THE UNIVERSITY of ADELAIDE

School of Economics

Working Papers

ISSN 2203-6024

The Importance of Punishment Substitutability in Criminometric Studies

Eugene Braslavskiy
School of Economics
University of Adelaide

Firmin Doko Tchatoka
School of Economics
University of Adelaide

Virginie Masson
School of Economics
University of Adelaide

Working Paper No. 2019-2
April 2019

The Importance of Punishment Substitutability in Criminometric Studies

Eugene Braslavskiy[*], Firmin Doko Tchatoka[†], and Virginie Masson[‡]

January 25, 2019

ABSTRACT

This study investigates the role of punishment substitutability in the empirical estimation of the economic model of crime. Using a dynamic panel data model fitted to a panel of Local Government Areas in New South Wales, Australia, we evaluate the effects of financial penalties and imprisonment on the crime rate. Our results show that crime is clearly a dynamic phenomenon, and that failure to incorporate both financial penalties and imprisonment can lead to a misspecified model. Furthermore, our results vary significantly for different crime categories, highlighting the importance of analysing specific crime categories separately.

Keywords: crime, deterrence, punishment, panel data, aggregation bias.

JEL classification: K14, C23, C26, C51.

[*]School of Economics, The University of Adelaide

[†]Corresponding author contacts: School of Economics, The University of Adelaide, 10 Pulteney St, Adelaide SA 5005, AUSTRALIA. Tel:+618 8313 1174, Fax:+618 8223 1460; e-mail: firmin.dokotchatoka@adelaide.edu.au

[‡]School of Economics, The University of Adelaide

I Introduction

Following the seminal papers of Becker (1968) and Ehrlich (1973), a significant body of empirical literature has arisen testing whether there is evidence of a deterrent effect on crime from the criminal justice system. While the deterrent effects of imprisonment and capital punishment have attracted significant attention–see e.g. Dezhbakhsh et al. (2003), Katz et al. (2003), and Donohue and Wolfers (2009) for capital punishment, and Kelaher and Sarafidis (2011), Wan et al. (2012) for imprisonment, these studies usually consider individual punishment types in isolation, thus ignoring the existence of some substitutability between them. Moreover, dealing with individual punishment types in isolation overlooks the significant role played by minor punishments, such as financial penalties; see Gordon and Glaser (1991).

In this paper, we consider a model of crime with substitutability between punishments[1], and provide results for a greater number of crime categories than is commonly considered in the literature. While the inclusion of punishment substitutability allows for interaction between the *criminal justice variables*,[2] the analysis of disaggregated crime data highlights the variation in results across crime categories. The importance of using the most exhaustive data on criminal sanctions was emphasized by Mustard (2003), who argues that omitting any of the criminal justice variables can lead to a misspecified model. This recommendation was adopted by Cornwell and Trumbull (1994), however Cherry and List (2002) showed that their results suffered from aggregation bias, as crime categories were pooled into a single decision model.

Our study focuses on New South Wales (NSW), Australia, where the *Crimes (Sentencing Procedure) Act* 1999 grants courts a broad discretion to impose fines in place of imprisonment, without limiting the nature of offences for which fines may be imposed.[3] In 2013 for example, a total of 39,226 people were fined in NSW for a criminal offence, compared to 9,503 who received a prison sentence [see NSW Bureau of Crime Statistics and Research (2015b)]. Using data constructed from recorded offences, we show that fines and imprisonment exert different effects on different crimes. For instance, fines act as a deterrent for violent crimes, while they do not for property offences. As evidence that substitutability between punishments matters, we show that

[1]Data availability restricts us to imprisonment and fines. Other punishment types such as parole and probation are thus not included in this study.

[2]Criminal justice variables consist of the probabilities of arrest, conviction, imprisonment, and being fined, as well as the average fine and average prison sentence.

[3]In the 1990s and early 2000s there was movement towards developing a system of guideline judgments in some Australian jurisdictions. This approach was 'most enthusiastically' embraced by the courts in New South Wales– see the link [18] at *http://www.loc.gov/law/help/sentencing-guidelines/australia.php#_ftn18*. The first formal guideline judgment in Australia was issued by the New South Wales Court of Criminal Appeals in 1998 in the Jurisic case– see the link [19] at *http://www.loc.gov/law/help/sentencing-guidelines/australia.php#_ftn19*. Following that decision and a second guideline judgment issued by the Court in 1999, the Chief Justice of New South Wales advocated for the use of such judgments in a speech at a national conference of judges.

the results by Kelaher and Sarafidis (2011) and Wan et al. (2012), who consider imprisonment in isolation, are significantly altered when fines are accounted for. In particular, the probability of imprisonment is no longer significant for either property or violent offences.

To the best of our knowledge, Wolpin (1978), Cherry (2001), Cherry and List (2002) and Braslavskiy (2015) are the only studies that consider a broader specification of the model of crime, and the importance of substitutability between imprisonment and fines. While Wolpin (1978) focuses on time series analysis, Cherry (2001) and Cherry and List (2002) use a static panel data framework. By contrast, we use a dynamic panel data model and exploit the system GMM procedure in Arellano and Bond (1991)[4] to account for reverse causality between the crime rate and the criminal justice variables. This problem of reverse causality in empirical studies of the economic model of crime was detected by Ehrlich (1973), and was also given prominence in 1978 by the US National Research Council; see Fisher and Nagin (1978). To address the issue, the two-stage least squares (2SLS) method is often used, with the first notable exception by Fajnzylber and Loayza (2002) who implemented the Generalised Method of Moments (GMM).[5] Infrequent use of the GMM procedure in most empirical studies is due to the fact that static specifications are often retained, so the system GMM estimation is not warranted. We argue that a static specification is inappropriate as changes in law enforcement policies require time to have their full effect, due to habit formation and cost adjustment.[6]

In Australia, Withers (1984) was the first study of the economic model of crime, and his work was later extended by Bodman and Maultby (1997). While novel in terms of applications to Australian data, neither paper specifies the complete economic model of crime as we do. In particular, neither of them emphasise the importance played by the probabilities of arrest and the conditional probability of conviction if arrested, as suggested by Ehrlich (1975).[7] Tait (2001) provides one of the most significant analyses of the effects of non-custodial penalties in Australia, but his study focuses on recidivism. More recently, Moffatt and Poynton (2007) used a methodological approach similar to Tait (2001) in a study of recidivism of driving offenders.

The remainder of this paper is structured as follows. We begin with the empirical specification of the model in Section 2. Results are presented in Section 3. Discussion and concluding remarks are provided in Section 4.

[4] Also, see Arellano and Bover (1995), and Blundell and Bond (1998).

[5] See Hansen (1982).

[6] See Kelaher and Sarafidis (2011).

[7] To mirror the American judiciary system, Ehrlich (1975) also includes the conditional probability of being executed if convicted.

II Empirical Specification

II.1 Model Specification and Identification Strategy

Let $CrimeRate_{it}$ denote the crime rate at time t in Local Government Area (LGA) i.[8] We consider the following specification:

$$\begin{aligned}\ln(CrimeRate_{it}) = \alpha_i + \phi_t + \delta \ln(CrimeRate_{it-1}) + \beta_1 \ln(arrest_{it}) + \beta_2 \ln(conviction_{it}) + \\ \beta_3 \ln(prison_{it}) + \beta_4 \ln(fine_{it}) + \beta_5 \ln(AverageSentence_{it}) + \\ \beta_6 \ln(AverageFine_{it}) + \gamma Controls_{it} + \epsilon_{it},\end{aligned} \tag{1}$$

where $ln(.)$ is the natural logarithm and α_i is the LGA fixed effect. The presence of the lag of the dependent variable in the right-hand side of (1), $CrimeRate_{it-1}$, accounts for the time required for changes in law enforcement policies to take effect. The term ϕ_t denotes time-specific fixed effects that control for any shock which may have affected the crime rate in all LGAs in year t. Following Ehrlich (1975), the right-hand side of (1) also includes the probability of arrest ($arrest_{it}$), the probability of conviction conditional on having been arrested ($conviction_{it}$), the probability of imprisonment conditional on having been convicted ($prison_{it}$), and average prison sentence ($AverageSentence_{it}$). To account for substitutability between imprisonment and fines, we also include the probability of being fined conditional on having been convicted ($fine_{it}$) and average fine ($AverageFine_{it}$).

As standard in the literature, we add some control variables for each LGA and each time period to account for variations in the socio-economic and demographic factors which may have an impact on the crime rate. This includes gender and age, the proportion of Indigenous people in the population,[9] population density,[10] levels of income and education,[11] and the proportion of the population working in particular industries.[12]

[8] The crime rate is considered for specific crime categories as defined in Section II.2.

[9] See Sjoquist (1973), Thaler (1977), Cornwell and Trumbull (1994), Entorf and Spengler (2000), Buonanno and Montolio (2008) for the inclusion of the percentage of ethnic minorities. See Moffatt and Poynton (2007) for the inclusion of ethnic minority in an Australian context.

[10] See e.g. Sjoquist (1973), Thaler (1977), Cornwell and Trumbull (1994) and Moody and Marvell (2010).

[11] See e.g. Sjoquist (1973), Thaler (1977), Lochner and Moretti (2004), Fajnzylber and Loayza (2002), and Moody and Marvell (2010).

[12] Cornwell and Trumbull (1994) disaggregate the income variable into the wage level for different industry categories. However, given data limitations the closest we only include the proportion of the population working in particular industries to control for the socio-economic composition of the population; see e.g. Moody and Marvell (2010).

The inclusion of gender and age in the regressions is justified by the fact that in 2013-14, there were 95,080 male offenders in NSW, compared to only 25,294 female offenders, and the most frequent age brackets for male offenders were 15-19 and 20-24 years of age (Australian Bureau of Statistics (2015a)), also see e.g. Thaler (1977), Cornwell and Trumbull (1994), Entorf and Spengler (2000), and Buonanno and Montolio (2008).

Model (1) is usually referred to as *a dynamic panel or autoregressive panel* model because a lagged dependent variable appears in the right-hand side. If $|\delta| < 1$ and ϵ_{it} are non-autocorrelated for all i, it can be shown using the first-difference of (1) that $\ln CrimeRate_{it-2}$ is correlated with $\Delta \ln CrimeRate_{it-1} = \ln CrimeRate_{it-1} - \ln CrimeRate_{it-2}$ but not with ϵ_{it-1}. Therefore, the second lag (i.e., $\ln CrimeRate_{it-2}$) and all subsequent superior lags are valid instruments for $\Delta \ln CrimeRate_{it-1}$, as long as the residuals do not exhibit second order autocorrelation [see Verbeek (2008, Eq.(10.47)). Also, $\ln CrimeRate_{it-2}$ should be highly correlated with $\Delta \ln CrimeRate_{it-1} = \ln CrimeRate_{it-1} - \ln CrimeRate_{it-2}$ by definition (strong instrument). This suggests one can consistently estimate the parameter of (1) by an instrumental variables approach after first-differencing the model. However, due to the presence of the specific effect α_i in the first-difference model, a simple IV estimator, such as the one proposed by Anderson and Hsiao (1981), suffers from large variances over a wide range of values for δ if T is small and other control variables are added to the model [see Arellano (1989)], as is the case in (1). Arellano and Bond (1991) develop a GMM framework and we use this method to estimate (1). There are two types of GMM in their paper. The first is call 'first-difference GMM' and is based on the first-difference model obtained from (1) upon assuming that all other right-hand side variables are strictly exogenous, i.e., they are uncorrelated with ϵ_{it}. In this case, the first-difference of these strictly exogenous variables act as their own instruments in the first-difference equation, thus in the GMM implementation. However, it is often the case that the strict exogenous assumption breaks down, in which case some control variables are *predetermined*. By predetermined variables, we mean that the current and lagged of these variables are uncorrelated with current error terms, but they can be correlated with past error terms, thus inducing an additional problem if the first-difference GMM is employed. The solution to this issue is the second GMM method in Arellano and Bond (1991). This GMM is called 'system GMM' and is implemented by using all the moment conditions of the first-difference GMM, along with the additional moment conditions that current and lagged of the predetermined variables are uncorrelated with current error terms, so no restriction such as strict exogeneity is imposed. The paper implements the latter GMM method.

As stated above, one of the main advantages of this method is its ability to produce instrumental variables within the system of moment conditions, by exploiting the time series nature of the data. It also requires no distributional assumptions and offers flexibility in the specification of variables as exogenous, endogenous, or predetermined. We specify all the criminal justice variables as predetermined to account for possible reverse causality between them. As identified by Kelaher and Sarafidis (2011), even if there was no reverse causality in the model, the variables would not be exogenous due to ratio bias from the construction of the different probabilities entering in the right hand side of Eq. (1). Among the control variables, income, population density, proportion of young males between 15-24, and education are specified as predetermined, with the rest of the variables being treated as strictly exogenous. The system GMM method also has all the advantages of standard panel data methods, including the controlling of unobserved heterogeneity. This is particularly important in the context of the economic model of crime, as there is likely to be a significant measurement error in the criminal justice variables, as well as in the crime rate, which cannot be controlled for using only cross sectional econometric techniques.

With regards to the appropriate choice of instrumental variables, we follow the approach recommended by Roodman (2009). For the lag of the dependent variable, the second lag and all subsequent lags are valid instruments, as long as the residuals do not exhibit second order autocorrelation. For the predetermined variables, the first lag and all subsequent lags are theoretically valid instruments. For the strictly exogenous variables, these variables act as instruments to themselves, so one moment condition is formed for each. A robust model is then estimated using the Windmeijer (2005) correction to obtain standard errors that are not downward biased in finite samples. For all regressions, the robust Hansen (1982) test statistic is reported, along with the Arellano and Bond (1991) test for autocorrelation of orders 1 and 2.[13]

II.2 Data

Data on recorded offences were obtained from the NSW Bureau of Crime Statistics and Research website,[14] and regional population statistics are published by the Australian Bureau of Statistics (ABS).[15] The data span a period of 13 years between 2001 and 2013, and cover 152 Local Government Areas (LGAs) in NSW. The crime rates for the various crime categories were computed by dividing the number of recorded offences by the population within each LGA.

[13]When conducting the Hansen test, rejection of the null hypothesis indicates that the overidentifying restrictions are not valid and that the instruments do not satisfy the orthogonality condition. However, when there are many weak instruments, the power of the Hansen (1982) test is significantly weakened, and it is less likely that the test will reject the validity of the overidentifying restrictions (Staiger and Stock (1997)). Hence, very large and unrealistic p-values can also indicate a problem with the model specification (a telltale sign is a p-value equal, or close, to 1 − Roodman (2009)). With regards to the autocorrelation test, it is expected that there would be first order autocorrelation in the model, but any autocorrelation of higher order also indicates model misspecification.

[14]NSW Bureau of Crime Statistics and Research (2015b).

[15]Australian Bureau of Statistics (2015b).

We consider various levels of disaggregation of crime categories. We start with all offences aggregated into one crime rate variable, and then consider property and violent offences, which are the most common crimes studied in the literature. Property offences consist of robberies and thefts, fraud and deception offences[16] while violent offences include assaults, homicides, robberies and sexual offences. Following Wan et al. (2012), we counted robbery as both a property and violent offence because it is an acquisitive crime that, by definition, involves actual or threatened violence. Double-counting these offences is unlikely to alter the quality of our results as robbery averaged 8.7 per cent of recorded violent incidents and only 2.35 per cent of recorded property offences over the period. Finally, we consider the components of property and violent offences separately, along with drug offences, public order offences, harassment and threatening behaviour[17] and weapons offences.[18]

The criminal justice variables consist of the probabilities of arrest, conviction, imprisonment, and being fined, as well as the average fine and average prison sentence. Following Cornwell and Trumbull (1994), we construct proxies of the probabilities of arrest, conviction, imprisonment, and being fined based on the observed data on arrests, convictions, and penalties.

The probability of arrest is proxied by the ratio of arrests to recorded offences within each LGA. An arrest is defined as an incident where a person of interest was proceeded against by NSW Police in relation to a criminal incident. These data are categorised by the most serious offence of the person of interest, in the month during which the proceeding began. Two alternative determinations of the appropriate LGA for each arrest were available: the first based on the residence of the offender, and the second based on the location of the criminal incident. To be consistent with the remaining criminal justice variables, we use the measure based on the residence of the offender.[19] The probability of conviction is proxied by the ratio of convictions to arrests within each LGA. To be recorded as a conviction, a guilty finding must be the result of the criminal proceeding in question. Where the conviction is for multiple offences, it is recorded under the principal offence of the offender, which is defined to be the most serious offence they committed. The LGA was determined solely based on the residence of the offender at the time of the court case, as it was not possible to match convictions to the location of the criminal incident based on the available data. The probability of imprisonment is proxied by the ratio of persons imprisoned to persons convicted within each LGA. Once again, this is based on the residence of the offender at the time of the court case. The probability

[16]For simplicity, we refer to theft, fraud and deception offences, as theft for the remainder of the paper.

[17]For clarity purposes, we refer to the category of harassment and threatening behaviour as harassment for the remainder of the paper.

[18]Homicides and robberies contain a large proportion of zeros, and are thus omitted from the analysis of disaggregated results.

[19]The measure based on the location of the incident is used as a robustness check following estimation; see Section A.3.3 of the Appendix.

of being fined is similarly proxied by the ratio of persons fined to persons convicted.

The two measures of severity of punishment used are the average sentence length for those receiving a sentence of imprisonment, and the average fine received for those sentenced to a fine. These measures were obtained by dividing the sum of the lengths of all prison sentences by the number of offenders imprisoned, and by dividing the total value of all fines received by the number of offenders receiving a fine. The average fine is in nominal dollars and has not been adjusted for inflation. We think that controlling for year fixed effects captures the impacts of inflation.[20]

By construction, the probability proxies are not all between 0 and 1, and thus cannot be interpreted as true probabilities, as discussed by Mustard (2003). A proxy can also sometimes be undefined due to a null denominator. In these cases, the observation is treated as missing. It has been suggested in the literature that the probabilities should incorporate the time lags between the commission of a crime and the subsequent arrest, between the arrest and conviction, and between the conviction and sentencing outcome respectively; see e.g. Fisher and Nagin (1978), and Durlauf and Rivers (2010). Since most criminal incidents are cleared well within a year in NSW,[21] we think that a dynamic panel specification (inclusion of the lag of the dependent variable as a regressor in the model) can alleviate this issue.

We use other control variables[22] constructed from the past three Australian Censuses in 2001, 2006, and 2011.[23] As a measure of income, we use the median household income, which covers all occupied private dwellings within a LGA. It is not adjusted for inflation, to be consistent with the available data on average fines. However, as it is unlikely that household income remains constant between censuses, the missing observations are approximated by assuming a constant annual growth rate between the known observations. A measure of population density is obtained by dividing the population of a LGA[24] by the geographic area (sq kms) of the LGA from the 2011 census. The proportion of males living in a LGA, obtained by dividing the total number of males by the total population in each census year, is specifically controlled for. Furthermore, a separate variable for the proportion of young males between the ages of 15 and 24 was constructed in a similar way. We also control for the proportion of individuals identifying as Indigenous, obtained by dividing the total number of Indigenous people by the total population in each LGA. As a measure of the level of education, we consider the proportion of people with a non-school qualification in each LGA. To control for the socio-economic status of a LGA, we consider the proportions of individuals working in two industries,

[20] As a robustness check, all regressions were also run with a correction for inflation. The results obtained were unchanged from those without the correction; see Tables 32−34 in the Appendix.

[21] NSW Bureau of Crime Statistics and Research (2015a).

[22] Summary statistics for these control variables are presented in Table 19 in the Appendix.

[23] Australian Bureau of Statistics (2011).

[24] Australian Bureau of Statistics (2015b).

namely manufacturing and construction.[25] These typical blue-collar industries are considered the most likely to be valid indicators of the socio-economic status of a LGA; see e.g. Moody and Marvell (2010).

III Results

The results of the estimations include both the short run and long run estimates of the coefficients[26] of the criminal justice variables. We first consider all offences, property and violent offences. We then look at thefts, assaults, and sexual offences, along with drug offences, public order offences, harassment and weapons offences, separately.

III.1 Persistence of Crime

Table 1 presents the estimated coefficients on the lag of the dependent variable and the corresponding adjustment speed of the crime rate.[27] The adjustment speed of the overall crime rate is less than 0.7. It is even lower than 0.5 when considering either property or violent offences. With the exception of weapons offences, where the adjustment speed of the crime rate is very close to one, the adjustment speed is always below 0.7 (drug offences, public order offences and harassment) or 0.5 (thefts and sexual offences). These results are encouraging from a policy perspective, as lower speeds of adjustment suggest that criminals may be more responsive to changes of the judiciary system. It is then possible to influence the crime rate through the criminal justice variables. Hence, incorporating a dynamic process is crucial to understand the factors that determine the crime rate in both the short and long run.

III.2 Aggregated vs. Disaggregated Crime Categories

When looking at the criminal justice variables, the deterrent effects vary whether we consider aggregated or disaggregated crime categories. Table 2 reports the results when all offences are aggregated into one crime rate, and for property offences and violent offences. Firstly, we observe that the estimated effects in long

[25] With the relatively limited sample in our possession, the inclusion of all industries results in instrument proliferation, thus biasing the estimates.

[26] Estimated long run coefficients are given by $\frac{\widehat{\beta}_j}{1-\widehat{\delta}}$, where $\widehat{\beta}_j$ is the estimated short run coefficient and $\widehat{\delta}$ is the estimated coefficient on the lag of the dependent variable. Their standard errors are given by $\sqrt{\frac{\widehat{\beta}_j}{(1-\widehat{\delta})^2}var(\widehat{\beta}_j) + \frac{\widehat{\beta}_j^2}{(1-\widehat{\delta})^4}var(1-\widehat{\delta}) + \frac{2\widehat{\beta}_j}{(1-\widehat{\delta})^3}cov(\widehat{\beta}_j, 1-\widehat{\delta})}$ (Kelaher and Sarafidis (2011), and Wan et al. (2012).

[27] The adjustment speed is approximated by $1-\widehat{\delta}$, where $\widehat{\delta}$ is the estimated coefficient on the lag of the dependent variable.

Table 1: Persistence and adjustment of the crime rate

	All Offences	Property Offences	Violent Offences
(Obs.)	(N=1989)	(N=1989)	(N=1989)
$\ln(\text{CrimeRate})_{-1}$	0.38 (0.07)***	0.57 (0.08)***	0.51 (0.09)***
Adjustment	0.61	0.42	0.48

	Assault	Drug offences	Public order offences	Theft
(Obs.)	(N=1989)	(N=1974)	(N=1989)	(N=1989)
$\ln(\text{CrimeRate})_{-1}$	0.58 (0.09)***	0.33 (0.07)***	0.36 (0.10)***	0.55 (0.08)***
Adjustment	0.41	0.66	0.64	0.45

	Harassment	Weapons Offences	Sexual Offences
(Obs.)	(N=1986)	(N=1981)	(N=1980)
$\ln(\text{CrimeRate})_{-1}$	0.30 (0.12)***	0.07 (0.07)***	0.55 (0.12)***
Adjustment	0.69	0.92	0.44

The variable presented in this table is the natural logarithm of the lagged crime rate, $\ln(\text{CrimeRate})_{-1}$. Adjustment is computed by subtracting the estimated coefficient of the lag variable to one. Significance levels are reported as follows: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Newey-West heteroskedasticity and autocorrelation consistent standard error estimates are presented in $(\cdot)$.

run and short run differ for most of the criminal justice variables. More precisely, the long run estimates of the coefficients are higher (in absolute terms) than the corresponding short run ones. This is not surprising as the estimated persistence $\hat{\delta}_j$ is positive and less than one in all cases, and the long run estimates in each case is given by $\hat{\beta}_j/(1-\hat{\delta}_j)$, where $\hat{\beta}_j$ is the estimated short run coefficient. Also, the estimated coefficients on the probability of arrest and the probability of conviction are negative. This is consistent with theory, and also corroborates the results of previous studies. Strikingly, when property and violent offences are considered separately, the inclusion of the probability of being fined becomes significant. However, a higher the probability of being fined encourages crime on property offences. This means that a higher probability of being fined is seen as reflecting a lenient criminal justice system for property offences. There can then be an adverse effect from raising the probability of being fined if this simply spurs on further crime to be committed. Furthermore, the average fine deters criminals from committing crimes, except for property offences. Interestingly, the probability of imprisonment and severity of custodial punishments never appears to have any effect on the crime rate. In particular, the probability of being fined is more effective than the probability of imprisonment at reducing violent crimes, although higher fine amounts dampen the deterrent effect.

Tables 3 and 4 provide the results when the crimes are disaggregated into more specific crime categories, of which there are seven in total. When analysing disaggregated crime categories, we only considered one type of punishment for harassment and weapons offences (fines only), and for sexual offences (imprisonment only), as shown in Table 4. This is due to missing data as very few people received a particular type of punishment. For all seven crime categories, the estimated coefficients on the probability of arrest and the probability of conviction are significant and negative. The probability of imprisonment has a deterrent effect only for drug offences and sexual offences. It also deters theft but not in the long run. The probability of being fined has a significant effect on harassment, weapons offences, drug offences, and theft. However, for theft and drug offences, a higher probability of being fined increases the crime rate. This result is rather intuitive as the monetary rewards expected from drug offences or theft can compensate for the fine received, while being imprisoned clearly interferes with the criminals' business model. Assaults and public order offences are not affected by either the probability of imprisonment or the probability of being fined. With regards to punishment severity, only the average prison sentence for public order offences is significant. Neither the average prison sentence nor the average fine is significant for any of the other offences, which is consistent with the theory that criminals react more to the likelihood of punishment than its severity.

Table 2: All offences, property offences and violent offences

(Obs.)	All offences (N=1989)		Property offences (N=1989)		Violent offences (N=1989)	
	Short run	Long run	Short run	Long run	Short run	Long run
$\ln(\text{CrimeRate})_{-1}$	0.38		0.57		0.51	
	(0.07)***		(0.08)***		(0.09)***	
Probability variables						
ln(arrest)	-0.56	-0.90	-0.30	-0.72	-0.33	-0.67
	(0.09)***	(0.09)***	(0.04)***	(0.18)***	(0.08)***	(0.25)***
ln(conviction)	-0.10	-0.16	-0.24	-0.58	-0.38	-0.78
	(0.05)*	(0.09)*	(0.03)***	(0.15)***	(0.05)***	(0.22)***
ln(prison)	0.03	0.05	-0.02	-0.04	-0.02	-0.05
	(0.02)	(0.04)	(0.01)	(0.04)	(0.01)*	(0.03)
ln(fine)	0.07	0.12	0.04	0.10	-0.04	-0.09
	(0.05)	(0.08)	(0.01)**	(0.04)**	(0.01)***	(0.04)**
Average level variables						
ln(AverageSentence)	0.02	0.04	-0.00	-0.00	0.00	0.00
	(0.02)	(0.04)	(0.01)	(0.03)	(0.01)	(0.02)
ln(AverageFine)	-0.09	-0.15	-0.00	-0.01	0.05	0.11
	(0.03)***	(0.05)***	(0.01)	(0.04)	(0.02)*	(0.06)*
Other controls	Yes		Yes		Yes	
Year dummies	Yes		Yes		Yes	
Specification tests	P-value		P-value		P-value	
Hansen test	0.15		0.16		0.21	
AR(1)	0.00		0.00		0.00	
AR(2)	0.30		0.51		0.74	

The variables presented in this table are: the natural logarithm of the lagged crime rate, $\ln(\text{CrimeRate})_{-1}$; the natural logarithm of the probability of arrest, ln(arrest); the natural logarithm of the probability of conviction conditional on having been arrested, ln(conviction); the probability of imprisonment conditional on having been convicted, ln(prison); the natural logarithm of the probability of being fined conditional on having been convicted, ln(fine); the natural logarithm of the average prison sentence, ln(AverageSentence); and the natural logarithm of the average fine, ln(AverageFine). Other controls indicate whether variables constructed from past Australian Censuses were included. Significance levels are reported as follows: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Newey-West heteroskedasticity and autocorrelation consistent standard error estimates are presented in $(\cdot)$.

Table 3: Disaggregated offence categories

(Obs.)	Assault (N=1989)		Drug offences (N=1974)		Public order offences (N=1989)		Theft (N=1989)	
	Short run	Long run	Short run	Long run	Short run	Long run	Short run	Long run
$\ln(\text{CrimeRate})_{-1}$	0.58		0.33		0.36		0.55	
	(0.09)***		(0.07)***		(0.11)***		(0.08)***	
Probability variables								
ln(arrest)	-0.31	-0.76	-0.99	-1.48	-0.86	-1.35	-0.30	-0.66
	(0.08)***	(0.28)***	(0.10)***	(0.24)***	(0.10)***	(0.29)***	(0.04)***	(0.14)***
ln(conviction)	-0.38	-0.93	-0.62	-0.93	-0.35	-0.55	-0.23	-0.52
	(0.04)***	(0.26)***	(0.08)***	(0.19)***	(0.07)***	(0.17)***	(0.03)***	(0.12)***
ln(prison)	0.00	0.01	-0.06	-0.10	-0.02	-0.04	-0.02	-0.05
	(0.01)	(0.03)	(0.02)***	(0.03)***	(0.01)	(0.03)	(0.01)*	(0.03)
ln(fine)	-0.02	-0.05	0.23	0.35	0.01	0.01	0.03	0.07
	(0.01)	(0.04)	(0.09)**	(0.15)**	(0.06)	(0.09)	(0.01)*	(0.04)*
Average level variables								
ln(AverageSentence)	-0.01	-0.03	-0.00	-0.00	-0.05	-0.08	0.00	0.00
	(0.01)	(0.03)	(0.02)	(0.03)	(0.02)**	(0.04)**	(0.01)	(0.02)
ln(AverageFine)	0.02	0.05	-0.11	-0.17	-0.06	-0.09	-0.00	-0.01
	(0.02)	(0.05)	(0.08)	(0.12)	(0.05)	(0.09)	(0.01)	(0.04)
Other controls	Yes		Yes		Yes		Yes	
Year dummies	Yes		Yes		Yes		Yes	
Specification tests	P-value		P-value		P-value		P-value	
Hansen test	0.38		0.59		0.85		0.31	
AR(1)	0.00		0.00		0.00		0.00	
AR(2)	0.42		0.27		0.13		0.72	

The variables presented in this table are: the natural logarithm of the lagged crime rate, $\ln(\text{CrimeRate})_{-1}$; the natural logarithm of the probability of arrest, ln(arrest); the natural logarithm of the probability of conviction conditional on having been arrested, ln(conviction); the probability of imprisonment conditional on having been convicted, ln(prison); the natural logarithm of the probability of being fined conditional on having been convicted, ln(fine); the natural logarithm of the average prison sentence, ln(AverageSentence); and the natural logarithm of the average fine, ln(AverageFine). Other controls indicate whether variables constructed from past Australian Censuses were included. Significance levels are reported as follows: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Newey-West heteroskedasticity and autocorrelation consistent standard error estimates are presented in $(\cdot)$.

Table 4: Effects on harassment, weapons offences, and sexual offences: only one type of punishment included

	Harassment		Weapons offences		Sexual offences	
(Obs.)	(N=1986)		(N=1981)		(N=1980)	
	Short run	Long run	Short run	Long run	Short run	Long run
$\ln(\text{CrimeRate})_{-1}$	0.30		0.07		0.55	
	(0.12)**		(0.07)		(0.12)***	
Probability variables						
ln(arrest)	-0.15	-0.21	-0.27	-0.29	-0.25	-0.56
	(0.04)***	(0.07)***	(0.05)***	(0.06)***	(0.06)***	(0.20)***
ln(conviction)	-0.12	-0.18	-0.24	-0.26	-0.25	-0.55
	(0.03)***	(0.06)***	(0.03)***	(0.05)***	(0.04)***	(0.18)***
ln(prison)					-0.08	-0.18
					(0.02)***	(0.09)*
ln(fine)	-0.06	-0.08	-0.09	-0.10		
	(0.02)***	(0.03)**	(0.03)***	(0.04)***		
Average level variables						
ln(AverageSentence)					-0.00	-0.01
					(0.01)	(0.04)
ln(AverageFine)	0.01	0.02	0.01	0.01		
	(0.01)	(0.02)	(0.02)	(0.02)		
Other controls	Yes		Yes		Yes	
Year dummies	Yes		Yes		Yes	
Specification tests	P-value		P-value		P-value	
Hansen test	0.41		0.11		0.37	
AR(1)	0.02		0.00		0.00	
AR(2)	0.17		0.41		0.49	

The variables presented in this table are: the natural logarithm of the lagged crime rate, $\ln(\text{CrimeRate})_{-1}$; the natural logarithm of the probability of arrest, ln(arrest); the natural logarithm of the probability of conviction conditional on having been arrested, ln(conviction); the probability of imprisonment conditional on having been convicted, ln(prison); the natural logarithm of the probability of being fined conditional on having been convicted, ln(fine); and the natural logarithm of the average prison sentence, ln(AverageSentence); and the natural logarithm of the average fine, ln(AverageFine). Other controls indicate whether variables constructed from past Australian Censuses were included. Significance levels are reported as follows: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Newey-West heteroskedasticity and autocorrelation consistent standard error estimates are presented in $(\cdot)$.

III.3 The Importance of Fines

To allow for comparisons, a model similar to those considered by Kelaher and Sarafidis (2011) and Wan et al. (2012) was estimated. To do so, we removed the probability of being fined, the average fine, and all control variables with the exception of income. The model is estimated for the aggregated offence categories, as in Kelaher and Sarafidis (2011) and Wan et al. (2012). The results of the estimation are presented in Table 5.

Including fines into the model significantly alters the results. When fines are included, the probability of imprisonment is no longer significant for property offences. In addition, we observe that fines do have a significant effect on both property and violent crimes, whether this be in the form of a deterrent effect, as for violent offences, or the opposite effect, as for property offences. Finally, the adjustments of 0.75 for the overall aggregated crime rate, 0.54 for property offences and 0.57 for violent offences are consistently higher than the ones we obtained when fines are included (0.61, 0.42 and 0.48 respectively). Hence, including fines decreases the persistence of the crime rate. This is quite intuitive as the role played by fines in determining the crime rate is significant.

III.4 Robustness Checks

The details of all regressions run in this section are presented in Section A.3 of the Appendix. The results confirm that aggregating all crime categories into one crime rate is not appropriate, as the results are sensitive to changes in the control variables (Appendix Table 35). Results for property offences are robust to variations in control variables (Appendix Table 36), however the probability of being fined is only significant at the 10% level when the alternative measure of arrests[28] is used in the construction of the variables (Appendix Table 45). For violent crimes, the results are robust to changing the arrest measure (Appendix Table 45), but the significance of fines is sensitive to the removal of control variables (Appendix Table 37).

Results for assaults, drug offences, theft, and harassment are robust to the removal of various combinations of the control variables, as well as to the use of the alternative arrest measure; see Appendix Tables 38-39, 41-42, and 46-47 . Results for public order offences are robust to the use of the alternative measure of arrest (Appendix Table 46), however the average sentence is sensitive to changes in the control variables (Appendix Table 40). The results for weapons offences are robust to both the alternative arrest measure and changes in the control variables (see Appendix Tables 47 and 43), although the Hansen test statistic sometimes falls too close to rejection of the null, which raises concerns regarding the validity of the specifi-

[28] The alternative measure of arrests is based on the location of the incident rather than on the residence of the offender.

Table 5: All offences, property offences and violent offences without fines

(Obs.)	All offences (N=1989)		Property offences (N=1989)		Violent offences (N=1989)	
	Short run	Long run	Short run	Long run	Short run	Long run
$\ln(\text{CrimeRate})_{-1}$	0.24		0.45		0.42	
	(0.08)***		(0.13)***		(0.11)***	
Probability variables						
ln(arrest)	-0.59	-0.79	-0.23	-0.43	-0.38	-0.67
	(0.09)***	(0.08)***	(0.09)**	(0.13)***	(0.06)***	(0.19)***
ln(conviction)	-0.06	-0.08	-0.16	-0.31	-0.38	-0.65
	(0.07)	(0.10)	(0.08)**	(0.13)**	(0.05)***	(0.19)***
ln(prison)	-0.00	-0.00	-0.12	-0.22	-0.02	-0.03
	(0.02)	(0.03)	(0.04)**	(0.06)***	(0.01)	(0.02)
Average level variables						
ln(AverageSentence)	-0.00	-0.01	-0.03	-0.07	-0.00	-0.00
	(0.03)	(0.03)	(0.03)	(0.06)	(0.01)	(0.02)
Other controls	Income only		Income only		Income only	
Year dummies	Yes		Yes		Yes	
Specification tests	P-value		P-value		P-value	
Hansen test	0.04		0.20		0.15	
AR(1)	0.00		0.00		0.00	
AR(2)	0.02		0.67		0.59	

The variables presented in this table are: the natural logarithm of the lagged crime rate, $\ln(\text{CrimeRate})_{-1}$; the natural logarithm of the probability of arrest, ln(arrest); the natural logarithm of the probability of conviction conditional on having been arrested, ln(conviction); the probability of imprisonment conditional on having been convicted, ln(prison); and the natural logarithm of the average prison sentence, ln(AverageSentence). Other controls indicate whether variables constructed from past Australian Censuses were included. Significance levels are reported as follows: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Newey-West heteroskedasticity and autocorrelation consistent standard error estimates are presented in $(\cdot)$.

cation when certain control variables are removed. Regarding sexual offences, most of the results are robust to control variable changes, although the significance of the lagged crime rate is sensitive to the choice of control variables (Appendix Table 44). The significance of the probability of arrests is also sensitive to the alternative measure of arrests, where it is only significant at the 10% level (Appendix Table 47).

As a further robustness check, we also used the difference GMM method to estimate the model.[29] As expected, this method does not perform well. We get the unexpected result that the lag of the dependent variable is not significant for many of the crime categories. Also, the probability of arrest is sometimes not significant, which is inconsistent with both the theoretical model and most previous results.

IV Discussion and Concluding Remarks

The results of this study emphasise the importance of disaggregation, and shed more light on the mentality of different types of offenders with respect to how they perceive different types of punishments. It clearly appears that higher probabilities of arrest and conviction have a deterrent effect, which is consistent across all the crime categories. However, different types of criminals seem to respond differently to the actual punishment they receive, thus emphasising the importance of including fines into the model. Perhaps the most striking result is for drug offences, where both the probabilities of being imprisoned and being fined are statistically significant, and yet the coefficients have opposite signs, suggesting that drug offenders do not treat fines as being a serious enough punishment to be deterred by them. This might be explained by the possibility that drug offenders simply treat fines as a price to pay for an otherwise desirable action, rather than as retribution for their criminality. Hence, receiving a fine may actually encourage more crimes to be committed (Gneezy and Rustichini (2000), and Kurz and Fonseca (2014)). The positive sign for property offences could be explained in a similar way.

While fines may have a deterrent effect on violent crime, and in particular on harassment and weapons offences, imprisonment seems to have a deterrent effect on drug crime, and sexual offences. However, some crimes such as assault and public order offences do not appear to respond to the probabilities of being imprisoned or fined at all. Interestingly, an experimental study by McFatter (1982) showed that subjects were more inclined to value certain levels of fines as being more suitable to accomplish the goal of deterrence for certain offences, rather than long prison sentences. Other studies have shown that both prison inmates and members of the general public rank alternative punishments, such as sufficiently high levels of fines or intensive probation, as being harsher than spending time in jail (see eg. Erickson and Gibbs (1979), McClelland and G.P. (1985), Petersilia and Deschenes (1994), and Spelman (1995)).

[29]See the results in Tables 48-50 in the Appendix.

The severity of the punishment, however, does not appear to deter crime. Indeed, the severity of prison sentences only appears to have a significant effect on public order offences. Although the expected severity of punishment is a theoretically relevant variable, some studies show that criminals pay less regard to the severity than they do to the probability of being punished. On the one hand, people may simply have inaccurate perceptions of the true level of punishment. For example, in a relatively recent study in the US, Kleck et al. (2005) show that there is generally no association between perceived and actual punishment levels. Kleck and Barnes (2008) show that this is true not only at the individual level, but also at the aggregate level. In particular, their results indicate that even average perceptions of punishment severity can actually be far from the reality. Therefore, changes to the severity of punishment may well have no significant deterrent effect if people's perception of the severity remains unaltered. In addition, the severity of punishment is unlikely to deter because the benefits of criminal offending are immediate, while the cost from punishment is imposed later in the future. Studies which have analysed the perception of offenders towards various levels of punishment severity tend to suggest that offenders may have quite high discount rates, and so would not be significantly deterred from an increase in punishment severity (see e.g. Spelman (1995)). Such discounting, also known as 'present bias', could also explain why people are more likely to react to the immediate probability of being caught, more so than the probability of actually being punished. Furthermore, other explanations such as the availability heuristic (Jolls et al. (1998), and Kahneman and Tversky (2000)), and the hot-cool theory of decision making (Van Gelder (2013)), provide additional insight into why we see a greater deterrent effect from the more initial stages of the criminal justice process. These theories would suggest that potential offenders would react more to things that are at the forefront of their mind at the time they commit the act, without fully considering the consequences of their actions, and so the remote threat of punishment is not something that enters into the decision making process.

In terms of appropriate policy design, this study confirms the previous result of Kelaher and Sarafidis (2011), and also Wan et al. (2012), that a greater emphasis should be placed on increasing the probabilities of arrest and conviction in order to fight crime. As a greater probability of being imprisoned or fined does not appear to deter all crimes, it is our suggestion that criminal justice policies cannot take a 'one-size-fits-all' approach. Appropriate deterrence strategies need to be designed for each offence category individually and using imprisonment as the main policy instrument should be deemed inadequate.

References

Anderson, T. W. and C. Hsiao (1981). Estimation of dynamic models with error components. *Journal of the American statistical Association 76*(375), 598–606.

Arellano, M. (1989). A note on the anderson-hsiao estimator for panel data. *Economics Letters 31*(4), 337–341.

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

Australian Bureau of Statistics (2011). 2011 census community profiles (cat. no. 2003.0). `http://www.abs.gov.au/websitedbs/censushome.nsf/4a256353001af3ed4b2562bb00121564/communityprofiles`.

Australian Bureau of Statistics (2015a). Recorded crime – offenders, 2013-14 (cat. no. 3218.0). `http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/3218.02013-14?OpenDocument`.

Australian Bureau of Statistics (2015b). Regional population growth, australia (cat. no. 4519.0). `http://www.abs.gov.au/AUSSTATS/abs@.nsf/allprimarymainfeatures/DA308C67766C3735CA257751001BD477?opendocument`.

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

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

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

Braslavskiy, E. A. (2015). The importance of punishment substitutability in criminometric studies. The University of Adelaide, Australia. `https://digital.library.adelaide.edu.au/dspace/bitstream/2440/108498/2/02WholeECONHon.pdf`.

Buonanno, P. and D. Montolio (2008). Identifying the socio-economic and demographic determinants of crime across spanish provinces. *International Review of Law and Economics 28*, 89–97.

Cherry, T. (2001). Financial penalties as an alternative criminal sanction: Evidence from panel data. *Atlantic Economic Journal 29*(4), 450–458.

Cherry, T. and J. List (2002). Aggregation bias in the economic model of crime. *Economics Letters 75*, 81–86.

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

Dezhbakhsh, H., P. Rubin, and J. Shepherd (2003, August). Does capital punishment have a deterrent effect? new evidence from postmoratorium panel data. *American Law and Economics Review 5*(2), 344–376.

Donohue, J. and J. Wolfers (2009). Estimating the impact of the death penalty on murder. *American Law and Economics Review 11*(2), 249–309.

Durlauf, S.N. Navarro, S. and D. Rivers (2010). Understanding aggregate crime regressions. *Journal of Econometrics 158*, 306–317.

Ehrlich, I. (1973). Participation in illegitimate activities: A theoretical and empirical investigation. *The Journal of Political Economy 81*(3), 521–565.

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

Entorf, H. and H. Spengler (2000). Socioeconomic and demographic factors of crime in germany: Evidence from panel data of the german states. *International Review of Law and Economics 20*, 75–106.

Erickson, M. and J. Gibbs (1979). On the perceived severity of legal penalties. *The Journal of Criminal Law and Criminology 70*(1), 102–116.

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

Fisher, F. and D. Nagin (1978). *Deterrence and Incapacitation: Estimating the Effects of Criminal Sanctions on Crime Rates*, Chapter On the feasibility of identifying the crime function in a simultaneous model of crime rates and sanction levels, pp. 361–399. National Academy of Sciences, Washington DC.

Gneezy, U. and A. Rustichini (2000). A fine is a price. *Journal of Legal Studies 29*, 1–17.

Gordon, M. and D. Glaser (1991). The use and effects of financial penalties in municipal courts. *Criminology 29*(4), 651–676.

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

Jolls, C., C. Sunstein, and R. Thaler (1998). A behavioral approach to law and economics. *Stanford Law Review 50*(5), 1471–1550.

Kahneman, D. and A. Tversky (2000). *Choice, Values, Frames.* Cambridge University Press, London.

Katz, L., S. Levitt, and E. Shustorovic (2003). Prison conditions, capital punishment and deterrence. *American Law and Economics Review 5*, 318–343.

Kelaher, R. and V. Sarafidis (2011). Crime and punishment revisited. MPRA Paper, University Library of Munich, Germany.

Kleck, G. and J. Barnes (2008). Deterrence and macro-level perceptions of punishment risks: Is there a "collective wisdom"? *Crime and Delinquency 59*(7), 1006–1035.

Kleck, G., B. Sever, S. Li, and M. Gertz (2005). The missing link in general deterrence research. *Criminology 43*(3), 623–660.

Kurz, T. Thomas, W. and M. Fonseca (2014). A fine is a more effective financial deterrent when framed retributively and extracted publicly. *Journal of Experimental Social Psychology 54*, 170–177.

Lochner, L. and E. Moretti (2004). The effect of education on crime: Evidence from prison inmates, arrests, and self-reports. *The American Economic Review 84*(1), 155–189.

McClelland, K. and A. G.P. (1985). Factor analysis applied to magnitude estimates of punishment seriousness: Patterns of individual differences. *Journal of Quantitative Criminology 1*(3), 307–318.

McFatter, R. (1982). Purposes of punishment: Effects of utilities of criminal sanctions on perceived appropriateness. *Journal of Applied Psychology 67*(3), 255–267.

Moffatt, S. and S. Poynton (2007). The deterrent effect of higher fines on recidivism: Driving offences. *NSW Bureau of Crime Statistics and Research: Crime and Justice Bulletin 106*, 1–15.

Moody, C. and T. Marvell (2010). On the choice of control variables in the crime equation. *Oxford Bulletin of Economics and Statistics 72*(5), 696–715.

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

NSW Bureau of Crime Statistics and Research (2015a). NSW criminal court statistics. `http://www.bocsar.nsw.gov.au/Pages/bocsar_media_releases/2014/mr_ccs_2013.aspx`.

NSW Bureau of Crime Statistics and Research (2015b). NSW recorded crime statistics. `http://www.bocsar.nsw.gov.au/Pages/bocsar_crime_stats/bocsar_detailedspreadsheets.aspx`.

Petersilia, J. and E. Deschenes (1994). What punishes? inmates rank the severity of prison vs. intermediate sanctions. *Federal Probation 58*(1), 3–8.

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

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

Spelman, W. (1995). The severity of intermediate sanctions. *Journal of Research in Crime and Delinquency 32*(2), 107–135.

Staiger, D. and J. Stock (1997). Instrumental variables regression with weak instruments. *Econometrica 65*, 557–586.

Tait, D. (2001). The effectiveness of criminal sanctions: A natural experiment. *report to the Criminology Research Council 33/96-7*, 1—61.

Thaler, R. (1977). An econometric analysis of property crime: Interaction between police and criminals. *Journal of Public Economics 8*, 37–51.

Verbeek, M. (2008). *A guide to modern econometrics.* John Wiley & Sons.

Wan, W., C. Moffatt, S. Jones, and D. Weatherburn (2012). The effect of arrests and imprisonment on crime. *NSW Bureau of Crime Statistics and Research: Crime and Justice Bulletin 158*, 1–18.

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.

Wolpin, K. (1978). An economic analysis of crime and punishment in england and wales, 1894−1967. *Journal of Political Economy 86*(51), 815–840.