–298
- Arellano M, Bover O (1995) Another look at the instrumental variable estimation of error-components models. J Econom 68:29–51
- Australian Bureau of Statistics (2006) Statistical geography volume 1 Australian standard geographical classification (ASGC). ABS Catalogue No. 1216.0. https://www.ausstats.abs.gov.au/ausstats/ subscriber.nsf/0/3E15ACB95DA01A65CA2571AA0018369F/\$File/12160_2006.pdf
- Becker GS (1968) Crime and punishment: an economic approach. J Polit Econ 76(2):169–217
- Blumstein A (1974) Seriousness weights in an index of crime. Am Sociol Rev 39:854–864
- Blumstein A, Cohen J, Nagin DS (eds) (1978) Deterrence and incapacitation-estimating the effect of criminal sanctions on crime rates. National Academies Press, Washington
- Blundell R, Bond S (1998) Initial conditions and moment restrictions in dynamic panel data models. J Econom 87(1):115–143
- Bodman PM, Maultby C (1997) Crime, punishment and deterrence in Australia: a further empirical investigation. Int J Soc Econ 24:884–901
- Bowsher CG (2002) On testing overidentifying restrictions in dynamic panel data models. Econ Lett 77:211–220

Bun MJG (2015) Identifying the impact of deterrence on crime: internal versus external instruments. Appl Econ Lett 22(3):204–208

- Bun MJG, Kiviet JF (2006) The effects of dynamic feedbacks on LS and MM estimator accuracy in panel data models. J Econom 132:409–444
- Bun MJG, Sarafidis V (2015) Chapter 3: The Oxford handbook of panel data. In: Baltagi B (ed) Dynamic panel data models. Oxford
- Bun MJG, Sarafidis V (2015) Dynamic panel data models. In: Baltagi B (ed) The Oxford handbook of panel data. Oxford University Press, Oxford, pp 76–110
- Chalfin A, McCrary J (2017) Criminal deterrence: a review of the literature. J Econ Lit 55(1):5-48
- Cherry TL, List JA (2002) Aggregation bias in the economic model of crime. Econ Lett 75:81-86
- Corman H, Mocan HN (2000) A time-series analysis of crime, deterrence, and drug abuse in New York city. Am Econ Rev 90(3):584–604
- Cornwell C, Trumbull WN (1994) Estimating the economic model of crime with panel data. Rev Econ Stat 76(2):360–366
- Dills AK, Miron JA, Summers G (2008) What do economists know about crime? National Bureau of Economic Research Working Paper Series, No. 13759
- Durlauf SN, Nagin DS (2011) The deterrent effect of imprisonment. In: Cook P, Ludwig J, McCrary J (eds) Controlling crime. University of Chicago Press, Chicago
- Ehrlich I (1973) Participation in illegitimate activities: a theoretical and empirical investigation. J Polit Econ 81(3):521–565
- Ehrlich I (1975) The deterrent effect of capital punishment: a question of life and death. Am Econ Rev 65(3):397–417
- Entorf H, Spengler H (2008) Is being 'soft on crime' the solution to rising crime rates? Evidence from Germany. IZA Discussion Paper No. 3710, Germany
- Entorf H, Spengler H (2015) Crime, prosecutors, and the certainty of conviction. Eur J Law Econ 39:167–201 Fajnzylber P, Lederman D, Loayza N (2002) Inequality and violent crime. J Law Econ 45(1):1–39
- Freeman RB (1999) The economics of crime. In: Ashenfelter OC, Card D (eds) Handbook of labor economics 3, Elsevier, Amsterdam, pp 3529–3571
- Griliches Z, Hausman J (1986) Errors in variables in panel data. J Econom 31:93-118
- Hansen LP (1982) Large sample properties of generalized methods of moments estimators. Econometrica 50(4):1029–1054
- Harris RDF, Tzavalis E (1999) Inference for unit roots in dynamic panels where the time dimension is fixed. J Econom 91:201–226
- Heller NB, McEwen JT (1973) Applications of crime seriousness information in police departments. J Crim Justice 1:241–253
- Im KS, Pesaran MH, Shin Y (2003) Testing for unit roots in heterogeneous panels. J Econom 115:53-74
- Johnson R, Raphael S (2012) How much crime reduction does the marginal prisoner buy? J Law Econ 55(2):275–310
- Kleibergen F, Paap R (2006) Generalized reduced rank tests using the singular value decomposition. J Econom 133:97–126
- Klick J, Tabarrok A (2005) Using terror alert levels to estimate the effect of police on crime. J Law Econ 48(1):267–279
- Levitt SD (1996) The effect of prison population size on crime rates: evidence from prison overcrowding litigation. Q J Econ 111(2):319–351
- Levitt SD (1997) Using electoral cycles in police hiring to estimate the effect of police on crime. Am Econ Rev 87(3):270–290
- Levitt SD (1998) Why do increased arrest rates appear to reduce crime: deterrence, incapacitation, or measurement error? Econ Inq 36:353–372
- Levitt SD (2002) Using electoral cycles in police hiring to estimate the effect of police on crime: reply. Am Econ Rev 92(4):1244–1250
- Liedka RV, Piehl AM, Useem B (2006) The crime-control effect of incarceration: does scale matter? Criminol Public Policy 5(2):245–276
- Lofstrom M, Raphael S (2016) Prison downsizing and public safety: evidence from California. Criminol Public Policy 15(2):349–365
- Marvell T, Moody C (1994) Prison population growth and crime reduction. J Quant Criminol 10(2):109-140
- Marvell TB, Moody CE (1996) Specification problems, police levels, and crime rates. Criminology 34(4):609–646

Mustard DB (2003) Reexamining criminal behavior: the importance of omitted variable bias. Soc Sci Q 61(1):23–43

Myers SL Jr (1980) Why are crimes underreported? What is the crime rate? Does it "really" matter? Rev Econ Stat 85(1):205–211

Nagin DS (1978) General deterrence: a review of the empirical evidence. In: Blumstein A, Cohen J, Nagin DS (eds) On the feasibility of identifying the crime function in a simultaneous model of crime rates and sanction levels. National Academies Press, Washington

Nagin DS (2013) Deterrence in the twenty-first century. Crime Justice 42(1):199-263

Nagin DS, Cullen FT, Jonson CL (2009) Imprisonment and reoffending. Crime Justice 38(1):115-200

Pesaran MH, Smith RP (1995) Estimating long-run relationships from dynamic heterogeneous panels. J Econom 68:79–113

Pyle D (1984) Combating crime: do the police give value for money? Public Money 4(1):27-30

Raphael S, Stoll MA (2009) Why are so many Americans in prison? In: Raphael S, Stoll MA (eds) Do prisons make us safer? The benefits and costs of the prison boom. Russell Sage Foundation, New York Roodman D (2009) A note on the theme of too many instruments. Oxf Bull Econ Stat 71:135–158

- Sampson RJ, Cohen J (1988) Deterrent effects of the police on crime: a replication and theoretical extension. Law Soc Rev 22(1):163–189
- Sjoquist DL (1973) Property crime and economic behavior: some empirical results. Am Econ Rev 63(3):439-446
- Spelman W (2000) The limited importance of prison expansion. In: Blumstein A, Wallman J (eds) The crime drop in America. Cambridge University Press, Cambridge, pp 97–129
- Sundt J, Salisbury EJ, Harmon MG (2016) Is downsizing prisons dangerous? The effect of California's realignment act on public safety. Criminol Public Policy 15(2):315–341
- Trumbull WN (1989) Estimations of the economic model of crime using aggregate and individual level data. South Econ J 56(2):423–439
- Vold GB, Bernard TJ, Snipes JB (2002) Theoretical criminology, 5th edn. Oxford University Press, Oxford Walmsley R (2009) World prison population list. Technical report, King's College London
- Weatherburn D (1994) Sentence disparity and its impact on the NSW district criminal court. General report series no. 34. NSW Bureau of Crime Statistics and Research, Sydney. http://www.bocsar.nsw.gov.au/ Documents/r34.pdf
- Weatherburn D (2016) "Rack 'em, pack 'em and stack 'em": decarceration in an age of zero tolerance. Curr Issues Crim Justice 28(2):137–156
- Wilson J, Hernstein R (1985) Crime and human nature. Simon and Schuster, New York
- Windmeijer F (2005) A finite sample correction for the variance of linear efficient two step GMM estimators. J Econom 126:25–51
- Withers G (1984) Crime, punishment and deterrence in Australia: an empirical investigation. Econ Rec 60(2):176–185
- Witt R, Witte A (2000) Crime, prison, and female labor supply. J Quant Criminol 16(1):69-85
- Wolpin KI (1978) An economic analysis of crime and punishment in England and Wales, 1894–1967. J Polit Econ 86(5):815–840
- Wolpin KI (1980) A time series-cross section analysis of international variation in crime and punishment. Rev Econ Stat 62(3):417–423
- Ziliak JP (1997) Efficient estimation with panel data when instruments are predetermined: an empirical comparison of moment-condition estimators. J Bus Econ Stat 15:419–431
- Zimring FE, Hawkins GJ (1973) Deterrence-the legal threat in crime control. University of Chicago Press, Chicago

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Maurice J. G. Bun¹ · Richard Kelaher² · Vasilis Sarafidis³ · Don Weatherburn⁴

- ¹ University of Amsterdam and De Nederlandsche Bank, P.O. Box 15867, 1001 NJ Amsterdam, Netherlands
- ² Discipline of Business Analytics, The University of Sydney Business School, Cnr Abercrombie and Codrington St., Darlington, NSW 2006, Australia
- ³ Department of Econometrics and Business Statistics, Monash University, Caulfield East, VIC 3145, Australia
- ⁴ NSW Bureau of Crime Statistics and Research, GPO Box 31, Sydney, NSW 2001, Australia